2010

The interactions between sentence complexity, working memory, and additional working memory load: an on-line measure

Christy Seidel
Louisiana State University and Agricultural and Mechanical College, cseide1@tigers.lsu.edu

Recommended Citation

Seidel, Christy, "The interactions between sentence complexity, working memory, and additional working memory load: an on-line measure" (2010). *LSU Master's Theses*. 2626.
https://digitalcommons.lsu.edu/gradschool_theses/2626

THE INTERACTIONS BETWEEN SENTENCE COMPLEXITY, WORKING MEMORY, AND ADDITIONAL WORKING MEMORY LOAD: AN ON-LINE MEASURE

A Thesis

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Arts

in

The Department of Psychology

by
Christy Seidel
B.S., Louisiana State University, 2007
May, 2010

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This experiment takes an on-line look at syntactical complexity, external loads, and working memory, and how the three influence one another. Based on off-line data looking at the interactions between these three factors, we have discovered that syntactic complexity and span have main effects, while the effect of load is most interestingly seen in a three way interaction representing the hardest possible combination of factors. Through this new design, we were able to see whether the off-line results of sentence processing are replicated with an on-line measure. Our new off-line findings replicate past results, which show that working memory impacts off-line sentence processing. However, we fail to see working memory play a major role in on-line sentence processing. This paper will give a background on the available literature, a brief summary of the off-line findings from the prior experiment, and our new off-line and on-line results for this experiment.

1. INTRODUCTION

1.1 An Introduction to Relatives and Clefts

This experiment investigates various factors that impact sentence comprehension. Specifically, three factors will be looked at. The first factor is internal processing demands/syntactic complexity (the inherent processing demands of a sentence itself). The second is working memory (the ability one has to manipulate information) (Baddeley, 1974; Cowan, 1999), which is argued to factor in off-line measures of processing (Waters & Caplan, 2001), and perhaps in on-line measures as well (King & Just, 1991). The third factor is external load (the additional tasks being accomplished during sentence processing). The sentences that will be used in this experiment are subject (e.g. "It is the doctor that contacts the patient") and object (e.g. "It is the doctor the dentist contacts") clefts. Clefts are related to relative clauses, which simply lack the "it is" preface, and much research has examined the processing of subject and object relatives.

In fact, one can consider a cleft to derive from a relative clause. The transformation from relative clauses into clefts is simply an example of internally manipulating the sentence. For example, to transform the subject relative clause (The doctor who contacts the dentist) into a subject cleft (It is the doctor that contacts the dentist), one only needs to preface the subject with the phrase "it is the". Thus, the transformation from an object relative clause "the doctor who the dentist contacts" would result in the object cleft "it is the doctor that the dentist contacts". Combinations of linked relative clauses, such as "the boy that the girl saw hit the man", do exist, although they are rare in the literature (Sheldon, 1977). But they too can be transformed from their relative clause states into cleft structures. For example, the double relative clause "the doctor who visits the patient who contacts the dentist" can be transformed into the cleft + relative structure "it is the doctor that visits the patient that contacts the dentist". In most of these transformations from linked relative clauses to cleft + relative clause structures, the meaning of the sentence is retained except for subject cleft plus subject relative clause, such as "it is the doctor that visits the patient that contacts the dentist". In this case, the subject of the second verb is ambiguous and can apply to both the doctor and the patient—that is, either the doctor or the patient could be the one contacting the dentist. In this way, clefts, because of the "it is the…" preface, have the potential to inject ambiguity into the interpretation of the sentence, and, as a result, may require more capacity to comprehend than relatives. However, since clefts are so similar to relatives in their basic structure, whatever research has been done on relative clauses should also be quite applicable to clefts.

1.2 An Introduction to Internal Processing Demands

Targeting internal load, and manipulating internal information, is one way of transforming "low load" sentences into those which can be considered comparatively "high load" or vice versa(Gordon, Hendrick, Johnson, 2001; King & Just, 1991). Internal load of a sentence can be manipulated in one of two ways without adding new information. The first way involves manipulating the sentence itself. One way to do this is to rearrange the order of the pre-existing words in the sentence to convey a different meaning, such as reordering the words in a subject cleft (e.g. It is the doctor that contacts the dentist) to form an object cleft (e.g. It is the doctor that the dentist contacts) (Gordon et al., 2001). What this rearrangement does is change the structure and syntax of the sentence itself. Research has shown that the object cleft is harder to process, and therefore takes more working memory resources than the subject cleft (Andrews, Birney, & Halford, 2006; Gordon, et al., 2001). This finding has shown that in both on-line measures, such as reaction times which measure processing as it occurs during stimulus presentation, and off-line measures (Gordon et al., 2001). In efforts to create a more complex sentence, one can look at which sentences are hard to process in their natural form. Research points to the fact that clefts, particularly object clefts, are among the most complex sentence structures (Leech, Aydelott, Symons, Carnevale, & Dick, 2007). In fact, adult-like performance on object clefts is not seen in children before nine-years-old, whereas mastery of subject clefts is seen around seven-years-old (Leech, et al., 2007).

Another way to manipulate the sentence without adding information is to focus on simply replacing one word type with another. For example, research shows that the difference in processing complexity between subject and object relatives can be decreased by replacing the second noun presented with a proper name (e.g. "the barber that admired Joe climbed the mountain") or the pronoun "you" (e.g. "the barber that admired you climbed the mountain") (Gordon et al., 2001). Contrarily, processing complexity can be increased when pragmatic cues helping to interpret the subject and object roles of the sentence are eliminated and replaced with less telling word choices(e.g. "the robber that the fireman detested watched the program" is harder than "the robber that the fireman rescued stole the jewelry")(King & Just, 1991). These experiments show that internal manipulations to the sentences themselves, whether by changing syntax or other non-syntactic manipulations, can change a sentence's processing demands, either increasing or decreasing them.

Another way of manipulating a sentence is through the addition of new information, whether it be a prepositional phrase, a descriptive clause, or even an extra clause (Bock & Miller, 1991; Hartsuiker & Barkhuysen, 2006). This is especially true when this "new information" is interjected between the subject and verb of the original sentence, and further, when a new noun is introduced (Bock & Miller, 1991; Hartsuiker & Barkhuysen, 2006). Also, internal manipulations involving the addition of new information are most successful when they include longer phrases over shorter phrases (Hartsuiker & Barkhuysen, 2006). Obviously, there are a number of ways in which one can contribute to the complexity of a sentence by contributing to the information that is already present.

Thus, one can theorize about the resulting effectiveness in enhancing the internal load by attaching a relative clause onto the first cleft. This will create four possible cleft plus relative structures, such as subject cleft plus subject relatives, subject cleft plus object relatives, object cleft plus subject relatives, and object cleft plus object relatives. These cleft plus relatives are then hypothesized to be "high load" counterparts to "low load" single object and subject clefts. With these four combinations, either the first or second verb in each structure can be questioned to analyze comprehension. The first verb will always only concern the initial cleft. Thus, by questioning the first verb, we can see how comprehension is impacted by the additional relative, which we suspect will increase internal load. Questioning the second verb examines the ambiguous subject cleft + subject relative structure, where participants may prefer one correct answer over another, depending on the overall load of the task.

In conclusion, sentences themselves can be divided into sentences with low and high processing demands, and further, low load sentences can be molded into high load sentences either through changing syntax or through the addition of new information.

1.3 An Introduction to Working Memory's Role in Sentence Processing

Working memory is the ability to store and manipulate information, and individual differences in working memory ability may be a vital factor in being able to comprehend sentences, particularly when under some form of external load (Vos, Gunter, Schriefers, Friederici, 2001a). In many studies, participants can be classified as either "high span" (having higher working memory abilities) or "low span" (having lower working memory abilities) depending on the amount of information they can manipulate and recall. In one experiment, conducted by Gunter, Jackson, and Mulder (1995), neural activity measured during syntax-based tasks showed that tasks with high processing demands, such as processing a sentence with a subordinate clause that interrupts the main clause and its ending, proved to be more difficult than tasks with lower processing demands overall, in which the subordinate clause precedes the main clause. However, participants with lower working memory abilities, tended to be particularly susceptible to these high load tasks as compared to high span participants (Gunter et al., 1995). One flaw in this study was that participants in the study were divided into high and low spans based on age, with the high spans being the younger participants, with a mean age of 20.5 years, and the low spans being those who were older, with a mean age of 57.5 years (Gunter et al., 1995). While these results may be valid, despite grouping span by age, for future research, I would argue that forming groups based on a working memory task would contribute more to the acceptability of the finding, since lower performance of the older participants could be due to other cognitive declines that occur with increased age, and not specifically working memory.

Other research has looked at the impact of working memory span on relative clause comprehension with no additional loads, and again, high spans perform better than low spans on both subject and object relatives (King & Just, 1991). Working memory is especially important in comprehending complex sentences with multiple layers, including 5- role object relatives (and clefts)(e.g. “the clown that the teacher that the actor liked watched laughed”) and 5- role subject relatives (and clefts) (e.g. “the actor liked the teacher that watched the clown that laughed”) (Andrews et al., 2006). In addition, working memory has been shown to interact with internal and external loads.

1.4 Working Memory and Internal Load

As discussed earlier, internal load can be seen as the manipulation of present information or the addition of new information. In the case of syntax, we can look at the differences between performance on object clefts and subject clefts and what role an individual’s working memory plays. Research points out that object clefts are inherently harder to process than subject clefts (Gordon et al., 2001), but when adding working memory as a factor, we see that low span individuals suffer more on the object clefts than the high span individuals (King & Just, 1991). This signifies that low span individuals are more susceptible to these internal sentence manipulations while high spans remain immune to some degree. Neural evidence also lends further support for this finding that low span participants suffer more in comprehending complex sentences (Vos et al., 2001a). Yet, working memory interacts not only with the internal load of a sentence, but also with any external factors which may compete for resources needed to comprehend a sentence.

1.5 An Introduction to External Load

To decrease the cognitive resources available to comprehend cleft sentences, one can introduce a supplementary external load. Past research has looked at sentence comprehension under load, but types of loads used vary, as well as, arguably, the complexity of these loads. In King and Just’s (1991) work, participants were presented with 1, 2, or 3 critical sentences in a set via a word-by-word visual moving window. After the last sentence of the set, participants were then prompted to recall the last word of every sentence for that set, which would have been one, two, or three final words (King & Just, 1991). Gordon, Hendrick, and Levine (2002) had participants remember 3 words while they read the target sentence, also in an on-line measure. Vos, Gunter, Schriefers, and Friederici (2001a), split the additional working memory load into high and low versions. The low version involved keeping only one word in mind while reading the sentence, while the high version involved retaining three words (Vos et al., 2001a). Other experiments have similarly used target words to remember during on-line reading measures (Vos, Gunter, Kolk, & Mulder, 2001b), as well as sentence production tasks (Hartsuiker & Barkhuysen, 2006) as their working memory load of choice.

The main complaint here is the complexity of such a task. While in all cases, having to remember words, or monitor for words during reading a sentence, impact performance (King &

Just, 1991; Gordon et al., 2001; Vos et al., 2001a; Vos et al., 2001b; Hartsuiker & Barkhuysen, 2006), the fact remains that simple rehearsal is not as complicated as having to manipulate additional information (D'Esposito, Postle, Ballard, Lease, 1999). Also, it is worthy to note that none of these experiments use a visual task for their external load, and thus, to be comprehensive, it would be interesting to see what interactions arise from combining a non-linguistic visual working memory load with an auditorily presented sentence comprehension task. That is, can syntactic processing be impacted by an additional load of any modality, or must it be a load with a linguistic component?

1.6 Predictions for Linguistic vs. Non-linguistic Loads Under Various Models of Working Memory

Indeed, different schools of thought predict different outcomes for sentence processing when combined with linguistic versus non-linguistic load. Below, I discuss four different theories of how different types of load should impact linguistic comprehension. First, given its syntactic component, Caplan and Waters (1999) argue that language should not be treated as a regular "auditory stimuli," such as a working memory digit span task, which lacks the syntactic element that language embodies. Waters and Caplan (1996a) further hypothesize that due to the difference between language and other auditory stimuli lacking syntactical elements, when given a situation in which the two are being processed at once, no interference between the two will occur, and performance will not decrease. Second, contrary to the beliefs of Waters and Caplan, Just and Carpenter (1992) argue that all forms of auditory information, regardless of the presence or absence of syntax, are essentially the same, and, thus, language and other auditory information *will* interfere with one another, decreasing performance. No specific hypothesis concerning potential interactions between visual loads and language processing were made by either Waters and Caplan or Just and Carpenter.

Thirdly, according to Baddeley, working memory is divided into different components, the visuospatial sketchpad and the phonological loop, designed specifically to individually address visual and auditory information, respectively (Baddeley, 1974). Given the separate stores for processing the different forms of information, it is theorized that visual and auditory information will not interfere with one another, but will only interfere with another task of the same nature, which draws on the same pool of resources (Baddeley, 1974). Thus, in the case of language processing under load, Baddeley's model would hypothesize that only an additional auditory task will hinder performance, while an additional visual task will have no effect (Baddeley, 1974). However, the fourth-component of Baddeley's model, the episodic buffer, may also play a role in external load interference (Baddeley, 2001). Because certain information presented visually may be recoded auditorily, this allows for a potentially multimodal stimulus best dealt with in the capacity of the episodic buffer.

Finally, a different view of working memory comes from Cowan's embedded processes model (Cowan, 1999). According to this model, processing information relies on the focus of attention, and all stimuli, visual or auditory, draw from one common pool of resources (Cowan, 1999). Thus, auditory and visual stimuli will interfere with one another (Cowan, 1999), although Cowan acknowledges that stimuli of the same nature will interfere to an even greater extent (Cowan & Morey, 2007).

In conclusion, we have four prominent schools of thought, each of which argues a separate point. Waters and Caplan (1996a) believe that language and non-syntactical auditory stimuli are inherently different and will not interfere with one another. Contrarily, Just and Carpenter (1992) disagree and state that language and non-syntactical auditory stimuli will interfere. Baddeley (1974) and Cowan (1999)'s models support Just and Carpenter's belief, but diverge on the topic of how a visual stimuli will interfere with one that is auditory. On this note, Baddeley's (1974) model predicts that no interference will occur if the visual and auditory stimuli stay within the visual and auditory modalities, respectively. However, some interference may arise if there is multi-modal encoding within the episodic buffer (Baddeley, 2001). Contrarily, Cowan (1999) argues that interference will occur since a pool of shared resources is being drained.

1.7 Working Memory and External Load

Generally speaking, when an external load is added to a task, one of two things can occur. The first potential outcome is that an individual has the necessary capacity to meet the new demands, with the additional task not showing any visible detriment to performance (Hartsuiker & Barkhuysen, 2006). The second potential outcome is that the new demands require more resources than are available, resulting in a performance drop, even for tasks for which performance was relatively high before the external load addition (King & Just, 1991). Typically speaking, high span individuals are those who are better equipped to handle additional external tasks, while low spans' performance suffers more due to interference (Vos et al., 2001a). This would particularly be the case when the external load task is added onto sentences which are already deemed complicated, since those sentences, sans load, already result in lower performance for low spans (Vos et al., 2001a). On the other hand, we see that not only low spans' performance drop when both internal and external load is high. In such cases, when the required working memory capacity extends past what is available, high spans' performance will drop to low spans' level (Kane & Engle, 2000), showing that there is a certain point when high span individuals reach floor performance.

1.8 Previous Research Reexamined and Past Hypotheses

By discussing three factors that can impact sentence comprehension: internal load, working memory, and external load (Vos et al., 2001a), a series of hypothesis can be generated about how the three will interact in a comprehensive study. Vos et al. (2001a) already outlined their views- i.e., that the most difficult sentences, when combined with the most demanding external loads, will be challenging for both high and low spans, but for low spans in particular. Based on this, last year, an experiment was completed that encompassed all of the elements presented above (Seidel & McDonald, 2008). The experiment itself consisted of auditorily presented single or cleft plus relative clause structures, followed by a visually presented off-line comprehension question asking for the subject of one of the actions mentioned (Seidel & McDonald, 2008). In the load conditions, the sentences and questions were flanked by either visual or auditory stimuli, which were judged to be the same or different (Seidel & McDonald, 2008). Figure 1 below outlines the layout of the experiment for all three loads.

Working memory scores were obtained after the experiment (Seidel & McDonald, 2008). Thus, the design afforded a comprehensive look at the interactions between internal load (subject versus object clefts and single clefts versus clefts plus relative clauses), working memory (low span versus high span), and external load (no load versus auditory load versus visual load.

It was hypothesized that there would be a main effect of syntactic complexity, or internal load, with object clefts being harder than subject clefts, and also clefts plus relatives being harder than single clefts (Seidel & McDonald, 2008). Results supported our hypothesis. Data showed a main effect of syntactic complexity, with object clefts being harder than subject clefts (Seidel & McDonald, 2008), which lends support to the previous findings of this sort (Andrews et al., 2006; Gordon et al., 2001). Also, we found that clefts plus relative clauses tended to be more difficult than single clefts (Seidel & McDonald, 2008). These results can be seen when comparing the averages in Figure 2, looking at single clefts, to those in Figure 3, looking at cleft + relatives.

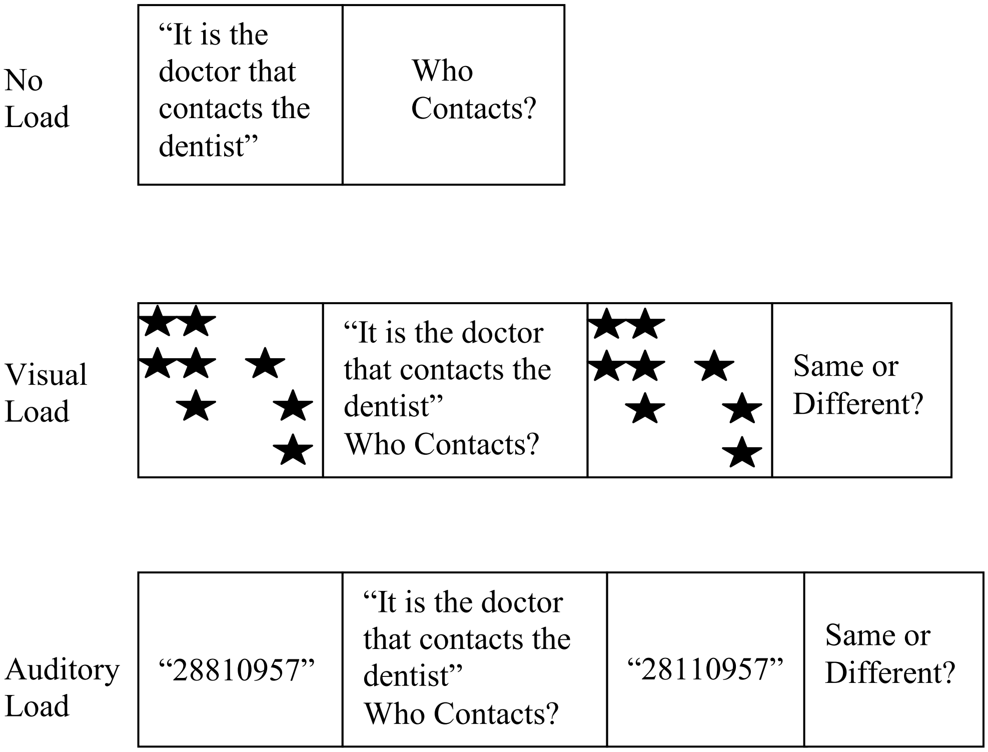


Figure 1: Layout of Seidel & McDonald's (2008) Experimental Conditions

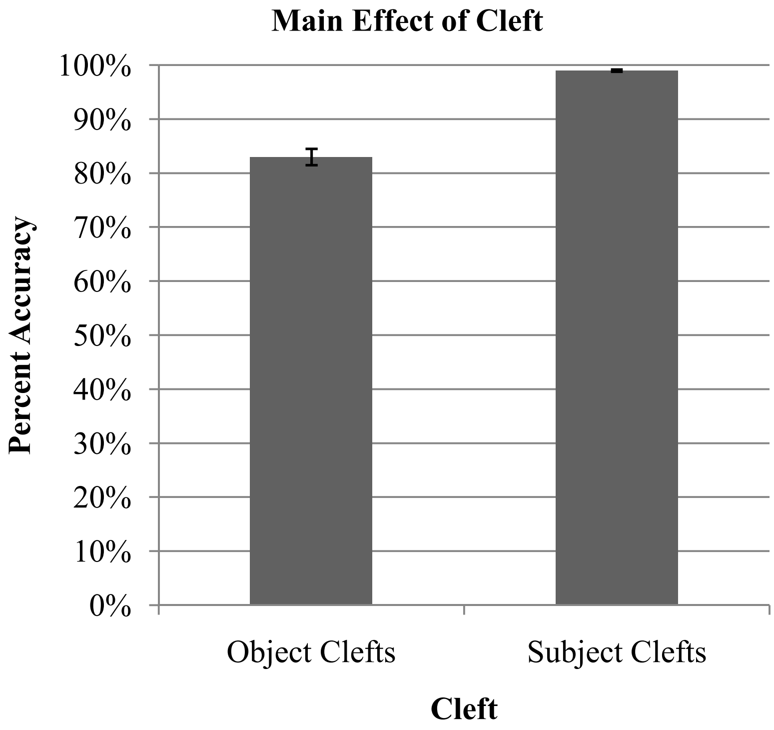


Figure 2: Main Effect of Cleft Type for Single Clefts (Seidel & McDonald, 2008)

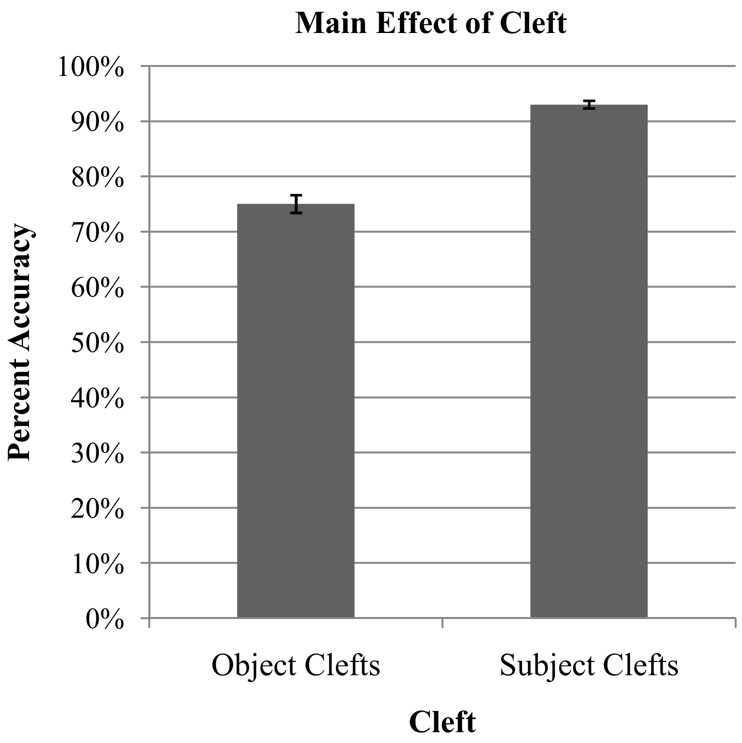


Figure 3: Main Effect of Cleft Type for Cleft + Relative Structures (1^{st} Verb) (Seidel & McDonald, 2008)

We hypothesized a main effect of working memory span, with high spans outperforming low spans (Seidel & McDonald, 2008). Data supported this hypothesis as well (Seidel & McDonald, 2008).

Further, we hypothesized a main effect of external load, with load conditions being harder than the no load condition, but particularly, with the auditory load condition being harder than the visual load condition (Seidel & McDonald, 2008). The data can't find a main effect of load on sentence processing for single cleft comprehension, or for cleft plus relative clause comprehension when asking about the second verb. We did find a main effect of load on sentence processing for cleft plus relative structures when focusing on the first verb. However, the trend was the reverse of what was expected, such that the auditory condition showed the highest performance, and significantly higher than the no load condition. To explain this result, we believe that the combination of the more difficult sentences, specifically those beginning with an object cleft, and the auditory load proved too complex for individuals to be able to access the first noun presented and consider it as an option. As a result, answers defaulted to the second noun presented, which happened to be correct. Thus, two of the three hypotheses concerning main effects were supported in the way we predicted (Seidel & McDonald, 2008).

