About time we make sense: Distinct neural processes engaged during temporal Sequencing and coherence building in discourse

An honors thesis for the Department of Psychology

Wonja M. Fairbrother

Tufts University, 2011

Table of Contents

Abstract 3
Introduction 4
Methods 19
Results 29
Discussion 42
References 54
Appendix A 61
Appendix B 62

Abstract
In the real world, causes always come before effects, while in communication, events can be described in either this canonical temporal order using causal connectors such as “and so”, or in non-canonical order using connectors like “because”. Using event-related potentials (ERPs), we determined whether the canonical sequencing of events influences the establishment of causal coherence, or vice versa, during online discourse processing. Two-clause cause/effect sentences were created, in which we fully crossed temporal sequencing of events with discourse coherence, yielding four experimental conditions (example set: Fred was hungry [and so/*because] he ate…; Fred ate [*and so/because] he was hungry…”). 32 participants read these sentences, presented word-by-word (450ms, ISI: 100ms), and made acceptability judgments at the end of each sentence. Participants’ working memory spans were assessed with an Automated Reading Span task. At anterior electrode sites, ERPs to critical words (“ate/hungry”) in clauses appearing in non-canonical sequence evoked a larger negativity between 400-500ms than in clauses appearing in canonical temporal sequence. At centro-parietal sites, ERPs to critical words in incoherent clauses evoked a larger negativity between 300-500ms (an N400 effect) and a larger positivity between 700-900ms (a P600 effect) than in coherent clauses. A main effect of working memory span between 200-300ms after the critical word and sentence-final word did not interact with either coherence or canonicity. Together, these results suggest that during discourse comprehension, establishing the temporal sequencing of events and establishing their causal coherence are driven by distinct neural mechanisms and are not influenced by individual differences in working memory span.

Introduction

Understanding causal relationships between events, actions, and states is crucial for the interpretation of text and discourse. The flexible nature of language allows humans to communicate causes and effects in two different ways: in canonical or chronological order (cause -> effect) or in non-canonical or non-chronological order (effect <- cause). Despite the order in which the events are communicated, we are effortlessly able to understand the order in which the events actually happened. The causally related events are “aligned” in long-term semantic memory in the correct chronological order (Anderson, Garrod, & Sanford, 1983; Carreiras, Carreido, Alonso, & Fernandez, 1997; Mandler, 1985; Radvansky, Zwaan, Federico, & Franklin, 1998; van der Meer, Beyer, Heinze, & Badel, 2002; Zwaan, 1996). One main question arises about this process: Are the mechanisms that establish causal coherence and temporal ordering across events distinct or do they interact? The present study investigates the time-course of establishing causal coherence during comprehension of sentences that introduce events in and out of canonical order, and whether individual differences in working memory span affect these processes.

Humans are able to understand language rapidly and fluidly by integrating large amounts of information conveyed through connected clauses and sentences. There is substantial behavioral evidence that comprehenders utilize information about causal relations during discourse processing to establish coherence. Causally related sentences in discourse facilitate processing; they are read faster than causally unrelated sentences (Keenan, Baillet, & Brown, 1984; Haviland & Clark, 1974; Zwaan, 1996), and probe words that follow them are recognized (McKoon & Ratcliff, 1989), named (Klin, 1995; Potts, Keenan, & Golding 1988), and lexically categorized (Potts et al., 1988) faster. Moreover, both comprehending and producing causally coherent ideas are signs of healthy language development (Katz & Brent, 1968; Donaldson, Reid, & Murray, 2005; French 1988). It has also been suggested that when events are communicated in canonical order, their processing is facilitated compared to non-canonical order (Mandler, 1985; Katz & Brent, 1968).

Behavioral studies of causal comprehension

Behavioral studies on language have been helpful in exploring what affects causal comprehension. Multiple sources of information such as context, memory, and semantic knowledge influence word-by-word processing in different ways (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; MacDonald, Pearlmutter, & Seidenberg, 1994; Altmann & Steedman, 1988). The studies discussed here focus on two specific sources of information that influence causal processing: the presence of causal connectives and the order of mention of events.

The presence of temporal or causal connectives in between clauses and sentences speeds their processing (De Vega, 2005; Gernsbacher, 1997) and decreases variability of interpretation of the narrative (Segal, Duchan, & Scott, 1991). Behavioral evidence supports the role of connectives as local and global cohesion markers as well as contributors to mental model formation and deictic shifts (Segal, Duchan, & Scott, 1991). Conversely, connectives that are inappropriate to discourse context increase reading time (De Vega, 2005). Different causal connectors bring different meanings to inter-clausal relationships, and the different connector used between clauses can significantly alter the meaning of a sentence.

Readers also use the temporal order in which events occur and are introduced in discourse to build causal coherence. Events communicated in past tense without a connector tend to be interpreted sequentially (De Vega, 2005), and we know, due to semantic knowledge, that causes happen before their effects. To use temporal order to establish coherence, comprehenders must access semantic information about what typically causes and results from the events in question. For example, because we know that hunger causes eating and eating relieves hunger, we understand that the sequence *Fred was hungry -> he ate a meal* makes sense, but the sequence *Fred ate a meal -> he was hungry* does not. Mandler (1985) found that matching and mismatching orders of mention of causally connected events did not affect comprehension speed. He therefore suggested that a mental model of temporal order for causally related events already exists in semantic knowledge and does not need to be built when comprehending them. In other words, humans may already have an existing framework of the order in which causally related events occur based on our knowledge of contingency relations.

Behavioral evidence illustrates that causal connectives can either speed up or slow down comprehension depending on its appropriateness, and that introducing events in and out of canonical order has no effects on the total time it takes to comprehend them. However, behavioral studies are quite limited in their ability to demonstrate the time-course of such a rapid process as discourse comprehension (Myers & Duffy, 1990; Myers et al., 1987; Keenan et al., 1984; Haviland & Clark, 1974, McKoon & Ratcliff, 1986, 1989, Singer & Halldorson, 1996; Singer, 1993; Singer et al., 1992), let alone the interpretation of causal propositions. Instead, a temporally accurate neuroimaging method such as event-related potentials (ERPs) can be useful in elucidating this issue.

Event Related Potentials

Event-Related Potentials (ERPs) measure neural activity on-line, and can yield information about the processes that build meaning word-by-word as they unfold. ERPs reflect electrophysiological activity measured at the surface of the scalp that can be time-locked to specific words within a sentence and averaged to reduce noise. The difference in amplitude of the same ERP between two experimental conditions indicates which condition was associated with more brain activity or more neural processing. Importantly, this processing difference can be measured and detected even before a behavioral response occurs. Three ERP components are of particular interest in language research and specifically in this study – the N400, P600 or late positivity, and the anterior negativity.

The N400 component is a negative-going waveform that usually peaks around 400 milliseconds after the onset of a stimulus. It is thought to reflect the ease of processing of an incoming word given its preceding context and semantic knowledge (Kutas, Van Petten, & Kluender, 2006). The N400 effect is known as the attenuation of its amplitude to an incoming word that fits well semantically in its context compared to a word that does not fit well into its context. This effect has been shown using semantic priming paradigms

(Bentin, McCarthy, & Wood, 1985; Rugg, 1985), sentence stems (Kutas & Hillyard, 1980, 1984), and the entire discourse (van Berkum et al., 1999) as context for an incoming word. The amplitude of the N400 is modulated by both the entire situation model built up by a word's preceding context (Ditman, Holcomb, & Kuperberg, 2007, 2008; Nieuland & Kuperberg, 2008; Kuperberg, Paczynski, & Ditman, 2010; Chwilla, Kolk, & Vissers, 2007; Otten & van Berkum, 2007; van Berkum et al., 1999), as well as the lexico-semantic relationships between a word and the content words of its preceding context (Ditman et al., 2007; Otten & van Berkum, 2007; Ledoux, Camblin, Swaab, & Gordon, 2006; Van Petten, 1993; Nieuwland & van Berkum, 2005). Because of this dual sensitivity, the N400 is a useful index of how these two aspects of language interact to affect semantic processing of words.

The P600s or late positive components are a group of positive-going waveforms that peak any time between 500 and 900 milliseconds after stimulus onset, and are thought to reflect sentential anomalies that require additional processing. For example, syntactic errors (Hagoort, 1993; Osterhout & Holcomb, 1992), sentence complexity (Kuperberg, 2007), and semantic plausibility (van de Meerendonk, Kolk, Vissers, & Chwilla, 2010; Kuperberg, 2007), among other complex aspects of language like metaphors (De Grauwe, Swain, Holcomb, Ditman, & Kuperberg, 2010; Coulson & Van Petten, 2002), jokes (Coulson & Kutas, 2001), and emotional language (Holt, Lynn, & Kuperberg, 2009; van Berkum, Holleman, Nieuwland, Otten, & Murrel, 2009) have all been demonstrated to show P600 effects. Because these different aspects of language evoke positivities that differ in exact time-course and scalp distribution, the precise functional significance of late positivities is debated (Kolk & Chwilla, 2007; Kuperberg, 2007; Osterhout & Hagoort, 1999; Coulson, King, & Kutas, 1998). Late positivities likely represent multiple neurocognitive processes, but it is generally agreed upon that these processes involve some sort of active continued analysis (past the N400) of an incoming word with respect to its context (Kuperberg, 2007), stored semantic knowledge in long-term memory (Van Petten, Kutas, Kluender, Mitchiner, & Melsaac, 1991), and the discourse situation model (Burkhardt, 2006, 2007).

One theory of note that explains the functionality and interplay of the N400 and P600 is the processing stream theory (Kuperberg, 2007). This account alleges two competing neural processing streams: a semantic-memory based mechanism (violation costs reflected by the N400) and a combinatorial mechanism that assigns structure based on morphosyntactic, linguistic, and thematic constraints. A conflict between these two streams is reflected by the P600 or late positivity. Top-down executive processes and working memory likely control the interaction between these two mechanisms, and therefore the balance of the processing streams may be affected by individual differences in working memory capacity.

The anterior negativity, also referred to as the left anterior negativity (LAN) or early left anterior negativity (ELAN), is a negative-going wave peaking between 200 and 500ms after stimulus onset and distributed mainly in the anterior regions of the scalp. It occurs most often in response to linguistic violations of word-category or phrase structure rules (Hagoort, 2003; Friederici & Weissenborn, 2007; Friederici, 2002), and is part of a well-documented biphasic response (paired with the P600) to morphosyntactic agreement violations like phonotactic or word gender conflicts (Molinaro, Vespignani, & Job, 2008). The anterior negativity has been associated with storing fillers in working memory and retrieving them upon gap detection in filler-gap dependencies (Kluender & Kutas, 1993). The effect also appears to be sensitive to the canonicity of word order (deviations from the canonical sequence of subject, indirect object, and direct object) and parsing load in German (Rosler, Pechman, Streb, Roder, & Hennighausen, 1998), and also to non-linguistic cognitive sequencing of symbols (Hoen & Dominey, 2000). Additionally, there has been evidence that the amplitude of the anterior negativity effect modulated by complexity and load of morphosyntactic violations correlates with individual differences in verbal working memory span (Vos, Gunter, Kolk, & Mulder, 2001).

