GARDEN PATH TRIPARTITE COMPOUNDS

BY

HEATHER MAUCH

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Arts in Educational Psychology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Master's Committee:

Associate Professor Kiel Christianson, Chair
Assistant Professor Darren Tanner
Professor Jerome Packard

Abstract

Because it is a Germanic language, compounding in English is extremely productive. In German, compounds are always non-spaced (*die Waschmaschine* is the equivalent of *washing machine* in English) and could, in theory, be infinitely long. In English, however, compounds occur in both spaced (*washing machine*) and non-spaced (*background*) forms. Spaced compounds introduce a particular type of ambiguity because the reader must decide if the compound should be parsed as one lexical unit or as two (or more) separate units. A classic example of this is the sentence *Washing machines can be boring*. To a more extreme degree, spaced tripartite compounds such as *dog bite victim*, where the middle constituent is a homonymous verb-derived noun, present an even greater ambiguity since these tripartite compounds appear to follow the canonical NVN word order found in English. Previous research has shown that readers struggle with ambiguous or "garden path" sentences. Readers often maintain both the initial incorrect parse of the sentence as well as the final correct parse, resulting in a "good enough" representation of the sentence structure. To my knowledge, no research has analyzed spaced tripartite compounds with respect to Good Enough sentence processing. The present study utilizes eye-tracking in order to analyze how garden path tripartite compounds affect sentence processing.

Keywords: garden path sentences, good enough processing, ambiguity, spaced tripartite compounds

Table of Contents

Chapter I

Introduction and Literature Review

English, being a Germanic language, allows for extremely productive compounding. Common English compounds can be either spaced (*post office*, *high school*) or non-spaced (*airport, background*), or both in one compound (*peanut butter, sunflower oil*). In the current proposal, the focus is solely on spaced compounds due to their unique capacity for introducing ambiguity. For example, the compound *cardboard drink coaster production line*, while entirely valid, could cause confusion in readers because it is not until the final constituent, *line*, that readers are able to integrate the previous four constituents into the compound. Additionally, the second constituent, *drink*, is temporarily ambiguous because it could be either a noun or a verb. Because the space between the two compound constituents is not a reliable cue for the head of the compound in English, ambiguity arises when the reader initially parses the compound into its constituent lexical units, rather than as a single syntactic unit with one head noun. In this way, English could be compared to Chinese, which has no spaces between characters. Readers of Chinese must constantly be deciding where to combine characters into meaningful units, because there is no reliable visual cue to guide them (see Lin, Anderson, Ku, Christianson, & Packard, 2011).

In the present study, I sought to investigate the way in which readers interpret ambiguous tripartite compounds, such as *dog bite victim*. It is possible that, because these nouns mimic the subject-verb-object sentence construction commonly found in English, the reader might first attempt to interpret the middle constituent (*bite*) as the main verb. This temporary ambiguity could consequently lead to an incorrect sentence parse (i.e., a "garden path" sentence (Bever, 1970)), which would need to be subsequently revised in order to arrive at a correct interpretation.

Participants were asked to read such garden path tripartite compounds when embedded in different sentence positions and contexts. Eye movements and probe question accuracy were measured in order to determine if readers were being garden-pathed and whether or not they were able to recover and ultimately derive the correct interpretation licensed by the syntactic structure.

In this chapter, I will discuss the research literature that is relevant to the present study. I will discuss Good Enough processing theory, Local Coherence theory, and tasks that assess working memory and print exposure before giving a description of the present study.

Good Enough Processing and Local Coherence

So-called "garden path" sentences illicit an initial incorrect sentence parse up until the point the reader encounters a disambiguating word or other information. For example, *The horse raced past the barn fell* (Bever, 1970) is a garden path sentence. The reader typically begins to parse the sentence so that “raced” is the main verb, “horse” is the subject, and “the barn” is the object. However, once the reader encounters the word “fell,” which is the true main verb, syntactic reanalysis must occur so that the reader arrives at the intended parse of the sentence in which *raced* is the verb of a reduced relative clause: *The horse (that was) raced past the barn fell.*

Sentences such as the one from Bever (1970) are the impetus for research involving Good Enough processing. Christianson, Hollingworth, Halliwell, and Ferreira (2001) used sentences such as *While the man hunted the deer ran into the woods* to demonstrate the importance of both offline and online reading measures when studying sentence processing and comprehension. After reading sentences such as the above, readers were asked either (a) *Did the deer run into the* *woods?* (yes) or (b) *Did the man hunt the deer?* (no). While the majority of participants correctly answered yes for (a), many participants also incorrectly answered yes for (b). A separate experiment utilized reflexive absolute transitive (RAT) verb constructions, such as *While Anna dressed the baby played in the crib.* Participants were then asked either (a) *Did the baby play in the crib?* or (b) *Did Anna dress the baby?* Results replicated those from the earlier experiment, where the majority of participants correctly answered yes for (a), but also incorrectly answered yes for (b). The results of these experiments support the idea that readers of garden path sentences are not able to completely inhibit the initial sentence misparse, and instead seem to maintain both the initial misparse and the final correct parse – at least temporarily, if not in the long-term (see Slattery, Sturt, Christianson, Yoshida, & Ferreira, 2013). Ferreira, Bailey, and Ferraro (2002) framed the study by Christianson et al. (2001) in terms of what is now known as Good Enough processing theory, where the final interpretation of garden path sentences is not entirely consistent with the meaning licensed by the syntactic structure, but is instead merely "good enough" (see Christianson, Williams, Zacks, & Ferreira, 2006) to satisfy a plausibility heuristic (i.e., a sense that the reader knows what the sentence is saying based on knowledge of the world).

Ferreira (2003; see also Christianson, Luke, & Ferreira, 2010) conducted several experiments in order to further analyze the theory of Good Enough processing. The experiments made use of both passive constructions (e.g. *The dog was bitten by the man*) and object-cleft constructions (e.g. *It was the dog the man bit.*). Both of these sentence constructions disregard traditional thematic role assignment order (Agent-Verb-Patient) in favor of the reverse order (Patient-Verb-Agent). Both sentences are also implausible, though not impossible, since one would typically expect a dog to bite a man, rather than the reverse. When compared with their plausible counterparts (e.g. *The man was bitten by the dog* and *It was the dog that bit the man*), the implausible passive and object-cleft sentences resulted in lower comprehension question accuracy. Ferreira hypothesized that sentence implausibility as well as the non-canonical thematic role assignment were the major causes for the decrease in comprehension accuracy (see Christianson et al., 2010, for a replication). While the sentences used in these experiments were not garden path sentences, they still are relevant to the current proposal in that they demonstrate other ways in which misinterpretation can occur during sentence processing. Ferreira (2003) proposed the use of a plausibility heuristic in addition to the NVN heuristic proposed by Townsend and Bever (2001). These two heuristics are essentially semantic in nature. The NVN heuristic assumes that the first noun in a sentence is the subject (and therefore the agent) and the second noun is the object (and therefore the patient). The plausibility heuristic takes into account world-knowledge and overall semantic plausibility and constructs a parse that is most consistent with this knowledge. These two heuristics compete with the algorithmic parser, which is based purely morphosyntactic information. The algorithmic parser, if given ample sufficient time, will always yield the correct interpretation, because it constructs the only parse that the syntactic structure allows. However, when the outputs of the heuristic parse and heuristics yield different interpretations compared to one another, or different interpretations when compared to the algorithmic parse conflict, eventually one parse must win out and misinterpretation can occur if the winner is the heuristic parse.

The interesting aspect about the present target tripartite compounds is that they seem to mimic an NVN or agent-verb-patient structure. For the present study, I utilized tripartite compounds that consist of three nouns, e.g. *dog bite victim*, with the middle constituent being a homonymous verb-derived noun. Because *bite* looks and sounds the same when used as either a noun or a verb, it is possible that readers will attempt to parse the tripartite compound as an NVN structure rather than an NNN structure. These stimuli make use of such tripartite compounds embedded within full sentences, such as *Last week, the obese dog bite victim met the neighbor for coffee*. Thus, the lexical ambiguity of *bite* leads to a larger sentence-level ambiguity, because it alters the reader's initial parse of the full sentence. No other sentence processing studies that we are aware of have made use of this embedded tripartite compound structure.

In a study related to Good Enough processing, Tabor, Galantucci, and Richardson (2004) hypothesized that garden path sentences are first comprehended locally before they are comprehended globally. This phenomenon was called "merely local coherence." Sentences such as *The coach smiled at the player tossed a Frisbee by the opposing team* were used in a self-paced reading paradigm in order to explore this phenomenon. Tabor and colleagues predicted that readers would process the locally coherent clause *the player tossed a Frisbee* as an active clause rather than part of a reduced relative clause despite the fact that the structure of the first part of the sentence should unambiguously rule out a main verb interpretation of the locally coherent clause. This would result in increased reading times on the verb *tossed*, since it introduces ambiguity because the reader would have initially parsed the sentence as a simple statement: *The coach smiled at the player.* In fact, the results of the experiment did show increased reading times on *tossed*.

The position of the tripartite compounds used in the materials in the experiment reported here are manipulated between subject and object position. In subject position, the lexically ambiguous middle constituent (e.g., *bite*) leads to a syntactic ambiguity in the subject position, given that it could be interpreted as either a main verb or as part of the compound. In object position, the verb reading of the middle compound constituent should be ruled out by previous context. It is an open question, however, whether merely local coherence effects can also be observed in this type of sentence.

Working Memory and Print Exposure

Working memory is the cognitive system that regulates the processing of new and previously stored information, and allows an individual to hold task-relevant information in focus in the midst of distraction. Working memory span tasks, such as the reading span task (RSPAN), have been shown to positively correlate with reading comprehension performance (Daneman & Carpenter, 1980; Conway et al., 2005). While there are several variations of the task, typically an individual is asked to read a sentence and make a logical judgment such as whether or not the sentence made sense. After making this judgment, the individual is then presented with a letter of the alphabet and told he or she will need to remember it. After a varying number (usually two to six) of similar sentences and letters, the individual is asked to recall the letters in the exact order of presentation. The participant is assigned an RSPAN score based on the number of correctly recalled letter sequences.

Another task that could have implications for the present study is the author recognition task. For this task, individuals are presented with a list of names one-by-one and asked to decide if the name represents an author or not. This task is scored by taking the number of correctly identified authors (hits) and subtracting from it the number of names that were incorrectly identified as authors (false alarms). This score provides a measure of an individual's print exposure, which has been shown to correlate positively with reading-related performance tasks (Acheson, Wells, & MacDonald, 2008). In the present study, more experience with text could provide exposure to this sort of multi-part compounding, which is quite common in English writing; however, individual compounds of three or more constituents tend to be extremely infrequent, and even novel, as are most of the compounds used here.

It is possible that individual differences in working memory and/or text exposure could explain individual differences in comprehension of garden path tripartite compounds – and thus, individual differences in eye movement patterns. Because of this, these two cognitive tasks (RSPAN and ART) were conducted at the end of each experimental session. It could be that participants with greater working memory capacity and/or more print exposure might be better able to arrive at the correct sentence parse – or, perhaps they will revise the initial misparse earlier on.

The Present Study

The goals of the present study are severalfold. First, although compounding is extremely productive in English, I am not aware of any eye-tracking studies examining the processing of tripartite compounds. This study aims to begin to fill that gap. Second, by creating tripartite compounds with ambiguous N-V second constituents, I seek to examine a new type of garden path structure. Moreover, in order to correctly interpret this garden path structure, reanalysis of any initial misparse would require collapsing an initial multi-phrasal structure (NVN) into a single phrasal head (N). The extent to which such a revision might result in lingering good-enough-style misinterpretation is an open question. Third, by manipulating the argument position of the compound, I am able to test the generalizability of "merely local coherence" to a radically different structure from the only one in which it has been demonstrated so far (i.e., the reduced relative clause in object position).

To these ends, participants read novel tripartite compounds embedded within sentences. There were four conditions for the experiment, and thus each compound appeared in four different sentences. The compound appeared in either the subject position or the object position of the sentence. Additionally, the compound was embedded in either a past tense context (*Last week, the obese dog bite victim…*) or a present tense context (*Every week, the obese dog bite victim…*). The former could lead to a faux tense error if the reader were to interpret the middle compound constituent *bite* as a verb, while the latter could lead to a faux agreement error (a singular dog *bites* rather than plural dogs *bite*). Participant's eye-movements were recorded while the read and answered comprehension questions for each target sentence.

I predicted that there would be a significant difference in the reading times for the subject and object conditions. Specifically, I predicted that when the tripartite compounds appeared in the subject position, comprehension question accuracy would be lower when compared to the object condition. Take, for example, the hypothesized step-by-step parse of the sentence *Last week, the obese dog bite victim met the neighbor for coffee.* Readers could have initially encountered the middle constituent, *bite*, and incorrectly interpreted it as a verb with a faux tense error (according to this parse, the sentence should read, *Last week, the obese dog* bit *the victim*…). This would result in longer fixation durations on the second compound constituent (N2). However, instead of performing syntactic reanalysis at this stage, readers might interpret the sentence as containing a misspelling and continue reading the sentence and pursuing the initial misparse. Once the reader encounters the true verb of the sentence, *met*, he or she would be forced to reread and pursue the final correct parse, as it is the only one allowed by the input. Thus, longer fixation durations would be seen for the true main verb, as well as more regressions to previous sentence regions. When the tripartite compounds occurred in the object position, there were two predicted possible outcomes. One prediction was that, for a sentence such as *Last week, the neighbor met the obese dot bite victim for coffee,* readers would exhibit no error signal. This prediction was based on the fact that the reader had already encountered the true verb, *met*, before encountering the tripartite compound. Thus, under this assumption, readers should have had much less trouble with the sentence conditions where the tripartite compound was in the object position (resulting in higher question accuracy rates). A second possibility, though, was that readers would still show reading patterns similar to the subject conditions with increased fixation durations on the N2 and increased regressions back to the main verb in particular. This situation would be consistent with Tabor et al.'s (2004) local coherence, as the locally coherent structure of *dog bite victim* would then be interfering with the existing globally coherent structure set up by the previous sentence regions.

With respect to the agreement and tense error conditions, I predicted that participants would have more difficulty when a faux tense error was introduced (e.g., when the sentence began in the past tense) due to the fact that there was more interfering information between the start of the sentence and the ambiguous middle constituent of the compound. Conversely, participants would have less trouble with the faux agreement error, because they would simply have to refer back to the first constituent of the tripartite compound in order to ensure that the faux verb (e.g., *bite*) agreed with the preceding noun (e.g., *dog*). This prediction is consistent with the theory of local coherence, since the faux agreement error is a more locally coherent lure.