We also predicted a series of interactions between the internal load, working memory, and external load (Seidel & McDonald, 2008). Concerning the interactions, we predicted two way interactions between sentence type and external load, as well as between sentence type and working memory span, and one three-way interaction between sentence type, external load, and span. In each case, we expected load and span to impact the processing of object structures more than subject structures. The results were mixed. Results showed that syntactic complexity did not interact with external load in lower level interactions, but did interact with span, as shown in both Figures 4 and 5.

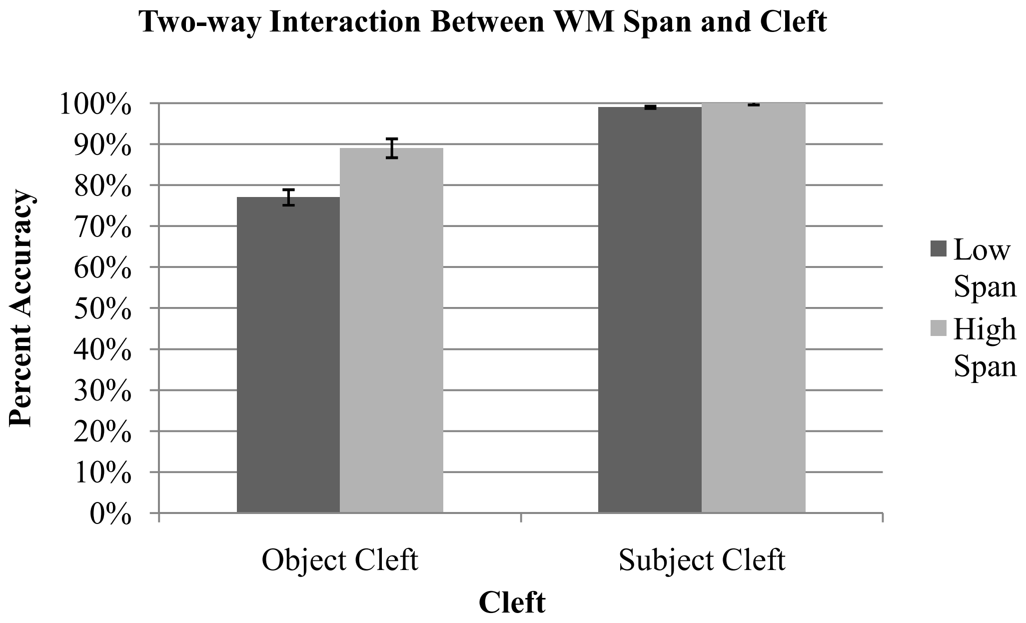


Figure 4: Two-way Interaction Between WM Span and Cleft for Single Cleft Structures (Seidel & McDonald, 2008)

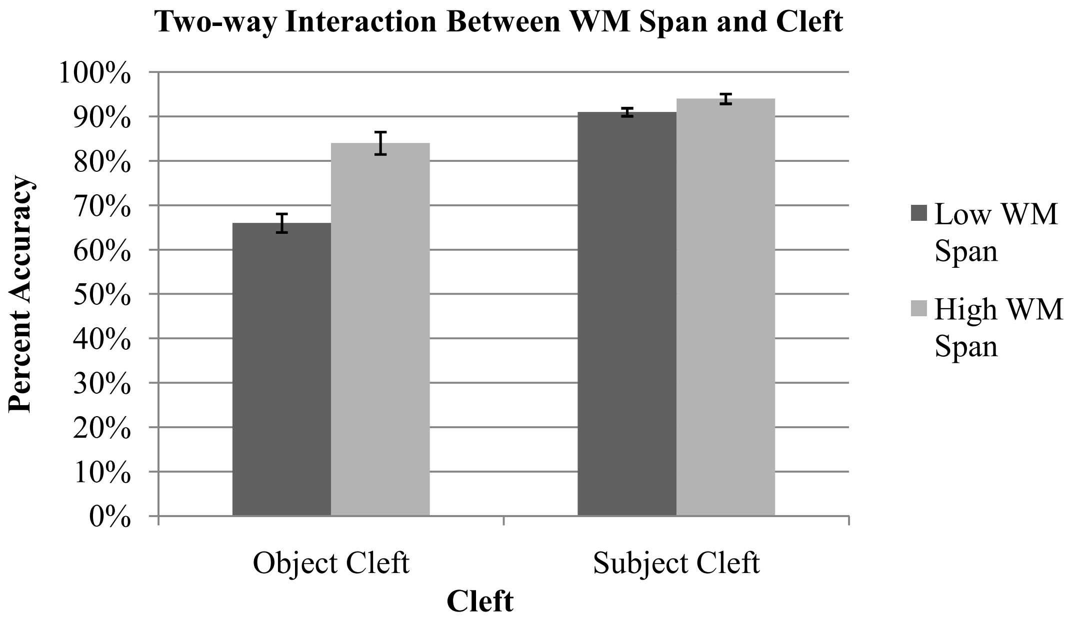


Figure 5: Two-way Interaction Between WM Span and Cleft for Cleft + Relatives (1st Verb) (Seidel & McDonald, 2008)

Specifically, high and low span individuals performed similarly on subject clefts, but on object clefts, high span individuals outperformed low spans (Seidel & McDonald, 2008). This shows that higher span individuals are more capable to meet the higher demands of the challenging sentence structures.

There was a significant four-way interaction of interest between WM span, cleft, relative, and external load in the cleft plus relative clause structure, when questioning the second verb (Seidel &

McDonald, 2008). What drove this interaction was the object-subject cleft structure (e.g. "It is the doctor that the patient visits the contacts the dentist"). This can be seen in Figure 6.

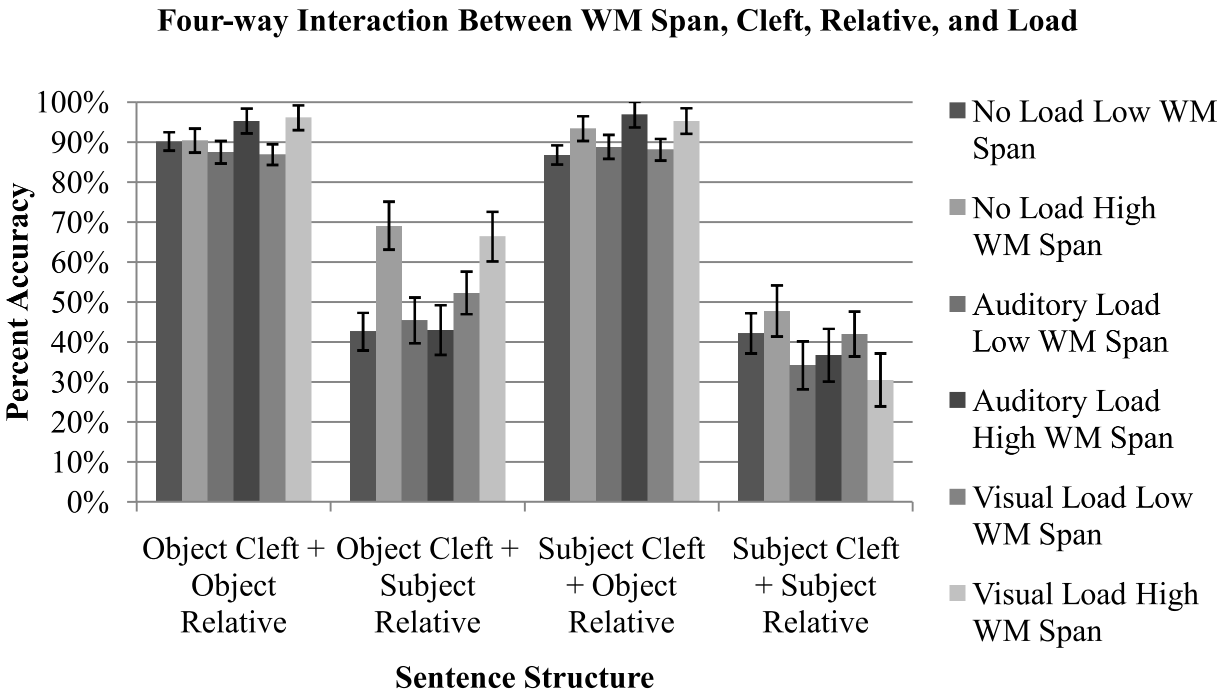


Figure 6: Four-way Interaction Between WM Span, Cleft, Relative, and Load for Cleft + Relatives (2nd Verb) (Seidel & McDonald, 2008)

When asking, "who contacts?" about this sentence, one must reach back to the beginning of the sentence for the correct answer, "the doctor". Given the necessity to hold this long stretch of information in mind, we can see why this sentence structure is the most complex—certainly more so than a single cleft or even a cleft plus relative clause in which one has to only go back one noun for the correct answer. Thus, given the demands of the sentence alone, it comes as no surprise that performance on this structure was much lower than the others. When the performance of this sentence was divided by the external load groups and participants' span, the interaction emerged. What we saw was that high span participants performed better than low span participants in the no load and visual load conditions. However, in the auditory condition, both high and low spans performed equally poorly. To explain this, we believe that high spans have the capacity necessary to hold onto information the sentence gives, including information revealed early on, such as the role of "the doctor". That is why high span individuals were able to answer correctly when the load was comparatively low (no load, visual load). However, when the load's demands increased (auditory load), it became more difficult to recall the correct answer—"the doctor"—from earlier in the sentence, and so participants reverted to answering with the noun farthest back that they could remember "the patient". Low spans, we speculate, use a process-of-elimination of sorts. In all conditions, they were unable to reach as far back in the sentence to access the correct answer "the doctor." However, they knew that the last noun they heard, "the dentist," was incorrect since it was the last piece of information revealed. Thus, they chose "the patient" by default. This explained the lower performance in the no load and visual load conditions in which the high spans perform better. Also, it tells us that perhaps it takes a certain level of difficulty within the sentence itself before load begins have an effect.

When looking at the results of these data and comparing it to the theories of syntactic processing and external load conflict, Baddeley's model seems to emerge as the best supported (Seidel & McDonald, 2008). The data showed that when load finally began to interfere with sentence comprehension, it was auditory load, and not visual load, that drove the interaction.

One criticism of the above visual condition, and why perhaps it did not have the desired effect on sentence comprehension, lies within the way it was presented. In the visual condition, eight stars were presented simultaneously on the screen for a given amount of time. The reason for choosing this type of visual load was hopefully to minimize the potential of internally recoding the visual stimulus linguistically, and keep the visual load as pure of a visual load as possible. The visual condition, although having eight separate visual stimuli, was presented at once, it could be argued that participants were encoding the visual stimuli as one "whole" pattern as opposed to the intended eight individual pieces. This would be an inaccurate counterpart to the auditory condition, which presented eight digits one after the other. Obviously, the eight digits in the auditory condition cannot be perceived as a whole given its sequential nature. Thus, to make a more accurate visual parallel, a second visual condition was constructed in which the stars were presented sequentially in their positions (Seidel & McDonald, 2008). This, however, might also make this additional task more spatial in nature given that the start stimuli flash sequentially in different locations.

Separate analyses were run comparing the first visual condition to the second visual condition, and comparing the auditory condition to the second visual condition (Seidel & McDonald, 2008). Results showed that sentence performance under the second, sequential visual condition load was comparably more similar to the original visual condition rather than the auditory condition. This showed that, while there are obvious criticisms to the original visual condition, the results obtained through the first visual condition should not be discarded or discredited.

1.9 The Debate Between On-line and Off-line

As mentioned earlier, all of the data gathered in the previously discussed experiment was off-line—i.e., it involved answering a comprehension question after the sentence was processed. However, are on-line or off-line data more appropriate in determining how working memory impacts sentence processing? On-line measures highlight the processing that occurs *during* initial stimulus presentation. Because of this, it could be argued that on-line measures are more appropriate. In off-line measures, processing is usually tested by a comprehension question after stimulus presentation, which could reflect processing that potentially occurred after the stimulus (Waters & Caplan, 2001). However, on-line measures of sentence comprehension frequently use moving window paradigms, which usually splice sentences, and thus may be considered unnatural measures (Ferreira, Henderson, Anes, Weeks, & McFarlane, 1996; Waters & Caplan, 2001). This comes from the fact that we hear, or see, sentences in one natural stream. Also, at least during reading on-line measures, our eyes have a tendency to scan back in the sentence, and this possibility is eliminated when only one word, or phrase, is presented at a time (Ferreira et al., 1996). However, on-line measures come with the benefits of knowing reaction times, and this can give insight as to precisely which are the more challenging parts of a sentence, as well as, overall, which sentences are more complex. Regarding this, on-line data show that object clefts are read overall slower than subject clefts (Gordon et al., 2001).

Depending on how sentences are spliced, different "critical words" (Gordon et al., 2001) or "regions" (Gordon et al., 2002) are identified. In subject clefts, the first critical word is identified as the "logical object of the relative clause" and the second critical word is the verb of the main clause (Gordon et al., 2001). Therefore, in the sentence, "The banker that the praised the barber climbed the mountain," the first critical word would be "barber" and the second would be "climbed" (Gordon et al., 2001). In object clefts, the first critical word is the "verb of the embedded clause" and the second critical word is, like the subject relatives, the verb of the main clause (Gordon et al., 2001). So, considering the sentence, "The banker that the barber praised climbed the mountain," the first critical word would be "praised" and the second critical word would be "climbed" (Gordon et al., 2001). When comparing the first and second critical words for subject and object relatives, in both bases, object relatives lead to longer reading times (Gordon et al., 2001). Clefts were also spliced into three regions, with only one region differing between subject and object clefts (Gordon et al., 2002). For example, the subject cleft "it was the dancer that liked the fireman before the argument began" given in the experiment, would be spliced "it was the dancer that/liked the fireman/ before the argument began" (Gordon et al., 2002). Its object cleft counterpart, "it was the dancer that the fireman liked before the argument" would therefore be spliced "it was the dancer that/ the fireman liked/ before the argument began" (Gordon et al., 2002). Results showed that reading time was longer for the second and third regions in object clefts compared to subject clefts (Gordon et al., 2002). This shows that the critical, second, region is more complicated in object clefts than subject clefts, and the difficulty in comprehension continues throughout the rest of the sentence (Gordon et al., 2002).

Although the majority of found literature does not take additional loads into consideration, in the few that do, we see that the higher additional external load, monitoring for three words, leads to slower reaction times than the lower additional load, monitoring for only one word (Vos et al., 2001b). Within the on-line measures, when WM span is added as a factor in addition to external loads, we see a trend that approaches significance—that memory load affects reading times only for those who are low span (King & Just, 1991).

A few studies have used both on-line and off-line measures. Of particular interest are the experiments of Waters and Caplan (2001) which compared the results of their on-line and off-line measure. In this experiment, Waters and Caplan (2001) tested five age groups: individuals from 18 to 30 years, 50 to 59 years, 60 to 69 years, 70 to 79, and those 80 years and older. Participants underwent a battery of neuropsychological tests, and a working memory test, previously used in earlier research (Waters, Caplan, & Hildebrandt, 1987, Experiment 2A; Waters & Caplan, 1996b; Waters & Caplan, 2001). In the working memory measure, participants were asked to make acceptability judgments about a series of cleft sentences in a set, and later recall final words for each sentences (Waters & Caplan, 2001). It should be noted that the main experiment, looking at processing, also has participants make acceptability judgments of relatives (Waters & Caplan, 2001). Thus, the working memory measure and main experiment are very similar in nature (Waters & Caplan, 2001). What Waters and Caplan's (2001) measure of working memory showed was a main effect of age, such that the younger individuals aged 18 to 30 years had a significantly higher span than older individuals, 60 years of age and older (Waters & Caplan, 2001). The more interesting finding, however, lies in the discrepancy of results between the on-line and off-line measures as it pertains to age, and working memory.

Where one would assume that on-line and off-line measures would, and ought to, support one another, parroting the same results in different ways, Waters and Caplan (2001)'s data suggest otherwise. Waters and Caplan (2001) conclude that age was correlated with the off-line measures of the more complex sentences. For example, older participants were shown to be less accurate in making acceptability judgments about the sentences they heard, specifically for subject clefts, object-subject relatives, and subject-object relatives (Waters & Caplan, 2001). Also, there is an interesting interaction to note for the more complex object-subject relatives and subject-object relatives. Specifically, object-subject clefts, with one critical area, have higher acceptability judgments than subject-object relatives, with two critical areas, and this is primarily true of the younger participants, aged 18-30 years (Waters & Caplan). This demonstrates that age interacts with the off-line measure (Waters & Caplan, 2001). Further analyses revealed that this result could be due in part to differences in working memory ability (Waters & Caplan, 2001). However, age was not correlated with the on-line measures that measure sentence complexity (Waters & Caplan, 2001). Specifically, even though older participants listened longer in general, there was no significant increase in listening times for older participants during the critical areas of the hardest sentences (Waters & Caplan, 2001). When looking at working memory instead of age, the same trend is revealed, showing that working memory is also not correlated with the on-line measure specific to the critical areas (Waters & Caplan, 2001). Overall, these results suggest that on-line and off-line measures may be sensitive to different things within comprehension, at least in terms of age, and perhaps working memory. Therefore, it is vital to test the previous hypotheses set in Seidel and McDonald's (2008) off-line study in an on-line environment to see if there is consistency in results between the two designs.

Most of the experiments on cleft and relative comprehension using on-line measures have implemented visual moving windows, both self-paced and computer-paced (Gordon et al., 2001; Gordon et al., 2002; King & Just, 1991; Vos et al., 2001a; Gunter, Jackson, & Mulder, 1995; Mak, Vonk, & Schriefer, 2008). More unusual are experiments which present the sentences in an auditory form, particularly through an auditory moving window (AMW), as with Waters and Caplan's (2001) design, discussed above. This then leads to the decision of whether to use a visual or an auditory based moving window design in creating an on-line measure of sentence processing. Considering the main effect of span we found (Seidel & McDonald, 2008), which Waters and Caplan (2001) might argue to be only evident in the off-line measure, it only makes sense to parallel Waters and Caplan's (2001) design by using an auditory moving window (AMW). An additional benefit of using the AMW is that it would allows us to continue to use our old stimuli for a more direct comparison. Our four-way interaction between WM span, cleft, relative, and load (Seidel & McDonald, 2008), as discussed above, also argues against Waters and Caplan's (1996a) theory that non-syntactic auditory stimuli will not interfere or impact sentence processing. To best argue against these views, as well as add support to the findings from our off-line measure, an on-line measure must be taken, and done in a way which would receive the least amount of criticism from the opposing viewpoints. Thus, a hybrid design was proposed for the following experiment: one that used the same sentences and same loads as Seidel and McDonald (2008), but presented the sentences and spliced them akin to Waters and Caplan's (2001) method.

1.10 New Hypotheses Proposed

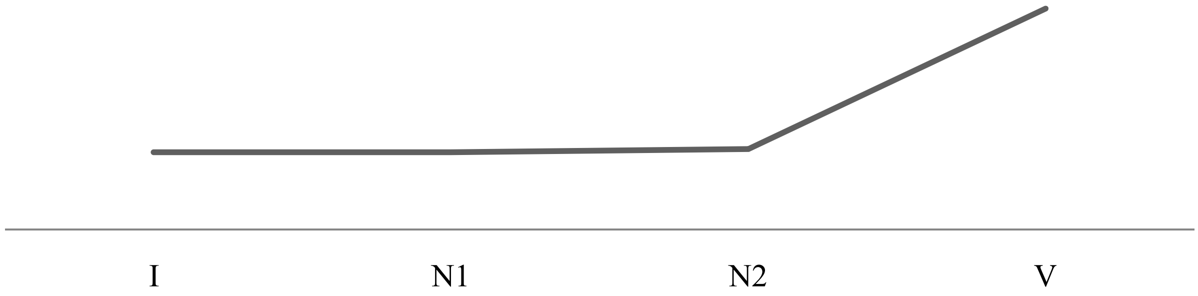
Using the data gathered from the off-line measure as preliminary data of sorts, I will keep my same set of hypotheses generated for the off-line measure and exclude any which were not significantly supported by data across the design. To restate, I hypothesize a main effect of syntactic complexity, with object clefts proving harder than subject clefts, and clefts plus relative clauses proving harder than single clefts. I also propose a main effect of working memory, with high span individuals outperforming low span individuals. Concerning low level interactions, I propose a two-way interaction between syntactic complexity and span, with object clefts and clefts plus relative clauses being harder than subject clefts or single clefts, respectively, but particularly for low span individuals. Finally, I hypothesize a four-way interaction between WM span, cleft, relative, and external load, hoping to replicate, if not extend, the interaction found between load and WM span on the object-subject cleft sentence focusing on the second verb.

On-line measures draw their conclusions of complexity and sentence processing from reading time and reaction time data. Therefore, I hold that these hypotheses, stated above, will be supported through longer listening time on critical fragments. Waters and Caplan (2001) spliced their single cleft sentences into four main parts: the introduction, the first noun phrase, the second noun phrase, and the verb. For example, a subject cleft would be spliced, "It was/ the food/ that nourished/ the child" (Waters & Caplan, 2001). The critical area for subject clefts is the second noun phrase, or “the child” (Waters & Caplan, 2001). The critical area for the object clefts, contrarily, is the verb phrase (Waters & Caplan, 2001). Therefore, in this sentence, “It was/ the woman/ that the toy/ amazed”, the verb phrase “amazed” was the critical region (Waters & Caplan, 2001). In the case of double relatives, or in our case, cleft and relatives, the critical area will be determined by the structures themselves. Waters and Caplan (2001) focused on two double relative structures, an object-subject relative and a subject-object relative, and showed in both cases, the third noun phrase was a critical area. However, in the case of the subject-object relative, the first verb also showed a peak in longer listening times (Waters & Caplan, 2001). So, for the subject-object relative sentence “The man/ that the fire/ injured/ called/ the doctor”, “injured” and “the doctor” would be the critical areas, given that they labeled as the first verb and third noun phrase, respectively (Waters & Caplan, 2001). I also predict these trends in my parallel object cleft plus subject relative structure and subject cleft plus object relative structure. However, for the remaining subject cleft + subject relative structure, I predict peak listening times to emerge at the third noun phrase. For the object cleft plus object relative structure, I predict critical areas to be located at the first and second verb phrases. I am basing these predictions on the patterns seen for the other two structures in Waters and Caplan’s work (2001). In general, higher times are predicted for end of sentence wrap-up effects, and at the verbs for object structures. A layout of the critical areas, where peak listening times are predicted to emerge, is shown below.