ERP studies of causal comprehension

There have been several relevant ERP studies that investigated the buildup of causal coherence between sentences, which are discussed below. The majority of these studies examined the degree of causal relatedness between sentences and measured ERPs to inference-confirming words or clauses. Here, four relevant studies are described in particular that offer evidence of the roles of inter-word lexico-semantic relationships and the whole situation model. Differences and time-courses of the resulting waveforms in these studies can inform us about when semantic information about the context and situation model is available in comprehending causal discourse, and the level and type of violation necessary for additional neural analysis.

In a recent study by Yang, Perfetti, and Schmalhofer (2007), participants read two causally related sentences, with the first content word of the second sentence being either the same as (in the referentially explicit condition; [a]) or semantically related to (in the referentially paraphrased condition; [b]) the final word of the first sentence. There was also a condition with the final word of the first sentence omitted (in the bridging inference condition; [c]). (Example set: *After being dropped from the plane, the bomb hit the ground* [a]*and exploded./*[b]*and blew up./*[c]*{n/a}. The* <u>*explosion*</u> *was quickly reported to the commander.*) ERPs measured at the critical word (*explosion* in the above example) in the second sentence showed a larger N400 effect in the bridging inference condition compared to the other two conditions. The authors attributed these results to a facilitation of the inference in the referentially explicit and paraphrased conditions due to lexical priming. This conclusion lends support to the role of lexico-semantic connections in building causal relationships but does not provide any evidence for the role of the situation model in this process.

Another ERP study in 2007, by Burkhardt, used a slightly different manipulation: participants read two-sentence scenarios in which the first sentence was varied on the explicitness of the described event. (Example set: *Yesterday a student was* [a]*shot/*[b]*killed/*[c]*found dead downtown. The press reported that the* <u>*pistol*</u> *was probably from army stocks.*) ERPs were measured to the critical word (*pistol* in the above example) in the second sentence that made clear the details of the described event. When the critical word was encountered, the amount of information that the participants needed to integrate into the discourse model was different dependent on the condition. Thus the authors proposed that the effects on the critical word had more to do with the discourse model than the semantic relations between words. Rather than an N400 effect, there was a P600 or late positivity effect to the critical word with incrementally increasing amplitudes in the less explicit conditions. These findings suggest that discourse model processing is qualitatively different from lexico-semantic processing, and that adding new information to the situation model requires later neurocognitive processes distinct from those associated with the N400.

Most recently, Kuperberg, Paczynski, and Ditman (2011) used ERPs to examine causal coherence building across three-sentence scenarios in which the final sentence was highly causally related ([a]), intermediately related ([b]), or causally unrelated ([c]) to the first two sentences. (Example set: [a]*Jill had very fair skin. She forgot to put sunscreen on. She had* <u>*sunburn*</u> *on Monday./*[b]*Jill had very fair skin. She usually remembered to wear sunscreen. She had* <u>*sunburn*</u> *on Monday./*[c]*Jill's skin always tanned well. She always put on sunscreen. She had* <u>*sunburn*</u> *on Monday.*) Importantly, the sentences were matched in all conditions on lexico-semantic relations between content words using a Latent Semantic Analysis (LSA; Landauer & Dumais, 1997). This allowed them to examine the time-course of causal coherence building strictly at the situation model level and not at the level of lexical relatedness. There was a larger N400 to critical words in causally unrelated condition compared to the other two conditions, and at electrode sites excluding the midline, the N400 to intermediately related critical words fell in between those to highly related and unrelated critical words. Interestingly, there was no difference in the P600/late positivity time window across conditions. These findings indicate that when lexico-semantic relationships are matched, situation model coherence can impact word-by-word discourse comprehension in the earlier N400 time-window. These results finding an N400 to difficulty in situation model integration differ from Burkhardt's (2007) finding of a P600 to integration difficulty. However, there are several differences between these two studies that may account for this, including the number of sentences in the stimuli, the placement of the critical word, and the variance in semantic relatedness of the stimuli. In any case, a clear story between these ERP components, the coherence of the discourse model in causal scenarios, and lexico-semantic relationships across words is yet to be determined.

Another feature of interest that may influence causal coherence is the order in which causally related events are mentioned. The behavioral study by Mandler (1985) discussed above showed that presenting events in non-canonical order has no behaviorally measurable effect on comprehension or reading speed. However, ERPs may be able to detect any behaviorally non-measurable differences that might exist between orders of mention. Munte, Schiltz, and Kutas (1998) had participants read sentences that described two non-causal events in either chronological (canonical; *After the scientist submitted the paper, the journal changed its policy.*) or non-chronological (non-canonical; *Before the scientist submitted the paper, the journal changed its policy.*) order. ERPs measured across the entire sentence showed a divergence in the left hemisphere within 300ms after the beginning of the sentence. Waveforms to *Before* sentences were more negative than *After* sentences, and this divergence continued until the end of the sentences. Because the two conditions differed only in their initial word, the syntax and lexico-semantic relationships were held constant and any effects between them would involve computation of the situation model and semantic knowledge stored in memory. This study indicated that presenting events out of chronological order (*Before X, Y*) requires additional processing on the discourse model level than presenting events in chronological order (*After X, Y*). However, since the ERPs were measured over the entire sentence and there were no mid-sentence critical words at which canonicity was established, it is difficult to speak on the time-course of this process in a precise manner from the results of this study.

Effects of working memory on language comprehension

We know that humans comprehend language very rapidly, but there exists substantial debate as to whether individuals differ in the rapidness of comprehension or the order in which different elements of language are processed. There are several views on how individual working memory span affects language comprehension. First, Just and Carpenter (1992) claim that working memory relates to the ability to process multiple sources of information in real-time comprehension. Second, Caplan and Waters (1999) argue that individuals vary on general, non-language-specific working memory, which only appears when readers must perform an additional task like recall or plausibility judgment. Lastly, MacDonald and Christiansen (2002) argue that individual differences are due to different amounts of exposure to language, and that comprehension can improve with practice. In this study, we are interested in how variation of working memory affects the process of determining the coherence and canonicity of a causal sentence.

Nakano, Saron, and Swaab (2010) found an interesting pattern of ERP results modulated by working memory span. Participants were presented with single-sentence stimuli belonging to one of three levels of plausibility with respect to thematic relations and animacy (*The* [a]*dog*/[b]*poet*/[c]*box is biting the mailman*.). ERPs were measured to the first noun, the verb, and the second noun and participants were split into high- and low-span groups based on reading span and listening span tasks (Daneman & Carpenter, 1980). In the *poet* condition, the second noun (*mailman*) elicited an N400 effect for both memory span groups due to conflict with store semantic knowledge. In the *box* condition, however, the initial noun (*box*) elicited an anterior negativity only in the high-span group, and the verb (*biting*) elicited an N400 in the low-span and a P600 in the high-span group. The authors suggest that high-span readers are immediately sensitive to animacy information (as seen by the anterior negativity to the inanimate subject noun) and that thematic roles are integrated into the discourse model earlier, resulting in a delayed, P600 reanalysis effect at the verb. In contrast, the low-span group did not use the animacy information at the inanimate subject noun, and when they encountered the verb, the newly recognized conflict in semantic knowledge caused an N400 effect. This can be explained in terms of Kuperberg's (2007) processing stream theory, discussed above – the differences in the allocation of attention early on in the sentence may have affected how the processing streams interacted later on, resulting in different ERP effects at the verb.

The Munte, Schiltz, and Kutas (1998) study discussed above also measured participants' verbal working memory span with a version of the Automated Reading Span

Task (Daneman & Carpenter, 1980). They found that the difference between the *Before* and *After* sentence waveforms was more pronounced in participants with higher working memory spans. This suggests that increased working memory capacity may enhance sensitivity to the canonicity of events in discourse. A number of other ERP studies have found effects of individual differences in working memory span on language comprehension as well (Bornkessel, Fiebach, & Friederici, 2004; Vos & Friederici, 2003).

In sum, these ERP studies have shown that causal relatedness can affect comprehension at both the lexical level and the situation model level, that ordering events out of canonical order requires additional computation, and that individual working memory span can influence how and when various aspects of language are processed. However, some gaps in the causal coherence research remain. There have been few studies examining causality within sentences (versus across sentences), and even fewer using ERPs to investigate them. To our knowledge, there have also been no ERP studies looking at whether the order in which causal events are introduced affects discourse comprehension, the effects of causal connectors, or individual working memory span and the buildup of causal coherence specifically.

<u>The present study and hypotheses</u>

The present study aims to address causal coherence, canonicity of causal events, presence of causal connectors, and individual working memory span in terms of comprehending causal, single sentences. We investigated the effect of coherence and canonicity on the time-course of single-sentence causal comprehension using ERPs, and whether this process is affected by individual differences in working memory span. The causal connectors *because* and *and so* were used in our stimuli to explicitly violate causal coherence in incoherent conditions and to reduce variability in interpretation. We measured ERPs to a critical word in the latter half of sentences that clearly determined the canonicity and coherence of the sentence when it was reached and not before that point. This provided a precise measuring point at which effects of coherence and canonicity take place. To ensure any effects we see are on the level of the situation model, we matched all stimuli for lexico-semantic associations based on LSA scores.

There were three questions this study hoped to address. First, can we replicate the effect of causal coherence that was found by Kuperberg et al. (2011) despite close lexico-semantic associations between words? This would be reflected by a larger N400 effect to incoherent critical words than to coherent critical words. If the conflict between causal incoherence and matched lexico-semantic relations requires additional processing, this will be reflected by a larger P600 effect to incoherent than to coherent critical words. Although there was no such P600 effect to the causally incoherent (versus coherent) critical words in the study by Kuperberg et al. (2011), no explicit causal connectors were included and so it may have been possible to make inferences at the critical words even in the causally unrelated scenarios. In the present study, the inclusion of the causal connectors explicitly rendered these scenarios highly implausible. Also of note, the critical words in the Kuperberg et al. (2011) study were either in mid-sentence or sentence-final position, while all of the critical words in this study were mid-sentence. This allowed us to distinguish ERP effects on the coherence- and canonicity-determining critical word and ERP effects on the sentence-final word that follows.

Second, can we find an effect of canonical ordering at the critical word and/or the sentence-final word? If so, what type of effect appears, and does it interact with the causal coherence of the sentence? If comprehending events presented in non-canonical order taxes more working memory resources than comprehending those presented in canonical order, then we expect a larger sustained anterior negativity to non-canonical critical words than to canonical critical words (Munte, Schiltz, & Kutas, 1998). If the mechanism that determines event order interacts with the one that determines causal coherence, then we expect the canonicity of the events to impact the ERP effects we see to coherence. If these mechanisms are distinct, there should be no difference in coherence effects to critical words between canonical and non-canonical scenarios.