Chapter II

The Experiment

The purpose of this experiment is to determine how garden path tripartite compounds are processed. The tripartite compounds introduce a lexical ambiguity that gives rise to a syntactic ambiguity, the effects of which should be evident in participants' eye-movement records via increased reading time measures and decreased comprehension question accuracy. Thus, the current experiment will provide valuable insight regarding the way we process a new type of garden path sentence construction. The measures related to the recover from the garden path, and accuracy of comprehension, will inform the literature related to the extent to which the language processing system relies on good-enough, or underspecified, representations.

Method

Participants. Sixty-seven participants from the University of Illinois at Urbana-Champaign were recruited for the experiment and compensated with either $7 or 1 course credit. All participants were native speakers of English with normal or corrected-to-normal vision.

Materials. Forty target tripartite compounds were used, each occurring in either the subject position or object position of a sentence, preceded by a sentence context that suggests either a faux tense error or a faux agreement error (e.g. *Last week, the obese dog bite victim…* suggests a faux tense error if the participant interprets *bite* as the main verb, whereas *Every week, the obese dog bite victim…* suggests a faux agreement error between the subject (*dog*) and supposed verb (*bite*). In addition to the 40 target sentences, there were 100 filler sentences consisting of a mixture of subject relative, object relative, and compound sentence structures. See the Appendix for a full list of target and filler stimuli.

In order to justify the use of only two answer options for the comprehension question (see item 1e below), frequencies of the adjective-noun pairings were analyzed for all 40 tripartite compounds via a Google search. For instance, the frequency of *obese dog* (26,900 hits), *obese bite* (2130 hits), and *obese victim* (849 hits) were all obtained by searching for each word pair (in quotation marks) using Google's search engine. This was done for all compounds, and then each frequency was log transformed. Two-tailed paired t-tests for the adj-N1 pair (*M* = 9.81, *SD* = 1.80) and the adj-N3 pair (*M* = 9.54, *SD* = 1.97) were not significant, *t*(39) = 1.77, *p*=.08, which means that the two sets of frequencies were not significantly different from one another list-wise. A t-test for the adj-N1 pair and the adj-N2 (*M* = 7.69, *SD* = 2.07) pair was significant, *t*(39) = 10.34, *p*<.001, and results were similar for the t-test for the adj-N3 pair and the adj-N2 pair *t*(39) = 8.96, *p*<.001, suggesting that the frequencies of the adj-N1 and adj-N3 pairings were significantly different from the adj-N2 pair list-wise. This ensured that the likelihood of the answer being *obese dog* or *obese victim* were relatively equal, and more likely than the answer being *obese bite*, which was not offered as an answer choice.

Norming. Twenty-one participants were asked to norm the tripartite compounds for ease of imageability on a 7-point Likert scale, with 1 being "very difficult to imagine" and 7 being "very easy to imagine." The purpose of this process was to ensure that all tripartite compounds were equi-plausible – or, if implausible, to ensure that they were equi-implausible. Constructing the target compounds proved to be difficult, and norming was conducted to ensure that the compounds did not appear overly contrived or unimaginable to participants. All compounds fell within 2.5 standard deviations of the mean (*M* = 5.16, *SD* = .82) and thus all tripartite compounds remained in the experiment.

Design. The experiment utilized a 2x2 repeated measures design. The first factor was tripartite compound position (subject [a-b] vs. object [c-d]) and the second factor was the faux error type (tense [a, c] vs. agreement [b, d]). An example comprehension question is given in (1e).

1. (a) Last week, the obese dog bite victim met the neighbor for coffee.

   (b) Every week, the obese dog bite victim meets the neighbor for coffee.

   (c) Last week, the neighbor met the obese dog bite victim for coffee.

   (d) Every week, the neighbor meets the obese dog bite victim for coffee.

   (e) What was obese? (options: the dog; the victim)

The motivation for these factors was the prediction that participants would have more difficulty with compounds in the subject position when compared to the object position, since in subject position the ambiguous middle constituent (e.g. *bite*) will occur before the main verb. The motivation for the agreement and tense error conditions was the prediction that readers might have more difficulty when there was a faux tense error, since they will have to refer all the way back to the beginning of the sentence as opposed to referring back to the previous word for the faux agreement error. In accordance with Christianson et al. (2001), a comprehension question was asked following the presentation of each target sentence, and it was always about the adjective just prior to the tripartite compound, e.g., *What was obese?* Participants were asked to make a forced-choice response between either the first noun (N1) in the compound (*dog*) or the third noun (N3) in the compound (*victim*). The answer was always the N3, since the right-most noun in English compounds is always the syntactic head. The question was always a what-question for the target items. Each filler was followed by a comprehension question similar to those used for the target sentences, but the filler questions consisted of who, what, when, and where questions and they probed other aspects of the sentence rather than just the adjective-noun relationships. Comprehension questions for the filler items consisted of 45.8% who-questions, 41.4% what-questions, 5.7% when-questions, and 7.1% where-questions. How- and why-questions were not used because they required more complex answers than the experimental design allowed for.

There were four experimental lists, each containing one of the four conditions for each target compound counter-balanced across the four lists. The order of the presentation of the two forced-answer answer choices was counterbalanced within each list so as to avoid any button-press bias. Participants saw each compound only once.

Post-tests. Two additional tasks were used in order to assess working memory capacity and exposure to printed text, as both could be considered predicting factors in participant performance on the eye-tracking portion of the study. For instance, an individual with higher working memory capacity or more print exposure could have less trouble interpreting the tripartite compounds when in subject position. The reading span task (RSPAN) has been shown to demonstrate that working memory capacity correlates with reading comprehension, such that individuals with smaller reading spans (as measured by the task) performed much worse on a comprehension task than did readers with larger spans (Daneman & Carpenter, 1980). In this task, participants were first shown a sentence and asked to read it silently. Once done reading, the participant pressed a button and a new screen appeared asking the participant to decide if yes, the sentence made sense, or no, the sentence did not make sense. Each sentence was grammatically correct, but some sentences were semantically nonsensical. After making a decision, a new screen appeared showing a lowercase letter of the English alphabet. Participants were instructed to try to remember this letter because they would later be asked to recall it and input the letter(s) in sequence into the computer. The number of letters each participant was asked to recall (set size) ranged from three to six, with three items in each set size. For instance, there were three points in the experiment at which the participant was asked to report back three, four, five, or six letters. RSPAN scores were calculated by summing the number of perfectly recalled set sizes, with 54 being the maximum possible score. The author recognition task (ART), which uses signal detection theory, has been shown to demonstrate a positive correlation between print exposure and performance on reading-related tasks (Acheson et al., 2008). The individual was presented with a series of names one-at-a-time and asked to identify whether or not this was the name of an author. An individual's score on the task (d') was calculated by the standardized difference between a person's hit rate (the number of correctly identified authors) and a person's false alarm rate (the number of names that were incorrectly identified as authors).

Apparatus. Eye movements were recorded using a desk-mounted SR Research Ltd. EyeLink 1000 eye-tracker. Text was displayed using 14-point Courier New font on a computer monitor approximately 70 cm away from the participant. A desk-mounted chin and headrest was used to minimize head movements.

Procedure. Upon entering the room, participants were presented with the informed consent form. After signing, each participant was fitted to the chin and headrest and his or her eye was calibrated using a 9-point calibration screen. Viewing was binocular, but only the right eye was tracked. Each testing session involved an initial practice block consisting of six trials. The experimental portion of the testing session consisted of four blocks of 35 trials each. Calibration occurred at the start of each block, with a drift correction screen between each trial. Each trial began with a fixation dot on the left side of the screen in the position of the start of the upcoming sentence. Once the participant had fixated on the dot and pressed a button to ensure that calibration was still accurate, the participant saw a sentence appear on the screen in full. Once the participant finished reading the sentence, he or she pressed the appropriate button on a standard game controller to advance to the next screen. A new screen with a forced-choice comprehension question appeared, and participants indicated their answer choice using one of two buttons on the game controller. Then the next trial began, starting with the fixation dot. After completing all 140 trials of the eye-tracking portion of the experiment, participants were instructed to complete the ART and RSPAN tasks on a separate computer. Instructions were provided verbally and were also written on the screen. After completing these tasks, participants were debriefed and given either research credit or payment. The entire experiment took approximately 45 minutes to complete.

Results

Five participants were eliminated because they did not participate in the two post-tests, and twelve participants were eliminated because the experimenter could not obtain a good calibration. Only the participants who achieved 80% accuracy or higher on the filler items were used for final data analysis, resulting in 36 total participants. Filler items were used as the criterion because if target items were used, this would have potentially excluded participants who did poorly on the target items but performed well on the filler items. Because I am interested in whether or not accuracy suffered for the target items, I used accuracy on the filler items as the criterion for exclusion from further data analysis. Fixations that were not within the range of 80-1200 milliseconds were removed from data analysis, and fixations that fell within 0.5 degrees of the nearest fixation were merged into a single fixation. Trials where it was apparent that the participant pressed a button prematurely or where it was noted that the participant began talking were removed from data analysis. This resulted in .003% data loss.

Eye-movements were analyzed for four different sentence regions: the first word of the tripartite compound (N1), the second word of the tripartite compound (N2), the third word of the tripartite compound (N3), and the main verb of the sentence (MV). Six separate eye-movement measures were examined for each region: total reading time, first fixation duration, gaze duration, go-past time, and regression in count. Regression out count was initially explored, but none of the models turned out significant. *Total reading time*, or dwell time, is the sum if all fixations on a region during a single trial. *First fixation duration* is the duration of the first fixation on a target region. *Gaze duration*, or first run dwell time, is the sum of all fixations on a region from when it is first entered to when it is first exited in either direction. *Go-past time*, or regression path duration, is the sum of all fixations on a region from when it is first entered to when it is first exited to the right, thus including regressions to previous regions. *Regression in count* is the number of times an individual regressed into a region from the right (later regions of the sentence). *Regression out count* is the number of times an individual regressed out of a region to the left (earlier regions of the sentence). Mean reading times for each sentence region for each eye-movement measure are provided in Table 1.

Table 1

Mean reading times for each sentence region in each condition

	N1		N2		N3		MV	
Variable	Subject	Object	Subject	Object	Subject	Object	Subject	Object
First Fix	242 (100)	246 (103)	242 (102)	243 (102)	247 (109)	242 (101)	258 (107)	231 (91)
Gaze	296 (165)	285 (133)	303 (164)	283 (137)	326 (184)	313 (162)	317 (181)	283 (146)
Go-past	410 (325)	407 (338)	422 (316)	387 (323)	557 (541)	523 (597)	451 (470)	377 (365)
Total	671 (497)	598 (434)	662 (483)	547 (391)	661 (465)	585 (373)	570 (398)	610 (460)
Reg In Count	.78 (.98)	.52 (.79)	.62 (.85)	.41 (.62)	.37 (.61)	.32 (.55)	.41 (.66)	.55 (.74)

	N1		N2		N3		MV	
Variable	Agrmnt.	Tense	Agrmnt.	Tense	Agrmnt.	Tense	Agrmnt.	Tense
First Fix	244 (98)	244 (107)	245 (103)	241 (101)	248 (110)	243 (100)	249 (103)	240 (98)
Gaze	292 (150)	291 (151)	293 (151)	294 (153)	316 (176)	325 (170)	308 (168)	292 (162)
Go-past	409 (313)	410 (350)	414 (340)	396 (300)	560 (614)	521 (521)	422 (430)	406 (414)
Total	634 (477)	636 (459)	623 (475)	589 (410)	606 (416)	641 (433)	600 (427)	582 (434)
Reg In Count	.64 (.92)	.67 (.87)	.54 (.78)	.50 (.72)	.36 (.57)	.33 (.59)	.47 (.70)	.50 (.71)

Note: Standard deviations are shown in parentheses

The results of this experiment were analyzed using linear mixed effects models in the R statistical software package (R Foundation for Statistical Computing, version 3.1.2, 2014) using the lmer() function of the lme4 package (Bates, Maechler, Bolker & Walker, version1.1-7, 2014). For each target region (N1, N2, N3, and MV), a separate model was created for each reading time measure. For each reading time measure, three models were constructed: a simple model containing only compound position and error type as fixed effects, a model containing

ART as an additional fixed effect, and a model containing RSPAN as an additional fixed effect. ART score and RSPAN score were log transformed and mean-centered to reduce collinearity. When both RSPAN and ART were included simultaneously in a single model, troublesome 4-way interactions occurred that were non-interpretable, and many of the models failed to converge. Additionally, RSPAN and ART scores were correlated. Therefore, each post-test score was fitted separately to its own model for each eye-movement measure on each target word. The categorical fixed effects of compound position and error type were contrast coded so that object position = 1, subject position = -1, and agreement error = 1, tense error = -1. These contrasts were set for each target region data subset (N1, N2, N3, and MV). Best fit models were determined using a maximal random effects structure method (Barr, Levy, Scheepers, & Tily, 2013). P-values are not provided as output with the lmer() function for each model, thus t-values greater than or equal to 2 were considered significant (Baayen, Davidson, & Bates, 2008; Gelman & Hill, 2007). Binomial measures were analyzed using logit mixed effects models with compound position and error type as fixed effects. P-values are reported in this analysis, and thus were used to determine significance. For simplicity, only the models that showed significant results will be reported. If the model containing ART and RSPAN scores as fixed effects showed nothing different from the simple model, or nothing significant at all, they were abandoned in favor of the simple model.

Accuracy was analyzed using a logit mixed effect model, where 0 represented an incorrect response and 1 represented a correct response. There was no significant difference in comprehension question accuracy between the four conditions (all p-values were greater than 0.05). On average, participants were 88% accurate in the agreement error object position condition, 90.5% accurate in the agreement error subject position condition, 88.3% accurate in the tense error object position condition, and 89.9% accurate in the tense error subject position condition. There was no significant difference in total sentence reading time between the four conditions, with a mean of 6623.83 ms in the agreement error object position condition, a mean of 6346.3 in the agreement error subject condition, 6421.123 in the tense error object condition, and 6478.99 in the tense error subject condition.

Target region N1.

Total fixation duration. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for compound position, error type, and their interaction was constructed. This model showed a main effect of compound position, such that participants had longer total fixation durations on the N1 when the compound was in subject position (see Table 2). The model containing ART as a fixed effect also showed a main effect of compound position, and a 3-way interaction between ART score, compound position and error type approached significance (t = -1.62), but nothing else was significant. This model was abandoned in favor of the simple model.