A.) Subject Cleft

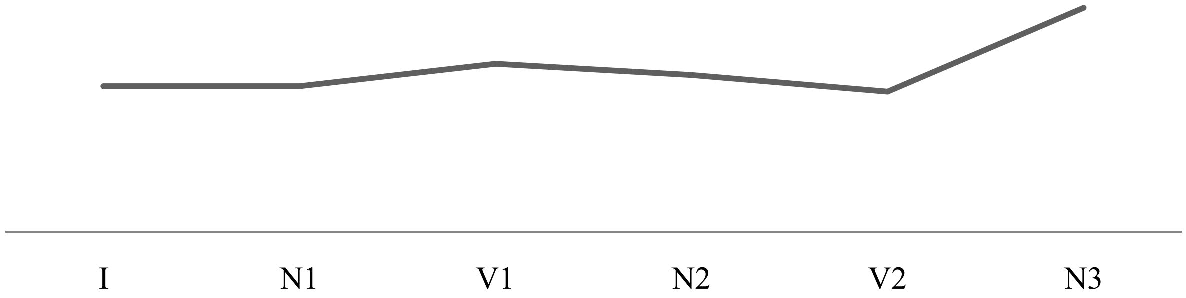
It is the doctor that contacts the dentist

B.) Object Cleft



It is the doctor that the dentist contacts

C.) Subject Cleft + Subject Relative

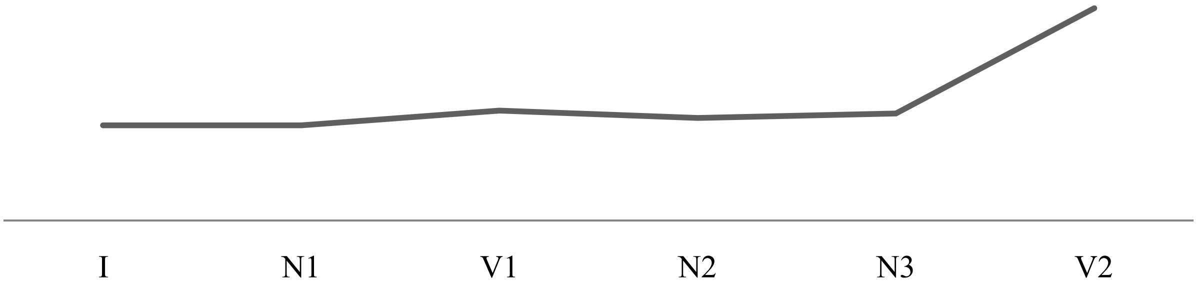


It is the doctor that visits the patient that contacts the dentist

Figure 7: Predicted Critical Areas for Each Sentence Type

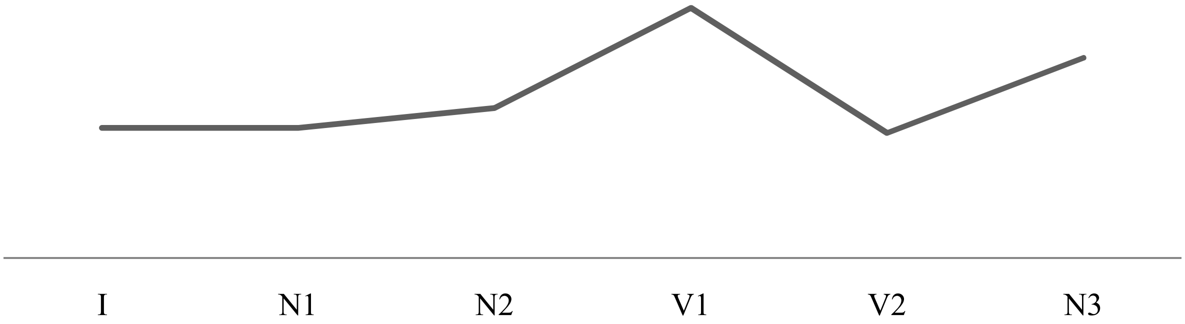
Figure 7 cont.

D.) Subject Cleft + Object Relative



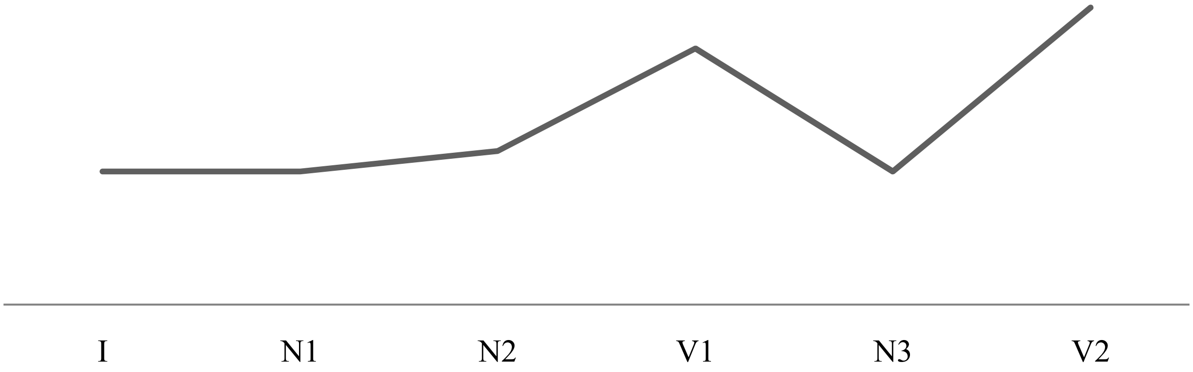
It is the doctor that visits the patient that the dentist contacts

E.) Object Cleft + Subject Relative



It is the doctor that the patient visits that contacts the dentist

F.) Object Cleft + Object Relative



It is the doctor that the patient visits that the dentist contacts.

2. METHODS

2.1 Power Analysis

For this new on-line experiment, the result we were most hoping to replicate would be the four-way interaction between the working memory span, cleft, relative, and losd condition when asking about the second verb (i.e. "it is the doctor that the patient contacts that visit's the dentist," "who visits?"), F(2,112)=3.688, p=.028.

In the original experiment (118 total participants: 35-auditory load, 38-visual load, 45-no load), the effect size (eta squared) for the interaction was .062 and power was .667. For this new experiment, we assumed that an equal number of participants to the old experiment would be sufficient to replicate the effects seen on the comprehension questions. However, we were unsure of how many participants it would take to see a significant on-line result. In determining what sample size would be needed for the on-line measure to replicate the off-line finding, the program G*Power 3.0.10 (Faul, Erdfelder, Lang, & Buchner, 2007) was used. Assuming a power of .8 and desiring a medium effect size of f=.25, a total sample size of 150 participants would be necessary. In comparison, if assuming a small effect size of .10, a total sample size would be 967 participants, and a large effect size of .40 would warrant a total sample size of 64.

2.2 Participants

Participants were all native English speakers and were recruited from the undergraduate student population via Louisiana State University's PAWS system, which allows students to enroll in psychology experiments for extra credit or course credit.

For the on-line condition, a total of 139 Louisiana State University undergraduates were run in one of three conditions: a no load condition (n=50), a visual load condition (n=46), or an auditory load condition (n=43). After data collection, the data were carefully scrutinized and certain question and phrase reaction times were marked unusable by the following criteria. First, all sentence phrase reaction times over five seconds, and all comprehension questions and load questions over 10 seconds, were marked as unusable to rule out outliers. These cutoffs were chosen since they eliminated only the most extreme outliers that most likely did not reflect true processing times. Second, any sentence or sentence fragments the experimenter marked during the experiment as being repetitively clicked through were discarded. Third, if a participant had to vocalize a comprehension question's answer more than once for the computer to accept, that reaction time was rendered unusable. For all conditions, any participant with 5% or more unusable data was excluded. By this criteria, we excluded two participants from the no load condition, four participants from the visual load condition, and five participants from the auditory load condition. Two additional participants from the no load condition were excluded due to complications during testing. Specifically, there was an initial sensitivity of the microphone, which reacted to environmental noises and interfered with the participant's vocalizations.

Therefore, data were collected for 139 participants, but analysis was only conducted on 126 participants (46 in the no load condition, 42 in the visual load condition, 38 in the auditory load condition).

Hearing was officially tested for 95 (75.4%) of the participants. Hearing was tested at 25dB for 500, 2000, 4000, and 6000 Hz. For those that were tested, 65.3% were able to hear all beeps at all frequencies, and 33% were unable to. Of the hearing trials that were missed, they were primarily missed at the 500Hz frequency (78.8%).

The high percentage of participants unable to pass the hearing test could be attributed to interfering distractor noises. These noises were unavoidable as the hearing test was not administered in a sound-proof room. We ran a one-way ANOVA using performance on the hearing screening as the independent variable (no hearing screening administered vs. fully passed hearing screening vs. hearing screening not fully passed) and all on-line and off-line data as the dependent variables. No significant differences between the groups were seen, so all participants were included in the analyses.

2.3 Stimuli

The 96 sentences used in the new on-line measure were identical to those used in the off-line measure. To recap, the sentences used for the off-line measure included single clefts and cleft plus relative clause sentences. There were 16 examples of each of the 6 structure types and all sentences were presented auditorily. Table 1 below provides examples of all the potential structures along with how they were spliced, described in more detail below.

Table 1: Example Cleft and Cleft + Relative Sentences With Splicing

Example Cleft and Cleft + Relative Sentences With Splicing	
Subject Cleft	It is/ the doctor/ that contacts/ the dentist. [I/N1/V/N2]
Object Cleft	It is/ the doctor/ that the dentist/ contacts. [I/N1/N2/V]
Subject-Subject Cleft	It is/ the doctor/ that visits/ the patient/ that contacts/ the dentist. [I/N1/V1/N2/V2/N3]
Subject-Object Cleft	It is/ the doctor/ that visits/ the patient/ that the dentist/ contacts. [I/N1/V1/N2/N3/V2]
Object-Subject Cleft	It is/ the doctor/ that the patient/ visits/ that contacts/ the dentist. [I/N1/N2/V1/V2/N3]
Object-Object Cleft	It is/ the doctor/ that the patient/ visits/ that the dentist/ contacts. [I/N1/N2/V1/N3/V2]

All sentences were written in present tense to add control, and only animate nouns were used to allow for reciprocal interpretation. To make the on-line and off-line measures as parallel as possible, all original auditory files were used, but were spliced by fragment to allow for the constructed auditory moving window (AMW) in SuperLab. For example, a single subject cleft will be spliced into “It is/ the doctor/ that contacts/ the patient”, separating the introductory phrases “it is” from the noun phrases and verb. Likewise cleft plus relative structures were similarly spliced. For example, participants heard “It is/ the doctor/ that the patient/ contacts/ that visits/ the dentist” as one of the object-subject clefts. This splicing is in line with how Waters and Caplan (2001) spliced their sentences for the AMW.

After the sentence task, a size judgment working memory test (Seidel & McDonald, 2008) was issued to determine an individual’s working memory span. This is the same working memory task used in the off-line experiment (Seidel & McDonald, 2008). In the size judgment task, lists of concrete physical objects ranging from 3 to 6 items were auditorily presented to be reordered in order of smallest object to largest object. For example, a participant will hear “cat, cherry, skyscraper”, he should respond “cherry, cat, skyscraper”. The ordering of these items is the same for each participant. The experimenter marked on an answer key the order that the participant recalled the items on the sheet for grading, as well as if a participant forgot a word, recalled a word not presented, or recalled a previously presented word. To allow for more sensitive scoring, participants were awarded one point for every word correctly remembered in the general correct order. In this way, instead of only being awarded for a completely remembered list in the correct order, partial credit could be garnered. We consider this measure of working memory to be appropriate for this experiment given its linguistic, but non-syntactic, nature. The size judgment task also arguably contains a spatial component because participants may visualize the items in the list to help in arranging them by size.

2.4 Procedure

To self-pace the sentences described above, participants were required to push the spacebar. The computer only accepted a button click after each phrase was presented in its entirety. This is different from Waters & Caplan's (2001) method where a button click could be accepted mid-phrase to progress onto the next segment. In the case that a participant engaged in repetitive clicking, the experimenter made a note for later exclusion of that particular phrase or sentence in the analyses.

The structure of the experiment itself remained true to its off-line counterpart. For a recap on how the off-line counterpart was constructed, refer again to Figure 1. Therefore, for the no load condition, the auditorily presented self-paced fragmented sentence was presented, followed by a visually presented comprehension question to answer aloud, to ensure comprehension and encourage participants to stay focused. The comprehension question asked for the subject of an action committed in the sentence ("Who contacts?"). In the case of clefts plus relative clauses, each verb was questioned an equal number of times. This also acted as an additional off-line measure of comprehension and gauge of sentence complexity. In the load conditions, either a warning beep (auditory condition) or 100 ms visual warning, consisting of a large black octagon, (visual condition), followed by a 100 ms pause, alerted the participant to the load presentation. The loads are described in more detail below. After the load presentation, the self-paced sentence in the AMW and comprehension question was presented as in the no load condition. After answering the comprehension question, a beep or 100 ms visual warning, followed by a 100 ms pause, preceded the second test stimulus for the load. Afterwards, the question "Same or Different?" concerning the loads appeared on the screen for participants to answer aloud. In all conditions, the experimenter circled all answers given by the participants on an answer sheet provided to the experimenter.

The auditory and visual load conditions were recreated to be identical to those in the off-line measure. The auditory condition was a series of 8 random numbers presented auditorily at an approximate averaged rate of 500ms per digit. Whole lists of numbers ranged from 3 to 5 seconds, with an approximate average of 4 seconds per series of numbers.

The original visual condition paralleled the auditory condition in number of star stimuli on the screen, with every assigned star location deriving from a specific number from the auditory load condition. Thus, the auditory digits determined potential locations for a star, with every number equaling two possible locations on a random 20 square grid created for this purpose. In the case that two or more of the same number were presented in the auditory condition, the star was arbitrarily assigned a new position. The star patterns with all 8 stars flashed on the screen for 500 ms.

After the original experiment, a second visual condition was created to better parallel the auditory condition, as discussed earlier. In this condition, all stars kept their original positions as in the first visual condition, but they flashed individually, one after the other, at a rate of 450 ms per star. It is this visual condition which was used in my on-line measure. This is because the second visual condition has a sequential element to it which more accurately parallels the auditory condition, which the original visual condition lacked.

As in the off-line measure, the size judgment working memory measure was issued after the on-line sentence comprehension task was completed.

3. RESULTS

3.1 Size Judgment Task

The way the size judgment task was scored allowed for a maximum score of 54 points, however, actual participants' scores ranged from 33 to 51. Both the mean and the median value were 44, and thus, 44 became the cut off value by which to split participants into high and low working memory groups. This is lower than the original off-line experiment in which the cut-off used was 45 (Seidel & McDonald, 2008). Individuals with scores 44 and below were labeled as having low WM spans (n=67), and participants scoring 45 and above were labeled as having high WM spans (n=59). Within the low WM span subgroup, the mean was 41.6, while for the high WM span subgroup, the new mean became 46.8.

The majority of the following analyses were done using this median split of working memory to divide participants into high and low subgroups. However, some analyses focused on participants with only extreme WM scores. In arriving to our extreme groups, all participants with WM scores of 44, 45, and 46 were excluded. By cutting participants with scores 44, 45, and 46, we eliminated the upper and lower boundaries of both the old off-line experiment (Seidel & McDonald, 2008) and the current experiment. A total of 39 participants were excluded, with now 34 individuals in the high WM span group and 53 in the low WM span group. Within the new low WM span subgroup, the mean became to 41.0, while for the high WM span subgroup, the new mean became 47.8.

3.2 Off-line Data

The off-line data was analyzed in two ways. The first way included all participants, using a median split of the working memory scores to define individuals as low or high span (low span: M=41.6, high span: M=46.8). The second way focused only on participants with extreme working memory scores (low span: M=41.0, high span: M=47.8) to verify our results. Overall, we see many but not all of the findings from Seidel & McDonald (2008) were replicated. Further, our data continued to support that working memory is evident in off-line processing.

One finding we hoped to replicate from the previous study (Seidel & McDonald, 2008) was a four-way interaction when looking at cleft + relative structures, focusing on the second verb. This interaction was between WM span, cleft, relative, and additional working memory load. To reiterate, this four-way interaction showed that high spans drop to low span levels of performance for the hardest sentence type (object cleft + subject relative when questioning the second verb) under the hardest additional working memory load (auditory load). To foreshadow our results, we were unable to replicate this specific interaction. There are possible reasons why we failed to replicate our earlier finding, which will be discussed in more detail later.

3.2.1 Off-line Data for All Participants: Single Clefts

Similar to the original experiment (Seidel & McDonald, 2008), we analyzed the simpler subject and object clefts first by means of a 2 (high WM span vs. low WM span) x 2 (subject cleft vs. object cleft) x 3 (no load vs. auditory load vs. visual load) ANOVA. Our results showed only two main effects, both of which were expected. The first main effect was of working memory, $F(1,120)=5.906$, $p<.05$. High spans (M=.920) outperformed low spans (M=.883). The second main effect was of cleft, $F(1,120)=139.901$, $p<.001$, showing object clefts (M=.814) were harder than subject clefts (M=.989). These results replicated Seidel & McDonald's (2008) earlier findings, with the exception that in the previous study, these two main effects were qualified by a two-way interaction. The previous interaction showed that low spans have more difficulty overall but particularly on object cleft sentence structures. This interaction was not replicated here, although the same numerical trend was evident.

3.2.2 Off-line Data for All Participants: Cleft + Relatives (1st Verb)

Next, we analyzed the cleft + relative structures when questioning the first verb with a 2 (high WM span vs. low WM span) x 2 (subject cleft vs. object cleft) x 2 (subject relative vs. object relative) x 3 (no load vs. auditory load vs. visual load) ANOVA. Three main effects were seen. The first main effect was of working memory, F(1,120)=26.320, p<.001, with high spans (M=.867) outperforming low spans (M=.763). Second, there was a main effect of cleft, F(1,120)=163.029, p<.001, with sentences being more difficult if they began with an object cleft (M=.724) instead of a subject cleft (M=.906). Finally, we saw a main effect of relative, F(1,120)=19.341, p<.001. Specifically, sentences ending with an object relative (M=.790) proved to be harder than sentences ending with a subject relative (M=.840). All of these main effects replicated the earlier study (Seidel & McDonald, 2008), with the exception that Seidel & McDonald (2008) also found a main effect of load.

In addition, two interactions were also seen. The first interaction was between cleft and relative, F(1,120)=7.867, p<.01.

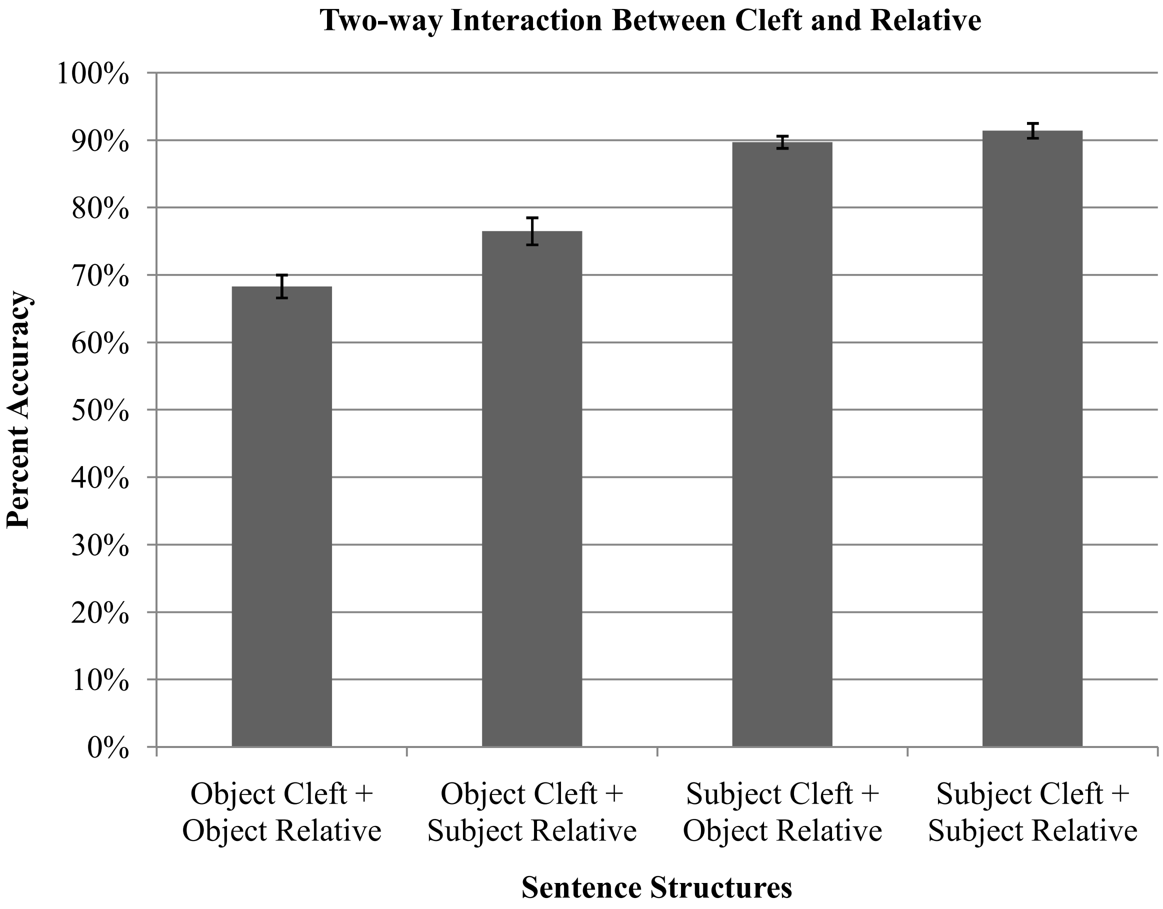


Figure 8: Two-way Interaction Between Cleft and Relative for Cleft + Relative (1st Verb) for All Participants

We saw, as expected, that the most difficult sentence was the object cleft + object relative (M=.683). The second hardest sentence was the object cleft + subject relative (M=.765), followed by the easier subject cleft + object relative (M=.897) and subject cleft + subject relative (M=.914). This was the same ordering of difficulty seen in Seidel & McDonald (2008). The second interaction was between WM span and cleft, F(1,120)=5.630, p<.05.

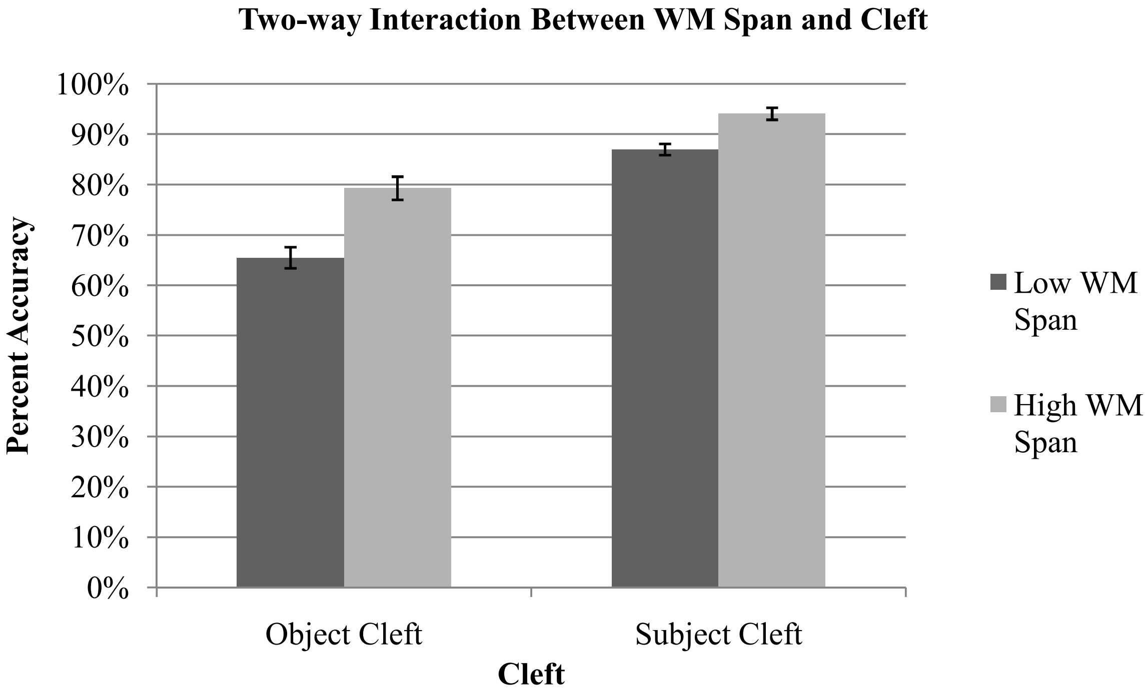


Figure 9: Two-way Interaction Between WM Span and Cleft for Cleft + Relatives (1st Verb) for All Participants

Replicating the trends from Seidel & McDonald (2008), we saw that sentences beginning with object clefts had poorer performance overall. However, for these structures, we saw low spans in particular showing more difficulty than their high span counterparts. Therefore, it was here that we saw working memory playing a role in off-line processing of syntactic complexity.