Finally, is there an effect of working memory span at the mid-sentence critical word and/or at the sentence-final word? And, if so, does this effect interact with the causal coherence or canonicity of the scenario? Based on previous research (Munte, Schiltz, & Kutas, 1998; Nakano, Saron, & Swaab, 2010), we expected to find an effect of working memory on the anterior negativity, and an interaction with canonicity of events such that the anterior negativity effect in non-canonical (versus canonical) critical words would be larger in low working memory span than in high working memory span participants.

Methods

Construction and characteristics of Stimuli

180 sets of two-clause sentences were developed, each with four conditions (sentence types). All sentences were in the past tense and contained, on average, 13.2 words. In each sentence, one clause described an event or state that caused a second event or state described in the other clause. In the *canonical coherent* sentences, the first clause described a cause, and the second clause described its effect, with the two described events or states unfolding in the sentence in the same order as they occur in the real world. The connector "and so", which encourages a forward causal inference, was used to render these clauses causally connected and coherent, e.g. "Fred was hungry and so he ate…". In the *non-canonical coherent* sentences, the first clause described an effect and the second clause described its cause, with the two events unfolding in the opposite order as they occur in the real world. The connector "because", which encourages a backwards causal inference, was used such that, once again, these clauses were rendered causally connected and coherent, e.g. "Fred ate because he was hungry…". Incoherent versions of each of these sentences were then constructed by swapping around the causal connectors, yielding two additional conditions: *canonical incoherent*, e.g. "Fred was hungry because he ate…" and *non-canonical incoherent*, e.g. "Fred ate and so he was hungry…". These manipulations gave rise to a 2 x 2 factorial design in which Canonical Sequence (canonical, non-canonical) was fully crossed with Causal Coherence (coherent, incoherent), see Table 1 for examples.

Table 1.

Sentence type (n=45 per condition)	**Example**	**Frequency of critical word‡**	**Length of critical word**	**LSA of critical word‡‡**	**Coherence rating***
Canonical Coherent	Fred was hungry and so he ate a meal that afternoon.	7.9 [2.0]	7.2 [2.1]	0.1 [0.1]	4.4 [0.3]
Canonical Incoherent	Fred was hungry because he ate a meal that afternoon.	7.9 [2.0]	7.2 [2.1]	0.1 [0.1]	1.8 [0.4]
Non-canonical Coherent	Fred ate a meal because he was hungry that afternoon.	8.2 [1.9]	6.9 [2.2]	0.1 [0.1]	4.4 [0.3]
Non-canonical Incoherent	Fred ate a meal and so he was hungry that afternoon.	8.2 [1.9]	6.9 [2.2]	0.1 [0.1]	1.8 [0.3]

The critical word in each of the example sentences is underlined here (although this was not the case in the experiment itself).

Means are shown with standard deviations in square brackets.

‡ Log transformed HAL word frequencies (Balota, et al., 2007) were taken from the English Lexicon Project database (http://elexicon.wustl.edu/). Some critical words did not exist in the HAL database and these were represented as null values in our calculations.

‡‡ Document-to-document LSA values were used to calculate semantic similarity values between the critical word and its preceding content words.

*Sentences were rated up to and including the critical word on a five-point scale from 1 (least plausible) to 5 (most plausible).

In each sentence, the critical word was defined as the content word that described the action or state described in the second clause of the sentence and which determined whether the sentence was coherent or incoherent. These are underlined in the examples above. Thus, one set of critical words appeared in the canonical sentences ("ate" in the example above) and another set of critical words appeared in the non-canonical sentences ("hungry" in the example above). In all sentences, the critical word appeared towards the end of the second clause but was never the final word of the sentence. 51 of the critical words were nouns, 77 were verbs and 52 were adjectives. 29 of the critical words (10.74%) were repeated once throughout the stimulus set, but lists were constructed such that the same participant never saw a critical word more than once (see below).

There was no difference in word frequency (Log-transformed HAL frequency; Balota, et al., 2007) or length (number of letters) between critical words appearing in the canonical and non-canonical sentences, all *t*s < 2.27, all *p*s $> .1$, see Table 1. A Latent Semantic Analysis (LSA, a measure of semantic relatedness) was carried out on a document-to-document basis (Landauer & Dumais, 1997; Landauer, Foltz, & Dumais, 1998), yielding Semantic Similarity Values between the critical words and preceding content words in each of the four experimental sentence types, see Table 1. A 2 (Canonical Sequence) x 2 (Causal Coherence) ANOVA showed no differences in these Semantic Similarity Values across the coherent and incoherent sentences, $F(1,179)=1.248$, $p=.265$, across the canonical and non-canonical sentences,

$F(1,179)=.379$, $p=.539$, and there was no interaction between Causal Coherence and Canonical Sequence, $F(1,179)=.001$, $p=.978$

Coherence ratings of sentences

A rating study was carried out in participants who did not take part in the ERP experiment to assess the plausibility of the sentences up until the point of the critical words. Each sentence was cut off after the critical word to create sentence stems, each followed by an ellipsis (…) indicating that the sentence continued after the critical word. These stems were divided into four counterbalanced lists using a Latin Square design, such that only one version (condition) of each sentence appeared per list, but so that all versions of a quadruplet appeared across all four lists. Each of these four lists was subdivided into half, giving rise to eight lists. The order of sentence-stems within each list was pseudo-randomized. The lists were presented using SurveyMonkey.com (one list per participant). Participants were recruited through online postings to Amazon Mechanical Turk (www.mturk.com), and were compensated $1.00 if they properly completed the survey, regardless of eligibility. However, only the data of participants who fulfilled the same eligibility requirements as the ERP study (described below) were analyzed: 177 participants in total (58 males; 119 females; average age: 23.4; at least 19 per list). Participants were asked to rate the sentence stems for how much they made sense (coherence) on a scale of 1 (least sense) to 5 (most sense). Each list contained four catch questions to ensure that participants were paying attention to the sentences (e.g. “This item is to make sure that bots are not completing this survey; please rate it a five . . .”).

A 2 (Causal Coherence) x 2 (Canonical Sequence) ANOVA confirmed that there was a significant difference in coherence ratings between the coherent and incoherent sentences, $F(1,179)=8971.056$, $p<.001$. There was, however, no difference in coherence ratings between the canonical and non-canonical sentences, $F(1,179)=.733$ p=.393, and there was no interaction between Causal Coherence and Canonical Sequence, $F(1,179)=.857$, $p=.356$, see Table 1 for coherence ratings.

Set-up of Experimental lists used in ERP study

90 sets of filler sentences were created to introduce variety in the structure and meaning of the sentences over the course of the experiment. Like the experimental sentences, these fillers also described events or states in the past tense and, on average, contained 13.4 words. Unlike the experimental sentences, however, they did not convey any explicit causal relationships. 50% of the fillers were incoherent, introducing a simple semantic anomaly at the point of the critical word. The filler sentences were constructed in quadruplets, in order that the same critical word could be counterbalanced across coherent and incoherent versions, e.g. "In his room the boy drew on/attached the construction paper with <u>markers</u>/<u>tape</u> all afternoon."). The two sets of critical words were matched on frequency, $t(89) = -1.663$, $p > .09$, length $t(89) = .271$, $p > .7$. The critical words of the fillers appeared towards the end of the sentence and were matched to critical words in the experimental sentences on length, $t(538) = 1.173$, $p > .6$ and frequency $t(500) = .262$, $p > .4$, but not on LSA values (document-to-document), $t(1078) = -2.937$, $p < .05$. In general, the LSA values for the fillers were higher (i.e., the content

words in the filler sentences were slightly more related to each other) than in the experimental sentences.

The experimental sentences were divided into four lists, counterbalanced using a Latin Square design. This ensured that, during the ERP experiment, only one participant would view one list and one version (condition) of each quadruplet of sentences, but all four conditions of each quadruplet would be seen across all participants. The filler sentences were also counterbalanced across the four lists such that each filler critical word appeared in both incoherent and coherent conditions across the four lists. The order of sentences within each list was pseudo-randomized so that no more than three sentences in succession belonged to the same condition, and so that each condition was represented equally across the beginning, middle, and end of each list. Each of the four lists included 180 experimental sentences (45 in each of the four conditions) and 90 filler sentences (45 in each of the two conditions).

ERP study

Participants

38 Tufts undergraduate students originally participated in the ERP study and 32 (13 male) participants were used in the final analysis (see below for reasons for exclusion). Eight participants viewed each of the four counterbalanced lists. All participants were right-handed, were native English speakers (having learned no other language before the age of 5), between the ages of 18 and 24, and had no history of psychiatric or neurological disorders. Prior to the ERP experiment, demographic

information was collected about each participant. Participants were paid for their participation and provided informed consent in accordance with the procedures of the Institutional Review Board of Tufts University.

Stimulus Presentation and Task

Participants sat in a comfortable chair in a dimly lit room. Stimuli were presented on a computer monitor, in white font, centered on a black background, and subtended at a visual angle of about 5°. Participants were randomly assigned to one of the four lists. Before starting the experiment, participants were given a practice round of 6 sentences (4 experimental and 2 fillers) that did not appear in the actual experiment, to ensure that they understood the task (see below).

All trials began with the word “READY”. The participant pressed a button on a gamepad to begin the trial. The sentence began with a fixation cross displayed for 500ms, followed by an inter-stimulus interval (ISI) of 100ms, followed by each word of the sentence presented individually for 450ms with an ISI of 100ms. The connector, “and so” was presented at once, as if it was one word. The final word of the sentence appeared on the screen for a longer duration of 750ms (400ms ISI).

After the final word of each sentence, a question mark appeared on the screen, cuing subjects to press one of two buttons on a game pad: “Yes” if the subject judged the sentence to make sense, or “No” if the subject judged the sentence not to make sense. Half of the participants pressed “Yes” with their left thumb and “No” with their right, and the other half did the opposite. Sentences were presented in six blocks of 45 sentences, with five breaks between blocks throughout the experiment.

ERP recording

The EEG response was recorded from 29 tin electrodes held in place by an elastic cap (Electro-Cap International, Inc., Eaton, OH; see Figure 1 for montage). Additional electrodes were placed below the left eye and at the outer canthus of the right eye to monitor vertical and horizontal eye movements. There were also two mastoid electrodes (A1, A2) and the EEG signal was referenced to the left mastoid online. The EEG signal was amplified by an Isolated Biometric Amplifier (SA Instrumentation Co., San Diego, California) with a band pass of 0.01—40 Hz. It was continuously sampled at 200Hz and the impedance was kept below 5kOhm for head electrodes, 2.5kOhm for mastoid electrodes and 10kOhm for eye electrodes.