Table 2

Total Fixation Duration on N1

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	628.28	44.72	14.05	
Compound	-34.47	11.5	-2.997	*
Error	-0.87	11.99	-0.07	
Compound x Error	0.57	11.47	0.05	

Note: * marks a significant *t*-value at $p < .05$

A separate model was constructed including compound position, error type, and RSPAN score as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for compound position, error type, and their interaction (see Table 3). This resulted in a main effect of RSPAN score such that higher RSPAN scores led to shorter total fixation durations, and a main effect of compound position such that readers showed longer reading times on the N1 when in subject position.

Table 3

Total Fixation Duration on N1 with RSPAN as a Fixed Effect

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	628.27	42.88	14.65	
RSPAN	-250.1	112.61	-2.22	*
Compound	-34.53	11.66	-2.96	*
Error	-0.81	12.02	-0.07	
RSPAN x Compound	-28.33	37.67	-0.75	
RSPAN x Error	-24.11	35.56	-0.68	
Compound x Error	0.51	11.23	0.05	
RSPAN x Compound x Error	-51.45	34.54	-1.49	

Note: * marks a significant *t*-value at $p < .05$

Regression-in count. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for compound position, error type, and their interaction was constructed. This model showed a main effect of compound position such that participants had more regressions into the N1 when the compound was in subject position (see Table 4). A separate model was created including ART score as a fixed effect, but this model showed only a main effect of compound position and ART did not contribute significantly to the model, and thus was abandoned in favor of the original model. The model containing RSPAN score as a fixed effect revealed only a main effect of compound position, and so it was abandoned in favor of the simple model, although a main effect of RSPAN score approached significance (t = -1.69).

Table 4

Regression In Count for N1

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	0.65	0.06	10.44	
Compound	-0.13	0.04	-3.65	*
Error	-0.01	0.02	-0.51	
Compound x Error	-0.004	0.03	-0.15	

Note: * marks a significant *t*-value at $p < .05$

Target region N2.

Total fixation duration. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for compound position, error type, and their interaction was constructed. This model showed a main effect of compound position, such that participants had longer total fixation durations on the N2 when the compound was in subject position (see Table 5).

Table 5

Total Fixation Duration for N2

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	599.52	38.72	15.48	
Compound	-59.31	14.7	-4.04	*
Error	18.14	12.27	1.48	
Compound x Error	-9.57	111.52	-0.83	

Note: * marks a significant *t*-value at $p < .05$

A separate model was constructed including compound position, error type, and ART score as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position and error type, and by-item random slopes for compound position and error type (see Table 6). This model resulted in a main effect of compound position such that participants spent longer looking at the N2 when the compound was in subject position. There was also a significant interaction between compound position and ART score, and a significant 3-way interaction between compound position, error type, and ART score. Participants with higher ART score (more print exposure) showed a decrease in average total fixation time on the N2, especially when it was in subject position and was a faux tense error. The model including RSPAN as a fixed effect showed a main effect of compound position and a marginally significant main effect of RSPAN (t = -1.73).

Table 6

Total Fixation Duration for N2 with ART as a Fixed Effect

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	599.43	37.95	15.8	
ART	-100.95	69.33	-1.46	
Compound	-59.17	14.27	-4.15	*
Error	18.26	11.74	1.56	
ART x Compound	48.49	24.46	1.98	*
ART x Error	-17.85	24.35	-0.73	
Compound x Error	-9.72	9.98	-0.97	
ART x Compound x Error	-46.1	21.72	-2.12	*

Note: * marks a significant *t*-value at $p < .05$

Gaze duration. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for error type was constructed. This model showed a main effect of compound position, such that participants had longer gaze durations on the N2 when the compound was in subject position (see Table 7).

Table 7

Gaze Duration for N2

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	292.28	7.6	38.47	
Compound	-9.99	4.96	-2.01	*
Error	-0.3	5.004	-0.06	
Compound x Error	-1.19	4.22	-0.28	

Note: * marks a significant *t*-value at *p* < .05

A separate model including compound position, error type, and RSPAN score as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, and by-item random slopes for compound position and error type. This model resulted in an interaction between RSPAN and error type (see Table 8), such that higher RSPAN scores resulted in longer gaze durations, primarily for the faux tense error condition (e.g., *Last week…*). A main effect of compound position approached significance.

Table 8

Gaze Duration for N2 with RSPAN as a Fixed Effect

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	292.43	7.36	39.74	
RSPAN	27.62	18.34	1.51	
Compound	-9.81	5.21	-1.88	
Error	-0.39	4.49	-0.09	
RSPAN x Compound	-11.44	16.41	-0.7	
RSPAN x Error	-28.65	13.19	-2.17	*
Compound x Error	-1.28	3.9	-0.33	
RSPAN x Compound x Error	8.61	13.13	0.66	

Note: * marks a significant *t*-value at *p* < .05

Go-past time. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, and by-participant random slopes for compound position and error type was constructed. In contrast with the other reading time measures, this model showed nothing significant, although a main effect of compound position approached significance (t = -1.64).

Another model was constructed including compound position, error type, and ART as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for compound position (see Table 9). This resulted in a significant 3-way interaction between ART score, compound position, and error type such that participants with higher ART score (more print exposure) showed a decrease in average total fixation time on the N2 when it was in object position and was a faux agreement error.

Table 9

Go-past Time for N2 with ART as a Fixed Effect

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	403.29	14.46	27.9	
ART	-12.86	24.14	-0.53	
Compound	-16.9	11.25	-1.5	
Error	8.68	10.67	0.81	
ART x Compound	-33.05	21.72	-1.52	
ART x Error	-7.93	22.97	-0.35	
Compound x Error	2.32	9.41	0.25	
ART x Compound x Error	-40.61	20.32	-1.99	*

Note: * marks a significant t-value at $p < .05$

A separate model included compound position, error type, and RSPAN as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position and error type, and by-item random slopes for compound position. This resulted in a marginally significant interaction between RSPAN and error type (t = -1.87) such that as RSPAN score increased, go-past times on the N2 also increased especially when there was a faux tense error. The main effect of compound position also approached significance (t = -1.49).

Regression-in count. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position and error type, and by-item random slopes for compound position, error type, and their interaction was constructed. This resulted in a main effect of compound position that, like previous models, indicated participants had on average more regressions into the N2 when the compound was in subject position (see Table 10). The model containing ART as a fixed effect also showed a main effect of compound position, but nothing else was significant and so the model was abandoned in favor of the simple model.

Table 10

Regression In Count for N2

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	0.52	0.05	11.48	
Compound	-0.11	0.03	-3.78	*
Error	0.02	0.02	0.91	
Compound x Error	-0.02	0.02	-0.78	

Note: * marks a significant t-value at $p < .05$

A separate model was constructed containing compound position, error type, and RSPAN as fixed effects, random intercepts by participant and item, by-participant random slopes for compound position and error type, and by-item random slopes for compound position, error type, and their interaction. This resulted in a main effect of compound position and an interaction between compound position and RSPAN score (see Table 11) such that higher RSPAN scores meant a decrease in the number of regressions into the N2, especially when the compound was in subject position. In accordance with the simple model, there were still overall more regressions into the N2 when the compound was in subject position.

Table 11

Regression In Count for N2 with RSPAN as a Fixed Effect

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	0.52	0.04	11.53	
RSPAN	-0.1	0.13	-0.79	
Compound	-0.1	0.03	-4.003	*
Error	0.02	0.02	0.94	
RSPAN x Compound	0.21	0.08	2.71	*
RSPAN x Error	0.06	0.06	0.87	
Compound x Error	-0.02	0.02	-0.78	
RSPAN x Compound x Error	-0.05	0.06	-0.75	

Note: * marks a significant *t*-value at $p < .05$

Target region N3.

Total fixation duration. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position and error type, and by-item random slopes for compound position, error type, and their interaction was constructed. This resulted in a main effect of compound position that, like previous models, indicated participants had on average more regressions into the N2 when the compound was in subject position (see Table 12).

Table 12

Total Fixation Duration for N3

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	619.84	37.06	16.73	
Compound	-37.76	16.02	-2.36	*
Error	-14.98	11.59	-1.29	
Compound x Error	-11.16	11.25	-0.99	

Note: * marks a significant *t*-value at $p < .05$

Another model containing compound position, error type, and ART score as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for compound position, error type, and their interaction. This model revealed main effects of compound position and ART score, as well as an interaction between ART and compound position such that total reading time was longer when the compound was in subject position, but higher ART scores resulted in a decrease in total fixation duration for the N3 of the compound in subject position in comparison to object condition (i.e., made subject position nearly as easy as object position). There was also an interaction between ART and error type (see Table 13) such that higher ART scores resulted in a greater decrease in total fixation duration for the faux tense error in comparison to the faux agreement error. A separate model containing RSPAN score as a fixed effect revealed only a main effect of compound position, consistent with the results of the simple model.

Table 13

Total Fixation Duration for N3 with ART as a Fixed Effect

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	619.99	35.71	17.36	
ART	-125.91	64.17	-1.96	*
Compound	-37.91	15.47	-2.45	*
Error	-15.04	10.68	-1.41	
ART x Compound	52.85	22.99	2.3	*
ART x Error	60.12	21.59	2.79	*
Compound x Error	-10.97	11.25	-0.96	
ART x Compound x Error	9.68	19.92	0.49	

Note: * marks a significant *t*-value at $p < .05$

Target region MV.

First fixation duration. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for compound position and error type was constructed. This revealed a main effect of compound position where participants had longer first fixation durations on average for the MV when the tripartite compound was in subject position (see Table 14). The model containing ART as a fixed effect also showed a main effect of compound position, but nothing else was significant and so the model was abandoned in favor of the simple model.

Table 14

First Fixation Duration for MV

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	244.69	4.14	59.08	
Compound	-13.55	3.08	-4.41	*
Error	4.63	2.89	1.6	
Compound x Error	3.2	2.96	1.08	

Note: * marks a significant *t*-value at $p < .05$

A separate model containing compound position, error type, and RSPAN as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position and error type, and by-item random slopes for compound position and error type revealed a main effect of RSPAN score and a main effect of compound position (see Table 15). This effect showed that participants had longer first fixation durations on the MV when the compound appeared in subject position, and participants with higher RSPAN scores also had longer first fixation durations on average on the MV region.

Table 15

First Fixation Duration for MV with RSPAN as a Fixed Effect

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	244.72	3.93	62.34	
RSPAN	24.61	12.25	2.01	*
Compound	-13.55	2.99	-4.52	*
Error	4.63	2.89	1.6	
RSPAN x Compound	-11.1	9.85	-1.13	
RSPAN x Error	-4.59	9.64	-0.48	
Compound x Error	3.2	2.6	1.23	
RSPAN x Compound x Error	1.96	8.67	0.23	

Note: * marks a significant *t*-value at $p < .05$

Gaze duration. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position and error type, and by-item random slopes for compound position, error type, and their interaction was constructed. The model revealed a main effect of compound position where readers had longer gaze durations on the MV when the tripartite compound was in subject position (see Table 16). The model containing ART as a fixed effect also showed a main effect of compound position, but nothing else was significant and so the model was abandoned in favor of the simple model.

Table 16

Gaze Duration for MV

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	299.1	10.45	28.62	
Compound	-16.89	5.15	-3.28	*
Error	7.53	4.73	1.59	
Compound x Error	5.32	4.79	1.11	

Note: * marks a significant *t*-value at $p < .05$

A separate model was constructed with compound position, error type, and RSPAN as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position and error type, and by-item random slopes for compound position, error type, and their interaction. This revealed a main effect of compound position such that participants had longer gaze durations on the MV when the compound was in subject position, and a main effect of RSPAN approached significance (*t* = --1.898). This is pattern is consistent with the one found for first fixation duration on the MV, such that higher RSPAN scores resulted in longer gaze durations on average. It is likely that with more participants, this effect would reach significance.

Go-past time. A simple model with compound position and error type as fixed effects, random intercepts for participant and item, by-participant random slopes for compound position, and by-item random slopes for compound position, error type, and their interaction was constructed. This model resulted in a main effect of compound position such that participants had longer go-past times on the MV when the compound was in subject position (see Table 17). The separate models containing ART and RSPAN as a fixed effects also showed a main effect of compound position, but nothing else was significant and so these models were abandoned in favor of the simple model.

Table 17

Go-past Time for MV

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	411.9	23.4	17.6	
Compound	-36.59	17.13	-2.14	*
Error	7.49	11.23	0.66	
Compound x Error	-1.09	11.77	-0.09	

Note: * marks a significant *t*-value at $p < .05$

Regression-in count. A simple model with compound position and error type as fixed effects, random intercepts for participants and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for compound position, error type, and their interaction revealed a main effect of compound position, such that there were more regressions into the MV when the compound was in object position (see Table 18). The separate models containing ART and RSPAN as a fixed effects also yielded a main effect of compound position with ART and RSPAN not contributing significantly to the model, thus these models were abandoned in favor of the simple model.

Table 18

Regression In Count for MV

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	0.48	0.04	10.81	
Compound	0.07	0.03	2.74	*
Error	-0.02	0.02	-0.87	
Compound x Error	0.01	0.02	0.37	

Note: * marks a significant *t*-value at $p < .05$

Discussion

Of the simple models constructed for each target region, all showed a main effect of compound position except for go-past time on the N2, which approached significance. Based on these results, it is clear that readers spent significantly longer reading the N1, N2, and N3 when the compound appeared in subject position. Participants also spent longer looking at the MV when the compound was in the subject position (and thus came before the MV). The simple models also revealed that individuals made more regressions into the N2 when it appeared in subject position, while there were more regressions into the MV when the compound appeared in object position. This pattern strongly suggests that readers do evaluate both the N2 and MV as candidates for main verb; however, when the compound was in object position, this competition was resolved significantly more quickly than when the compound was in subject position. Based on these results, it appears that processing was overall more difficult when the compound was in the subject position. These results are consistent with the prediction that individuals would have more difficulty with the compound in the subject position due to the tendency to interpret the N2 as the main verb of the sentence. Based on these simple models, there is not much evidence for local coherence effects when the compound was in object position. Nevertheless, the inflated rate of regressions-in on the MV region when the compound was in object position suggests that readers experienced interference from the locally coherent apparent NVN string on at least some trials.

While the above story is sweet and simple, it becomes more complicated when ART and RSPAN scores were included. For the models containing ART score as a fixed effect, total fixation duration on the N2 and N3 and go-past time on the N2 revealed significant interactions between ART and the other fixed effects. For the models containing RSPAN score as a fixed effect, gaze duration for the N2, regression in count for the N2, and total fixation duration on the N1 revealed main effects or interactions for RSPAN score and the other fixed effects. A main effect of RSPAN also approached significance for gaze duration on the MV, regression in count on the N1, and total fixation duration on the N2. An interaction between RSPAN and error type for go-past time on the N2 also approached significance. It is interesting to note that for reading time measures on each target region where the ART models yielded significant results, the RSPAN models did not and vice versa.