3.2.3 Off-line Data for All Participants: Cleft + Relatives (2nd Verb)

Finally, we analyzed the cleft + relative structures when questioning the second verb with a 2 (high WM span vs. low WM span) x 2 (subject cleft vs. object cleft) x 2 (subject relative vs. object relative) x 3 (no load vs. auditory load vs. visual load) ANOVA. This was where we hoped to encounter the four-way interaction between WM span, cleft, relative, and additional load. When we analyzed this structure, we saw the same three main effects appear as when focusing analysis on the first verb. First, there was a main effect of working memory, F(1,120)=6.132, p<.05, with high spans (M=.688), again, outperforming low spans (M=.647). Next, there was a main effect of cleft, F(1,120)=111.711, p<.001, with sentences starting with subject clefts (M=.601) being harder than sentences starting with object clefts (M=.733). While this initially may seem surprising, sentences beginning with a subject cleft forced participants to go back further in the sentence for a possible answer, increasing processing demands. Consider the subject cleft + subject relative, "It is the doctor that contacts the patient that visits the dentist". To answer the question, "who visits," participants may have tried to remember back to the first noun presented, "the doctor," for a possible answer. Likewise, there was a main effect of relative F(1,120)=1037.706, p<.001, such that sentences ending in a subject relative (M=.401) were harder than sentences ending in an object relative (M=.934). Sentences ending with an object relative were easier because the answer to the verb in question was found at the very end of the sentence. Also, the answer directly preceded the verb in question.

These main effects were qualified by a two-way interaction between cleft and relative, F(1,120)=77.724, p<.001.

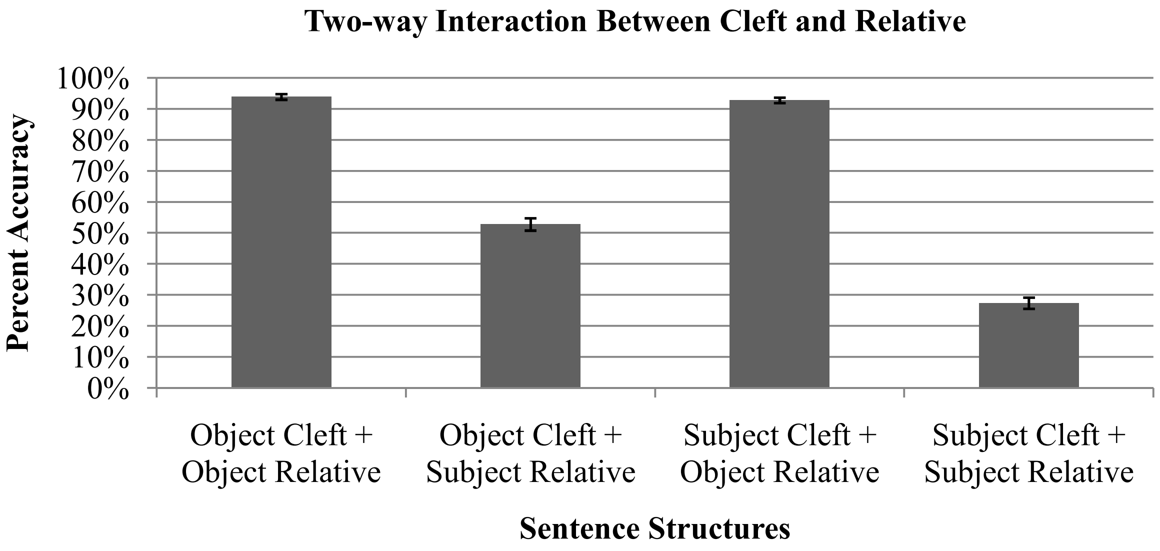


Figure 10: Two-way Interaction Between Cleft and Relative for Cleft + Relatives (2nd Verb) for All Participants

We replicated the trend of sentence difficulty seen in Seidel & McDonald (2008). Specifically, we saw that the ambiguous subject cleft + subject relative structures had the worst performance (M=.274), followed by object cleft + subject relatives (M=.528), and finally by subject cleft + object relatives (M=.928) and object cleft + object relatives (M=.939), which were not significantly different from one another. Even though the subject cleft + subject relative structure was seen as the most difficult, note from earlier that this sentence structure was ambiguous. Our scoring only considered the first noun presented in the sentence as correct, thus true percent accuracy for this structure was not reflected. However, because two of the three nouns presented were potential correct answers, it is difficult to discriminate guessing and legitimate comprehension of the sentence. Therefore, the true hardest sentence structure here was the object cleft + subject relative.

A second two-way interaction was seen, this time between WM span and cleft, F(1,120)=7.082, p<.01, which also supported working memory's role in off-line measures.

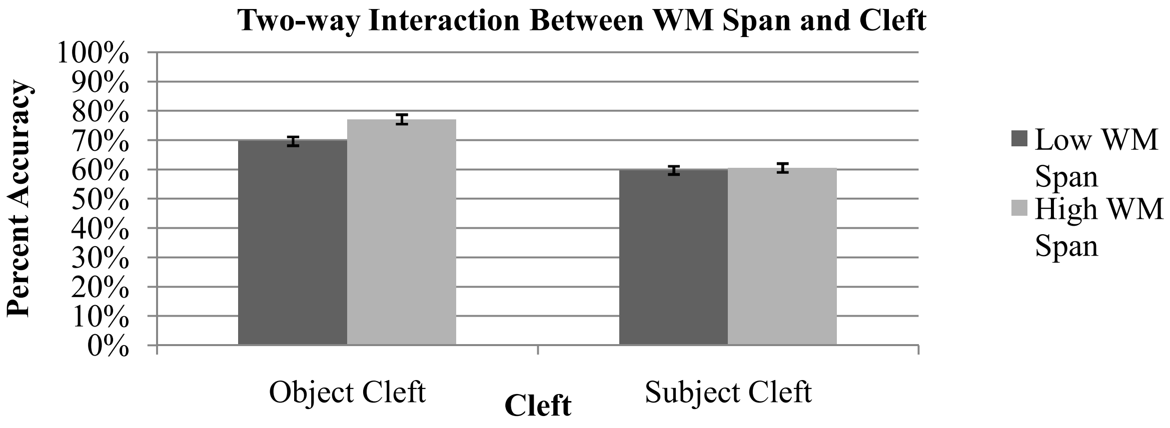


Figure 11: Two-way Interaction Between WM Span and Cleft for Cleft + Relatives (2nd Verb) for All Participants

In this interaction, also replicating an earlier finding (Seidel & McDonald, 2008), we saw distinct performance differences between high and low spans for sentences beginning with object clefts but not subject clefts.

Finally, a three-way interaction emerged. It was for the object cleft + subject relative sentence structure that Seidel & McDonald (2008) found the four-way interaction between WM span, cleft, relative, and load. While this four-way interaction was not replicated in the current experiment, a three-way interaction between WM span, cleft, and relative, F(1,120)=3.912, p=.05, did emerge.

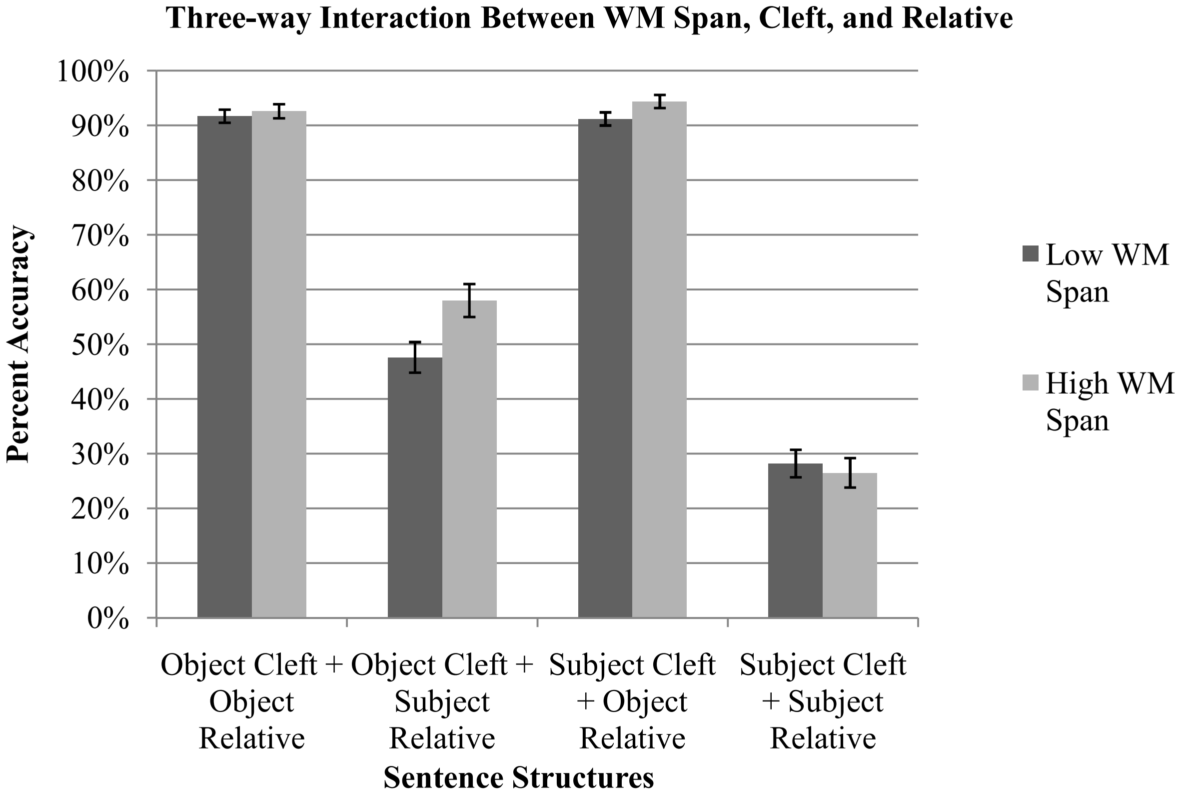


Figure 12: Three-way Interaction Between WM Span, Cleft, and Relative for Cleft + Relatives (2nd Verb) for All Participants

Except for the ambiguous subject cleft + subject relatives, low spans tended to perform poorer than high spans on all cleft + relative structures. This was particularly the case for the most difficult object cleft + subject relatives, where low spans performed markedly lower. The trends of this three-way interaction mirrored those of the three-way interaction, later qualified by the four-way interaction, found by Seidel and McDonald (2008). The lack of the four-way interaction shows that even when adding the hardest additional load to the difficult object cleft + subject relative, high spans still arrived at the correct answer almost as frequently as in the other two conditions. Even though a four-way interaction was not found, our three-way interaction provided further evidence for the role of working memory in off-line syntactic processing by showing performance gaps in comprehension driven by WM span. The on-line data analysis later clarifies why this may be the case.

As a general recap of the off-line findings, when including all participants, we saw that the major findings from Seidel & McDonald (2008) were replicated, with the exception of the four-way interaction. Table 2 below summarizes the current findings and those of Seidel & McDonald (2008).

Table 2: Comparison of Seidel & McDonald's (2008) Earlier Off-line Findings With the Current Off-line Findings

	Seidel & McDonald (2008) Off-line Results			New Experiment - All Participants		
	Single Clefts	Cleft + Rel (1st Verb)	Cleft + Rel (2nd Verb)	Single Clefts	Cleft + Rel (1st Verb)	Cleft + Rel (2nd Verb)
Main effect of Cleft	YES	YES	YES	YES	YES	YES
Main effect of Relative		YES	YES		YES	YES
Main effect of WM Span	YES	YES	YES	YES	YES	YES
Main effect of Load		YES				
Cleft x Rel		YES	YES		YES	YES
Cleft x WM	YES	YES	YES		YES	YES
Cleft x Load			YES			
Rel x WM						
Rel x Load						
WM x Load						
Cleft x Rel x WM			YES			YES
Cleft x Rel x Load						
Cleft x WM x Load			YES			
Rel x WM x Load						
Cleft x Rel x WM x Load			YES			

Further, we found repeated evidence that working memory is a factor in off-line measures of syntactic processing. To add strength to these results, we then analyzed the data focusing only on participants with extreme working memory span scores.

3.2.4 Off-line Data for Participants With Extreme WM Span Scores

When looking at the extreme WM span participants (Low WM Span: n=53, High WM Span: n=34), we saw the same results as when including all participants, with two exceptions. The first exception was seen when analyzing the cleft + relative structures focusing on the first verb. In this case, we no longer saw the two-way interaction between cleft and relative. The second exception was seen when analyzing the cleft + relative structures focusing on the second verb. Specifically, we no longer saw the three-way interaction between WM span, cleft, and relative. Otherwise, all findings were replicated when focusing on only participants with extreme working memory scores.

For an overview of the differences between analyzing the off-line results with all participants and those with extreme WM scores, see Table 3 below.

Table 3: Summary of Off-line Findings for All Participants and Those With Only Extreme WM Scores

	For All Participants			Participants with Extreme WM Scores		
	Single Clefts	Cleft + Rel (1st Verb)	Cleft + Rel (2nd Verb)	Single Clefts	Cleft + Rel (1st Verb)	Cleft + Rel (2nd Verb)
Main effect of Cleft	YES	YES	YES	YES	YES	YES
Main effect of Relative		YES	YES		YES	YES
Main effect of WM	YES	YES	YES	YES	YES	YES
Main effect of Load						
Cleft x Rel		YES	YES			YES
Cleft x WM		YES	YES		YES	YES
Cleft x Load						
Rel x WM						
Rel x Load						
WM x Load						
Cleft x Rel x WM			YES			
Cleft x Rel x Load						
Cleft x WM x Load						
Rel x WM x Load						
Cleft x Rel x WM x Load						

The reason we failed to reproduce these findings could be attributed to a decrease in power due to a loss of participants included in the analysis.

In short, when looking at the off-line data, a number of our hypotheses were supported. First we found main effects of working memory with high spans showing higher performance than low spans. We also found main effects of syntactic complexity. For example, we confirmed that object clefts are harder than subject clefts. And finally, we had interactions between working memory and syntactic complexity, specifically when analyzing the cleft + relative structures (1st and 2nd verbs). The curious finding from the off-line data was the fact that we did not replicate our earlier four-way interaction (Seidel & McDonald, 2008). However, examining the on-line data might help explain why. One possibility to be considered is that the additional button pushing involved in the self-paced listening task forces more on-line processing than would occur naturally. This additional on-line processing may have given participants the needed boost to arrive to the correct answer of even the most difficult sentences under the most difficult load. That is, even when comprehending object cleft + subject relatives under the visual WM load, participants still could answer the comprehension question due to the additional on-line processing afforded by self-pacing the sentence.

3.3 Analysis for Comprehension Question Reaction Times

The next set of analyses focused on the comprehension question reaction times. When plotting the comprehension question reaction times by load condition and span, special attention should be focused on the three hardest sentences. As a reminder, the hardest sentence when questioning the first verb is object cleft + object relative. When questioning the second verb, the hardest sentences were the object cleft + subject relative and the ambiguous subject cleft + subject relative.

3.3.1 Analysis for Comprehension Question Reaction Times for Single Clefts

Single cleft structures were analyzed in a 2 (high WM span vs. low WM span) x 2 (object cleft vs. subject cleft) x 3 (no load vs. visual load vs. auditory load) ANOVA. Only two main effects emerged. First, there was a main effect of WM span, F(1,120)=6.596, p<.05. As expected, high spans (M=1935.126ms) answered significantly quicker than low spans (M=2189.807ms). Second, there was a main effect of cleft, F(1,120)=84.221, p<.001. Comprehension questions about object clefts (M=2286.996ms) took longer to answer than questions concerning subject clefts (M=1837.936ms).

3.3.2 Analysis for Comprehension Question Reaction Times for Cleft + Relatives (1st Verb)

The comprehension question reaction times for the cleft + relative structures were analyzed by question type. First, the cleft + relative structures asking about the first verb presented were analyzed with a 2 (high WM span vs. low WM span) x 2 (subject cleft vs. object cleft) x 2 (subject relative vs. object relative) x 3 (no load vs. auditory load vs. visual load) ANOVA. Two main effects were found. The first main effect was of working memory span, F(1,119)=9.511, p<.01. Again, high spans (M=2082.800ms) answered more quickly than low spans (M=2398.846ms). The second main effect was of cleft, F(1,119)=49.505, p<.001. Sentences starting with object clefts (M=2382.026ms), rather than subject clefts (M=2098.620ms), led to longer comprehension question reaction times.

Three interactions also emerged. First, a two-way interaction between WM span and cleft, F(1,119)=4.414, p<.05 was seen.

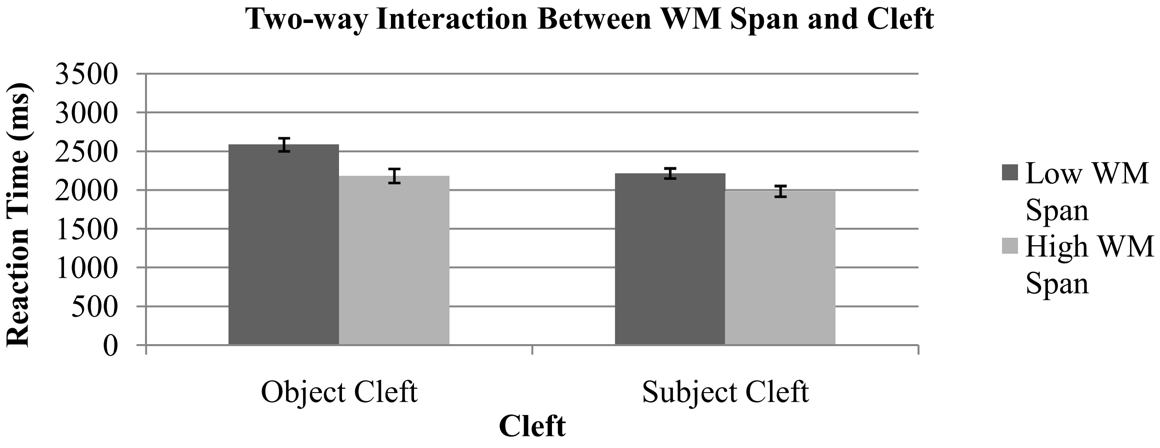


Figure 13: Two-way Interaction Between WM Span and Cleft for Comprehension Question Reaction Times for Cleft + Relatives (1st Verb)

Low spans, in particular, took significantly longer than high spans in answering comprehension questions for sentences starting with object clefts. Although here we witness an interaction between syntactic complexity and working memory, comprehension question reaction times are arguably an off-line measure of processing. This is because the sentence has already ended by the time the processing is occurring. The next interaction seen was a two-way interaction between relative and load, F(2,119)=3.449, p=<.05.

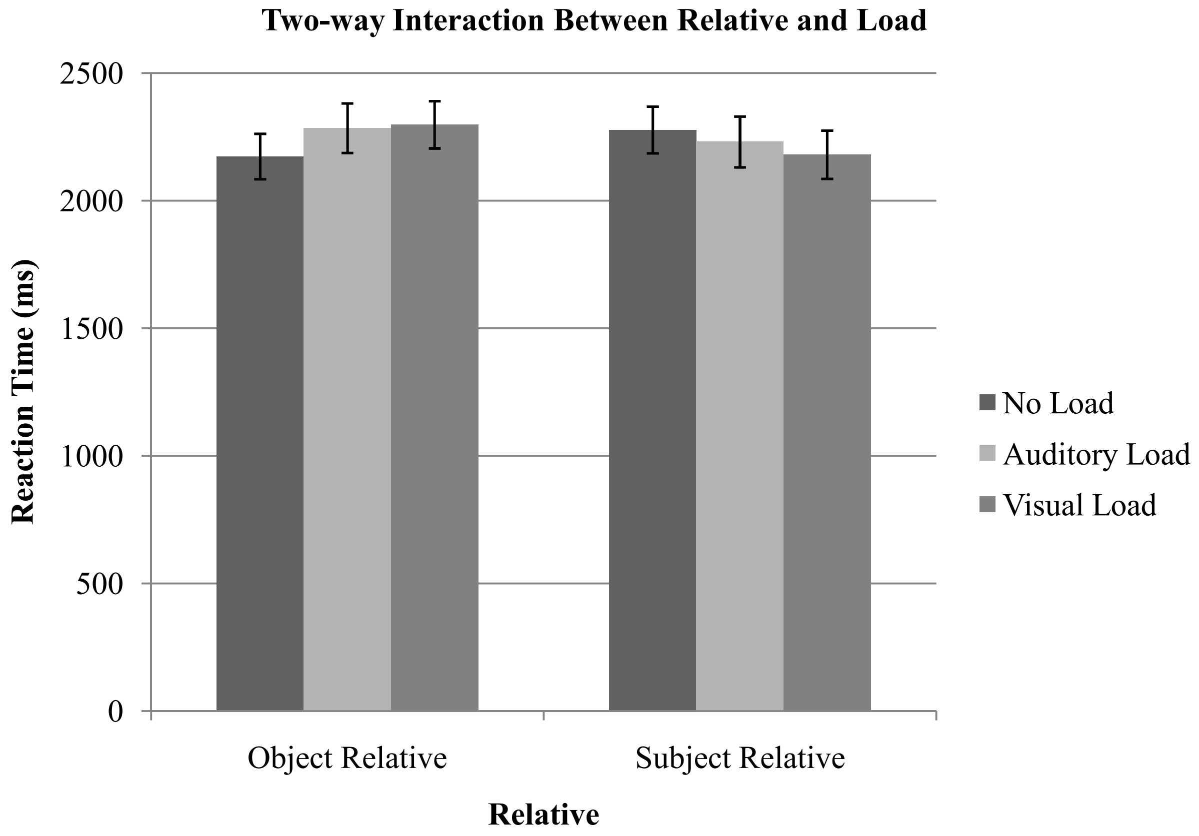


Figure 14: Two-way Interaction Between Relative and Load for Comprehension Question Reaction Times for Cleft + Relatives (1^{st} Verb)

Sentences ending in an object relative instead of a subject relative had longer comprehension question reaction times for both the auditory and visual load conditions. The reverse trend, however, was seen for the no load condition. In the no load condition, sentences ending in an object relative had shorter comprehension question reaction times. This is against what was expected. However, because this interaction focused on the relative, and the verb in question involved the cleft, this interaction was difficult to interpret. The third interaction was a four-way interaction between WM span, cleft, relative, and load, F(2,119)=3.392, p<.05, seen in Figure 15. Low spans in all conditions generally took longer than high spans in answering the comprehension question for all cleft + relative structures. However, in the visual condition, the reverse trend was seen for the object cleft + object relative structure. This was the hardest sentence structure when asking about the first verb. For this structure, high spans took longer than low spans in answering the comprehension question. This showed that to correctly arrive at the answer, high spans devoted more time to processing when the structure was difficult.