Additional Working memory tasks

After the ERP portion of the experiment, participants completed five behavioral tasks aimed at measuring different aspects of working memory and inhibitory control. These were: the Automated Reading Span (RSPAN) Task (Daneman & Carpenter, 1980; Conway et al., 2005; Unsworth, Heitz, Schrock, & Engle 2005), a Stroop Task with keyboard response (Ilan & Polich, 1999; MacLeod, 1991; Repovs, 2004), the AX-Continuous Performance Task (AX-CPT; Barch & Smith, 2008; MacDonald, 2008), the Stop-Signal paradigm (Verbruggen, Logan, & Stevens, 2008), and the Anti-Saccade task (Jamieson & Harkins, 2007; Manoach et al., 2002).

Written instructions were provided before each task, but the participant was instructed to ask the experimenter if they did not understand the task. Of most interest to the current study was the Automated Reading Span (RSPAN) Task (Daneman &

Carpenter, 1980; Conway et al., 2005). In the RSPAN task, participants read sets of unconnected sentences and judged the plausibility of each sentence. After participants made their judgment for each sentence, a random letter was presented. At the end of each sentence set, subjects are asked to recall the letters in the order they were presented. WM score was calculated as the percentage of letters correctly recalled (in the correct position) across all sets. The automated presentation procedure developed and programmed for Inquisit by Millisecond Software by Unsworth, Heitz, Schorck, and Engels (2005) was used, which allows the responses to be recorded on the computer (rather than by the experimenter) and creates greater consistency in presentation and recording.

ERP Data Analysis

Analyses were carried out on averaged waveforms time-locked to critical words and sentence-final words in all trials, binned by the coherence and canonicity of the trials. Only correctly answered trials free of ocular and muscular artifact were averaged for analysis. Of the thirty-eight subjects who initially participated, six participants were excluded because, after artifact rejection (due to blinks and blocking (3), alpha waves (1) or other technical problems (2)), fewer than 25 trials remained in at least one experimental condition to be averaged for analysis.

Averaged ERPs were quantified by calculating the mean amplitude (relative to a peristimulus baseline of -50ms to +50ms the onset of the target word) in the following time-windows: 200-300ms, 300-400ms, 400-500ms, 500-750ms, and 750-900ms. In order to examine how the modulation of the waveforms varied across the scalp surface, the scalp was subdivided into regions along the anterior-posterior distribution of the scalp surface, at both mid and lateral sites (each region contained 3 electrode sites; see Figure 1). Two omnibus analyses of variance (ANOVAs), one covering mid regions and another covering lateral regions across the scalp, were conducted for each time window. In the mid-regions ANOVA, Causal Coherence, Canonicity, and Region were within-subjects factors; in the peripheral regions ANOVA, Hemisphere was an additional within-subjects factor. Significant interactions involving the Region factor were parsed by carrying out ANOVAs at each three-electrode Region individually. The Greenhouse Geisser correction was applied to repeated measures with more than one degree of freedom (Greenhouse & Geisser, 1959) and a significance level of alpha=0.05 was used for all comparisons.

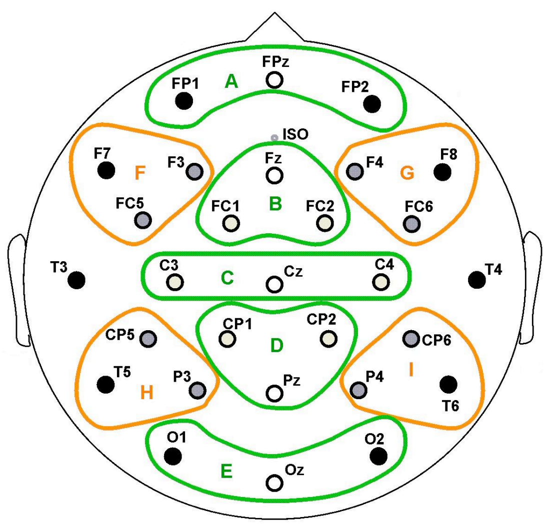


Figure 1. Electrode Montage and Regions for analysis. Omnibus Analyses of Variance (ANOVAs) covered a group of five mid regions (green) and a group of four peripheral regions (orange). Mid-regions: A: Prefrontal. B: Frontal. C: Central. D: Parietal. E: Peripheral. Peripheral regions: F: Left frontal. G: Right frontal. H: Left peripheral. I: Right peripheral.

Effects of Working Memory

Effects of working memory span were examined at two regions: the prefrontal region where we predicted anterior negativity effects, and the central region where we expected a robust N400 effect. In both regions, for each time window, we carried out 2 (Coherence) x 2 (Canonicity) x 2 (Working Memory) ANOVAs with Working Memory as a between-subject variable (high WM *versus* low WM, with groups defined through a median split on letter accuracy on the Automated Reading Span Task). We also carried out correlations as appropriate between WM and effects of interest within each of these regions. Analyses were carried out for both the critical word and at the sentence-final word.

Results

Behavioral data

During the ERP experiment all participants were fairly accurate in their acceptability judgments (average accuracy = 94.8%). There was no main effect of coherence or of canonicity for judgment accuracy, and no interaction (all Fs < 3.38, all ps > 0.07).

ERP data

Causal Connector (see Figure 2)

Approximately 14.3% of connector trials were rejected for artifact. Trial rejection for connectors did not differ across sentence type (all Fs < 1.19, all ps > 0.2).

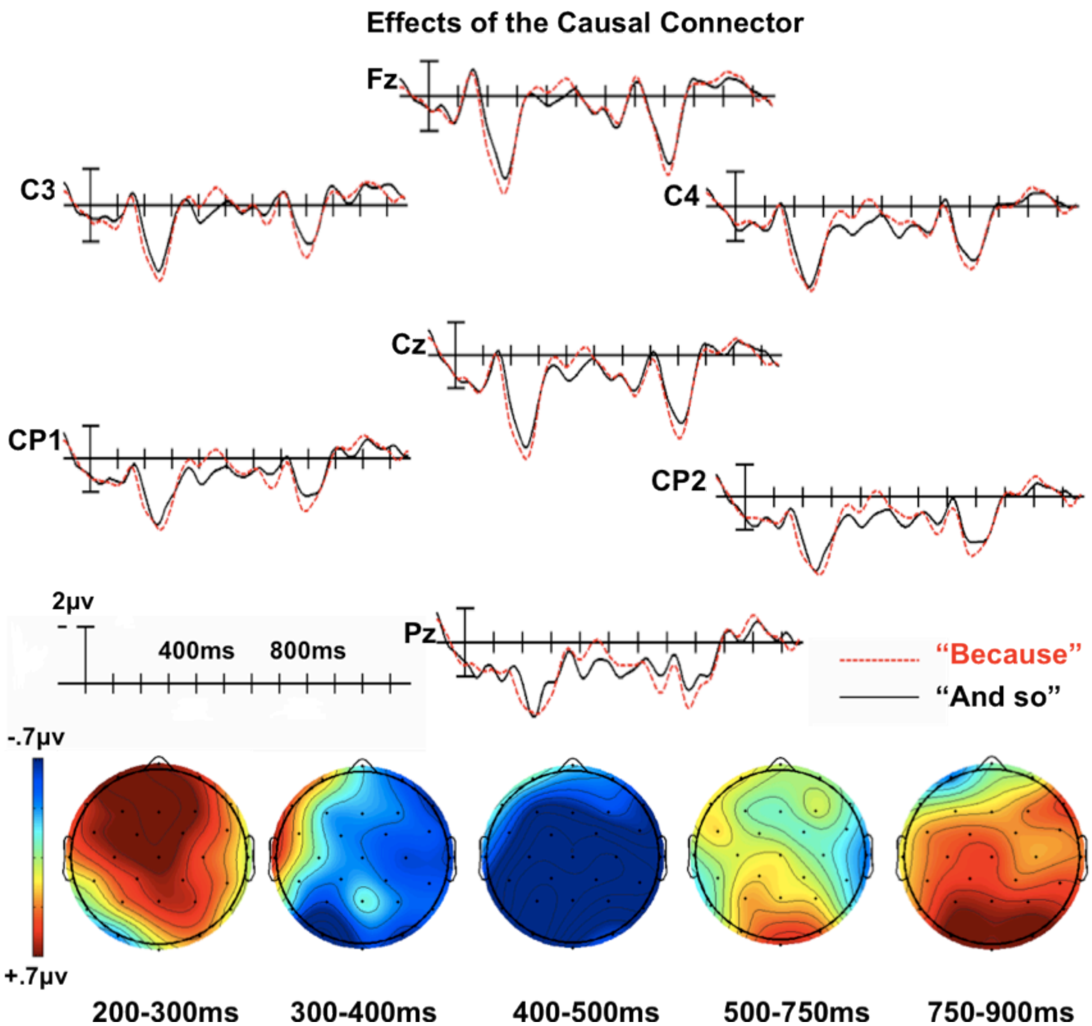


Figure 2. Grand-averaged waveforms to causal connectors "because" and "and so". Voltage maps comparing ERPs evoked by the connectors between 200 and 300 msec – the P200, 300 and 400 msec – the early N400 effect, 400 and 500 msec – the late N400 effect, 500 and 750 msec – the early P600 effect, and between 750 and 900 msec – the late P600 effect.

Between 200-300ms, 'because' evoked a slightly larger P2 than 'and so', reflected by effects of the Connector that approached significance in the omnibus mid-regions ($F(1, 31) = 3.54$, $p < 0.07$) and peripheral regions ($F(1, 31) = 3.49$, $p < 0.08$) ANOVAs. Between 300-400ms, 'because' evoked a slightly larger widely-distributed negativity than 'and so', again reflected by effects of the Connector that approached significance in the omnibus mid-regions ($F(1, 31) = 3.72$, $p < 0.07$) and peripheral regions ($F(1, 31) = 2.89$, $p < 0.1$)

ANOVAs. By 400-500ms, this negativity effect reached significance (mid-regions ANOVA: $F(1, 31) = 6.55$, $p < 0.02$; peripheral regions ANOVA: $F(1, 31) = 7.58$, $p < 0.02$), see Figure 2. No significant effects were seen from 500-750ms, but between 750-900ms, 'because' once again evoked a larger positivity than 'and so', particularly at occipital sites (mid-regions ANOVA Connector x Region interaction: $F(4, 124) = 3.29$, $p < 0.05$; peripheral regions ANOVA: main effect of Connector: $F(1, 31) = 4.41$, $p < 0.05$).

Critical word (see Table 2, Figure 3A, 3B)

Approximately 14.6% of critical word trials were rejected for artifact. There was an interaction between Coherence and Canonicity for trial rejection ($F(1, 31) = 6.951$, $p < 0.05$), but this was driven only by a slightly higher rejection rate in the coherent non-canonical sentences than in the other sentences (mean difference = 3.1%).