To ease interpretation, the results of the ART and RSPAN models were graphed in R using ddply() of the plyr package (Wickham, version 1.8.1, 2014) and ggplot() of the ggplot2 package (Wickham & Chang, version 1.0.0, 2014). Results were collapsed across participants and items, and the mean reading times for each condition were graphed based on ART or RSPAN score. I will first discuss the significant ART models, and then the significant RSPAN models.

ART models. For total fixation duration on the N2 target region, the ART model revealed a main effect of compound position (which was consistent with the simple model), an interaction between compound position and error type, as well as a 3-way interaction between ART, compound position, and error type (see Figure 1). Based on these results, it seems that participants with higher ART scores (more print exposure) showed a decrease in average total fixation time for the N2, especially when it occurred in subject position and was a tense error.

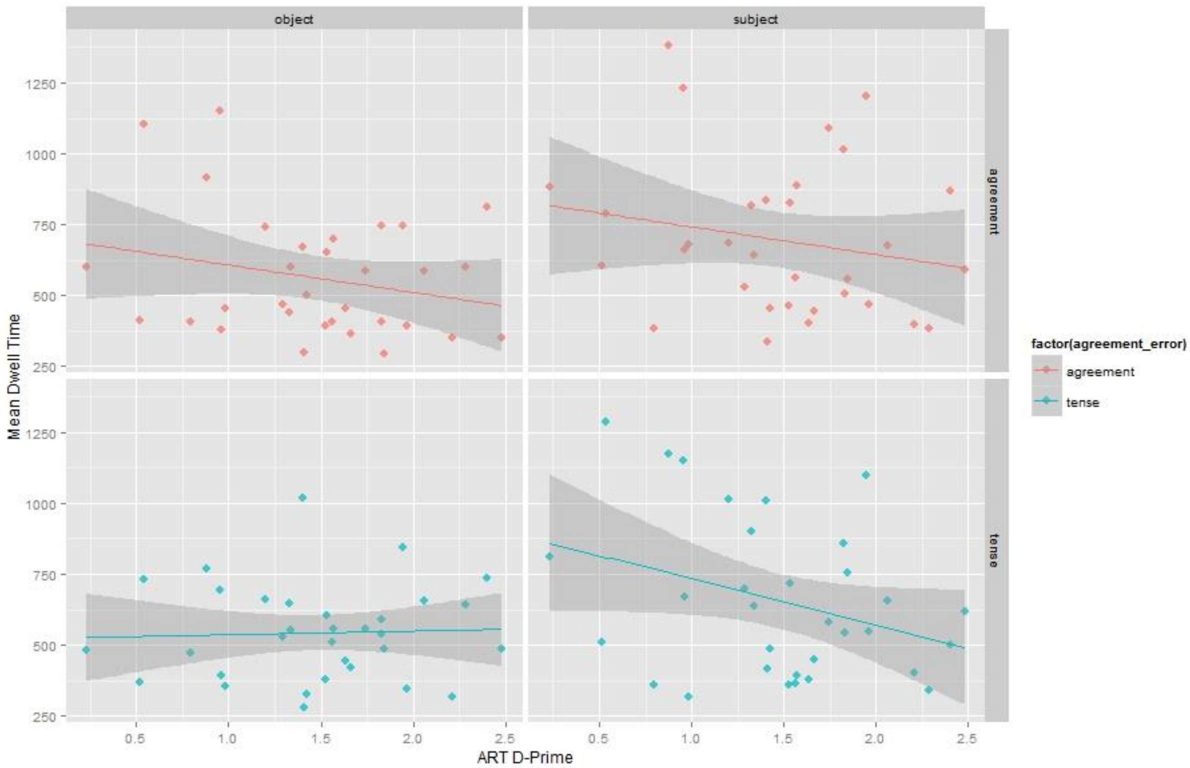

Figure 1. Total Fixation Duration for N2 with ART as Fixed Effect (facet)

The model for go-past time on the N2 revealed a 3-way interaction between ART score, compound position, and error type such that higher ART scores resulted in shorter go-past times on the N2 when the compound was in the object position and was a faux agreement error (see Figure 2).

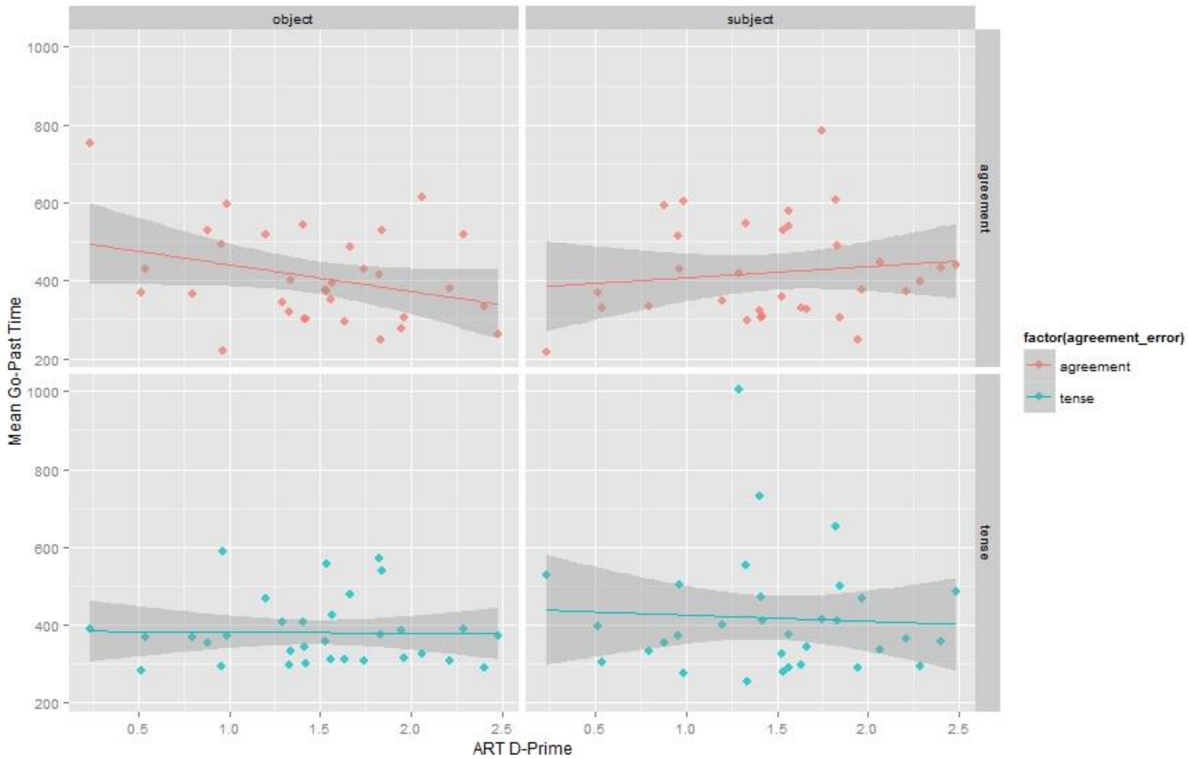

Figure 2. Go-Past Time for N2 with ART as Fixed Effect (facet)

The ART model for total fixation duration on the N3 yielded a main effect of compound position (which was consistent with the simple model), a main effect of ART score, an interaction between ART and compound position, and an interaction between ART and error type (see Figure 3). It appears that participants with higher ART scores (more text exposure) showed a decrease in average total fixation time on the N3, particularly when in the faux tense error condition (e.g., preceded by the *Last week…* sentence context) and the subject position condition. This could be because individuals with more print exposure are have had more experience with these types of compounds, which are frequent as a class, although any particular token of these compounds is extremely rare.

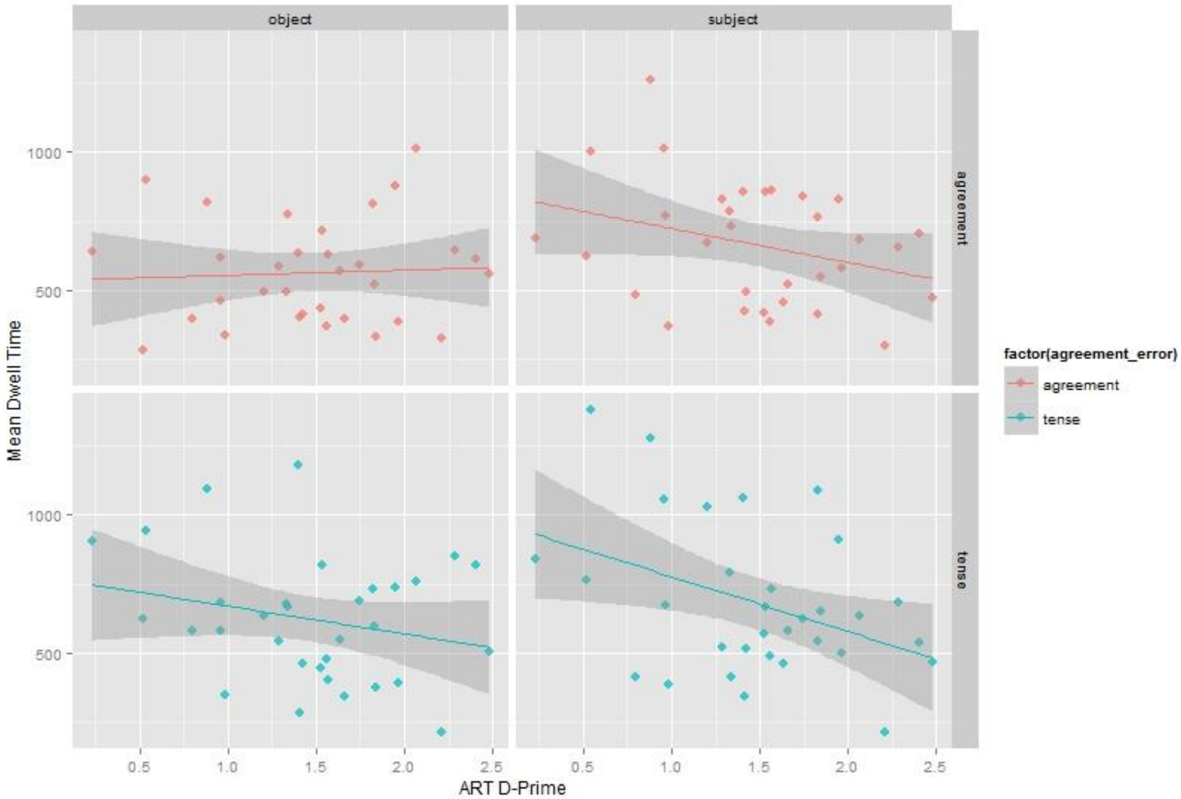

Figure 3. Total Fixation Duration for N3 with ART as Fixed Effect (facet)

RSPAN models. For total fixation duration on the N1 target region, the RSPAN model revealed a main effect of RSPAN score and a main effect of compound position that was consistent with the simple model. Figure 4 reveals that as RSPAN score increased, individuals spent significantly less total time fixating on the N1. I interpret this to mean that individuals with higher working memory capacity are at an advantage in that they either have an increased reading rate (Traxler et al., 2012), or they can more easily integrate written information into the sentence context.

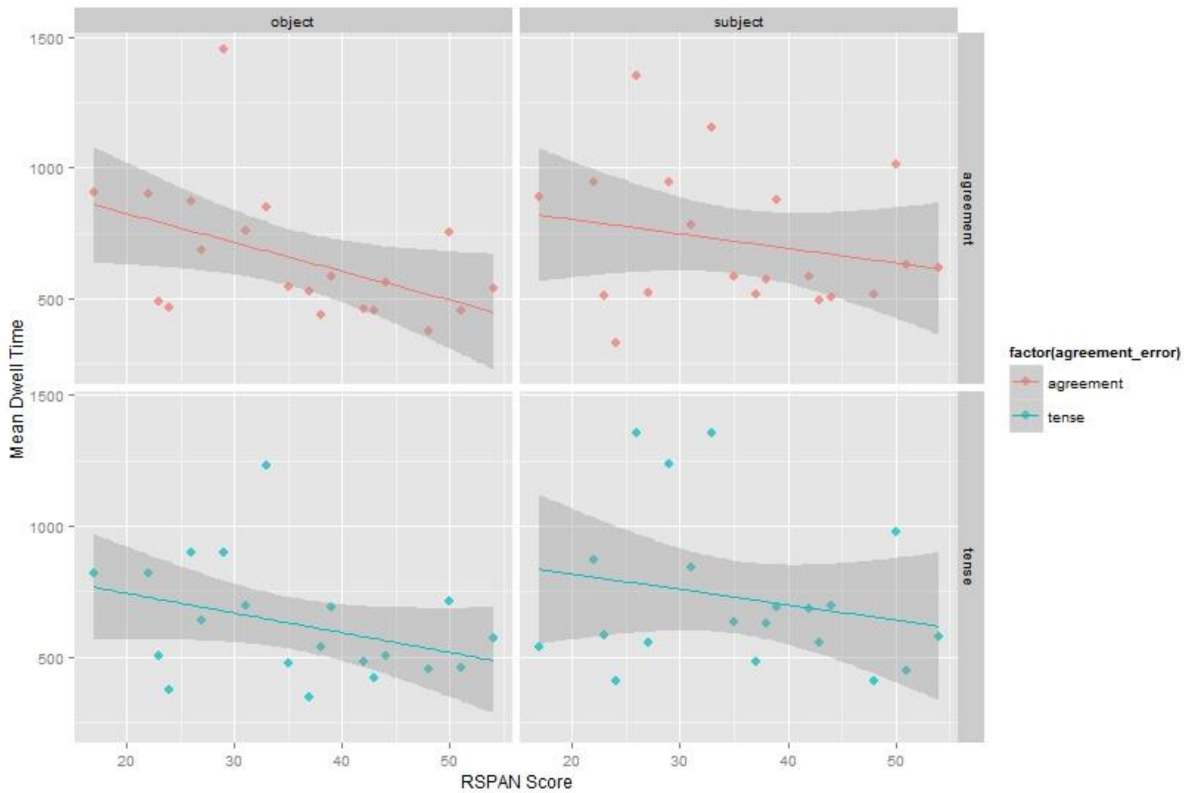

Figure 4. Total Fixation Duration for N1 with RSPAN as Fixed Effect

The RSPAN model for total fixation duration on the N2 yielded a main effect of compound position consistent with the simple model, but a main effect of RSPAN approached significance as well (see Figure 5). The pattern observed in the graph, while not significant, suggests the same as the graph for the N1: readers with higher RPSAN scores spent less total time fixating on the N2.