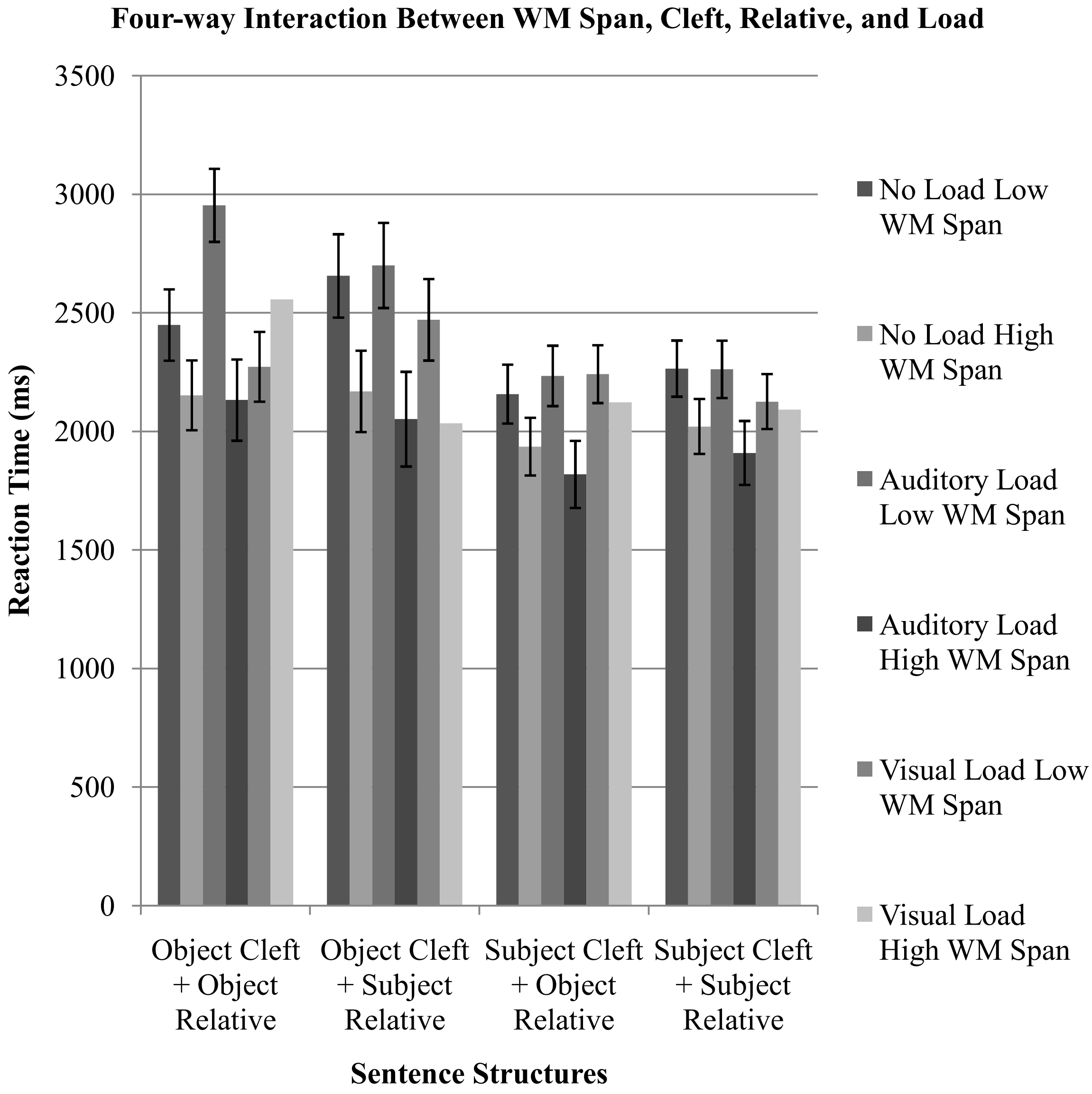


Figure 15: Four-way Interaction Between WM Span, Cleft, Relative, and Load for Comprehension Question Reaction Times for Cleft + Relatives (1st Verb)

3.3.3 Analysis for Comprehension Question Reaction Times for Cleft + Relatives (2nd Verb)

The comprehension question reaction times focusing on the second verb in cleft + relative structures were analyzed in a 2 (high WM span vs. low WM span) x 2 (object cleft vs. subject cleft) x 2 (object relative vs. subject relative) x 3 (no load vs. visual load vs. auditory load) ANOVA. One main effect emerged. Specifically, there was a main effect of relative, F(1,100)=166.944, p<.001. Sentences ending in a subject relative (M=2883.423ms) as opposed to an object relative (M=1878.813ms) led to significantly longer comprehension question reaction times. This finding was probably driven by the fact that the correct answer for structure ending in an object relative is found at the end of the sentence, directly before the verb in question.

In addition to the main effect of relative, two interactions were found. The first interaction was a two-way interaction between WM span and load, F(2,100)=3.148, p<.05.

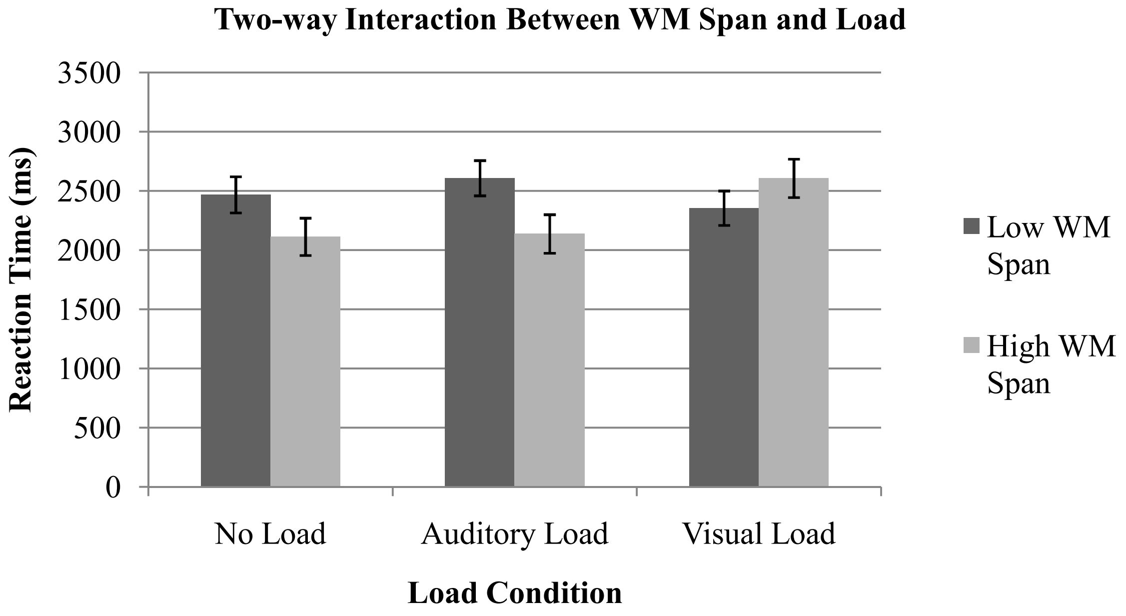


Figure 16: Two-way Interaction Between WM Span and Load for Comprehension Question Reaction Times for Cleft + Relatives (2nd Verb)

The two-way interaction was further developed with a three-way interaction between WM span, relative, and load, F(2,100)=5.315, p<.01.

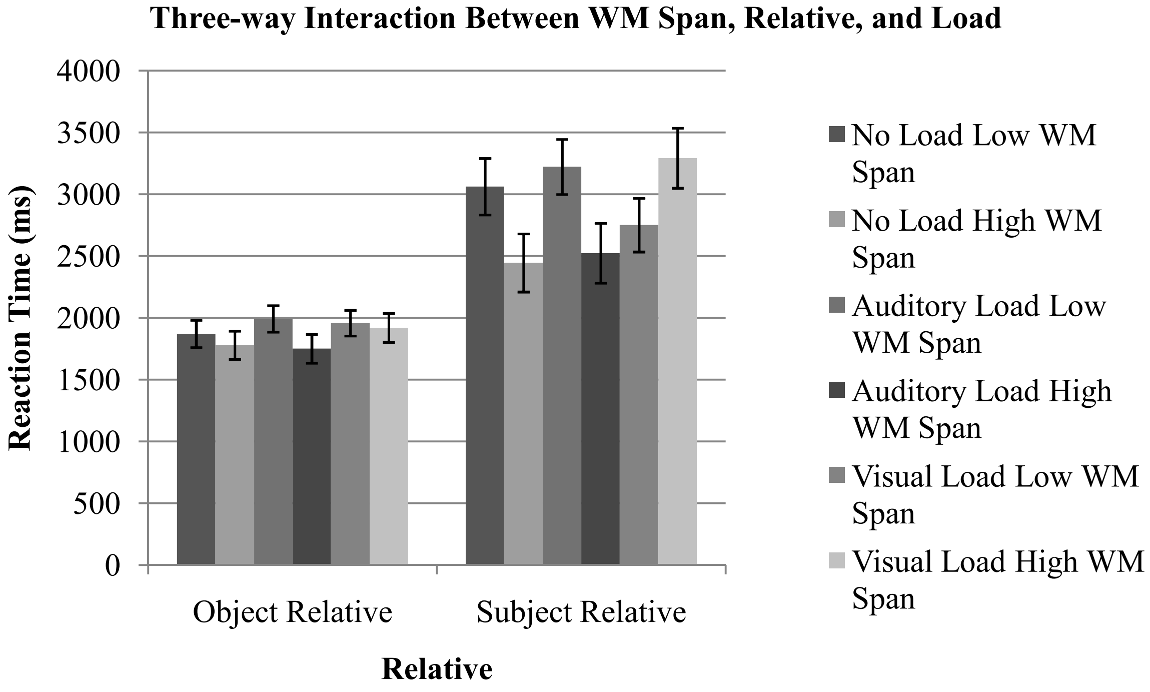


Figure 17: Three-way Interaction Between WM Span, Relative, and Load for Comprehension Question Reaction Times for Cleft + Relatives (2nd Verb)

In this interaction, we saw the reverse trend of high spans taking longer than low spans to answer the comprehension questions specific to the visual load condition. This reverse trend was seen for sentences ending in a subject relative. It seems that when under a high amount of working memory strain from both the sentence and the load, high spans were able to arrive to the correct answer, but they required additional processing time to do so. For a recap of all findings from the on-line critical area and comprehension question reaction times analyses, see Table 4 below.

Table 4: Summary of Critical Area and Comprehension Question Reaction Time Findings

	Single Cleft Critical Area	Single Cleft Comp Question RT	Cleft + Rel 1st Critical Area	Cleft + Rel 2nd Critical Area	Cleft + Rel Comp Question 1 RT	Cleft + Rel Comp Question 2 RT
Main effect of Cleft	YES	YES	YES	YES	YES	
Main effect of Relative			YES			YES
Main effect of WM Span	YES	YES	YES		YES	
Main effect of Load	YES		YES			
Cleft x Rel						
Cleft x WM					YES	
Cleft x Load			YES			
Rel x WM			YES			
Rel x Load					YES	
WM x Load						YES
Cleft x Rel x WM			YES			
Cleft x Rel x Load						
Cleft x WM x Load						
Rel x WM x Load						YES
Cleft x Rel x WM x Load					YES	

3.4 On-line Data

The on-line reaction time data were analyzed from sentences that had correct comprehension answers. Before beginning the analysis, the processing time for each individual sentence fragment was computed. To do this, the actual length of each auditory clip was subtracted from the participants overall listening time for each fragment. Thus, we could see how long a participant took before choosing to advance the sentence.

I predicted that longer listening times would be seen for object cleft and object relative constructions than for corresponding subject cleft and subject relative constructions. Figure 7 shows the predicted listening time patterns for each sentence type. In addition, if my hypotheses were supported, I expected to see an interaction between working memory and syntactic complexity. I expected to see, overall, low spans listening markedly longer than high spans, but particularly for the harder constructions. Further, compared to the rest of the sentence, I expected the critical areas would show increases in listening times, but particularly for low span individuals. Specifically, we identified two types of critical areas: the end of the cleft, identified as phrase “4” in Figures 20 thru 26 below, and the end of the sentence, identified as phrase “6”. In the case of the single clefts, there is only one critical area, given that the end of the cleft and the end of the sentence would be the same. All other sentences would have two critical areas. However, for sentences beginning with a subject cleft, we expected to only really see one peak in listening time at the end of the sentence. Contrarily, for sentences beginning with an object cleft, we expected to see a peak in listening time at the end of the cleft as well as the end of the sentence. Figure 18 gives an example of what data would look like provided an interaction with working memory span is present for object cleft comprehension. Figure 19 demonstrates what the data were expected to resemble if an interaction with working memory was not found for an object cleft, reminiscent of Waters and Caplan’s findings (2001).

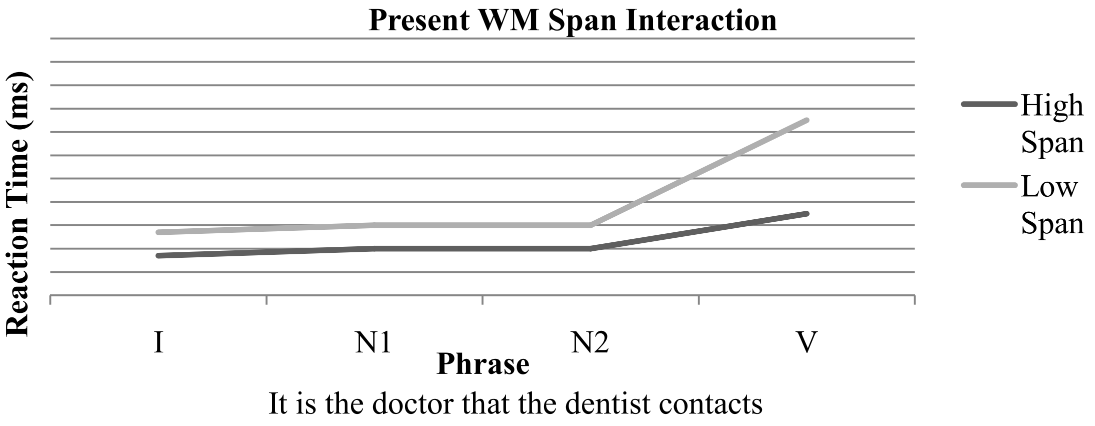


Figure 18: Hypothesized Interaction With WM Span for Object Cleft Comprehension

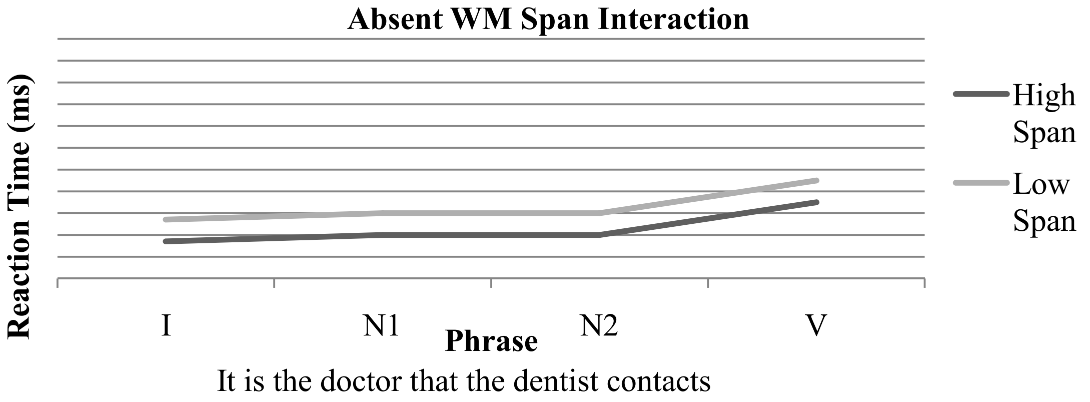
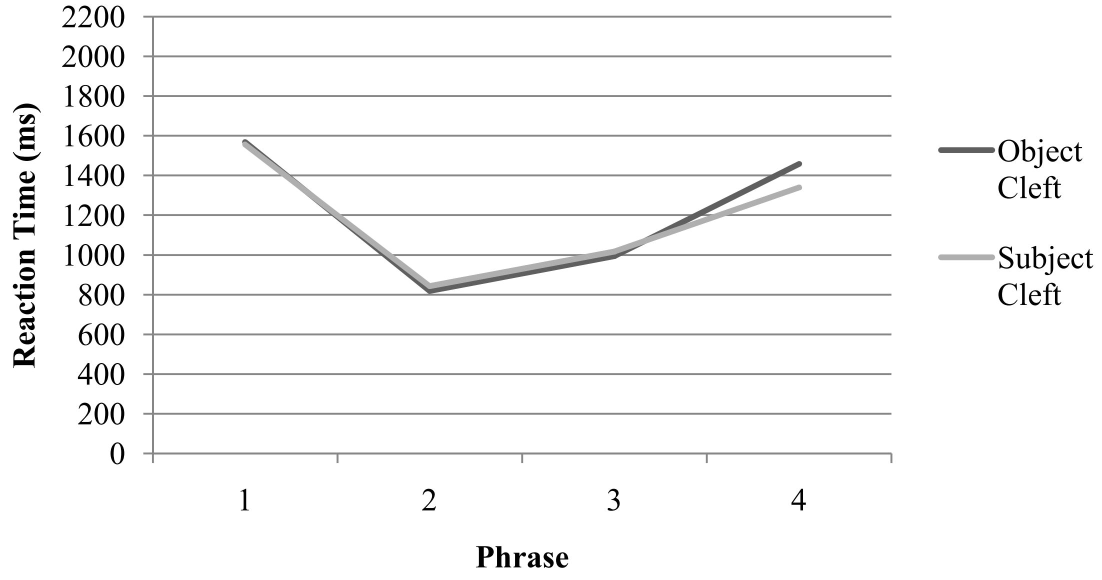


Figure 19: Object Cleft Comprehension Without WM Interaction

Before turning to the analyses, a general sense of the data can be gathered through inspection of graphs of the listening times. The listening times by sentence type, collapsed over WM span and load are shown below in Figure 20.

A.) **Single Cleft Sentence Trajectories**



B.) **Cleft + Relative Sentence Trajectories**

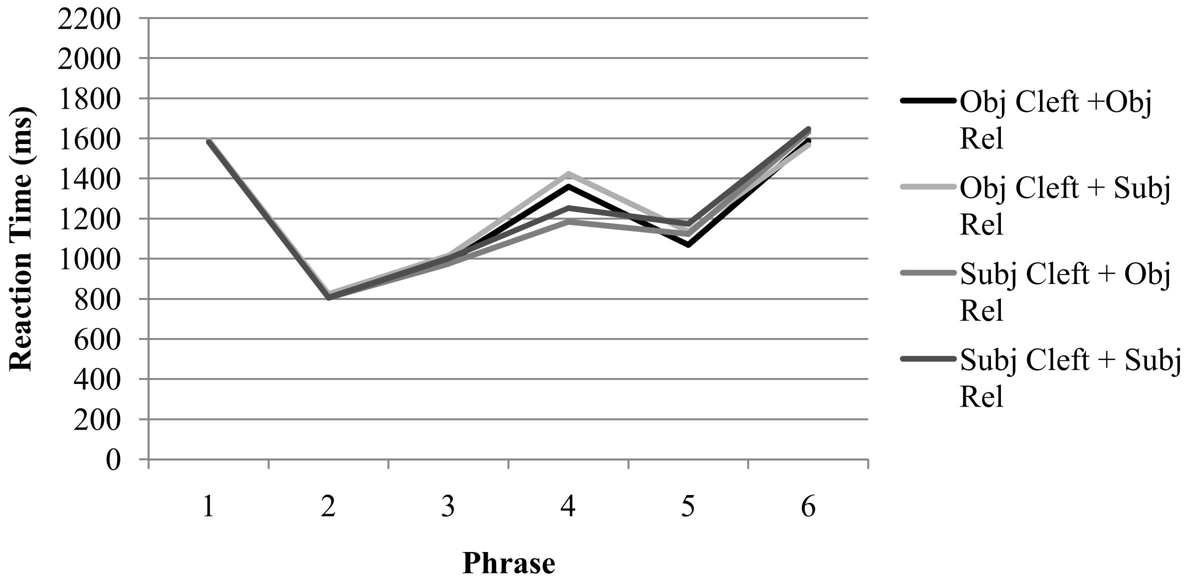


Figure 20: General Sentence Trajectories For All Sentences, Collapsed Over WM Span and Load

Figures 21 thru 26 below show this same data for each working memory span group (high WM span vs. low WM span) by external memory load condition (no load vs. auditory load vs. visual load) for each sentence structure.