200-300ms

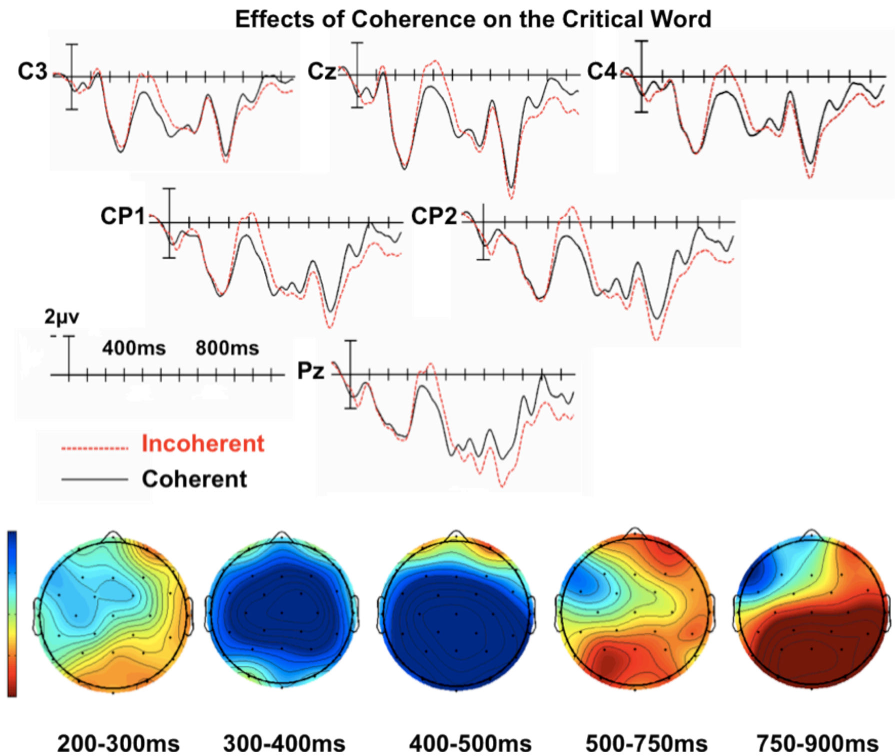
The omnibus ANOVAs failed to reveal any effects of either Coherence or Canonical Ordering in the 200-300ms time windows encompassing the P2 component (all Fs < 1.7, all ps > 0.2).

300-400ms

There was a clear N400 effect to causally incoherent (versus coherent) critical words, see Figure 3A. In this early N400 time window (300-400ms), this was reflected by main effects of Coherence in both the mid-regions and peripheral regions ANOVAs. This early N400 effect was widespread but maximal in frontal, central and parietal regions (Coherence x Region interaction in the mid-regions ANOVA, with significant effects at all regions except the prefrontal and occipital regions).

Table 2. Statistical effects of Canonical Ordering and Coherence at various time intervals time-locked to the critical word.

			300-400ms		400-500ms		500-750ms		750-900ms	
			Effects of Coherence							
	Effect	***df***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
Mid-regions omnibus ANOVA	Co	1,31	16.90	**0.000**	18.50	**0.000**	0.44	0.510	6.27	**0.020**
	CoxR	4,124	5.08	**0.010**	13.30	**0.000**	1.67	0.200	6.43	**0.000**
Prefrontal	Co	1,31	0.64	0.430	0.30	0.590	N/A		0.22	0.640
Frontal	Co	1,31	15.00	**0.000**	16.50	**0.000**			0.01	0.940
Central	Co	1,31	24.10	**0.000**	31.00	**0.000**			3.52	0.070
Parietal	Co	1,31	12.60	**0.000**	22.30	**0.000**			11.20	**0.000**
Occiptal	Co	1,31	1.81	0.190	11.20	**0.000**			20.80	**0.000**
Peripheral regions omnibus ANOVA	Co	1,31	13.70	**0.001**	20.00	**0.000**	0.17	0.680	3.43	0.070
	CoxRxH	1,31	0.03	0.850	6.22	**0.020**	16.90	**0.000**	10.80	**0.000**
Left Frontal	Co	1,31	N/A		6.22	**0.020**	2.83	0.100	3.28	0.080
Right Frontal	Co	1,31			4.64	**0.040**	0.48	0.500	0.54	0.470
Left Parietal	Co	1,31			16.50	**0.000**	1.94	0.170	13.50	**0.000**
Right Parietal	Co	1,31			17.50	**0.000**	0.71	0.410	11.60	**0.000**
			Effects of Canonicity							
	Effect	***df***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
Mid-regions omnibus ANOVA	Ca	1,31	0.75	0.390	0.02	0.900	0.14	0.710	1.18	0.290
	CaxR	4,124	0.43	0.620	4.38	**0.020**	0.74	0.440	3.61	**0.030**
Prefrontal	Ca	1,31	N/A		5.53	**0.030**	N/A		7.00	**0.010**
Frontal	Ca	1,31			0.26	0.610			1.22	0.280
Central	Ca	1,31			0.05	0.830			0.18	0.670
Parietal	Ca	1,31			0.66	0.420			0.00	0.960
Occipital	Ca	1,31			3.18	0.080			0.01	0.910
Peripheral regions omnibus ANOVA	Ca	1,31	0.24	0.630	0.02	0.880	0.02	0.880	1.00	0.320
	CaxR	1,31	0.70	0.410	4.77	**0.040**	0.39	0.540	4.34	**0.046**
Left Frontal	Ca	1,31	N/A		1.31	0.260	N/A		2.10	0.160
Right Frontal	Ca	1,31			0.52	0.470			2.39	0.130
Left Parietal	Ca	1,31			0.26	0.620			0.01	0.910
Right Parietal	Ca	1,31			1.28	0.270			0.08	0.780



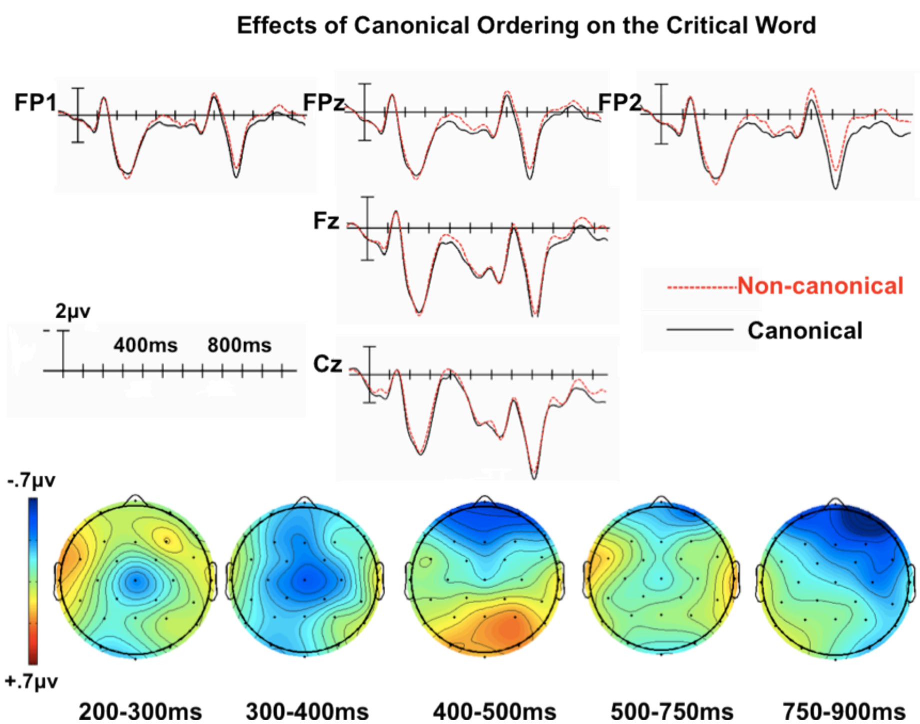


Figure 3A. Grand-averaged waveforms to critical words in coherent and incoherent sentences. Voltage maps comparing ERPs evoked by the critical word between 300 and 400 msec – the early N400 effect, 400 and 500 msec – the late N400 effect, 500 and 750 msec – the early P600 effect, and between 750 and 900 msec – the late P600 effect.

Figure 3B. Grand-averaged waveforms to critical words in canonical and non-canonical sentences. Voltage maps comparing ERPs evoked by the critical word between 300 and 400 msec, 400 and 500 msec – the anterior negativity effect, 500 and 750 msec, and 750 and 900 msec.

There were no effects of Canonical Ordering and no interactions between Canonical Ordering and Coherence within this 300-400ms time window.

400-500ms

In the 400-500ms time window, the N400 effect to causally incoherent (versus coherent) critical words was again reflected by main effects of Coherence in both the mid-regions and peripheral regions ANOVAs. Again, the effect was again widespread but maximal at centro-posterior sites (interactions between Coherence and Region, with follow-ups revealing significant effects in all regions except the prefrontal region). The effect also was also slightly right lateralized (a three-way interaction between Coherence, Region and Hemisphere in the peripheral regions omnibus ANOVA, with follow-ups revealing effects in all peripheral regions, but largest in the right parietal region).

Within this 400-500ms time window, effects of Canonical Ordering began to emerge: critical words in non-canonical sentences evoked a larger frontally distributed negativity than in the canonical sentences, see Figure 3B. This was reflected by interactions between Canonical Ordering and Region in both the mid-regions and peripheral regions ANOVAs, with follow-ups revealing effects of Canonical Ordering only in the prefrontal region.

Finally, there were interactions between Canonical Ordering, Coherence and Region that approached significance in the mid-regions ANOVA ($F(4, 124) = 2.66$, $p < 0.07$) and reached significance in the peripheral regions ANOVA ($F(1, 31) = 4.32$, $p < 0.05$). Follow-ups revealed significant N400 effects to incoherent (versus coherent) critical words in both canonical and non-canonical sentences (significant in all regions, all *F*s > 5.1, all *ps* < 0.04, except the prefrontal region for both and the occipital central region for non-canonical). However, the magnitude of the N400 effect was slightly larger in the canonical than in the non-canonical sentences in all individual regions. Direct comparisons between waveforms to canonical and non-canonical critical words, examining the coherent and incoherent sentences separately, revealed a larger N400 amplitude in centrl regions to coherence critical words in non-canonical scenarios (versus canonical), and in the anterior central region to incoherent critical words in non-canonical scenarios (versus canonical).

500-750ms

The only effect that reached significance in this time-window was a three-way interaction between Coherence, Region and Hemisphere in the peripheral regions omnibus ANOVA, $F(1, 31) = 16.9$, $p < .01$, but follow-ups failed to reveal effects of Coherence at any individual peripheral region (all *F*s < 2.9, all *p*s > 0.1).

750-900ms

Here, a robust positivity effect was seen to critical words in incoherent (versus coherent) sentences. This effect was most prominent in posterior regions (Coherence x Region interactions in both mid-regions and peripheral regions ANOVAs, with follow-ups revealing effects in parietal and occipital regions), and was larger over the left than the right hemisphere (Coherence x Hemisphere and Coherence x Hemisphere x Regions interactions in the peripheral regions ANOVA).