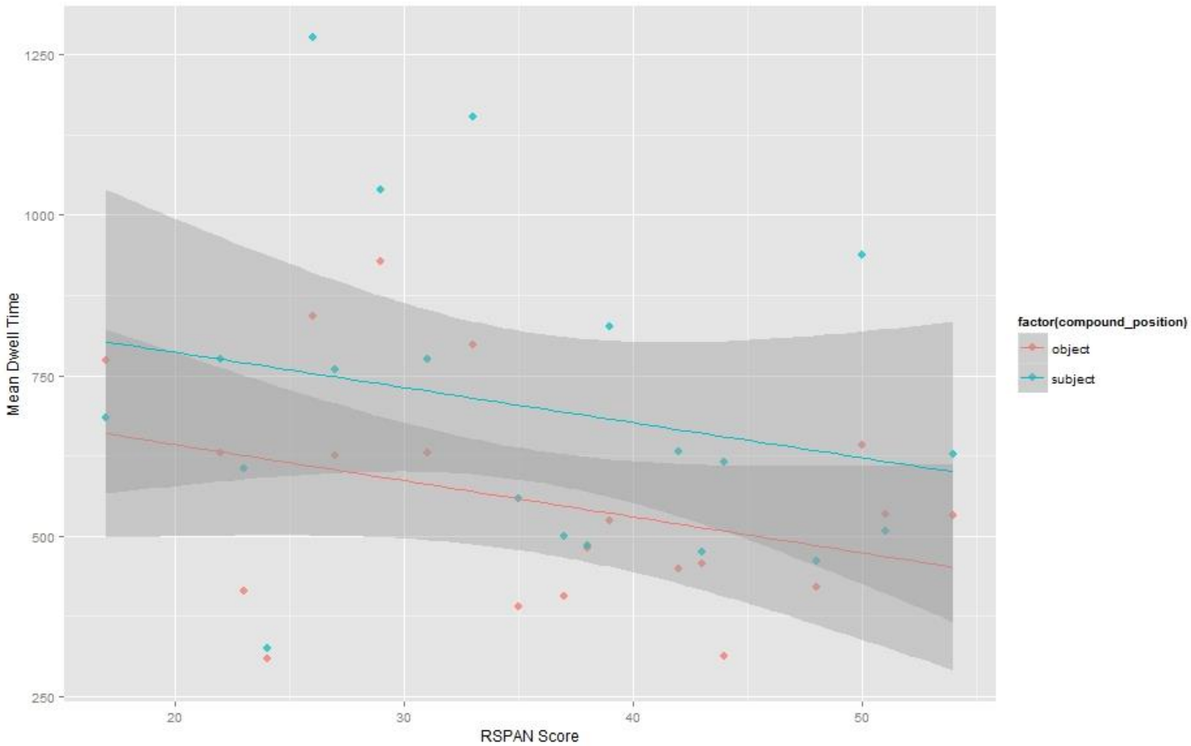

Figure 5. Total Fixation Duration for N2 with RSPAN as Fixed Effect

The RPSAN model for gaze duration on the N2 revealed a marginally significant main effect of compound position and an interaction between RSPAN and error type (see Figure 6). At first, these results were counter to what I was expecting. I predicted that I would see higher RSPAN scores resulting in shorter average reading times, meaning that individuals with higher working memory (HWM) have an easier time interpreting the compounds. However, upon further reflection, it appears that individuals with higher RSPAN scores are better able to recognize the faux tense error when they finally get to the N2. The HWM individuals are expecting a verb that matches in tense to the beginning of the sentence (e.g., *Last week…*) and therefore they are boggling *more* (not less) because they are likely interpreting the N2 as the MV and recognizing the faux tense error. So, the HWM individuals are more disrupted by the faux-tense error than the low working memory (LWM) individuals. This is especially true for the subject position/faux tense error condition, where the individual has not yet encountered the true MV by the time he/she reads the tripartite compound.

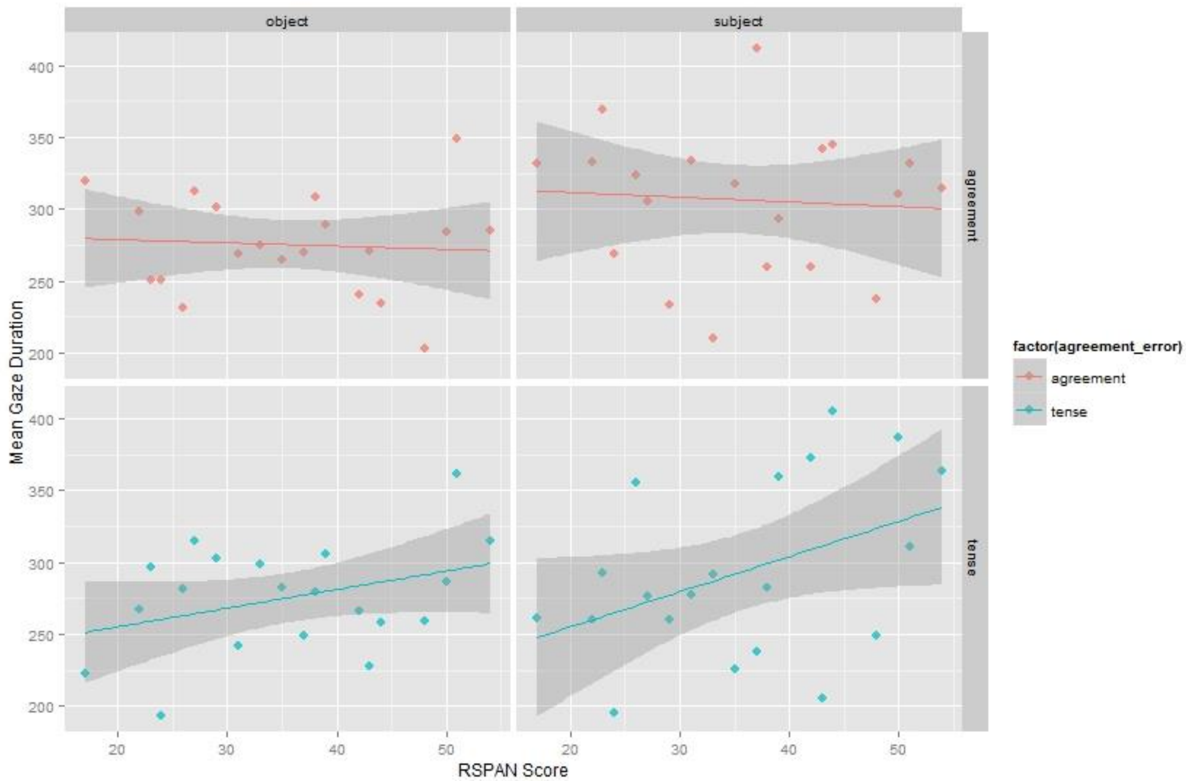

Figure 6. Gaze Duration for N2 with RSPAN as Fixed Effect (facet)

It could be that we don't see as large of an effect/difference between HWM and LWM groups for the agreement error condition because those sentences were not as tightly controlled. The sentences that begin with *Every weekend…* can lead to a number of tenses and possible constructions of the N2 (e.g., *Every week, the dog bites/bit*). It could therefore be the case that any effect was washed out, or participants simply were not as confused by the N2 because they had not predicted which specific verb form they should encounter in the upcoming sentence regions (cf. Luke & Christianson, 2015).

The RSPAN model for Go-past time on the N2, while not significant, showed similar results to gaze duration: a main effect of compound position approached significance, as did an interaction between RSPAN and error type (see Figure 7). Again, HWM individuals were disrupted more by the faux tense error (elicited, presumably, by the *Last week…*sentence contexts) because, I suspect, they are better able to remember it. On the other hand, LWM individuals seem to be less disrupted (resulting in shorter fixation durations), perhaps because they do not remember the first part of the sentence or did not use it to generate any specific predictions about the form of the upcoming verb.

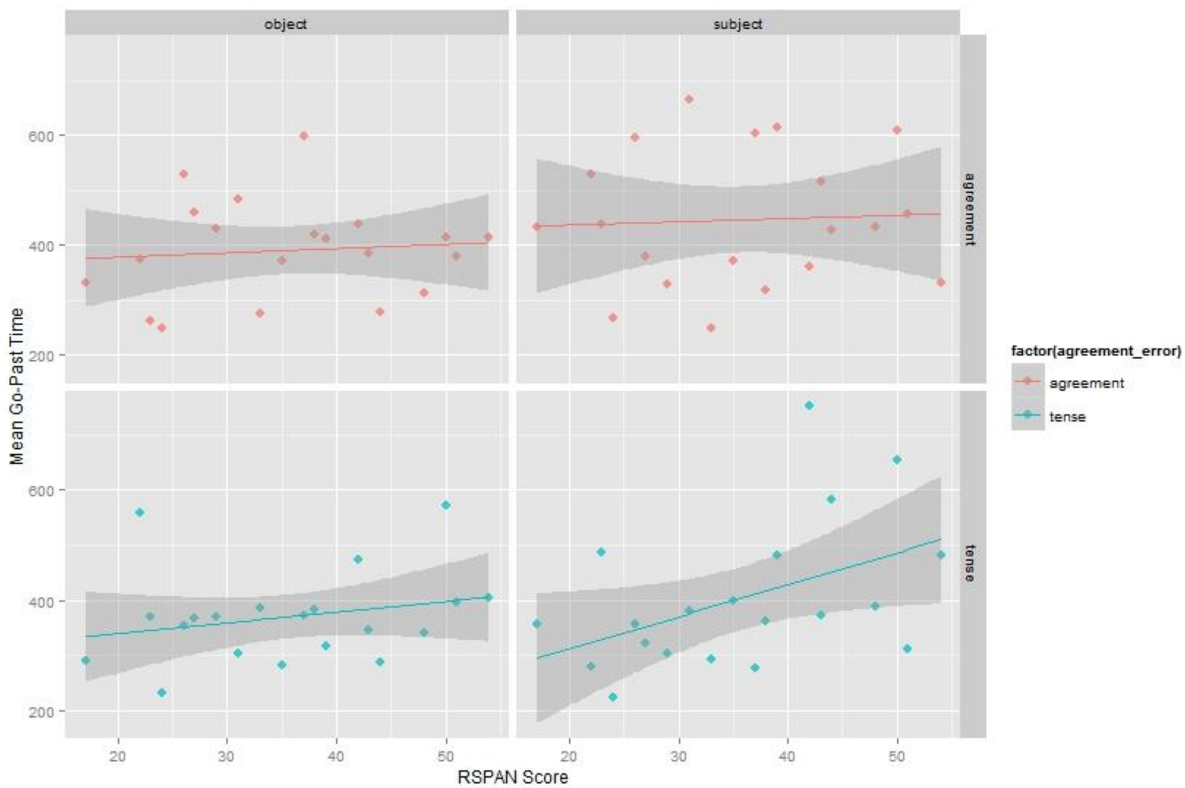

Figure 7. Go-Past time for N2 with RSPAN as Fixed Effect (facet)

The RSPAN model for regression in count on the N2 revealed a main effect of compound position and an interaction between RSPAN and compound position (see Figure 8). As working memory increased, individuals made fewer regressions into the N2 in subject position, perhaps because they were able to keep the N2 in focus in memory and therefore did not need to physically look back to refer to it. I cannot say, therefore, that individuals did not have to refer back to the N2 whatsoever. It is unclear as to where the participants regressed into the N2 from, however, since there was nothing significant in the models that were constructed for regression-out count for the N3 or MV regions. However, it is clear that individuals, regardless of RSPAN score, made more regressions into the N2 in subject position as opposed to object position. HWM individuals, however, made fewer regressions into N2 than their LWM counterparts when the N2 was in subject position.

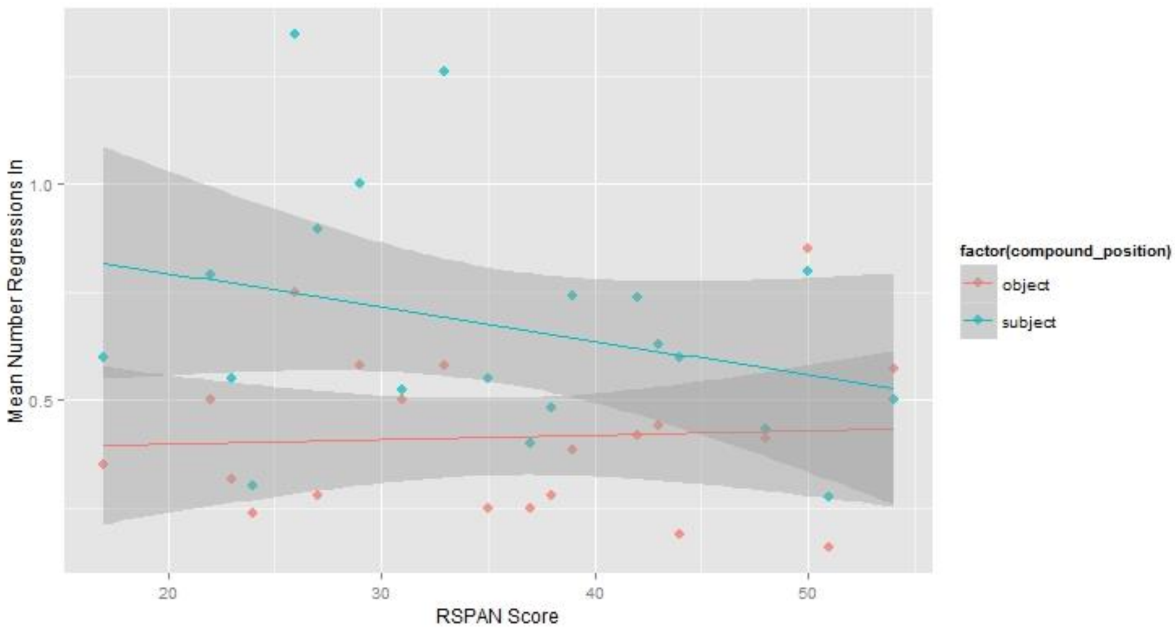

Figure 8. Regression in Count for N2 with RSPAN as Fixed Effect

The RSPAN model for first fixation duration on the MV showed a main effect of RSPAN score and a main effect of compound position. Higher RSPAN scores resulted in longer average reading times, regardless of condition (see Figure 9). It is interesting to note that this was the opposite effect than that observed for total fixation duration on the N1, where higher RSPAN scores meant shorter average reading times. This is likely because no error signal was expected on the N1 (since the reader had not yet reached the ambiguous N2), whereas an error signal is expected once the reader encounters the MV. Based on the results of this model, it seems that

HWM individuals were more sensitive to the error signal once they encountered the real MV of the sentence.