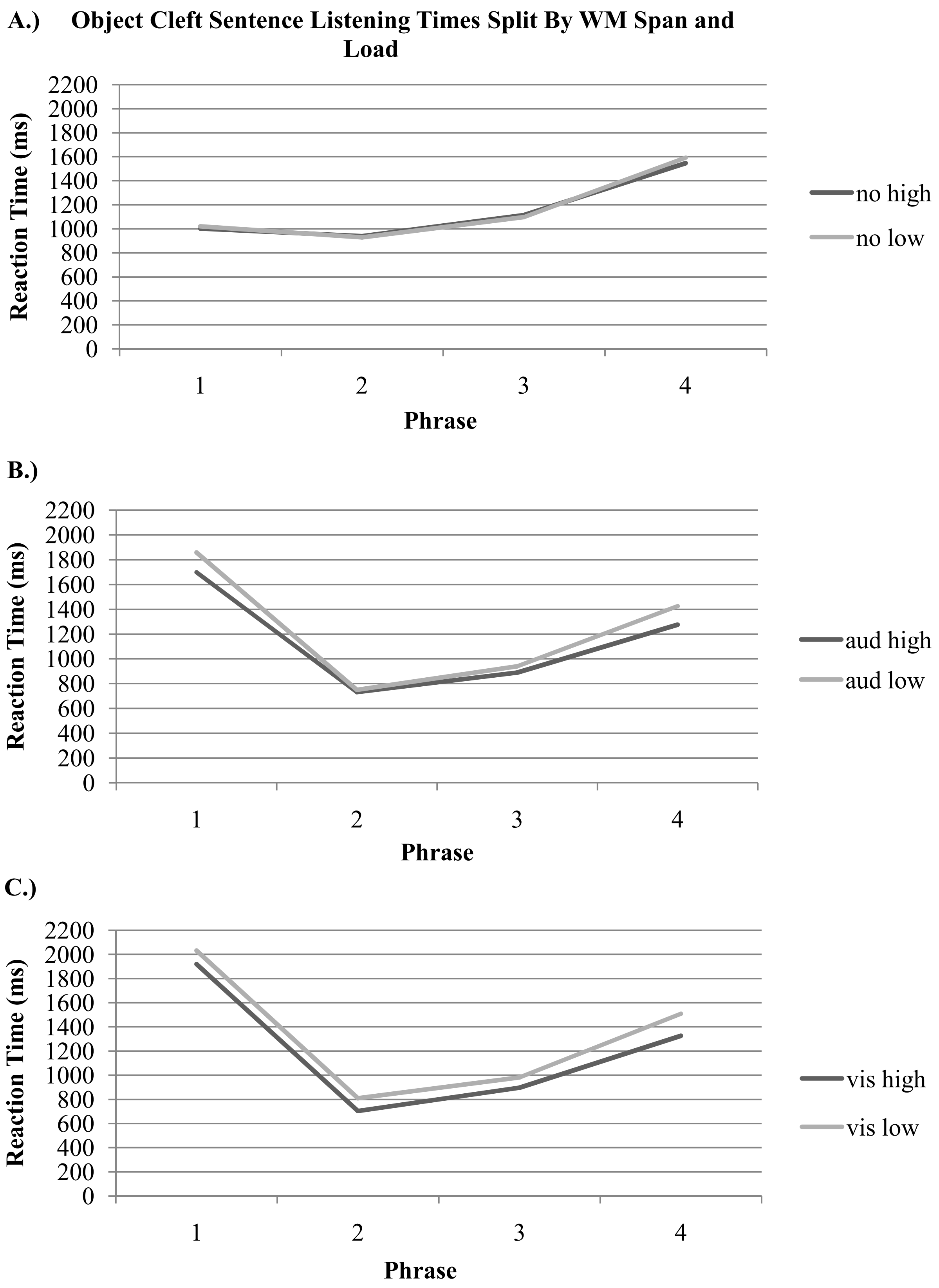


Figure 21: Object Cleft Sentence Listening Times Split by WM Span and Load

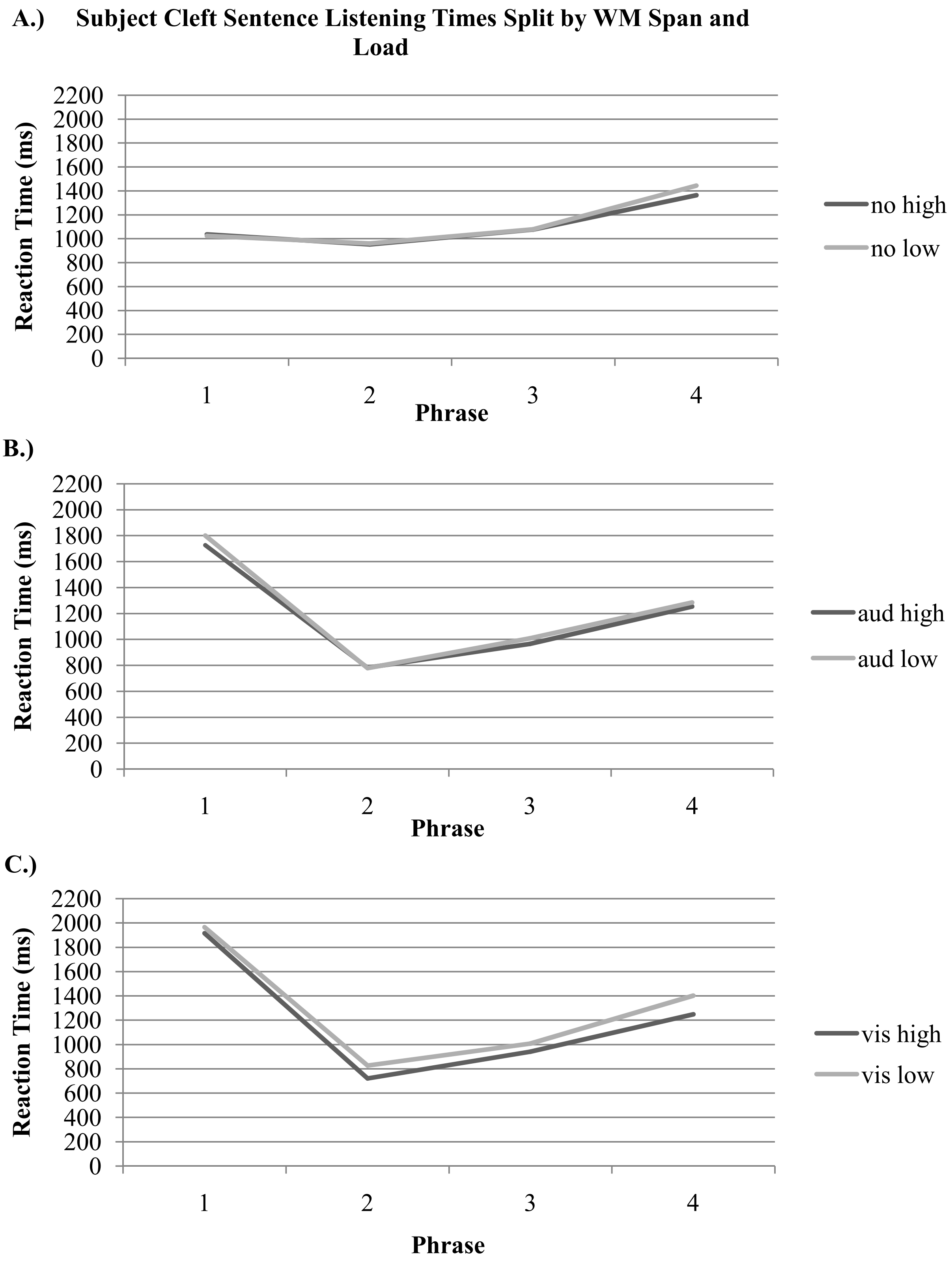


Figure 22: Subject Cleft Sentence Listening Times Split by WM Span and Load

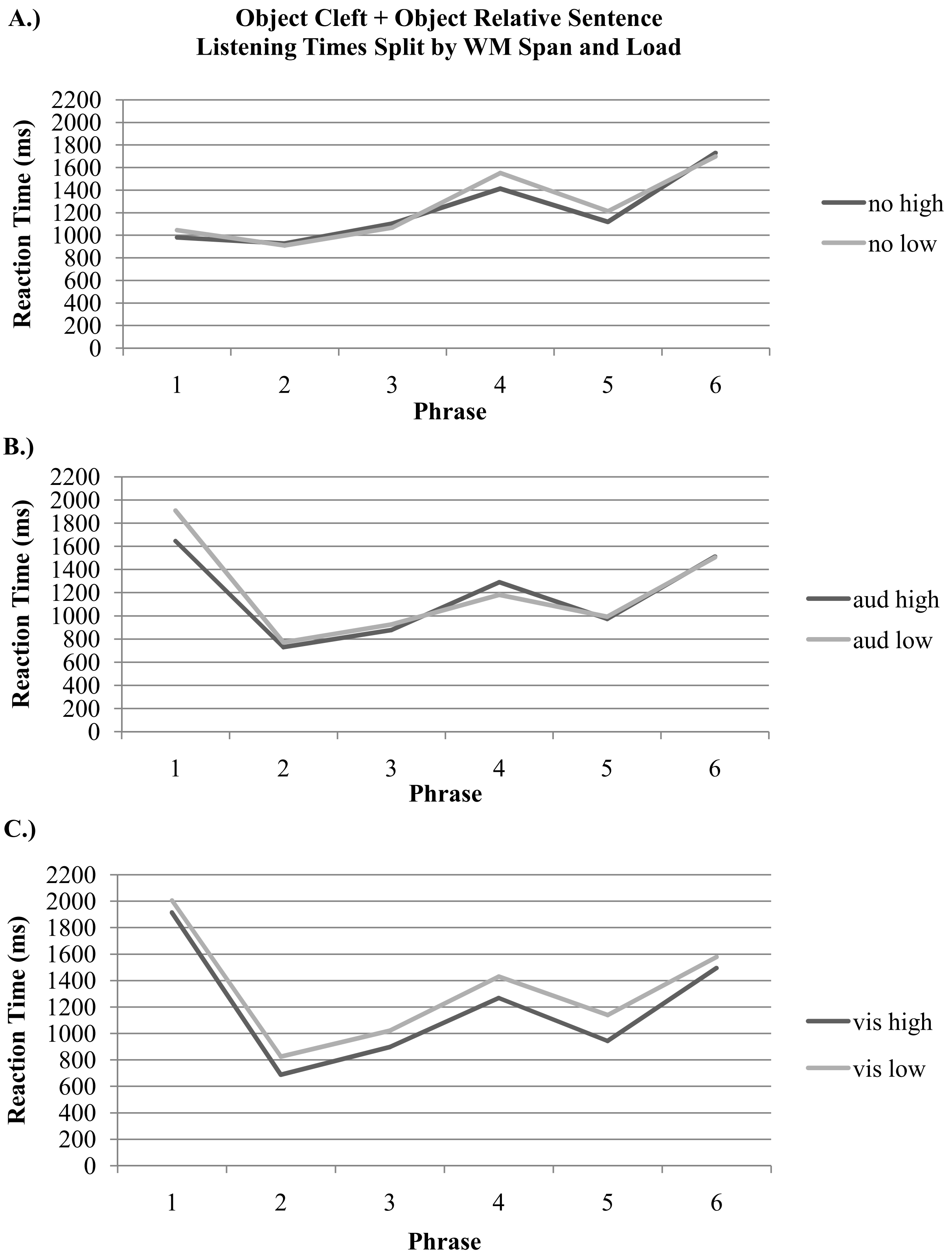


Figure 23: Object Cleft + Object Relative Sentence Listening Times Split by WM Span and Load

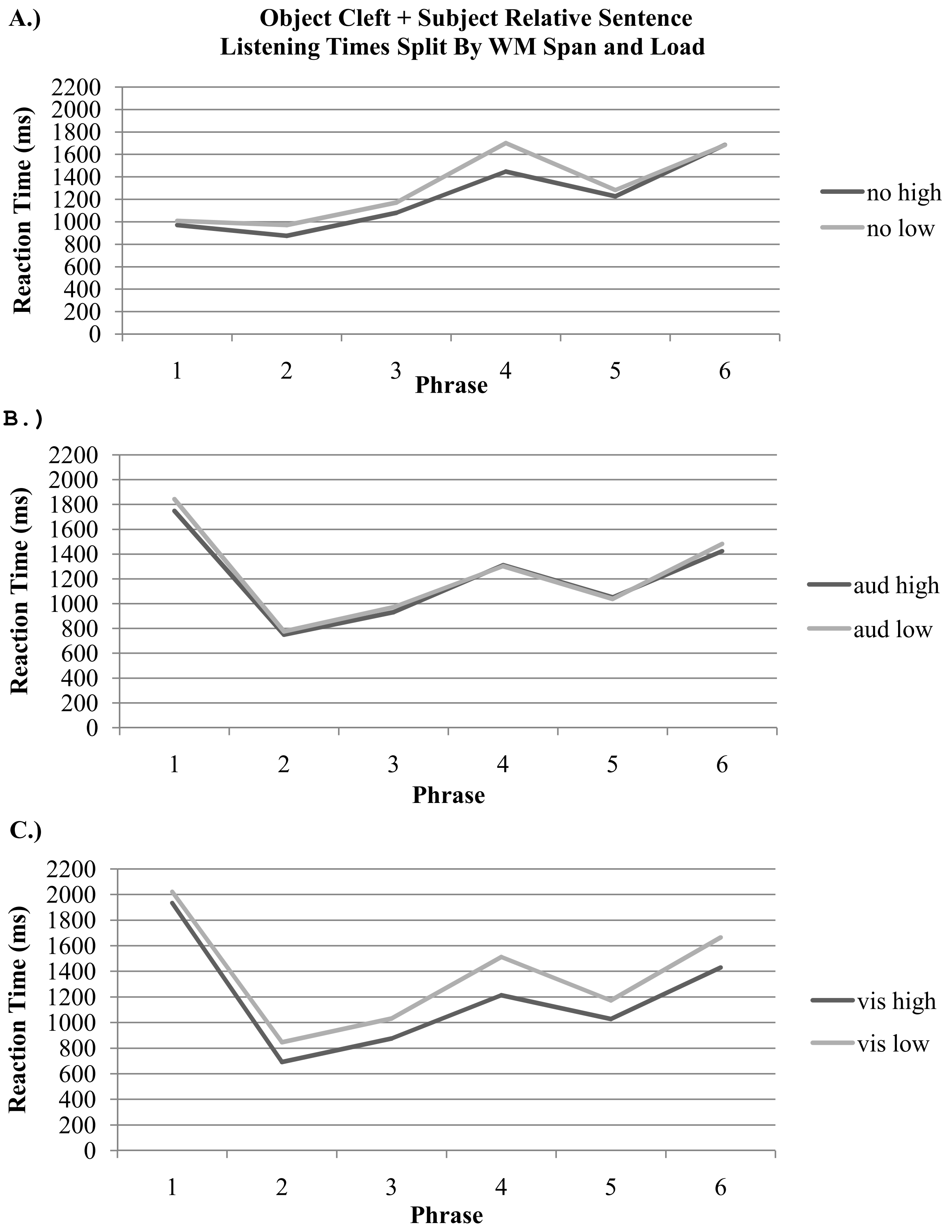


Figure 24: Object Cleft + Subject Relative Sentence Listening Times Split by WM Span and Load

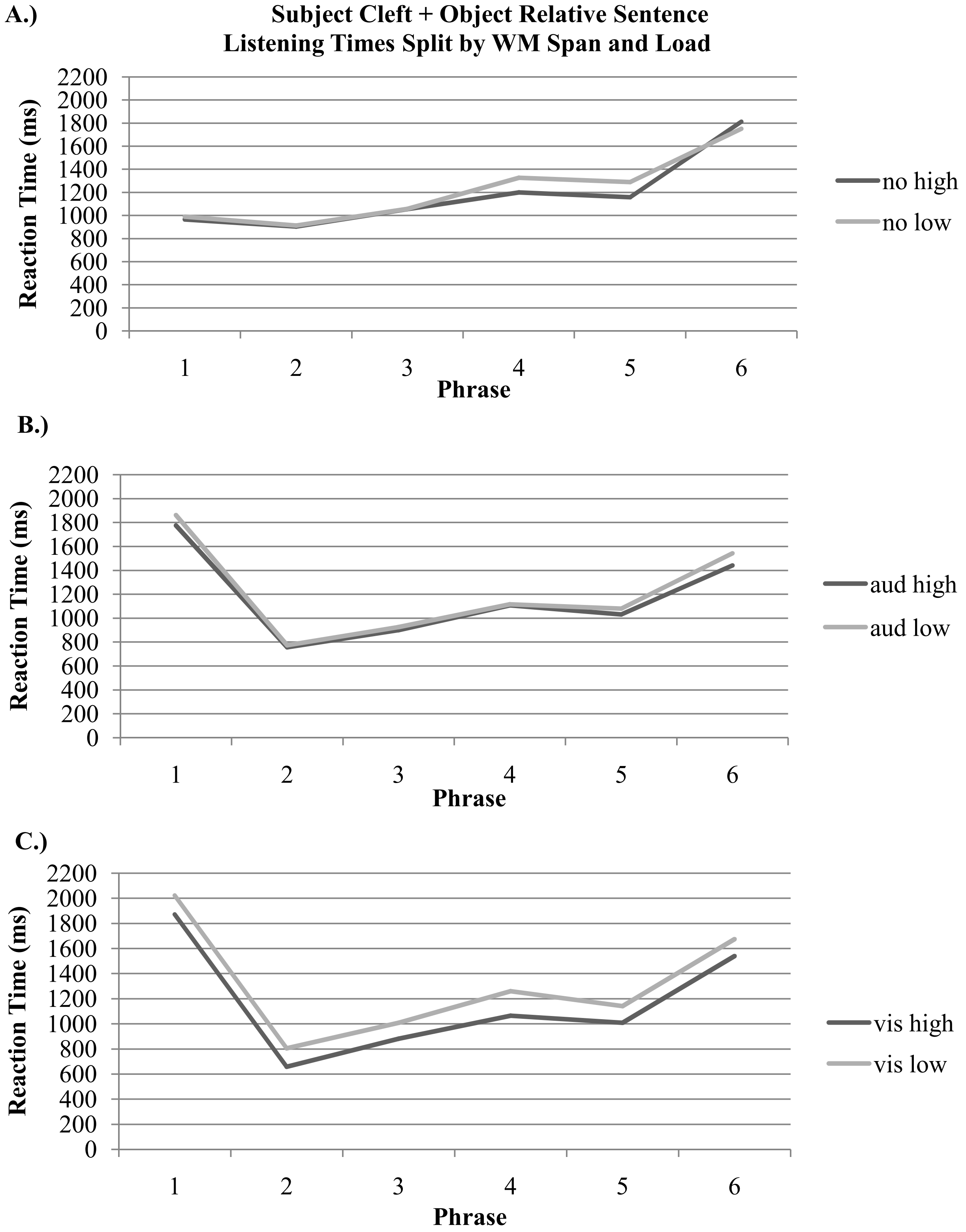


Figure 25: Subject Cleft + Object Relative Sentence Listening Times Split by WM Span and Load

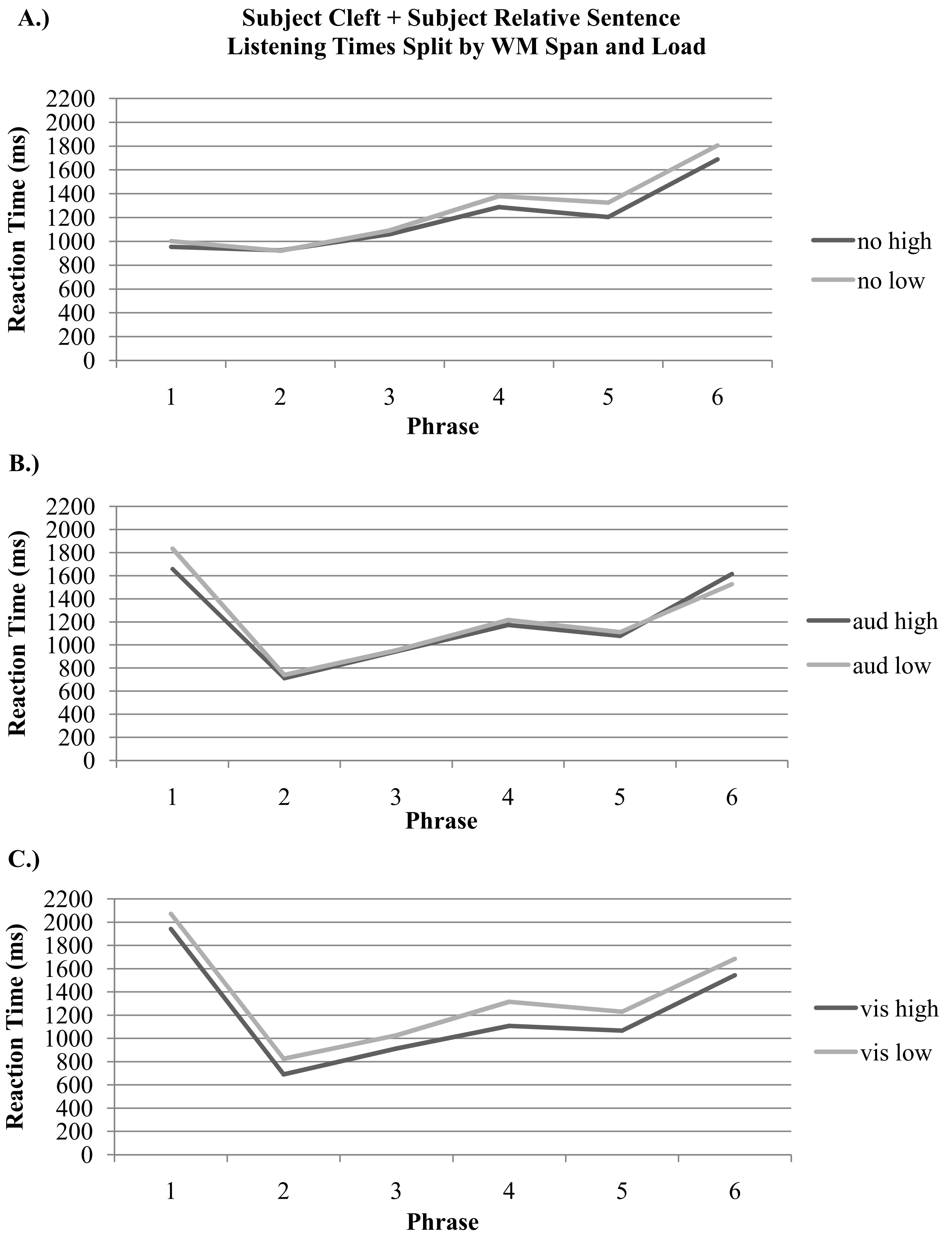


Figure 26: Subject Cleft + Subject Relative Sentence Listening Times Split by WM Span and Load

We were able to note a few general trends upon visually examining the graphs above. Concerning working memory's role, we saw spans listening longer to the sentence fragments than high spans for all sentence types. Also, when comparing the additional working memory load conditions, it was the trend of those in the no load condition to take longer during the sentences than in the auditory or visual load conditions. Individuals in the load conditions may have opted to click faster at the end of the sentence to reach the comprehension question quicker. In this way, the additional working memory load may have forced participants to choose to defer processing to off-line because it may be easier. Thus, as a general trend, the no load condition seems to show the most promise in offering evidence for working memory's role in on-line processing.

3.4.1 On-line Analysis for Sentence Listening Times: Critical Area of Single Clefts

For simplification, we honed our analysis in on the predicted critical areas—for single subject clefts versus object clefts, the critical area is phrase 4. We performed a 2 (low WM span vs. high WM span) x 2 (subject cleft vs. object cleft) x 3 (no load vs. auditory load vs. visual load) ANOVA. Three main effects emerged. First, there was a main effect of working memory span, F(1,120)=5.988, p<.05, with high spans (M=1335.648ms) being faster at the critical area than low spans (M=1443.403ms). Second, there was a main effect of cleft, F(1,120)=33.724, p<.001. As predicted, the critical area for object clefts (M=1446.353ms) had longer listening times than for subject clefts (M=1332.698ms). Finally, there was a main effect of load, F(2,120)=5.651, p<.01. Individuals in the no load condition (M=1486.796ms) took longer at the critical area when compared to individuals both in the visual load (M=1371.748ms) or auditory load conditions (M=1310.032ms). This perhaps reflects aforementioned processing strategies being used in the load conditions versus the no load condition. We will continue to see this trend in the other, harder sentence types as well, which mount evidence for this possibility.

3.4.2 On-line Analysis for Sentence Listening Times: First Critical Area of Cleft + Relative Structures

Data from cleft + relative structures focusing on the first verb and the second verb were combined across question type for the analysis. Because participants did not know which question they would be asked during the presentation of the cleft + relative structures, no difference should have been found in the critical area reaction times.

We analyzed the first critical area of the cleft + relative structures--the end of a cleft (phrase 4)--in a 2 (high WM span vs. low WM span) x 2 (subject cleft vs. object cleft) x 2 (subject relative vs. object relative) x 3 (no load vs. auditory load vs. visual load) ANOVA. We noticed four main effects and a series of interactions. First, there was a main effect of working memory span, F(1,101)=5.622, p<.05, with low spans (M=1357.770ms) having longer listening times than high spans (M=1239.926ms). Second, as expected, there was a main effect of cleft, F(1,101)=98.376, p<.001. Longer listening times were seen at the end of object clefts (M=1384.799ms) than for subject clefts (M=1212.907ms). Third, there was a main effect of relative, F(1,101)=26.909, p<.001. Longer listening times at the first critical area were seen for sentences ending in subject relatives (M=1330.629ms) compared to object relatives (M=1267.077). This is a curious finding since at the first critical period, the relative had not yet been revealed. Therefore, no main effects or interactions with relative should have occurred. Finally, there was a main effect of load, F(2,101)=5.788, p<.01. There were longer listening times for the no load condition (M=1413.261ms) when compared to both auditory (M=1211.982ms) and visual loads (M=1271.315ms).

There were also two interactions that involved the relative clause. There was a two-way interaction between WM span and relative, F(1,101)=6.266, p<.05, seen in Figure 27. Low spans had longer critical area listening times than high spans, but particularly when the sentence ended in a subject relative. There was also a three-way interaction between WM span, cleft, and relative, F(1,101)=4.473, p<.05, seen in Figure 28. In this interaction, we saw the biggest gap in listening times between the low and high WM span individuals exist for the object cleft + subject relative structure. Because no interactions with relative should exist, no real interpretation of these interactions is possible.