In addition to the effects of Coherence, effects of Canonical Ordering were also seen in this time window, with a larger frontally distributed negativity effect to critical words in non-canonical sentences than canonical sentences, see Figure 3B. This was reflected by interactions between Canonical Ordering and Region in both the mid-regions and peripheral regions ANOVAs, with follow-ups revealing effects of Canonical Ordering only in the most anterior prefrontal region.

Sentence-final word (see Table 3, Figure 4A, 4B)

Approximately 18.7% of sentence-final word trials were rejected for artifact. Trial rejection for sentence-final words did not differ across sentence type (all Fs < 0.8, all ps > 0.3).

200-300ms

In this early time window, there was a main effect of Coherence in the mid-regions omnibus ANOVA which reflected the beginning of the N400 effect to the sentence-final words in the causally incoherent (versus coherent) sentences (Figure 4A). This effect was maximal at more anterior regions (interactions between Coherence and Region approached significance in the mid-regions ANOVA, with follow-ups revealing effects in prefrontal, frontal and central regions).

There were also main effects of Canonical Ordering in both mid-regions and peripheral regions ANOVAs, reflecting a larger negativity to sentence-final words in non-canonical than canonical sentences . In the mid-regions ANOVA, the interaction between Canonical Ordering and Coherence again approached significance, with follow-ups again revealing effects in prefrontal, frontal and central regions.

Table 3. Statistical effects of Canonical Ordering and Coherence at various time intervals time-locked to the sentence-final word.

			200-300ms		300-500ms		500-750ms		750-900ms	
			Effects of Coherence							
	Effect	***df***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
Mid-regions omnibus ANOVA	Co	1,31	5.44	**0.030**	24.30	**0.00**	31.20	**0.000**	4.51	**0.040**
	CoxR	4,124	3.02	0.050	10.30	**0.00**	8.63	**0.000**	1.01	0.370
Prefrontal	Co	1,31	5.42	**0.030**	3.70	0.06	4.97	**0.030**	N/A	
Frontal	Co	1,31	5.50	**0.030**	20.50	**0.00**	23.00	**0.000**		
Central	Co	1,31	4.69	**0.040**	24.40	**0.00**	28.60	**0.000**		
Parietal	Co	1,31	4.14	0.050	30.10	**0.00**	39.10	**0.000**		
Occipital	Co	1,31	0.00	0.990	23.10	**0.00**	31.40	**0.000**		
Peripheral regions omnibus ANOVA	Co	1,31	1.86	0.220	19.00	**0.00**	19.10	**0.000**	0.49	0.490
	CoxR	1,31	2.78	0.110	7.10	**0.01**	8.36	**0.010**	2.12	0.160
Left Frontal	Co	1,31	N/A		8.61	**0.01**	4.15	0.050	N/A	
Right Frontal	Co	1,31			8.48	**0.01**	9.18	**0.000**		
Left Parietal	Co	1,31			24.90	**0.000**	19.20	**0.000**		
Right Parietal	Co	1,31			17.80	**0.00**	34.80	**0.000**		
			Effects of Canonicity							
	Effect	***df***	***F***	***p***	***F***	***p***	***F***	***p***	***F***	***p***
Mid-regions omnibus ANOVA	Ca	1,31	5.22	**0.030**	4.39	**0.04**	0.78	0.380	1.09	0.310
	CaxR	4,124	2.06	0.130	0.99	0.38	0.78	0.460	0.72	0.500
Prefrontal	Ca	1,31	5.50	**0.030**	1.55	0.22	N/A		N/A	
Frontal	Ca	1,31	4.88	**0.030**	4.40	**0.04**				
Central	Ca	1,31	4.57	**0.040**	4.23	**0.05**				
Parietal	Ca	1,31	1.94	0.170	3.21	0.08				
Occipital	Ca	1,31	2.08	0.160	1.63	0.21				
Peripheral regions omnibus ANOVA	Ca	1,31	4.30	**0.046**	3.30	0.08	0.00	0.950	0.39	0.540
	CaxR	1,31	2.78	0.105	7.10	**0.010**	8.36	**0.007**	2.12	0.160
Left Frontal	Ca	1,31	N/A		8.61	**0.01**	4.15	0.05	N/A	
Right Frontal	Ca	1,31			8.48	**0.01**	8.18	**0.01**		
Left Parietal	Ca	1,31			24.88	**0.00**	19.18	**0.00**		
Right Parietal	Ca	1,31			17.84	**0.00**	34.77	**0.00**		

*Although this effect did not reach significance, post-hoc follow-ups were carried out to explore the distribution of effects.

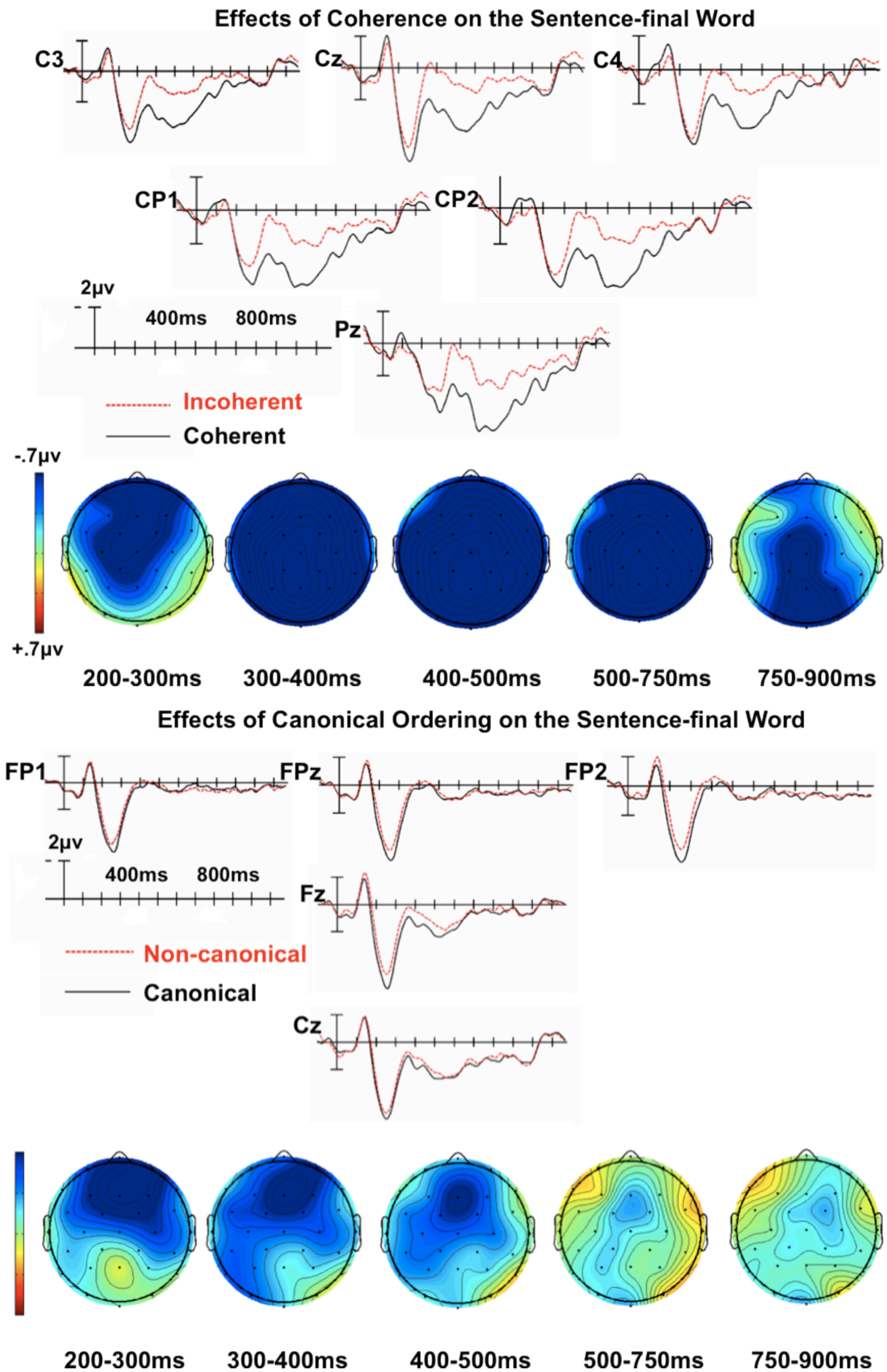


Figure 4A. Grand-averaged waveforms to sentence-final words in coherent and incoherent scenarios. Voltage maps comparing ERPs evoked by the sentence-final words between 200 and 300 msec, 300 and 500 msec – the N400 effect, 500 and 750 msec, and 750 and 900 msec.
Figure 4B. Grand-averaged waveforms to sentence-final words in canonical and non-canonical scenarios. Voltage maps comparing ERPs evoked by the sentence-final word between 200 and 300 msec, 300 and 500 msec – the anterior negativity effect, 500 and 750 msec, and 750 and 900 msec.

300-500ms

Within the main N400 time window, the effect on sentence-final words in the causally incoherent (versus coherent) sentences was widespread but largest posteriorly (main effects of Coherence and Coherence x Region interactions in both the mid-regions and peripheral-regions ANOVAs, with follow-ups showing effects in all but the prefrontal region), see Figure 4A and Table 3.

In addition, there were effects of Canonical Ordering which were significant in the mid-regions ANOVA and approached significance in the peripheral regions ANOVA (see Table 3) due to a larger negativity to non-canonical than canonical sentences. Examination of Figure 4B indicates that this effect was mainly frontally distributed (although there was no interaction between Canonical Ordering and Region, post-hoc analyses revealed significant effects only in frontal and central regions).

There was no interaction between Coherence and Canonical Ordering within this time window.

500-750ms

The negativity effect to incoherent (versus coherent) critical words continued into the 500-750ms time window where it was again reflected by main effects of Coherence and Coherence x Region interactions in both the mid-regions and peripheral regions ANOVAs, see Table 3. There were, however, no effects of Canonical Ordering in this time window, and no interactions involving Coherence and Canonical Ordering.

750-900ms

The negativity effect to incoherent (versus coherent) critical words further continued into the 750-900ms time window, reflected by main effects of Coherence in the mid-regions ANOVA.

Once again, there were no effects of Canonical Ordering. Nonetheless, there was a three-way interaction between Coherence, Canonical Ordering and Region, which approached significance in the mid-regions ANOVA ($F(4, 124) = 2.39$, $p < 0.1$) and reached significance in the peripheral regions ANOVA ($F(1, 31) = 4.74$, $p < 0.04$). Follow-ups at individual regions showed that, for the canonical sentences, the effect of Coherence was significant only in the occipital central region, whereas in the non-canonical sentences, the effect was significant in this region and as well as in the parietal central and left parietal peripheral region. Direct comparisons between ERPs to sentence-final words in canonical and non-canonical critical words, examining the coherent and incoherent sentences separately, revealed a larger negativity in central and frontal regions to coherent sentence-final words in non-canonical scenarios (versus canonical), and a larger negativity in frontal regions to incoherent sentence-final words in canonical scenarios (versus non-canonical).