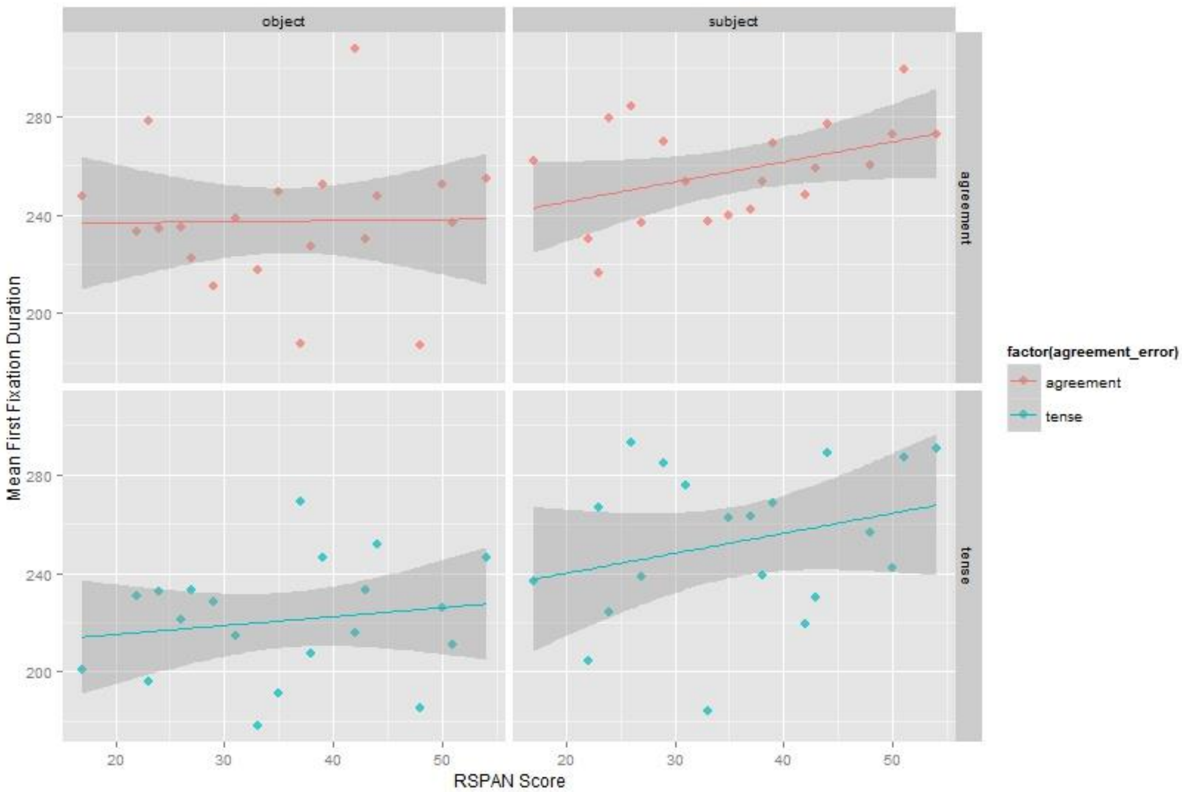

Figure 9. First Fixation Duration for MV with RSPAN as Fixed Effect (facet)

The RSPAN model for gaze duration on the MV showed a main effect of RSPAN score approaching significance (t = 1.9), and a main effect of compound position consistent with that found in the simple model (see Figure 10). These results mirror those of the first fixation duration for the MV. One plausible interpretation of these effects on reading times for the MV is as follows: If reading times on the MV index the relative difficulty of integrating the verb into the structure – e.g., matching it with its subject – then doing so when the subject is a tripartite compound required more processing effort than when it was a simple noun (i.e., when the tripartite compound was in object position). Readers with less working memory capacity may tend to "chunk" the three constituents of the compound into one NP quite quickly (as they should), reducing working memory load. On the other hand, readers with more working memory capacity may experience interference between the constituents of the compound, which might be represented in a less integrated representation for a longer period of time. Although this explanation is speculative, it is consistent with other recent counter-intuitive processing data related to working memory capacity. Swets, Desmet, Clifton, and Ferreira (2008) showed that low-capacity readers tend to attach ambiguous relative clauses to higher, more linearly distant attachment sites (high attachment), whereas high-capacity readers tend to attach more locally (low attachment). Swets et al. attributed this result to the same sort of chunking by low-capacity participants that is suggested here.

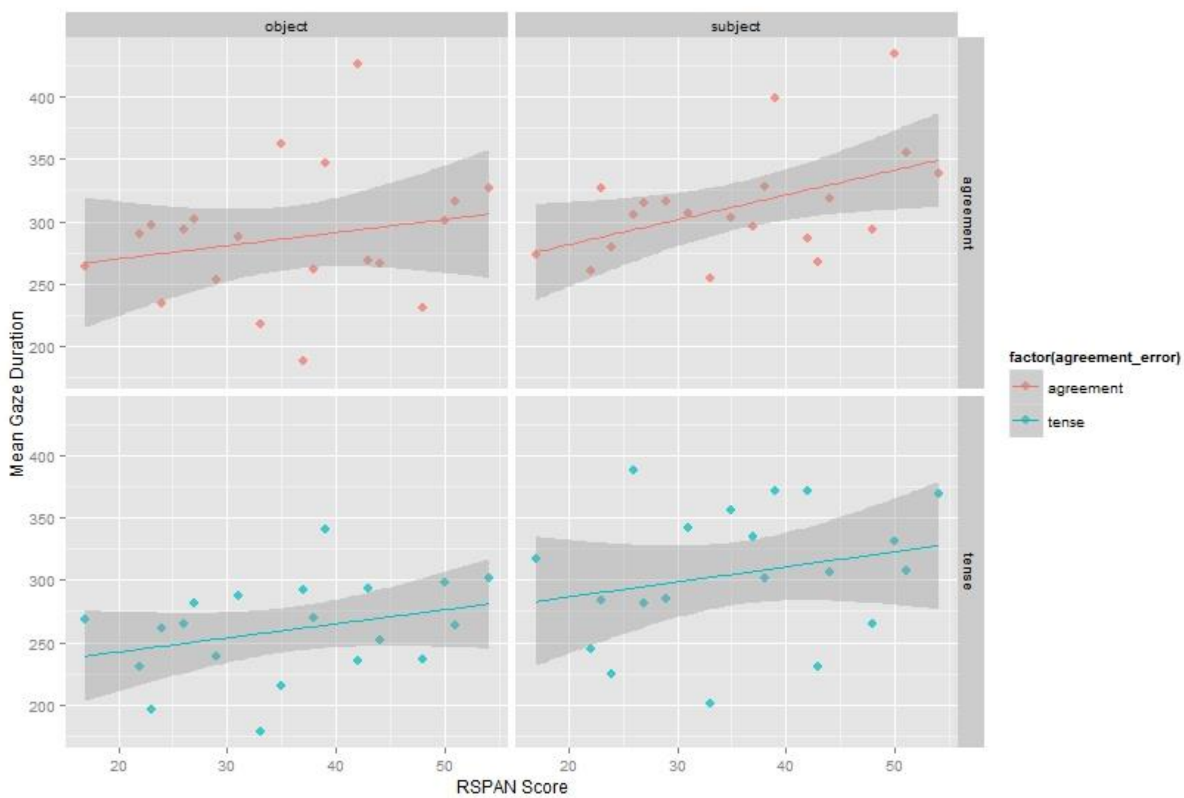

Figure 10. Gaze Duration for MV with RSPAN as Fixed Effect (facet)

With respect to the accuracy data, participants were highly accurate when answering the comprehension questions. I interpreted this to mean that, although readers are garden-pathed at the level of the compound when it appeared in subject position, they easily recovered by the time they are asked to answer the comprehension question. Thus, it seems that many individuals came away from each target sentence with the syntactically licensed interpretation – that *dog bite victim* is a compound word consisting of three nouns, with *victim* as the head noun.

A recent study by Cutter, Drieghe, and Liversedge (2014) might shed some light on what is (or is not) happening with the present study's tripartite compounds. In the 2014 study, the authors used spaced bipartite compounds in order to explore parafoveal preview effects. Previous studies have shown a preview benefit for the upcoming word (n+1) when fixated on the current word (n), resulting in shorter fixation times when the upcoming word is actually fixated (Schotter, Angele, & Rayner, 2012). Interestingly, Cutter et al. (2014) discovered preview benefit for word n+2 when words n+1 and n+2 were part of a spaced compound. Thus, it seems that preview benefit can extend beyond word n+1 when it is part of a spaced compound, or a single lexical unit. The authors also showed that the length and frequency of word n+1 (the first word of the compound) was not the determining factor for whether or not a preview benefit will occur for word n+2. This is an important finding, because previous studies had only found an n+2 preview benefit when word n+1 was short (three letters) and highly frequent. In sum, readers were able to process up to two words to the right of the currently fixated word when word n+1 was indicative of a compound, aided by the fact that word n+2 was highly predicted by word n+1. This, the authors theorized, indicated that the bipartite compounds used in the study were lexicalized, or accessible via a single lexical entry rather than via the two separate entries for each compound constituent.

But how do these results relate to the spaced tripartite compounds utilized in the present study? Cutter et al. note that compounds that are common should have a single lexical entry

(Bybee, 2006). However, the tripartite compounds used in the present study are highly novel, and, while not uninterpretable, are therefore not likely lexicalized. Cutter et al. state that further studies should explore what other multi-word units (MWUs) are lexicalized, and what exactly the stipulations are in order for MWUs to become lexicalized. The results of Cutter et al. (2014) revealed a pattern such that reading time measures decreased as the reader progressed through the compound, whereas the reading time measures for the present study do not seem to reveal the same trend (see Table 1). If anything, the opposite pattern is observed, such that reading times tend to increase from the N1 to the N2 and N3 constituents. This argues against the notion that these tripartite compounds are lexicalized – which is not surprising. However, the compounds used in the present study contain ambiguities that the bipartite compounds of Cutter et al. did not, so the two sets of stimuli are not directly comparable. It could be that, even if readers had a single lexical entry for *dog bite*, access to this entry was blocked by the fact that the N2 *bite* was likely processed as a verb rather than part of a bipartite compound. And, if the reader then processed the N3 as the object, it would not have been until the MV that the ambiguity of the tripartite compound was fully recognized and the syntactic parse revised.

The results provide limited evidence for a purely local coherence effect. There was a noticeable difference in eye-movement data for the compound in subject as opposed to object position, such that readers had more difficulty (longer reading times) for the compound in subject position. If, as I hypothesized early on, there was a local coherence effect, I would expect readers to have equal trouble with the compound regardless of sentence position – but this was not apparent in the data. However, it is interesting to note that there were more regressions into the MV when it appeared before the compound (e.g. the compound was in object position). This could indicate that readers did have some difficulty with the compound and thus had to refer back to the MV in order to ensure that they indeed had the correct sentence parse. Another possible explanation, however, is that words occurring earlier in a sentence are more likely to be regressed into when compared to words at the end of a sentence – which is a much less interesting explanation. If indeed the increased number of regressions into the MV when the compound is in object position indicates some level of processing difficulty, this would provide some evidence for local coherence as described by Tabor et al. (2004), and thus I cannot entirely rule out any local coherence effect in the current data.

It is possible that the increased reading times on the compound when in subject position are due simply to readers reading more slowly at the beginning of a sentence and speeding up as the sentence context becomes more constraining (and thus easier to predict upcoming input). Thus, a separate linear mixed effects model was created based on total fixation duration for each interest area in the entire subset of target sentences (not just the compound constituents and main verb). Ordinal sentence position, which varied from 1-12 (each target sentence was between 10 and 12 words long), was log transformed and mean-centered. This was then included as a fixed effect along with compound position and error type, with random intercepts for participant and item, by-participant random slopes for compound position, error type, and their interaction, and by-item random slopes for compound position, error type, and their interaction (see Table 19).

Table 19

Total Reading Time Based on Sentence Position

Parameters	Fixed effects			
	Estimate	*SE*	*t*	
Intercept	580.53	28.94	20.06	
Sentence Position	-17.41	5.18	-3.36	*
Compound	-0.7	6.33	-0.11	
Error	3.14	6.83	0.46	
Sentence Position x Compound	28.48	5.17	5.51	*
Sentence Position x Error	-5.99	5.17	-1.16	
Compound x Error	2.55	5.93	0.43	
Sentence Position x Compound x Error	-4.29	5.17	-0.83	

Note: * marks a significant *t*-value at $p < .05$

This resulted in a main effect of sentence position such that as sentence position increased, so did the total fixation duration on the word occupying that sentence position (see Figure 11). There was also an interaction between sentence position and compound position such that as sentence position increased, total reading time increased when the compound was in object position. The opposite was true, however, for the compound in subject position, such that as sentence position increased, total reading time decreased. This suggests that readers did on average speed up as they read the sentences, but *only* when the compound was in subject position; when it was in object position, reading slowed as the sentence progressed. This pattern is consistent with the premise that the compound required extra processing no matter where it occurred in the sentence. Total fixation durations on the sentence regions were longer at the end of the sentence when the compound also appeared toward the end of the sentence, but longer at the beginning of the sentence when the compound also appeared toward the beginning, thus suggesting that the main effect of compound position seen in the previously reported models is not likely driven by participants reading faster as they progress through the target sentences, but is indeed evidence of a garden path effect caused by the ambiguous tripartite compound.