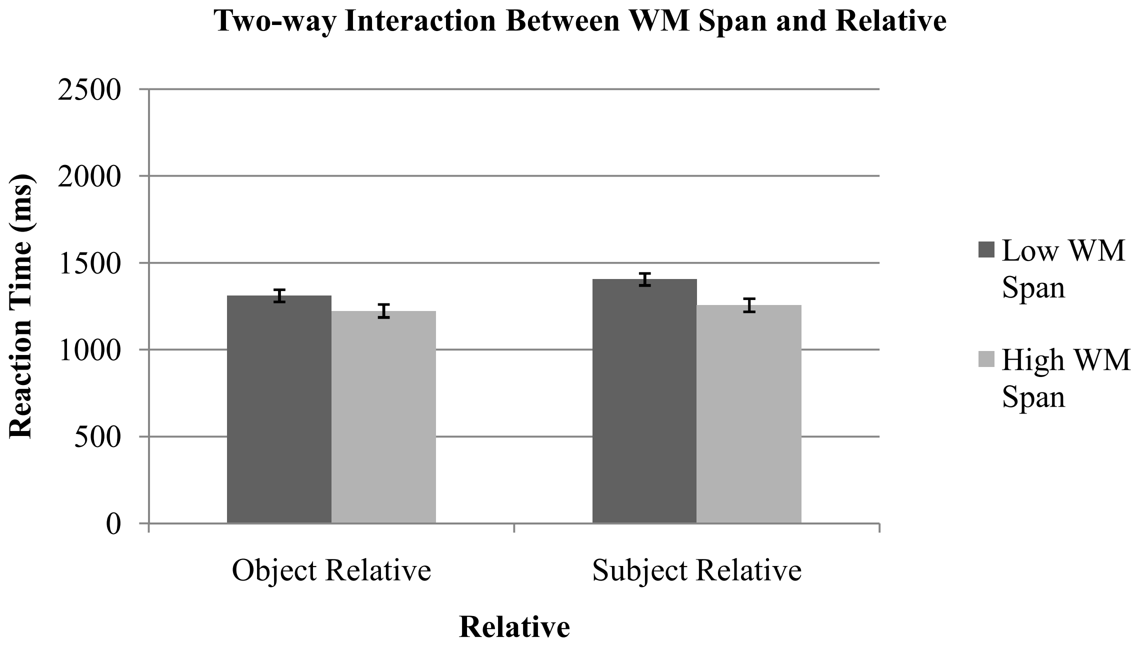


Figure 27: Two-way Interaction Between WM Span and Relative for First Critical Area

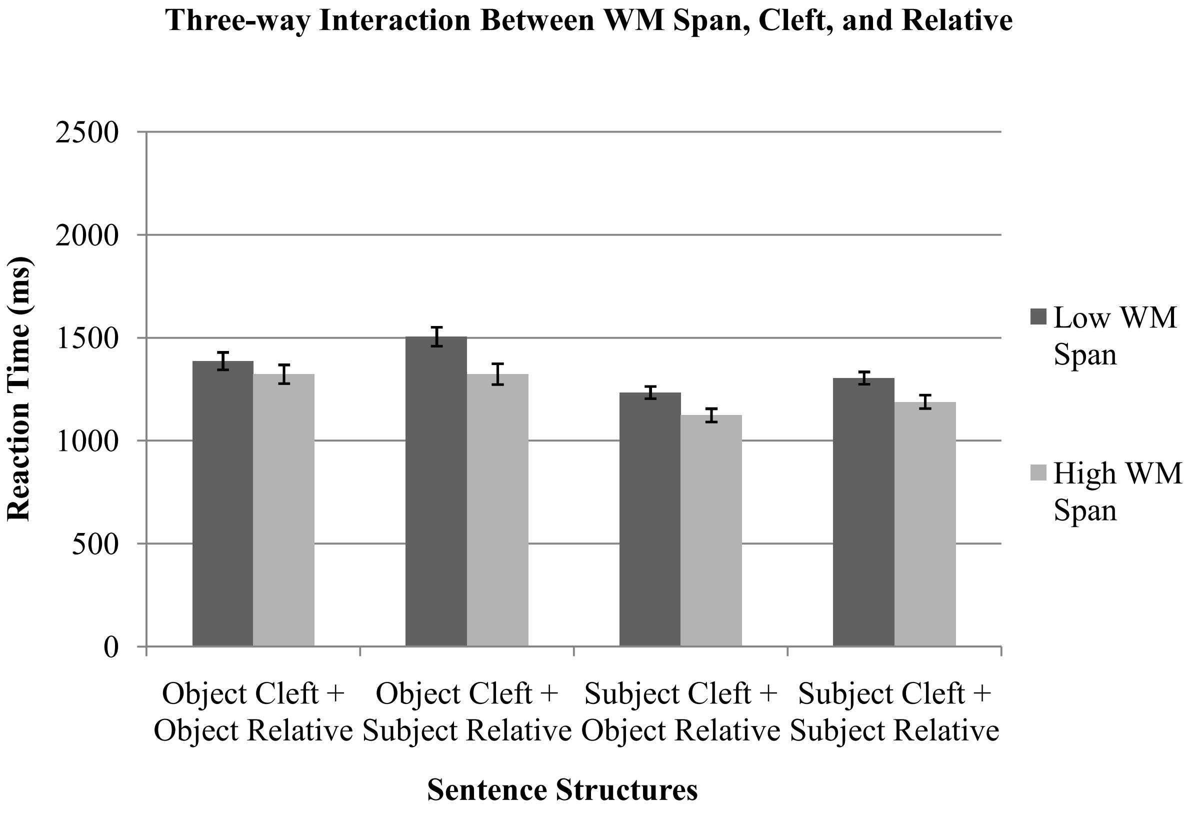


Figure 28: Three-way Interaction Between WM Span, Cleft, and Relative for First Critical Area

There was a two-way interaction between cleft and load, F(2,101)=3.524, p<.05, which could be explained by the aforementioned processing strategy.

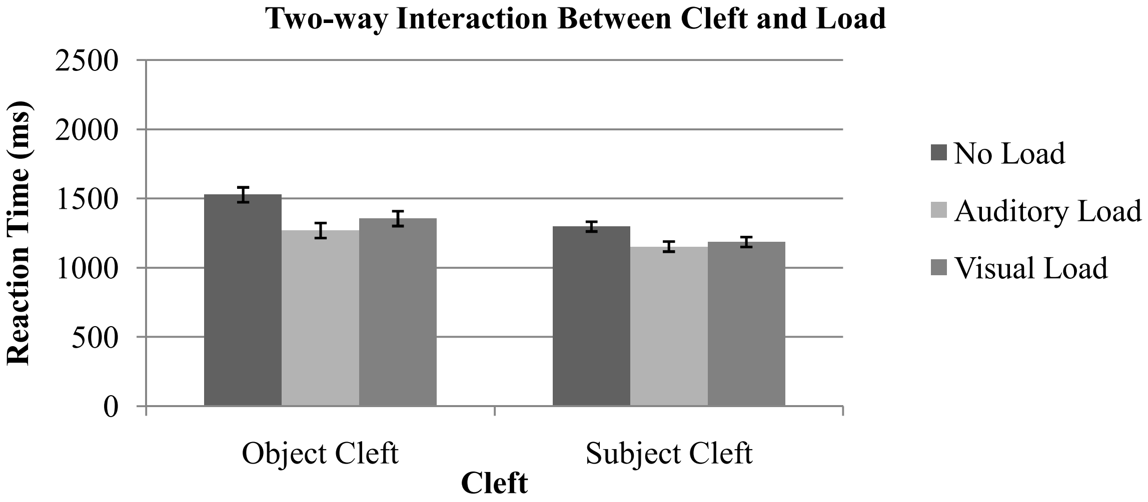


Figure 29: Two-way Interaction Between Cleft and Load for First Critical Area

In this interaction, we saw longer critical area listening times at the end of object clefts as opposed to subject clefts. However, the individuals in the no load condition, in particular, had exceptionally longer listening times at the end of object clefts. To some degree, online processing had occurred since syntactic complexity was impacting listening times. What is curious is the fact that the load conditions had shorter listening times when compared to the no load condition. This suggests that individuals in the load conditions were strained by the additional load and were clicking through the sentence faster, opting to process off-line.

3.4.3 On-line Analysis for Sentence Listening Times: Second Critical Area of Cleft + Relative Structures

The second critical area (phrase 6) of the cleft + relative structures was analyzed in a 2 (object cleft vs. subject cleft) x 2 (object relative vs. subject relative) x 2 (high span vs. low span) x 3 (no load vs. auditory load vs. visual load) ANOVA. The only significant result of this analysis was a main effect of cleft, F(1,101)=9.711, p<.005. The second critical area was longer when sentences began with subject cleft (M=1635.650ms) as opposed to an object cleft (M=1573.629ms). Most likely, this is driven by the ambiguous subject cleft + subject relative structure. This shows that how a sentence begins affects even the later critical areas. Because this was the only significant result of this analysis, this adds support to the idea that individuals are waiting until receiving the comprehension question before opting to process. Thus, we have support for a strategy to defer processing to off-line.

3.4.4 On-line Analysis for Sentence Listening Times: Overview

Overall, in analyzing the critical areas, we found that syntactic complexity impacted processing, as different structures led to longer listening times. Generally, object clefts led to longer listening times, with the exception of the effect of the object clefts on the second critical area listening time. Also, we found that low WM span individuals listen longer than high WM span individuals. However, more interesting than what was found in the critical area analyses is what was not found. Specifically, we did not find any interactions between syntactic complexity and working memory span. This suggests that working memory does not play a role in on-line processing.

3.5 On-line Analysis of Critical Areas for Only the No Load Condition Including All Participants

All of the previous analyses included load as a between subjects factor. However, it could be beneficial to focus our analyses on only the no-load condition. Because this condition has no additional load, it would be the least likely to be affected by some task performance strategy. In addition, when looking at the sentence trajectories divided by span and load earlier, the no load condition in particular appeared to be the most likely to show any possible interactions between sentence complexity and working memory. Refer back to Figures 18 thru 23. All no load participants were included in the following analyses, and the same types of analyses were run in analyzing the no load condition.

3.5.1 On-line Analysis for Sentence Listening Times of the No Load Condition: Critical Area of Single Clefts

First, we analyzed the critical area at the end of the single cleft structures. The same main effect of cleft emerged, F(1,44)=25.559, p<.001. However, there was no significant main effect of WM span, as seen when including the load conditions in the analysis. No main effect or interactions with working memory were found. Thus, we see that even for the simplest sentences in the no load condition, we fail to find working memory playing a role in on-line processing.

3.5.2 On-line Analysis for Sentence Listening Times of the No Load Condition: First Critical Area of Cleft + Relative Structures

Next, the first critical area of the cleft + relative structures was analyzed. When the results were compared to our earlier findings that included all load conditions, the two main effects of cleft, F(1,34)=52.230, p<.001, and relative, F(1,34)=10.109, p<.01, were replicated. The main effect of working memory and all subsequent interactions with working memory were not found. Thus, we fail to find an interaction with working memory.

3.5.3 On-line Analysis for Sentence Listening Times of the No Load Condition: Second Critical Area of Cleft + Relative Structures

Finally, we analyzed the second critical area of the cleft + relative structures. Our findings replicated the results found when all load conditions were included in the analysis. Specifically, only a main effect of cleft was found, F(1, 34)=4.256, p<.05. Again, working memory fails to emerge as a significant factor in on-line sentence processing, even in the no load condition.

3.6 Analysis of Comprehension Question Reaction Times for Only the No Load Condition Including All Participants

As when analyzing all conditions together, we now shift our attention to the comprehension question reaction time data for all participants in only the no load condition.

3.6.1 Analysis for Comprehension Question Reaction Times for Single Clefts in the No Load Condition

The comprehension question reaction times for the cleft structures was analyzed using the same statistical tests as before. Mirroring the earlier findings including all loads, only two main effects were found. First, there was a main effect of working memory, F(1,44)=6.531, p<.05. High spans (M=1840.657ms) answered the comprehension question faster than low spans (M=2184.155ms). Secondly, there was a main effect of cleft, F(1,44)=21.737, p<.001. Questions after object clefts (M=2192.325ms) took longer to answer than questions after subject clefts (M=1823.486ms).

3.6.2 Analysis for Comprehension Question Reaction Times for Cleft + Relatives (1st verb) in the No Load Condition

The comprehension question reaction times for the cleft + relative structures, focusing on the first verb, were analyzed. The same two main effects were found. First, there was a main effect of working memory span, F(1,43)=4.929, p<.05. High spans (M=2069.345ms) continue to answer more quickly than low spans (M=2381.690ms). There was also a main effect of cleft, F(1,43)=16.142, p<.001, with longer comprehension question reaction times resulting for sentences beginning with object clefts (M=2356.369ms) over subject clefts (M=2094.665ms).

3.6.3 Analysis for Comprehension Question Reaction Times for Cleft + Relatives (2nd verb) in the No Load Condition

Finally, the comprehension question reaction times for the cleft + relative structures, focusing on the second verb were analyzed. One significant main effect of relative, F(1,33)=70.075, p<.001, replicating earlier findings, was found. Also, there was a two-way interaction between WM span and relative, F(1,33)=5.607, p<.05.

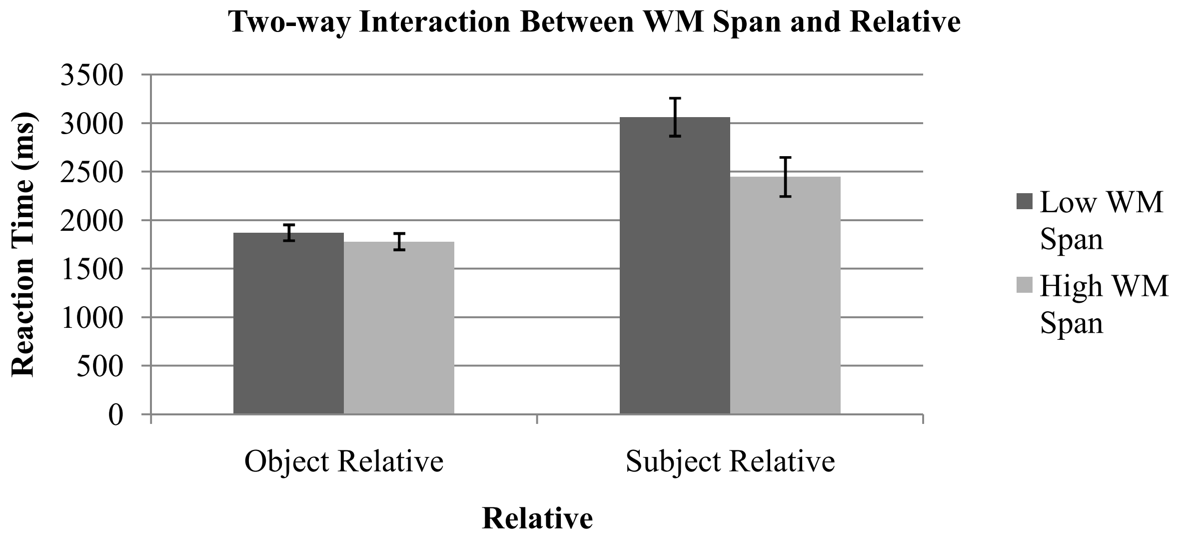


Figure 30: Two-way Interaction Between WM Span and Relative for Comprehension Question Reaction Times for Cleft + Relatives (2nd Verb) in the No Load Condition

Comprehension question reaction times for low spans were longer than high spans for sentences ending in a subject relative. However, high spans and low spans did not differ for object relatives. Even though we finally saw an interaction between WM span and syntactic complexity, comprehension question reaction times are arguably an off-line measures of processing.

3.7 On-line Analysis on Only the No Load Condition for Participants With Extreme WM Scores

Perhaps the reason WM did not emerge as a main effect or interaction was because participants were divided into high and low span groups by a median split. To try to verify our findings, the no load condition was then analyzed using only participants with extreme WM scores. However, by only looking at the extreme WM score cases of the no load condition, more participants were eliminated (Low Span: n=17, High Span: n=10). Results of these analyses showed that most findings were replicated. However, even when looking at only participants with extreme WM scores, no interactions between working memory and syntactic complexity emerged.

Of all the conditions, it could be assumed that the no load condition in particular was least effected by a clicking strategy. This is because the no load condition was free from additional WM strain which may, in turn, have encouraged participants to choose to process off-line. Therefore, we suspected that if any condition would show WM playing a role in on-line processing, it would have been the no load condition. In analyzing the no load condition, both including all participants and only those with extreme WM scores, we failed to see working memory impact processing. It is possible that WM simply does not play a role in syntactic processing. However, a clicking or processing strategy may still have existed, even for the no load condition since the sentences were inherently difficult in and of themselves. Or another explanation could be that it is possible that WM did in fact play a role, but not strongly enough to have been noticed with our methods.

3.8 On-line Analysis on Only the No Load Condition for First 1/3rd of Experiment for All Participants

As stated earlier, of all three conditions, it would be expected that the no load condition would be the least likely affected by a clicking strategy. If a clicking strategy existed in the no load condition as well, it probably would be most evident in the later trials of the experiment. Therefore, an analysis was conducted, honing in on the first one third of the trials in the no load condition. Again looking for any interactions between WM span and syntactic complexity, we examined all critical areas for all sentence structures. No significant interactions were found, possibly due to large variability in the reaction times. However, upon examining the numerical trend, we see that in general, particularly for the single clefts and the first critical area of the cleft + relative structures, low spans tend to show more difference in reaction time between object and subject structures than do high spans.

Table 5: Reaction Times for the Critical Areas of All Structures for the First 1/3rd of the No Load Condition Experiment

Single Clefts	Low WM Span	High WM Span
Object Clefts	1708.3ms Std Error: 79.1	1539.5ms Std Error: 79.1
Subject Clefts	1491.9ms Std Error: 56.2	1360.5ms Std Error: 56.2
Cleft + Rel-Critical Area 1	Low WM Span	High WM Span
Object Cleft + Object Relatives	1886.1ms Std Error: 97.4	1740.6ms Std Error: 97.4
Object Cleft + Subject Relatives	1481.0ms Std Error: 67.6	1427.9ms Std Error: 67.6
Subject Cleft + Object Relatives	1311.8ms Std Error: 62.9	1290.9ms Std Error: 62.9
Subject Cleft + Subject Relatives	1428.630 Std Error: 60.2	1447.870 Std Error: 60.2
Cleft + Rel-Critical Area 2	Low WM Span	High WM Span
Object Cleft + Object Relatives	1853.4ms Std Error: 119.9	1841.8ms Std Error: 119.9
Object Cleft + Subject Relatives	1837.1ms Std Error: 130.5	1842.7ms Std Error: 130.5
Subject Cleft + Object Relatives	1786.4ms Std Error: 111.3	1879.8ms Std Error: 111.3
Subject Cleft + Subject Relatives	2039.7ms Std Error: 141.0	1807.2ms Std Error: 141.0

3.9 Analyzing the Additional Working Memory Load: Performance on the Load Task

Now, we shift our attention to performance on the secondary load task. As a recap, participants, prior to hearing a cleft or cleft + relative sentence, were given either a visual or auditory stimulus to remember. This additional stimulus was either a series of 8 digits for the auditory condition or a pattern of 8 stars for the visual condition. After the participant heard a sentence and answered a comprehension question about that sentence, a second visual or auditory stimulus would appear. The additional WM load task was to identify whether the two strings of digits or star patterns were the same or different.

First, we analyzed the off-line performance (percent correct) of the additional load task. Multiple one sample t-tests were run on each sentence type within each load condition to determine if performance was above chance (50%). In the auditory load condition, performance on the load task for all the sentence types was above chance, with significance ranging from $p<.001$ to $p=.024$. In the visual load condition, performance was above chance for all sentence types save three: 1) object clefts + object relatives when asking about the second verb, $p=.153$, 2) subject cleft + object relatives when asking about the first verb, $p=.068$, and 3) subject cleft + object relatives when asking about the second verb, $p=.092$. This came as a surprise because it would be assumed that if performance on the additional load task would drop to chance levels, it would be on those trials coupled with the hardest sentences (object cleft + object relative structure when focusing on the first verb and the sentences ending with subject relatives when focusing on the second verb).

We ran a series of three analyses using the same designs to look at percentage of correct answers in the load conditions. The only difference in the analyses used was that the load factor consisted of two levels (auditory load vs. visual load) instead of three (no load vs. auditory load vs. visual load). We looked at performance on the load tasks coupled with single clefts, and then coupled with the cleft + relatives questioning the first and second verbs.

3.9.1 Additional Working Memory Load: Performance on the Load Task When Coupled With Single Clefts

We explored same-different judgments on the additional load task when paired with single clefts and no significant main effects or interactions were found.

3.9.2 Additional Working Memory Load: Performance on the Load Task When Coupled With Cleft + Relatives (1st Verb)

When looking at the same-different judgment performance coupled with the cleft + relative sentences questioning the first verb, a significant two-way interaction between WM span and relative, $F(1,76)=22.579$, $p<.001$.

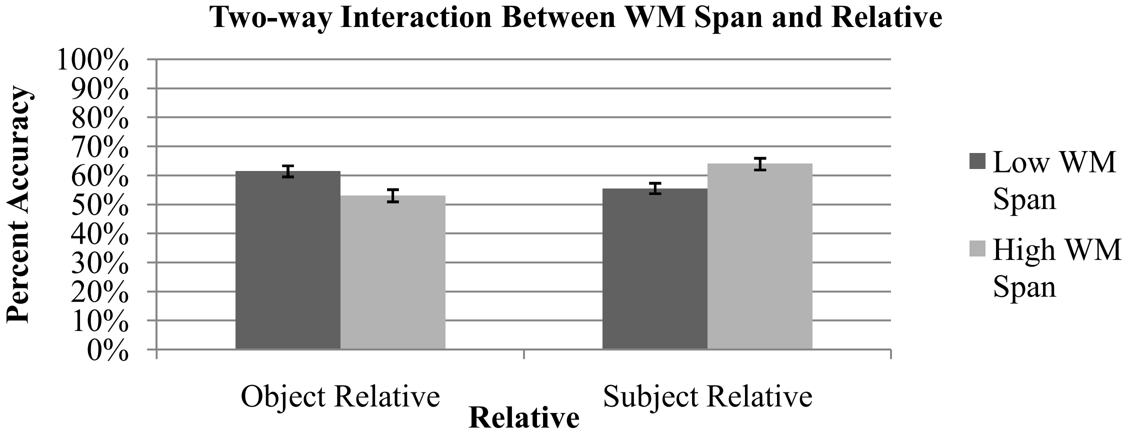


Figure 31: Two-way Interaction Between WM Span and Relative for Load Task Performance Coupled with Cleft + Relatives (1st Verb)

In this interaction, we saw high spans outperforming low spans when the sentence ended in a subject relative, but low spans outperformed high spans when the sentence ended in an object relative. While this interaction is puzzling, it may not be particularly meaningful. This is because the sentence comprehension question in focus revolved around the first verb, which concerned only the cleft, and the interaction involved only the relative.

3.9.3 Additional Working Memory Load: Performance on the Load Task When Coupled With Cleft + Relatives (2nd Verb)

Finally, we looked at the same-different judgment performance coupled with the cleft + relative structures focusing on the second verb. Again, only one significant result emerged. Specifically, there was a main effect of load, F(1,76)=11.374, p<.01, with higher performance on the same-different judgments seen in the auditory condition (M=.627) over the visual condition (M=.556). In addition to evidence seen earlier, this supports the idea that the visual condition was harder than the auditory condition.

When analyzing the load task performance, one might assume that a trade off was occurring between the comprehension question and the same-different judgments. To see if this was indeed occurring, the data was split by load condition and a bivariate correlation was run. Only two significant correlations, one in each of the load conditions, were found between accuracy of comprehension question and accuracy of same-different judgments. Both positive correlations showed the same value, r=.4, and were for the subject cleft + object relative (2nd verb). Thus, participants who correctly answered the comprehension question focusing on the subject cleft + object relative structure (2nd verb) were also likely to correctly answer the corresponding same-different judgment. The lack of other significant correlations tends to suggest no trade off is occurring.

3.10 Analyzing the Additional Working Memory Load: Same-Different Reaction Times on the Load Task

The reaction times for the same-different judgments was analyzed, in the same vein as earlier analyses, for the single clefts, and then for the cleft + relative structures questioning the first and second verb, respectively.

3.10.1 Same-Different Reaction Times on the Load Task for Single Clefts

When analyzing the load task reaction times for the single clefts, two main effects were seen. First, there was a main effect of cleft, F(1,76)=17.631, p<.001. As expected, same-different judgment RTs were longer when coupled with object clefts (M=1910.817ms) instead of subject clefts (M=1738.791ms). The second main effect was of load, F(1,76)=16.265, with reaction times being longer for the visual (M=2015.444ms) over the auditory load (M=1634.164ms).

3.10.2 Same-Different Reaction Times on the Load Task for Cleft + Relatives (1st Verb)

Next, the reaction times for the load task coupled with the cleft + relative structures, questioning the first verb, was analyzed. One significant main effect and two significant interactions emerged. First, there was a main effect of load, F(1,76)=6.206, p<.05. Individuals had longer same-different judgment RTs in the visual condition (M=2023.296ms) over the auditory condition (M=1748.101ms). Next, there was a two-way interaction between cleft and relative, F(1,76)=5.942,p<.05, seen in Figure 32. We saw the longest RTs coupled with the object cleft + subject relative sentence (M=1945.806ms), followed by object cleft + object relative (M=1908.234ms) and subject cleft + object relative structures (M=1909.724ms) equivalently, and finally by subject cleft + subject relative structures (M=1779.030ms). The fact that the longest RTs were coupled with the slightly easier object cleft + subject relative structure, instead of the harder object cleft + object relative, may suggest a trade off of sorts. For the hardest structure, individuals may have spent longer on the sentence comprehension question and less on the load judgment. For easier structures, however, participants might have exhibited the reverse trend. The second interaction is a four-way interaction between WM span, cleft, relative, and load, F(1,76)=5.045, p<.05, seen in Figure 33.