Effects of Working Memory

Critical words: We subdivided our participants into high and low WM span groups through a median split on letter accuracy on the Automated Reading Span Task, and carried out ANOVAs with WM as an additional between-subject variable in the prefrontal region. This analysis revealed a main effect of WM span on critical words in the 200-300ms time window (encompassing the P2 component), $F(1, 28) = 6.593$, $p < 0.05$. As shown in Figure 5A, this was due to a larger anteriorly-distributed positivity to critical words (collapsed across all sentence types) in high WM-span than low WM-span individuals. As show in Figure 5B, there were significant correlations between WM span and the amplitude of the P2 (again collapsed across all sentence types), r = 0.436, $p < 0.05$. There was also an interaction between WM span and Coherence, $F(1, 28) = 6.14$, $p < 0.05$; follow ups in high and low WM span groups suggested that this interaction arose because the negativity effect to incoherent (versus coherent) critical words began slightly earlier (within this 200-300ms time window) in the high-span than in the low-span group. No effects of WM were seen in any other time windows within this prefrontal region (all Fs < 2.16, all ps > 0.15). There were also no effects of WM in the centro-parietal region (all Fs < 0.22, all ps > 0.64).

Sentence-final words. Within the prefrontal region, there were again main effects of WM which this time approached significance ($F(1, 28) = 3.89$, $p < 0.1$). There were once again no interactions involving WM and Coherence and/or Canonical Ordering (all Fs $<$ 2.07, all ps > 0.16). Again, as shown in Figure 5, the main effect of WM was due to a larger anteriorly-distributed positivity to sentence-final words (collapsed across all sentence types) in high than low WM-span individuals. And, once again, effects of WM were confirmed by significant correlations between WM span and the amplitude of the P2 (again collapsed across all sentences), r = 0.444, $p < 0.05$. Again, there were no effects of WM on the sentence-final word in any other time window within this most anterior region (all Fs < 0.9, all ps > 0.35), or within the centro-parietal region (all Fs < 0.41, all ps > 0.53).

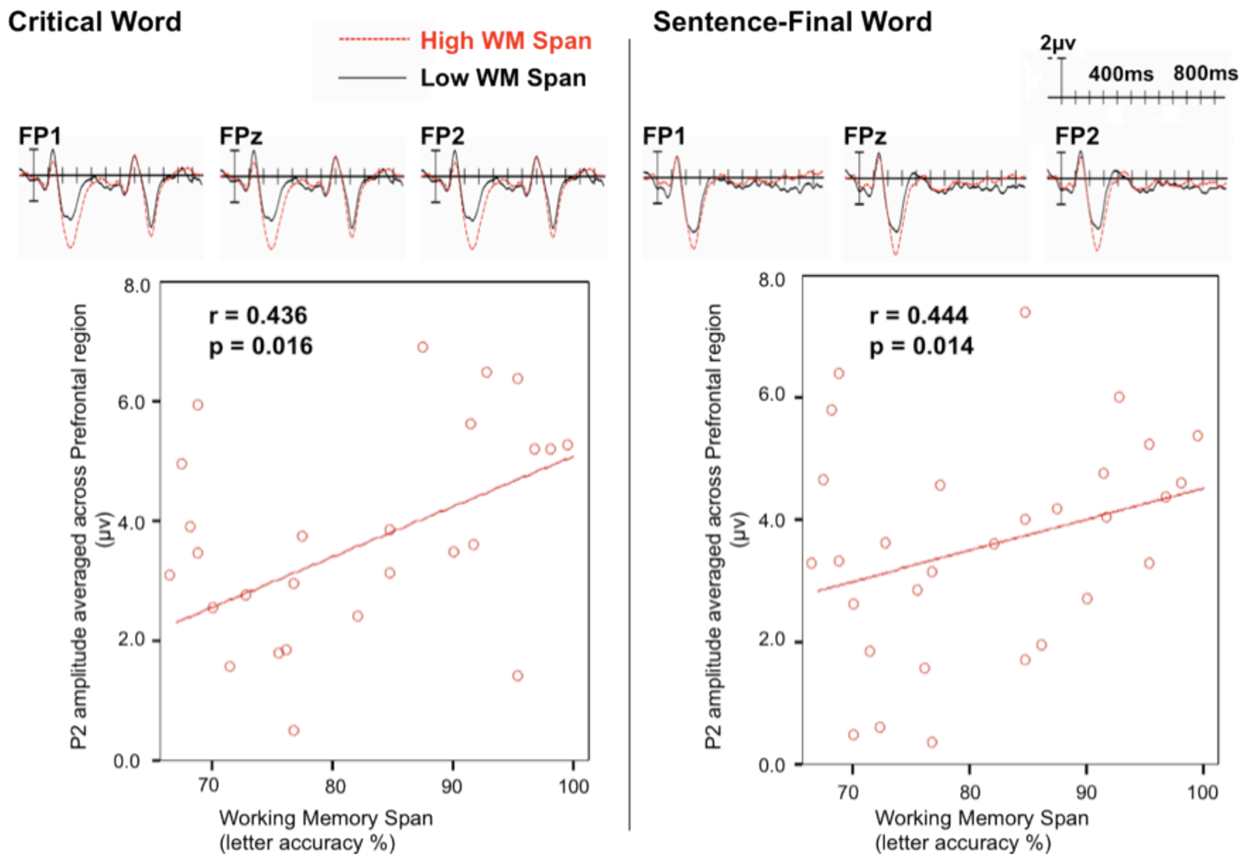


Figure 5. Top: Grand-averaged waveforms to critical words (left) and sentence-final words (right), collapsed across all four sentence types, in high Working Memory (WM) span participants (red dotted) and low Working Memory span participants (black solid). Bottom: Scatter plots illustrating significant correlations between the average amplitude of the P200 across the Prefrontal region evoked by critical words (left) and sentence-final words (right) and each participant's Working Memory span, measured by letter accuracy in the Automated Reading Span Task.

Discussion

The aim of the present study was to examine the time-course of establishing the coherence and canonicity of causally related events within sentences. There was a larger P2, N400, and P600 effects to the connector "because" than to "and so". There was an effect of coherence reflected in the N400 and late positivity to incoherent critical words and the sustained negativity to sentence-final words in causally incoherent compared to coherent sentences. There was also an effect of canonicity seen by a larger anterior negativity to non-canonical critical words and sentence-final words than to canonical ones. There was no interaction between the effect of coherence and the effect of canonicity. Lastly, there was an effect of working memory span at both the critical word and sentence final word such that high-span participants showed larger P200 amplitude than the low-span participants. This working memory affect was not modulated by either the coherence or canonicity of the sentences. Each of these findings will now be discussed in turn.

There were larger P2, N400, and P600 effects to the causal connector “because” compared to “and so”. The interpretation of this result is unclear, but it does suggest that there is some neural cost to encountering “because” since it carries information that the causal events being discussed are in non-canonical order. Munte, Schiltz, and Kutas (1998) found a larger negativity to “*Before X, Y*” sentences compared to “*After X, Y*” sentences starting at 300ms after the first word. This negativity starting in the N400 time-window may have appeared because “Before” also carries information that the events are described in reverse order from which they happened. However, in our study, the word “because” contains more letters than “and so” does, so it could merely be that the number of letters in the connector had driven the effect we saw.

We found a larger N400 to critical words in incoherent versus coherent scenarios. This finding is consistent with the findings of Kuperberg et al. (2011) and other past ERP studies (van Berkum et al., 1999; Camblin, Gordon, & Swaab, 2007) that indicate that, even when lexico-semantic associations between words are held constant, readers are immediately sensitive to incoherence in the discourse model. We were able to find this N400 effect specifically to mid-sentence critical words in causally incoherent discourse. The attenuation of the N400 to causally coherent vs. incoherent critical words is interpreted as a facilitation of processing due to a match between the situation model and the reader's stored semantic knowledge upon encountering the meaning of the critical word.

Notably, the N400 effect at the end of the sentence was sustained to incoherent versus coherent sentence-final words. This effect may have been maintained as a prolonged indicator of discourse incoherence, since reading a causally incoherent sentence in everyday life is generally rare. The sustained negativity may therefore reflect a continued suppression of the representation created by the sentence, as it is in conflict with the reader's semantic knowledge (Ye & Zhou, 2008; Pijnacker, Geurts, van Lambalgen, Buitelaar, & Hagoort, 2010). The lack of a P600 end-of-sentence effect may have resulted from encountering a period, signaling the end of the sentence and therefore eliminating the need for the type of additional neurocognitive processing reflected by the P600.

There was also a larger P600 to critical words in incoherent versus coherent scenarios, suggesting that additional neurocognitive processes are recruited to mid-sentence critical words when explicit causal connectors unambiguously violate coherence of the discourse model. Our findings differ from Kuperberg et al. (2011), who did not see a P600 effect of coherence. However, while all of the critical words in the present study occurred before the end of the sentence, in Kuperberg et al. (2011), about half of the critical words appeared before the end of the sentence and the other half were at the sentence-final position. Because of this, ERPs measured to critical words in Kuperberg et al. (2011) may have been confounded by end-of-sentence wrap-up effects.

Another important difference between the studies that could account for this P600 finding is the number of sentences in each stimulus. The stimuli in Kuperberg et al. (2011) were three-sentences, while the stimuli in the present study were single sentences. There have been a number of ERP studies examining causal coherence or relatedness that have not found P600 effects to causal incoherence, but they have all used multi-sentence stimuli (Kuperberg et al., 2011; St. George, Mannes & Hoffman, 1997; Yang, Perfetti, Schmalhofer, 2007). Perhaps breaking up the establishment of causal relationships among multiple sentences reduces the need for later neurocognitive processes (the P600) at a critical word. This ease of processing may be attributed to temporal factors (i.e., reading a single sentence takes less time than reading multiple sentences) or to wrap-up processing that happened when a period was encountered before the critical word (in these past studies, the critical word was always in the last sentence). Additionally, the present study had causal connectors in the single-sentence stimuli. These connectors carry information about causal direction of events and thus explicitly establish coherence or incoherence when the critical word is encountered. While in Kuperberg et al. (2011) there were three different levels of causal relatedness, in the present study there were only two levels: explicitly coherent or explicitly incoherent. Previous studies have found P600's to verbs in unambiguous semantic verb-argument violations such as "*Every morning the eggs would eat*…" (Kuperberg et al., 2003) and

"*The meal was* <u>*devouring*</u>*…*" (Kim & Osterhout, 2005). Thus, the unambiguous incoherence established by the critical word in the present study may have also contributed to the additional computation reflected in the P600 effect we found.