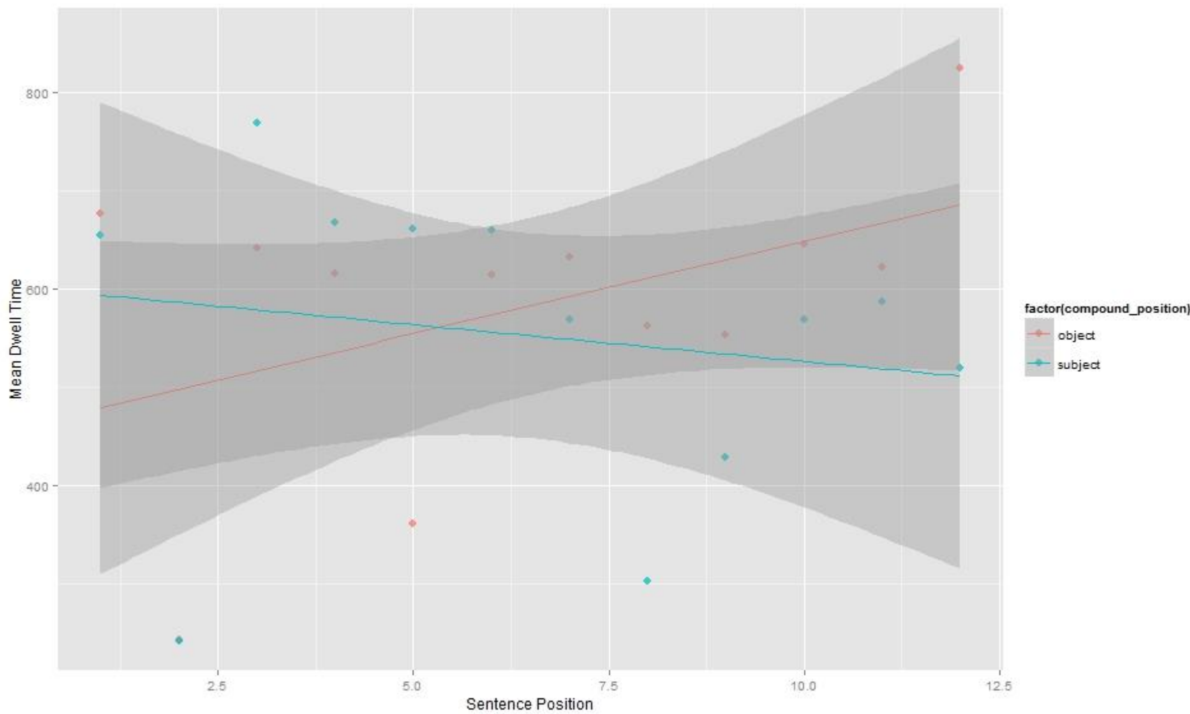

Figure 11. Total Reading Time Based on Sentence Position

Limitations and Future Directions

Due to the novel nature of the stimuli, there are some potential limitations to the current study. The exemplar compound I have chosen, *dog bite victim*, has a middle constituent that is an irregular verb. *Bite* is irregular in that it doesn't take the typical -ed past tense suffix, such as *attack* (see Appendix).

Overall, there were four N2 constituents whose verb forms were irregular and did not take the past tense -ed suffix. It is possible that readers treated these four irregular N2 constituents slightly differently than the regular N2 constituents. For instance, consider "Last year, the appealing food drive organization helped a friend in New Orleans," which was intended to introduce a faux tense error when the reader encountered the N2 *drive*. It is possible that it is easier to mistake the N2 constituent *attack* as being the main verb, because the reader could write it off as a spelling mistake such that the original author forgot to add the -ed ending. All that the reader needs to do is mentally insert the past tense suffix. However, with an irregular verb like *drive*, the reader needs to recognize that the past tense should be *drove*, which involves both a deletion and an insertion rather than a simple insertion. Thus, it could be that these two kinds of N2 constituents result in different behaviors in the eye-movement data.

Another potential limitation of the stimuli involves the sentence contexts. The sentences that began with the past tense (e.g., *Last week…*) and were constructed to induce a faux tense error licensed a limited number of verb tenses, most likely the simple past tense (e.g. *bit, attacked*). The sentences that began with the present tense (e.g. *Every weekend…*) were constructed with the intention that the reader would also expect a present tense verb that agrees with the preceding noun (e.g. *dog bite(s)*). However, *Every weekend* can elicit a number of different verb tenses, including past tense, present tense, and future tense, thus making the reader's predictions for these sentence contexts less constrained. This could have washed out any potential differences between the two error types.

Cutter et al. (2014) utilized a cloze task for their bipartite compounds in order to determine how predictable the second constituent was if given the first constituent. In fact, their stimuli were created so that the second constituent was highly predictable based on the first constituent, but not on sentence context. While the tripartite compounds for the present study were normed for imageability, it could be that a cloze task for the compounds would help determine if either the N2 or N3 constituents were highly predictable, since that might have had an effect on reading times

The results of the present study suggest that individuals with higher working memory have an advantage over their lower working memory counterparts when it comes to identifying the ambiguous tripartite compounds, but perhaps not in integrating the compound as a unified subject of the main verb. It would be interesting to conduct future experiments that assess whether individuals with higher RSPAN scores (and higher working memory capacity) are more productive when it comes to creating, rather than simply reading, novel compounds that are composed of three or more constituents. Participants with more text exposure, however, were more efficient at recovering from any garden-path effects when the compound was in subject position, likely due to their wider experience with compounding in English. As noted briefly above, including both ART and RSPAN into the same models proved problematic. An exhaustive examination of these and related individual difference measures is needed to determine their relative degrees of intercorrelation and which underlying mechanisms they are tapping into. Another interesting follow-up experiment could assess whether or not readers with higher RSPAN scores are disrupted by inconsistencies in sentences such as the following: *After Bill walked [into the room/outside], he wondered whether Marie still loved him, remembered their long days together in Paris, and how she wrinkled her nose at his awkward French accent, and then he looked up at the [gray clouds/gray ceiling] and cried*. It would be interesting to see if higher RSPAN individuals, when presented with inconsistent information would be more disrupted than their lower RSPAN counterparts because the higher RSPAN individuals are better able to recognize the inconsistency (depending on Bill's location).

Another expansion for the present study would be to extend it to include native Chinese speakers, since Chinese is a language that does not utilize spaces. This lack of spacing introduces a large degree of ambiguity, especially when it comes to compounding.

Conclusion

The results of the present study successfully addressed the goals laid out at the start. Eye-tracking was used to examine the processing of novel tripartite compounds, a garden path structure that had not previously been researched. Although the apparent NVN structure of the compounds introduced temporary ambiguity, participants were able to reanalyze an initial misparse of this structure into a single head noun quite quickly in order to come away with the correct interpretation. Although the data do not provide undeniable evidence for local coherence, there is limited evidence that suggests at least a partial local coherence effect. In summary, the unique stimuli used for this experiment introduce a brand new type of garden path sentence that allowed me to explore syntactic reanalysis in a novel way. We observed clear signs of garden-pathing, but little evidence of lingering difficulty or misinterpretation (good enough effects). Participants were temporarily confused by the tripartite compounds when in subject position, but quickly recovered so that they were highly accurate on comprehension questions that probed understanding of the compound. Furthermore, once the main verb was encountered, readers were much less likely to misinterpret the N2 constituent as the main verb. It could be that participants recover so easily from the ambiguous tripartite compounds because compounding is so productive in English already. The tripartite compound used in this experiment, *dog bite victim*, could be expanded to *dog bite victim treatment protocol developer*, and, while entirely novel, this compound can still be understood by the reader. Whether there are limitations on the processing and/or comprehension of similarly lengthy compounds is, however, a question for further research.

References

Acheson, D. J., Wells, J. B., & MacDonald, M. C. (2008). New and updated tests of print exposure and reading abilities in college students. *Behavior Research Methods*, *40*(1), 278-289.

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*, 390-412.

Barr, D., Levy, R., Scheepers, C., & Tily, H. (2013). Random-effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255-278.

Bever, T. G. (1970). The cognitive basis for linguistic structures. *1970*, 279-362.

Bybee, J. (2006). From usage to grammar: The mind's response to repetition. *Language*, 711-733.

Christianson, K., Hollingworth, A., Halliwell, J. F., Ferreira, F. (2001). Thematic Roles Assigned Along the Garden Path Linger. *Cognitive Psychology*, *42,* 368-407.

Christianson, K., Williams, C. C., Zacks, R. T., & Ferreira, F. (2006). Misinterpretations of garden-path sentences by older and younger adults. *Discourse Processes, 42*, 205-238.

Christianson, K., Luke, S. G., & Ferreira, F. (2010). Effects of Plausibility on Structural Priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36,* 538-544.

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic bulletin & review*, *12*(5), 769-786.

Cutter, M. G., Drieghe, D., & Liversedge, S. P. (2014). Preview benefit in English spaced compounds. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(6), 1778.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior*, *19*, 450-466.

Ferreira, F., Bailey, K. G. D., & Ferraro, V. (2002). Good-Enough Representations in Language Comprehension. *Current Directions in Psychological Science, 11,* 11-15.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47,* 164-203.

Ferreira, F., & Patson, N. (2007). The good enough approach to language comprehension. *Language and Linguistics Compass*, *1*, 71-83.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models.* New York: Cambridge University Press.

Lin, T., Anderson, R. C., Ku, Y., Christianson, K., & Packard, J. L. (2011). Chinese children's concept of word. *Writing Systems Research, 3*, 41-57.

Luke, S. G., & Christianson, K. (2015). Predicting inflectional morphology from context. *Language, Cognition and Neuroscience*, (ahead-of-print), 1-14.

Patson, N. D., Darowski, E. S., Moon, N., & Ferreira, F. (2009). Lingering misinterpretations in garden-path sentences: Evidence from a paraphrasing task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 280-285.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysics*, *74*(1), 5-35.

Slattery, T. J., Sturt, P., Christianson, K., Yoshida, M., & Ferreira, F. (2013). Lingering misinterpretations of garden path sentences arise from competing syntactic representations. *Journal of Memory and Language*, *69*(2), 104-120.

Swets, B., Desmet, T., Clifton, C., & Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory & Cognition*, *36*(1), 201-216.

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language, 50*, 355-370.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules* (Vol. 1950). Cambridge, MA: MIT Press.

Traxler, M. J., Johns, C. L., Long, D. L., Zirnstein, M., Tooley, K. M., & Jonathan, E. (2012). Individual differences in eye-movements during reading: Working memory and speed-of-processing effects. *Journal of Eye Movement Research*, *5*(1), 5.

Appendix

Below is a list of all of the target and filler sentences in the experiment.

Targets

Items a-d represent the following four conditions, respectively: subject position-tense error, subject position-agreement error, object position-tense error, object position-agreement error.

1.
a. Last year, the appealing food drive organization helped a friend in New Orleans.
b. Every year, the appealing food drive organization helps a friend in New Orleans.
c. Last year, a friend helped the appealing food drive organization in New Orleans.
d. Every year, a friend helps the appealing food drive organization in New Orleans.

2.
a. Last year, the brave professional divorce attorney sought out a psychiatrist in Tokyo.
b. Every year, the brave professional divorce attorney seeks out a psychiatrist in Tokyo.
c. Last year, the psychiatrist sought out the brave professional divorce attorney in Tokyo.
d. Every year, the psychiatrist seeks out the brave professional divorce attorney in Tokyo.

3.
a. Last night, the clumsy intruder attack dog observed the burglar in the backyard.
b. Every night, the clumsy intruder attack dog observes the burglar in the backyard.
c. Last night, the burglar observed the clumsy intruder attack dog in the backyard.
d. Every night, the burglar observes the clumsy intruder attack dog in the backyard.

4.
a. As usual, the compassionate business suit designer criticized the employee at work.
b. Without fail, the compassionate business suit designer criticizes the employee at work.
c. As usual, the employee criticized the compassionate business suit designer at work.
d. Without fail, the employee criticizes the compassionate business suit designer at work.

5.
a. Yesterday afternoon, the confused shark attack observer conversed with the officer outside.
b. Every afternoon, the confused shark attack observer converses with the officer outside.
c. Yesterday afternoon, the officer conversed with the confused shark attack observer outside.
d. Every afternoon, the officer converses with the confused shark attack observer outside.

6.
a. Last year, the cooperative guest lecture committee invited the doctor to Italy.
b. Every year, the cooperative guest lecture committee invites the doctor to Italy.
c. Last year, the doctor invited the cooperative guest lecture committee to Italy.
d. Every year, the doctor invites the cooperative guest lecture committee to Italy.

7.
a. Last week, the covert car crash investigator told the policeman what he knew.
b. Without fail, the covert car crash investigator tells the policeman what he knows.
c. Last week, the policeman told the covert car crash investigator what he knew.
d. Without fail, the policeman tells the covert car crash investigator what he knows.

8.
a. Last year, the distraught animal research scientist helped the veterinarian with his dog.
b. Every year, the distraught animal research scientist helps the veterinarian with his dog.
c. Last year, the veterinarian helped the distraught animal research scientist with his dog.
d. Every year, the veterinarian helps the distraught animal research scientist with his dog.

9.
a. Yesterday, the efficient bee sting analyst conversed with the patient from Japan.
b. Each day, the efficient bee sting analyst converses with the patient from Japan.
c. Yesterday, the patient spoke with the efficient bee sting analyst from Japan.
d. Each day, the patient speaks with the efficient bee sting analyst from Japan.

10.
a. Yesterday, the enthusiastic family trip planner arrived with a client at the airport.
b. Today, the enthusiastic family trip planner arrives with a client at the airport.
c. Yesterday, a client arrived with the enthusiastic family trip planner at the airport.
d. Today, a client arrives with the enthusiastic family trip planner at the airport.

11.
a. Last year, the established customer call center acquired another facility in India.
b. Every year, the established customer call center acquires another facility in India.
c. Last year, the facility acquired another established customer call center in India.
d. Every year, the facility acquires another established customer call center in India.

12.
a. Last night, the excited fox hunt leader informed the crew from Canada.
b. Each evening, the excited fox hunt leader informs the crew from Canada.
c. Last night, the crew informed the excited fox hunt leader from Canada.
d. Each evening, the crew informs the excited fox hunt leader from Canada.

13.
a. Last concert, the fearless punk rock musician sang with the audience during the chorus.
b. Every concert, the fearless punk rock musician sings with the audience during the chorus.
c. Last concert, the audience sang with the fearless punk rock musician during the chorus.
d. Every concert, the audience sings with the fearless punk rock musician during the chorus.

14.
a. Last week, the frightened child abuse victim met with the therapist in the office.
b. Every week, the frightened child abuse victim meets with the therapist in the office.
c. Last week, the therapist met with the frightened child abuse victim in the office.

d. Every week, the therapist meets with the frightened child abuse victim in the office.

15.
a. Last week, the industrious mouse trap builder disagreed with the manager about sales.
b. Every week, the industrious mouse trap builder disagrees with the manager about sales.
c. Last week, the manager disagreed with the industrious mouse trap builder about sales.
d. Every week, the manager disagrees with the industrious mouse trap builder about sales.

16.
a. Last week, the inexperienced horse race gambler paid his friend fifty dollars.
b. Every week, the inexperienced horse race gambler pays his friend fifty dollars.
c. Last week, the friend paid the inexperienced horse race gambler fifty dollars.
d. Every week, the friend pays the inexperienced horse race gambler fifty dollars.