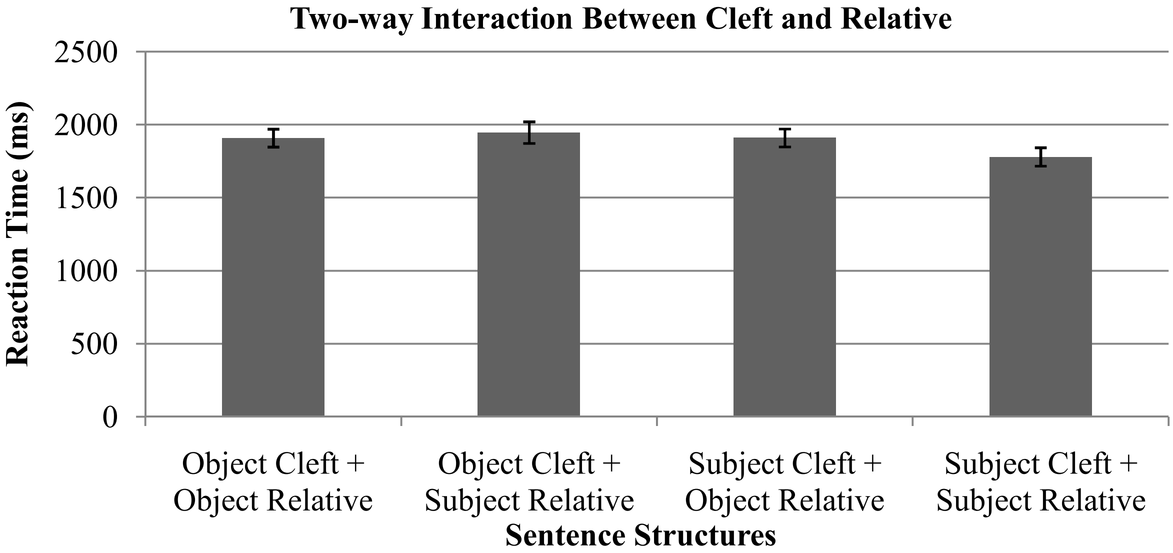


Figure 32: Two-way Interaction Between Cleft and Relative for Load Task Reaction Times Coupled with Cleft + Relatives (1st Verb)

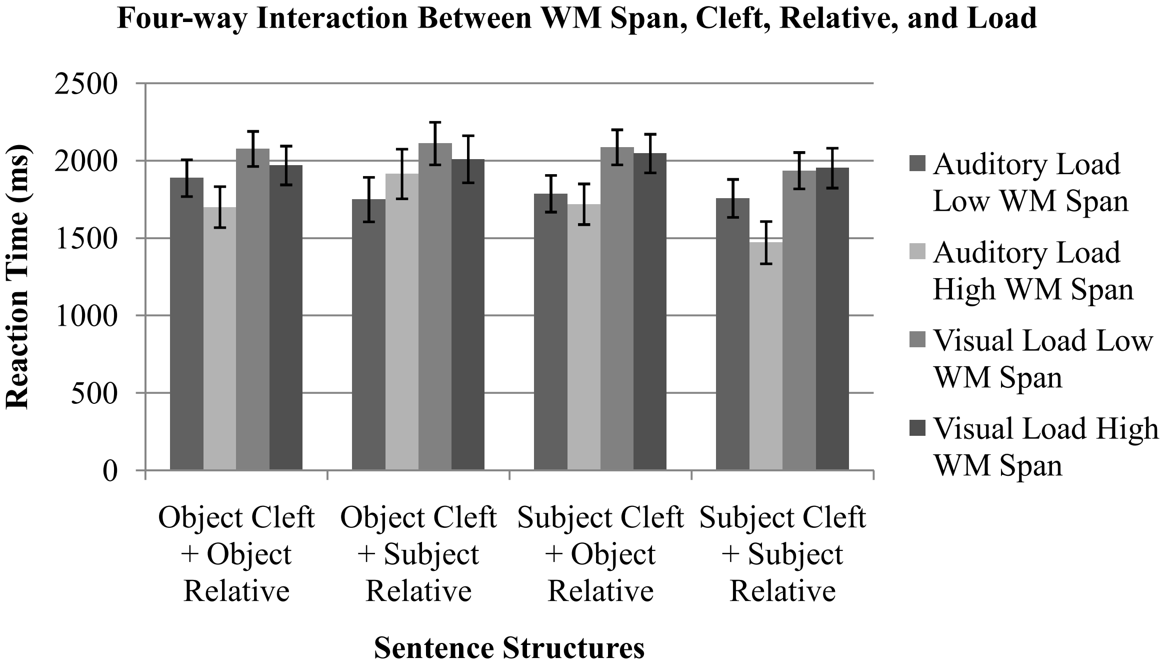


Figure 33: Four-way Interaction Between WM Span, Cleft, Relative, and Load for Load Task Reaction Times Coupled With Cleft + Relatives (1st Verb)

What is interesting to note from this reaction is how the high spans in the visual condition had longer reaction times than the low spans in the auditory condition. Again, this shows the degree of difficulty between the visual and auditory loads.

3.10.3 Same-Different Reaction Times on the Load Task for Cleft + Relatives (2nd Verb)

Lastly, a 2 (subject cleft vs. object cleft) x 2 (subject relative vs. object relative) x 2 (high WM span vs. low WM span) x 2 (auditory load vs. visual load) ANOVA was run to analyze the load reaction times paired with cleft + relative structures questioning the second verb. Only two main effects were seen. The first main effect was of relative, F(1,76)=12.265,p<.05. Sentences that ended with subject relatives (M=2035.060ms) instead of object relatives (M=1887.332ms) led to longer load task reaction times. This main effect is mostly likely is driven by the difficult object cleft + subject relative and ambiguous subject cleft + subject relative structures. Another possibility is that the sentence ending with an object relative was easier to answer, and this may have in turn carried over onto the load task. The second main effect was of load, F(1,76)=6.505, p<.05. Again, we consistently see the visual load same-different questions (M=2105.359ms) taking longer to answer than those concerning the auditory load (M=1817.034ms).

Overall, the visual load proved to be harder than the auditory load. When comparing the performance of the load tasks to chance, the visual condition was shown to be harder.
This finding was also seen both in terms of accuracy, specifically when analyzing percent correct for the cleft + relative structures (2nd verb), and in terms of reaction times for all sentence types.

There is the possibility that a trade of is occurring when looking at same-different judgment reaction times. Perhaps individuals spending more time answering the sentence comprehension question might be spending less time on the load task, or vice versa. To check this, the data was split by load condition and bivariate correlations were run. The correlations looked at comprehension question reaction times and same-different judgment reaction times for each sentence structure. In the auditory condition, six positive correlations were found. Both the single object clefts, r=.419, and subject clefts, r=.531, showed positive correlations. In addition, four positive correlations were seen concerning the cleft + relative structures: (1) Object cleft + object relatives (1st verb), r=.389, (2) object cleft + object relatives (2nd verb), r=.387, (3) object cleft + subject relatives (2nd verb), r=.324, and (4) subject cleft + subject relatives (2nd verb), r=.539.

In the visual condition, all sentence structures except for the subject cleft + object relative (2nd verb) showed similar significant positive correlations. Correlations ranged in value from r=.331, for the object cleft + object relative (2nd verb) to r=.586 for single subject clefts.

Because all correlations found between comprehension question reaction time and same-different judgment reaction time were positive, this suggests that no trade off was occurring. Thus, participants who answered comprehension questions quickly also answered same-different judgments quickly as well.

4. DISCUSSION

Through this experiment, and its predecessor (Seidel & McDonald, 2008), we hoped to shed light on both working memory and language comprehension, and how the two interact in off-line and on-line measures. Ideally, the results from the on-line measure would have mirrored those found in the off-line measures both from this experiment and its predecessor (Seidel & McDonald, 2008), adding support for the general interpretations behind the initial findings. However, it appears that the constructs measured through on-line and off-line data are different. This, too, is beneficial in guiding future research in how to interpret data, depending on whether the data are on-line or off-line. Below I will go over the general findings from both the off-line and on-line data.

4.1 Off-line Data Discussion

Some of the main findings from Seidel & McDonald (2008) were successfully replicated in the new off-line data. This included our two-way interactions with working memory and syntactic complexity. These interactions show that working memory ability does play a role in sentence processing off-line as expected. Again, for a detailed comparison of the results from Seidel & McDonald (2008) and this design, please refer to Figure 13. However, one earlier finding we were unable to replicate was a four-way interaction that showed that high span performance drops to the low span levels when trying to comprehend the hardest sentence type (object cleft + subject relatives when asking about the second verb) under the hardest load condition (auditory load for the original experiment). While we did not obtain a four-way interaction, when including all participants in the analysis, we did obtain a three-way interaction between WM span, cleft, and relative, when asking about cleft + relative structures (2nd verb). In this three-way interaction, we noticed poor performance for both the non-ambiguous object cleft + subject relative structures and for the ambiguous subject cleft + subject relative structures. Ignoring the subject cleft + subject relative structure and focusing on the object cleft + subject relative structure, we see that high spans still outperform low spans.

Another way to gauge difficulty of the comprehension questions is to look at the comprehension question reaction times for the cleft + relative structures (2nd verb). Again, while still not obtaining the four-way interaction between working memory, cleft, relative, and load, we do see a three-way interaction between working memory, relative, and load. The trend seen in the no load and auditory load conditions is that low spans take longer to answer the question ending in a subject relative when compared to their high span counterparts. However, for the visual condition (arguably the hardest condition), this trend is reversed. High spans in the visual condition take taking longer to answer the comprehension questions for sentences ending in a subject relative. Ignoring the ambiguous subject cleft + subject relative structure, we see that although high spans still are able to more frequently correctly answer the comprehension sentence of the object cleft + subject relative structure, they take exceptionally long to do so. This perhaps shows either one of two things. First, the modality of the load may not be the issue for taxing processing, but rather, the inherent difficulty of the load is. In the original experiment (Seidel & McDonald, 2008), we see high spans' performance drop to low spans' levels for this same sentence type when under the hardest condition (auditory load). The second possibility is perhaps participants are somehow recoding the new visual condition auditorily, thus activating the episodic buffer. This would mean we are continuing to see additional auditory interference, albeit internally generated, affecting processing. However, we can only speculate as to what is truly occurring.

To elaborate on the auditory load condition, remember that participants were required to remember series of numerical digits. However, one question that arises is whether all auditory stimuli are the same? Or is speech special? Would the results from this experiment be replicated if, instead of digits, different pitched beeps were played? In the real world, it is not uncommon for participants to have to remember a series of digits, whether it be a phone number or an address. Concerning the visual condition, it seems uncommon that individuals would need to remember sequential locations of images.

Therefore, perhaps to better parallel the unusual visual condition, non-speech stimuli should be used. If results were replicated, it would show that speech is not special, and that all auditory stimuli can be grouped together. If, however, results differed, perhaps then we could conclude that speech is inherently different from other auditory stimuli.

In making comparisons between the original experiment and the new experiment, we must also bear in mind that the two experiments have important differences. First, the new experiment adds additional button pressing to progress the sentence, which itself might add additional load. Second, the two experiments differ in the nature of the visual load presentation. The original experiment (Seidel & McDonald, 2008) had all stars presented at once, whereas the current experiment had stars presented in different positions, sequentially. Because the stars appear in different locations, it could be argued that the new visual condition has a spatial component. According to Baddeley's (1974) model of working memory, visual and spatial information are both entertained within the visuospatial sketchpad. According to Cowan's (1999) model, all stimuli pull from a common pool of resources. Therefore, even if the "visual condition" would perhaps be more accurately labeled the "spatial condition," the predicted hypotheses and conclusions would not be affected.

As the on-line data supports, and will be discussed below, some amount of on-line processing is occurring as the critical areas do show trends of increased listening times as proposed. Thus, this shows that not all processing is deferred to the comprehension question presentation. Because individuals are self-pacing through the sentence, th ey may, as a result, linger on each phrase longer than they would during a natural speech stream. Therefore, the act of self-pacing may facilitate a certain amount of on-line processing, in addition to the natural off-line processing, that otherwise would not occur. While this seems in contradiction of what a rapid clicking strategy would show, participants still must monitor for the end of a phrase before proceeding to the next. And, we argue, this monitoring could lead to slightly more on-line processing. Perhaps, then, it is this combination of on-line and off-line processing that affords high spans the necessary boost to correctly answer even the hardest questions.

4.2 On-line Data Discussion

Together, our findings seem to support the idea that working memory capacity affects processing in that low spans tend to take more time for processing overall than high spans. Sufficient evidence is lacking that shows the desired interaction between sentence complexity and working memory span, where low spans take *even more* time to process the sentence at the critical areas. This supports Waters & Caplan's (2001) belief that working memory does not play a role in on-line measures of sentence processing. However, since our predicted results of an interaction with working memory were not found, this may be the result of different participant strategies.

The task, even in the simplest condition, is inherently difficult. This on-line version is even more difficult than the original off-line version (Seidel & McDonald, 2008) because this new version additionally forces participants to click a button to self-pace the sentence. This means that participants must constantly be monitoring and anticipating the end of each phrase. This adds additional processing demands above and beyond passively listening to a sentence in its entirety. When adding a supplementary working memory load, external to the demands of sentence task itself, we believe that strategies may be formed. Such strategies would help retain the information necessary to correctly answer the comprehension questions being asked at the end. These strategies may include rapidly clicking through the whole sentence to quickly reach the comprehension question, and not lingering specifically on these critical areas. When the comprehension question is reached, participants might then rely on some auditory trace to recall the correct answer. If this is the case, it would explain why low spans are overall slower, and not exaggeratedly slower on the critical areas. This would be because the reaction times being gathered do not reflect the amount of time it takes to process each phrase. Rather, they reflect the amount of time it takes to react to realizing the phrase has ended.

These strategies may be implemented in all three load conditions of the experiment. However, we suspect that they would be more likely to occur in those conditions where an external load is added. Further evidence for this assumption arises when we notice that participants in the no load condition consistently have longer reaction times than those in both of the load conditions. This suggests that people in the no load condition are actually performing the task more in line with what the directions state: to listen to each phrase of the sentence, and click the spacebar when that phrase is understood. Contrarily, those in the visual and auditory load conditions seem to use a strategy whereby they wait until the comprehension question is revealed before beginning to try to really comprehend the sentence. Because of this, individuals in the load conditions seem to be less burdened by the additional loads. Rather, they seem to opt to click faster to reach the comprehension question and defer processing. In short, people in the load conditions rely on some auditory trace more so than actual accumulated comprehension of the sentence. Therefore, in trying to determine whether or not WM impacts these critical areas, we must look at the no load condition in isolation. This is because the no load condition appears to be the lease likely affected by some clicking strategy.

The results we got from the no load condition were not as promising. Focusing on the critical areas of all the sentence structures, we neither see main effects nor interactions with working memory. However, it should be remembered that our sample size is rather small. In particular, when analyzing the second critical area of the cleft + relative structures, only one main effect emerges, and we do not see any main effects of relative or working memory, or interactions with either of the above. The results coming from analyzing the no load condition focusing on the second critical area offers the greatest evidence against our hypothesis that working memory will play a role in on-line sentence comprehension. Not even relative emerges as a main effect. This may suggest that a clicking strategy may be evident in even the no load condition to a certain extent. Alternatively, this could also support the idea that working memory does not factor in on-line processing at all.

This experiment modeled after that of Waters and Caplan (2001). While the results of this experiment additionally seem to mirror those of Waters and Caplan (2001), a few criticisms of both designs that should be addressed. The main point to be addressed is the way in which the on-line (listening time) data, were gathered. In Waters and Caplan's (2001) design, sentences were spliced from a natural speech stream and participants were required to self pace by clicking a button. However, participants were not required to completely finish listening to one phrase before they were able to proceed to the next. The logic was that participants would be forced to listen to each clip instead of repetitively clicking through the sentence. If participants, however, did rapidly click through the whole sentence, they would arrive at the comprehension question having heard very little of the sentence, if anything. The problem with this is, by allowing participants to click mid-phrase, one might be tapping into something above and beyond listening time or comprehension time. If these sentences were counterbalanced and featured the same characters, as it was the case, participants over time, I assume, would become familiar with the sentences. Participants would become familiar with the potential subjects and objects and even verbs of the sentence, and perhaps be familiar with which combinations typically go together. So, when hearing the first few phonemes of a clip, participants may be able to guess "who" or "what" from past experience, and click the button. Thus, it would be tapping into some other form of memory driving the ability to "recall" past experiences.

On the other hand, our design forced participants to listen to each clip in its entirety before their reaction would be recorded. While this put all participants on even footing as far as allowing the whole clip to play, there is a trade off. The trade off is this might have opened up the possibility for clicking strategies to form. It is understandable how listening to the whole sentence without breaks is easier than having to monitor for the end of a clip. Therefore, as discussed above, one strategy that might form would be to continue repetitively clicking. In fact, certain participants seen repetitively clicking and not waiting for the end of a phrase during the experiment were marked by the experimenter and later discarded from analysis. This offers some form of observational support for the presence of a strategy.

By having the option to click repetitively and piece the sentence together as smoothly as possible for comprehension, participants can make a choice to defer processing until the comprehension question. The processing would then appear in the off-line measures taken instead of on-line.

Thus, although it appears that working memory does not play a role in on-line syntactic processing, a final judgment should be reserved until more conclusive data are available. The question remains: if we could force participants to process on-line, would we then see working memory play a greater role? Yet, how can one force online processing? Or must on-line processing be measured in a different way, such as, for example, with eye tracking? If forcing on-line processing were possible, would this task even be natural, and thus generalizeable to the field? If anything, the results from this experiment have opened the door to more questions, and more testable possibilities.

4.3 Limitations and Future Directions

While our design nicely parallels Waters & Caplan's work (2001) in using an auditory moving window, the visual moving window seems to be the standard in on-line data gathering. Because not many individuals have used the auditory moving window, there could be a discrepancy in the results gathered with one over the other. This concern has been addressed in the literature. It was shown that results obtained from the auditory moving window and the visual moving window are similar (Ferreira, et al., 1996). However, even though it appears that the VMW and AMW are equivalent, not a lot of literature is available comparing the two techniques. Future research could include doing an on-line visual moving window design to determine whether or not the AMW and the VMW are indeed parallel in the results they offer.

Other limitations and future directions are concerned with the load task. Note that load did not emerge as a major factor in the analyses. Perhaps this is because the load task is not on-line, as with some of the word-monitoring load conditions found in other literature (Vos, et al. 2001a; Vos, et al., 2001b). To address this issue, future research can diverge in one of two directions. The first direction is to go on a quest to make the off-line load tasks even more complex, keeping the original off-line nature of the task. However, given the existing poor performance on the load task, this might not be a successful adjustment. The second direction is to force the existing loads into an on-line form. Perhaps, we will see the previously predicted outcomes. For example, instead of asking participants to remember a pattern of stars presented before the sentence, one could have participants monitor the stars during the presentation of the sentence.

Concerning the main experiment itself, remember that it appeared that participants were deferring to off-line processing. Perhaps we are not seeing working memory factor in on-line processing because the task itself is too difficult. Therefore, it might be interesting to include different, easier sentence types into the experiment. In this way, we would be able to observe any differences in the amount of on-line processing done between sentence types of varying difficulty.

REFERENCES

Andrews, G., Birney, D., & Halford, G.S. (2006). Relational processing and working memory capacity in comprehension of relative clause sentences. *Memory and Cognition*, 34(6), 1325-1340.

Baddeley, A.D., & Hitch, G.J. (1974). Working memory. In G.A. Bower (Ed.) *The psychology of learning and motivation* (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A.D. (2001). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4(11), 417-423.

Bock, K. & Miller, C.A. (1991). Broken agreement. *Cognitive Psychology*, 23, 45-93.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences, 22*, 77-126.

Cowan, N. (1999). *An embedded-processes model of working memory*. In A. Miyake & P. Shah (eds.), Models of Working Memory: Mechanisms of active maintenance and executive control. Cambridge, U.K.: Cambridge University Press. (pp. 62-101)

Cowan, N., & Morey, C., (2007). How can dual-task working memory retention limits be investigated? *Psychological Science*, *18(8),* 686-688.

D'Esposito, M., Postle, B.R., Ballard, D., & Lease, J. (1999). Maintenance versus manipulation of information held in working memory: An event-related fMRI study. *Brain and Cognition*, 41, 66-86.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191.

Ferreira, F., Henderson, J.M., Anes, M.D., Weeks, P.A. Jr. & McFarlane, D.K. (1996). Effects of lexical frequency and syntactic complexity in spoken language comprehension: Evidence from the auditory moving window technique. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 324-335.

Gordon, P., Hendrick, R., & Johnson, M. (2001). Memory inference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6), 1411-1423.

Gordon, P., Hendrick, R., & Levine, W. (2002). Memory-load interference in syntactic processing. *Psychological Science*, 13(5), 425-430.

Gunter, T. C., Jackson, J. L., & Mulder, G. (1995). Language, memory, and aging: An electrophysiological exploration of the N400 during reading of memory demanding sentences. *Psychophysiology*, 32, 215-229.

Hartsuiker, R. & Barkhuysen, P. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, 21, 181-204.

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

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 336–358.

King, J., & Just, M.A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30, 580-602.

Leech, R., Aydelott, J., Symons, G., Carnevale, J., & Dick, F. (2007). The development of sentence interpretations: effects of perceptual, attentional, and semantic interference. *Developmental Science*, 10(6), 794-813.

Mak, W.M., Vonk, W., & Schriefers, H. (2008). Discourse structure and relative clause processing. *Memory and Cognition*, 36(1), 170-181.

Seidel, C.M. & McDonald, J.L. (2008). Effects of syntactic complexity, load, and span on cleft comprehension. Presented at Psychonomics Conference, Chicago, IL.

Sheldon, A. (1977). On strategies for processing relative clauses: A comparison of children and adults. *Journal of Psycholinguistics Research*, 6(4), 305-318.

Vos, S.H., Gunter, T., Schriefers, H., & Friederici, A.D. (2001a). Syntactic parsing and working memory: The effects of syntactic complexity, reading span, and concurrent load. *Language and Cognitive Processes*, 16(1), 65-103.

Vos, S.H., Gunter, T., Kolk, H., & Mulder, G. (2001b). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology, 38*, 41-63.

Waters, G.S. & Caplan, D. (1996a). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review*, *4*, 761-772.

Waters, G.S. & Caplan, D. (1996b). The measurement of verbal working memory capacity and its relation to reading comprehension. *The Quarterly Journal of Experimental Psychology*, 49A, 51-79.

Waters, G.S. & Caplan, D. (2001). Age, working memory, and on- line syntactic processing in sentence comprehension. *Psychology and Aging*, 16(1), 128-144.

Waters, G.S., Caplan, D., & Hildebrandt, N. (1987). Working memory and written sentence comprehension. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading* (pp. 531-555). London: Erlbaum.

VITA

Christy Seidel is a New Orleans native, born to parents Joseph and Lea Seidel. In 2007, she graduated from Louisiana State University with a Bachelor of Science degree in psychology, and minors in chemistry and Russian language. Pursuing her loves of language and psychology, she began to work under Dr. Janet McDonald during the fourth year of her undergraduate career. After graduating with her Bachelor of Science degree, she continued to work with Dr. Janet McDonald, pursuing her doctorate in cognitive psychology. Also, under Dr. Russell Matthews, Christy is pursuing a minor in industrial-organizational psychology. During her graduate career, she has presented research at three conferences: the annual meeting of the LSU Women and Gender Studies program, Baton Rouge, Louisiana, the 49th annual meeting of the Psychonomic Society, Chicago, Illinois, and the 50th annual meeting of the Psychonomic Society, Boston, Massachusetts.

Christy is in her third year of the doctoral program. She is currently working with Drs. Emily Elliott and Janet McDonald looking at the irrelevant speech effect on grammaticality judgment. In addition, she is part of a National Institutes of Health-funded grant research team headed by Drs. Janna Oetting, Janet McDonald, and Michael Hegarty, looking at Specific Language Impairment (SLI) and child dialect speakers.