17.
a. Last Christmas, the infamous winter coat designer didn't give her friend a gift.
b. Every Christmas, the infamous winter coat designer doesn't give her friend a gift.
c. Last Christmas, the friend didn't give the infamous winter coat designer a gift.
d. Every Christmas, the friend doesn't give the infamous winter coat designer a gift.

18.
a. Last month, the intellectual Asian badger enthusiast admired the zoologist at dinner.
b. Every month, the intellectual Asian badger enthusiast admires the zoologist at dinner.
c. Last month, the zoologist admired the intellectual Asian badger enthusiast at dinner.
d. Every month, the zoologist admires the intellectual Asian badger enthusiast at dinner.

19.
a. That night, the intelligent animal track expert followed the tiger to the meadow.
b. Every night, the intelligent animal track expert follows the tiger to the meadow.
c. That night, the tiger followed the intelligent animal track expert to the meadow.
d. Every night, the tiger follows the intelligent animal track expert to the meadow.

20.
a. Last year, the intrepid sea cave explorer brought the archaeologist for the dive.
b. Every year, the intrepid sea cave explorer brings the archaeologist for the dive.
c. Last year, the archaeologist brought the intrepid sea cave explorer for the dive.
d. Every year, the archaeologist brings the intrepid sea cave explorer for the dive.

21.
a. Yesterday, the lively magic trick performer baffled the audience at the auditorium.
b. Every day, the lively magic trick performer baffles the audience at the auditorium.
c. Yesterday, the audience baffled the lively magic trick performer at the auditorium.
d. Every day, the audience baffles the lively magic trick performer at the auditorium.

22.
a. Yesterday, the nasty fish farm manager visited his mother at home.

b. Every day, the nasty fish farm manager visits his mother at home.
c. Yesterday, the mother visited the nasty fish farm manager at home.
d. Every day, the mother visits the nasty fish farm manager at home.

23.
a. Last month, the new trailer park manager conferred with the expert during lunch.
b. Every month, the new trailer park manager conferred with the expert during lunch.
c. Last month, the expert conferred with the new trailer park manager during lunch.
d. Every month, the expert confers with the new trailer park manager during lunch.

24.
a. Last week, the obese dog bite victim met the neighbor for coffee.
b. Every weekend, the obese dog bite victim meets the neighbor for coffee.
c. Last week, the neighbor met the obese dog bite victim for coffee.
d. Every weekend, the neighbor meets the obese dog bite victim for coffee.

25.
a. Last period, the prolific writing test creator assessed the class during the break.
b. Every period, the prolific writing test creator assesses the class during the break.
c. Last period, the class assessed the prolific writing test creator during the break.
d. Every period, the class assesses the prolific writing test creator during the break.

26.
a. Yesterday, the punctual baby shower host discussed strategies with a friend in the cafe.
b. Every day, the punctual baby shower host discusses strategies with a friend in the cafe.
c. Yesterday, a friend discussed strategies with the punctual baby shower host in the cafe.
d. Every day, a friend discusses strategies with the punctual baby shower host in the cafe.

27.
a. Last month, the responsible public aid officer assisted the community in California.
b. Every month, the responsible public aid officer assists the community in California.
c. Last month, the community assisted the responsible public aid officer from California.
d. Every month, the community assists the responsible public aid officer from California.

28.
a. Yesterday, the shrewd criminal research analyst spoke with the psychologist via Skype.
b. Every day, the shrewd criminal research analyst speaks with the psychologist via Skype.
c. Yesterday, the psychologist spoke with the shrewd criminal research analyst via Skype.
d. Every day, the psychologist speaks with the shrewd criminal research analyst via Skype.

29.
a. Last night, the skeptical family film critic ate with his daughter at home.
b. Every night, the skeptical family film critic eats with his daughter at home.
c. Last night, the daughter ate with the skeptical family film critic at home.
d. Every night, the daughter eats with the skeptical family film critic at home.

30.
a. Last night, the skilled prostitute murder investigator met a friend at the bar.
b. Every night, the skilled prostitute murder investigator meets a friend at the bar.
c. Last night, a friend met the skilled prostitute murder investigator at the bar.
d. Every night, a friend meets the skilled prostitute murder investigator at the bar.

31.
a. Last week, the smart spider bite expert spoke with the class from Urbana.
b. Every month, the smart spider bite expert speaks with the class from Urbana.
c. Last week, the class spoke with the smart spider bite expert from Urbana.
d. Every month, the class speaks with the smart spider bite expert from Urbana.

32.
a. A month ago, the smooth oil spill team informed the city about the current methods.
b. Every month, the smooth oil spill team informs the city about the current methods.
c. A month ago, the city informed the smooth oil spill team about the current methods.
d. Every month, the city informs the smooth oil spill team about the current methods.

33.
a. Last weekend, the spotless hospital waste supervisor asked a nurse out on a date.
b. Every weekend, the spotless hospital waste supervisor asks a nurse out on a date.
c. Last weekend, a nurse asked the spotless hospital waste supervisor out on a date.
d. Every weekend, a nurse asks the spotless hospital waste supervisor out on a date.

34.
a. Last time, the stubborn gang arrest suspect exploded on the lawyer during the meeting.
b. Every time, the stubborn gang arrest suspect explodes on the lawyer during the meeting.
c. Last time, the lawyer exploded on the stubborn gang arrest suspect during the meeting.
d. Every time, the lawyer explodes on the stubborn gang arrest suspect during the meeting.

35.
a. Last year, the stylish elevator lift engineer instructed the apprentice on the machinery.
b. Every year, the stylish elevator lift engineer instructs the apprentice on the machinery.
c. Last year, the apprentice instructed the stylish elevator lift engineer on the machinery.
d. Every year, the apprentice instructs the stylish elevator lift engineer on the machinery.

36.
a. Last evening, the surprised skunk spray victim scared the woman in the yard.
b. As expected, the surprised skunk spray victim scares the woman in the yard.
c. Last evening, the woman scared the surprised skunk spray victim in the yard.
d. As expected, the woman scares the surprised skunk spray victim in the yard.

37.
a. The other day, the temperamental cattle feed inspector argued with the rancher at the bank.
b. Every day, the temperamental cattle feed inspector argues with the rancher at the bank.
c. The other day, the rancher argued with the temperamental cattle feed inspector at the bank.

d. Every day, the rancher argues with the temperamental cattle feed inspector at the bank.

38.
a. Last week, the tender ear implant surgeon talked to the nurse in the hospital.
b. Every week, the tender ear implant surgeon talks to the nurse in the hospital.
c. Last week, the nurse talked to the tender ear implant surgeon in the hospital.
d. Every week, the nurse talks to the tender ear implant surgeon in the hospital.

39.
a. Last time, the tiring student vote organization disagreed with the colleague in Alaska.
b. Every time, the tiring student vote organization disagrees with the colleague in Alaska.
c. Last time, the colleague disagreed with the tiring student vote organization in Alaska.
d. Every time, the colleague disagrees with the tiring student vote organization in Alaska.

40.
a. Last holiday, the trusted computer display manufacturer beat the new company in sales.
b. Every holiday, the trusted computer display manufacturer beats the new company in sales.
c. Last holiday, the new company beat the trusted computer display manufacturer in sales.
d. Every holiday, the new company beats the trusted computer display manufacturer in sales.

Fillers

1. The adept mapmaker from Greece excites the geographer with his skills.
2. The customer from Kentucky called the affable salesperson regularly.
3. The assistant embarrasses the stockbroker, who is apathetic, in the meeting room.
4. The wife that debates the case with the apprehensive judge is upset.
5. The ardent coach thanked the team on the bus after the game.
6. The beautiful woman laughs with her sister at the restaurant during lunch.
7. Last Sunday, the granddaughter baked a cake with the beloved grandmother in the kitchen.
8. The bold art director that the model took out for dinner every evening was snobby.
9. Last Friday, the bossy high school cheerleader went with the geek to lunch.
10. The bubbly hostess greets the patron at the front desk before brunch.
11. The partner that the cautious roofer scrutinizes is quiet.
12. This morning, the guest recognized the clever ticket checker from earlier.
13. Every day, the critical toolmaker monitors the welder as she works.
14. The cute kitten that wrestled the puppy on the couch during the party was grey.
15. The dapper pilot announces his arrival to the chipper flight attendant each morning.
16. Last winter, the delightful officer traveled with his good-looking wife to Europe.
17. As before, the desperate pest controller found the marketing company online.
18. The stagehand, who is devoted, supports the troupe during the production.
19. The man remembered the diligent record keeper, who was at the conference.
20. The hyperactive fan that the dreamy Broadway actor hugged was on stage.
21. The bartender that complained to the dull chemist about his life was depressed.
22. The stonemason, who was earnest, waited for the apprentice in the shop.
23. The good friend that the easygoing beauty consultant invited never showed up.
24. The emotional son that embraced his father at the airport was crying.
25. Yesterday, the empathetic funeral director talked with the family about the process.

26. The florist that thanks the employee for her kindness is grateful.
27. The patient that discussed the problem with the exuberant dentist was concerned.
28. The seamstress, who was fickle, separated from the sculptor from New York.
29. The flamboyant librarian interviewed the man in the lounge.
30. Every spring, the conceited director entertains the flawless actress on the beach.
31. The fretful deep sea diver joined the captain at the headquarters after each dive.
32. Every time, the green tree frog kisses the frigid princess on the lips.
33. Every week, the board member orders a meeting with the frugal executive in the office.
34. The generous diplomat invited the couple over to his home for dinner and drinks.
35. The dolphin that the gentle marine biologist passes by is in the aquarium.
36. The gregarious bilingual secretary that the nurse chatted with was eating lunch.
37. Every five minutes, the hairy captain calls the crewmember from the deck.
38. The son makes a peanut butter sandwich and the hard-working editor goes to work.
39. Every year, the humble metal engraver travels with his partner to Texas.
40. The imaginative animator that the boss waits to hear from is lazy.
41. The crew that the ingenious architect congratulated was arrogant.
42. The microbiologist that listened to the eager student's explanation was introspective.
43. The mother that hugs her joyful child at the playground is smiling.
44. Every month, the brother phones the lazy company president for a favor.
45. The widow that writes the lonely carpenter letters is anxious.
46. Last night, the loving caretaker read and the child went to sleep.
47. The prompt meter reader skips work and takes his sweet niece out for brunch.
48. The magnificent racecar driver admires the skydiver for her courage.
49. Yesterday, the jolly mentor amused the merry journeyman with a joke.
50. Last fall, the meticulous physicist collaborated with the astronomer in Maine.
51. The moody auto mechanic that the daughter drives home is exhausted.
52. The worker that the nasty auditor spied on was at the corporation.
53. The neat bookkeeper that the coworker borrows a pen from is handsome.
54. Yesterday morning, the bird chased the nimble cat around the yard.
55. Every sunrise, the obnoxious crow squawks at the sparrow in the tree.
56. The CEO ignores the obstinate broadcaster, who rants during the meeting.
57. Every spring, the woodsman wants the optimistic lumberjack to be at the competition.
58. The cheery advertiser welcomed the sign painter, who was outgoing, into her store.
59. Yesterday afternoon, the cowgirl drank a beer with the passionate cowboy at the bar.
60. The farmer that meets the patient horseshoe maker at the barn is old.
61. The patronizing hairdresser relayed to the customer what color hair dye to use.
62. The student that convened with the pedantic instructor was impatient.
63. The woman that the peppy aerobics instructor encouraged was at the party.
64. Every afternoon, the pioneering game programmer fits the intern with 3D goggles.
65. The nice copywriter that the pretty editor sat with was in the break room.
66. Every Wednesday, the famous sushi chef drives the polite waiter to work.
67. The pompous camera operator that the caterer asks out is confident.
68. Last weekend, the popular boy snuck the girl into the theater.
69. Last game, the powerful college scout observed the basketball player in the gym.
70. The well-known woman that upset the hotel receptionist was rude.
71. The quiet crane operator shows the novice employee the control panel.

72. The amateur bird watcher that the rude executive scoffs at is astonished.
73. The ruthless environmental planner assisted the company in forming a new agenda.
74. The reporter that questioned the sarcastic accountant on the news was intimidating.
75. The trembling expert witness entered the room and the secretive court clerk sat.
76. The selfish doctor that the family hates is in Europe.
77. Last night, the lovely bride danced with the sentimental groom in their room.
78. The technician that the serious pharmacist watches looks suspicious.
79. Every month, the shy gallery attendant helps the curator arrange the artwork.
80. Every time, the customer gives the sleepy clerk the money at the counter.
81. Last month, the sociable goldsmith entertained the curious visitor at the expo.
82. The traffic warden, who was stern, scolded a pedestrian on the corner.
83. The girl leaves work and the cheap disc jockey from Minnesota goes home.
84. The student that explained the problem to the strange statistician was angry.
85. Yesterday, the chubby toddler hid from the surprised mouse in the kitchen.
86. Last night, the sweaty prisoner stole from the prison guard in the hall.
87. The tenacious television journalist interrogates the reporter about the scandal.
88. Every game, the jealous teammate glares at the tense quarterback on the field.
89. Last night, the furious passenger navigated the calm taxi driver to the destination.
90. The retail merchandiser, who is thrifty, bargains with the distributor on a price.
91. The customer that waved to the trustworthy fishmonger at the market was friendly.
92. Last Sunday evening, the fat lady showed the unenthusiastic theater usher her ticket.
93. The unpredictable script supervisor works with the tense team to refine the project.
94. Last summer, the upbeat swim instructor taught her friend how to surf.
95. Every weekend, the friend trains while the vain lifeguard lounges on the beach.
96. Yesterday, the volatile freight exporter yelled at the inspector about the shipment.
97. The weepy artist that begged his sympathetic roommate for some cash was hungry.
98. The middle school principal that the weird teacher met with was cordial.
99. Last class, the trainee faced the witty safety officer during the exercise.
100. The zany author that refuses to listen to the editor is frustrating.

N2 Constituents and Past Tense Forms

Verb	Past Tense
*bite**	*bit*
drive	*drove*
sting	*stung*
feed	*fed*
attack*	attacked
divorce	divorced
suit	suited
lecture	lectured
crash	crashed
research*	researched
trip	tripped
call	called

hunt	hunted
rock	rocked
abuse	abused
trap	trapped
race	raced
coat	coated
badger	badger
track	tracked
cave	caved
trick	tricked
farm	farmed
park	parked
test	tested
shower	showered
aid	aided
film	filmed
murder	murdered
spill	spilled
waste	wasted
arrest	arrested
lift	lifted
spray	sprayed
implant	implanted
vote	voted
display	displayed

* marks a verb that was used twice
Italicized words mark irregular verbs