We found an anterior negativity effect to non-canonical versus canonical critical words and sentence-final words, consistent with our prediction. Contrary to our prediction that this negativity would be sustained and to the findings of Munte, Schiltz, and Kutas (1998), this effect diminished between 500-750ms and resurfaced between 750-900ms following the critical word (but not the sentence-final word). As described in the Introduction section, the anterior negativity is sensitive to the sequencing of words, phrases, and symbols in working memory (Hagoort, 2003; Friederici & Weissenborn, 2007; Kluender & Kutas, 1993; Hoen & Dominey, 2000; Vos, Gunter, Kolk, & Mulder, 2001). Our results showing a larger anterior negativity to critical words in non-canonical scenarios suggest that this negativity may also be involved in sequencing causal events and that this process requires working memory resources. This effect was seen in non-canonical scenarios because the causal connector "because" signaled that the first event was an effect of the second event and occurred afterwards in time. Therefore, in these scenarios, the first described event needed to be held in working memory until the second event was incorporated into the discourse model to be sequenced.

There was a larger anterior negativity between 400-500ms and between 750-900ms , but not between 500-750ms following non-canonical critical words (compared to canonical critical words). It is unclear whether the negativities in these two time-windows reflect the same process interrupted between 500-750ms, or whether they reflect different processes. Critical words in the present study determined both the causal coherence and canonicity of the sentence, so it is possible that processing between 500-750ms was involved in stabilizing the two ongoing processes of establishing discourse coherence and establishing canonicity. The negativity in the Munte et al. (1998) study may have been sustained without interruption because the events in their stimuli were not causally connected, while in the present study, the cognitive load of establishing coherence of the causal relationship between the described events may have temporarily drawn attention away from the continued sequence-determining process that took place in the Munte et al. (1998) study.

It is important to note that the N400 effect was located on the scalp centro-parietally, while the anterior negativity was located in the most frontal anterior region. The differing scalp distributions of the N400 and the anterior negativity suggest that there are distinct neural processes for building causal coherence and for building canonicity of events. This is further supported by the lack of statistical interaction of effects of coherence and canonicity both on the critical word and the sentence-final word. There was an interaction (significant at peripheral regions and approaching significance at midline regions) between coherence, canonicity, and region at the critical word between 400-500ms and direct comparison between canonical and non-canonical critical words revealed a slightly larger N400 effect to canonical than to non-canonical critical words. However, this difference was not large enough to result in a direct interaction between coherence and canonicity of scenarios. Accordingly, we determine that causal coherence of the discourse model did not significantly influence the sequencing effect – it occurred independent of discourse coherence.

Lastly, we found that individual differences in working memory span had an effect on the P200 component, contrary to our prediction of an effect on the anterior negativity. Participants with high working memory span had larger P2 amplitudes than did participants with low working memory span. Individual working memory span did not interact with coherence or canonicity of sentences, suggesting that working memory span does not directly influence ERP effects of causal coherence building. Furthermore, our finding of no effect of working memory but a significant effect of canonicity on the anterior negativity indicates that the type of working memory involved in determining canonicity of events is not modulated by individual differences, at least not as measured by the Automated Reading Span Task.

The P200 component is a positive-going waveform peaking around 200ms, typically in response to visual or auditory stimuli. There is agreement among researchers that the P200 reflects some feature of higher-order perceptual processing that is modulated by attention. However, the exact functionality that this component indexes is under debate because it can be modulated by a wide variety of cognitive tasks. The P200 has been seen as a pre-cognitive response to physical attributes of perceptual stimuli (Hampton & Weber-Fox, 2008; Roll, Horne, & Lindgren, 2010), and also appears to be related to working memory load processes in perceptual tasks, such as during long inter-stimulus intervals (Hampton & Weber-Fox, 2008), n-back tasks (Missionnier et al., 2003), and other memory tasks with visual or auditory stimuli (Lijffijt et al., 2009). In general, these studies have found that the greater the working memory load or difficulty of the perceptual task, the greater the amplitude of the P200. The ERP portion of the present study did not include a working memory task, but still we found modulation of the P200 component to critical words in all scenarios based on individual differences in working memory span. Therefore, we interpret higher working memory capacity as an automatic enhancement of pre-cognitive top-down attentional influences on perceptual processing of upcoming words in discourse.

<u>Future Directions</u>

We feel that our finding of a P600 effect in addition to an N400 effect to incoherence critical words is worth investigating further. Specifically, did the P600 effect emerge due to the reversibility of causal clauses? All of the incoherent sentences in the present study were "reversible": the sentence was coherent if the clauses swapped positions. There is some past evidence to semantic reversibility effects of the P600. Kolk, Chwilla, and van Herten et al. (2003) reported a P600 but no N400 effect to verbs such as "*fled/vluchtte*" in semantically reversible Dutch sentences such as "*The cat that from the mice <u>fled</u> ran across the room*/*De kat die voor de muizen <u>vluchtte</u> rende door de kamer*". Another Dutch study by Hoeks, Stowe and Doedens (2004) found again a P600 but no N400 to verbs like "*thrown/geworpen*" in sentences such as "*The javelin has <u>thrown</u> the athletes/De speer heft de atleten <u>geworpen</u>*". In both of these studies, the sentences were unambiguously incoherent, but recombining the lexical items such that the semantic roles were reversed (cat and mice; javelin and athlete in the above examples) made the sentence coherent. Attempting a reversal of these semantic roles may be involved in the processing reflected by the P600 in these studies (Casado, Martin-Loeches, Munoz, & Fernandez-Frias, 2005). Similarly, the P600 found in the present study may reflect attempts to construct causally coherent representations of the incoherent sentences by reversing the clauses.

We aim to address this question of reversibility and the P600 with a follow-up study using only non-canonical causal stimuli. We will use the coherent non-canonical and incoherent reversible non-canonical scenarios from this stimulus set and add two other types of scenarios: incoherent non-canonical unrelated (content words not matched for lexical semantic associations) and incoherent non-canonical irreversible (with content words matched for lexical semantic associations) (van Herten, Kolk, & Chwilla, 2005). If the P600 arose in the present study because a reversibility was detected, then we expect not to see a P600 effect to critical words incoherent non-canonical irreversible sentences. If instead, the P600 effect arose because of unambiguous causal incoherence due to explicit causal connectors, we do expect to see a P600 to incoherent non-canonical irreversible critical words.

Further, this study will investigate whether the P600 effect only sensitive to reversibility of clauses when lexico-semantic relations are matched across words. Does the P600 arises in cases of unambiguous causal incoherence only when words are matched for lexico-semantic associations, or in any sentence with explicit causal connectors creating discourse incoherence? If lexical semantic relatedness between words is required for this unambiguous causal violation to trigger the P600, then we expect not to see a P600 to critical words in incoherent non-canonical unrelated sentences. If semantic relatedness is not required for this P600 effect to clear-cut causal incoherence, then we do expect to see a P600 to critical words in these unrelated incoherent sentences.

There are several other open questions that future studies could help to elucidate. First, as discussed above, did the P600 arise because the coherence of the causal relationship was established within one sentence with explicit causal connectors unambiguously establishing discourse incoherence, rather than across multiple sentences? One way to explore this question using stimuli from the present study is to remove the causal connectors and break up the clauses into individual sentences.

Lastly, can an effect of working memory to causal coherence or canonicity of scenarios be found in ERPs? We demonstrated a strong main effect of working memory on the P200 predicted by performance on the Automated Reading Span Task, but surprisingly no interactions with discourse coherence or canonicity of events. However, there are different types of working memory that impact different types of cognitive processing (Kuperberg, 2007). As mentioned above, it is possible that the kind of working memory measured by the Automated Reading Span task is more sensitive to perceptual processing than to linguistic discourse-level processing. Perhaps individual differences in another measure of working memory span may interact with building coherence and/or ordering events in discourse.

In a recent study, Prat, Mason, and Just (2011) used fMRI to investigate individual working memory differences in neural correlates of causal inferencing. Their stimuli consisted of two-sentence stimuli in which causal relatedness (coherence) and presence of a causal connective (cohesion) were varied, and they used the Reading Span task (Daneman & Carpenter, 1980) to determine reading skill level. Skilled readers were more sensitive to coherence: they showed greater activation and synchronization to moderately related compared to distantly related scenarios. Less skilled readers were more sensitive to cohesion, showing greater activation and synchronization when causal connectives were present than absent. It is possible that we did not find an interaction between working memory and coherence or canonicity because all of our stimuli contained causal connectives. As suggested above, an ERP investigation with stimuli from the present study as they are, compared to the stimuli with the clauses broken up into individual sentences might yield working memory span (measured by reading span) interactions with coherence, cohesion, and perhaps canonicity.

Conclusions

There was an effect of coherence on the N400 and P600 at both the critical word and the sentence-final word, indicating that even when lexico-semantic relations across words are matched, causal incoherence of the discourse model affects neural processing at 300-500ms. When causal connectors unambiguously establish incoherence of the discourse model, additional neurocognitive processing is required past this N400 time-window, reflected by a P600/late positivity. There was also an effect of canonicity on the anterior negativity at both the critical word and the sentence-final word, suggesting that introducing causal events in non-canonical order requires some amount of working memory capacity. However, the effect was not continuously sustained past 500ms after the stimulus, possibly to due an attentional pull towards establishing causal coherence of the sentence. The effect of coherence and the effect of canonicity did not interact, indicating that the ordering of events does impact neural processing, but that this mechanism is distinct from the one that establishes causal coherence. Lastly, we found an effect of working memory span on the P200 component at both the critical and sentence-final words that did not interact with coherence or canonicity. This suggests that individual working memory capacity does not influence the mechanisms that establish temporal ordering and causal coherence, but rather may serve to enhance top-down attentional effects on perceptual processing of upcoming words in discourse.

References

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

Anderson, A., Garrod, S. C., & Sanford, A. J. (1983). The accessibility of pronominal antecedents as a function of episodic shifts in narrative text. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 35,* 427-440.

Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon project. *Behavior Research Methods, 39,* 445-459.

Barch, D. M. & Smith, E. (2008). The cognitive neuroscience of working memory: Relevance to CNTRICS and schizophrenia. *Biological Psychiatry, 64,* 11-17.

Bentin, S., McCarthy, G., & Wood, C. C. (1985). Event-related potentials, lexical decision and semantic priming. *Electroencephalography and Clinical Neurophysiology, 60,* 343–355.

Bornkessel, I.D., Fiebach, C.J., & Friederici, A. D. (2004). On the cost of syntactic ambiguity in human language comprehension: An individual differences approach. *Cognitive Brain Research, 21,* 11-21.

Burkhardt, P. (2006). Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. *Brain and Language, 98,* 159–168.

Burkhardt, P. (2007). The P600 reflects cost of new information in discourse memory. *NeuroReport, 18,* 1851–1854.

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking*. Journal of Memory and Language, 56,* 103–128.

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

Carreiras, M., Carriedo, N., Alonso, M. A., & Fernandez, A. (1997). The role of verb tense and verb aspect in the foregrounding of information during reading. *Memory & Cognition, 25,* 438-446.

Casado, P., Martin-Loeches, M., Munoz, F., & Fernandez-Frias, C. (2005). Are semantic and syntactic cues inducing the same processes in the identification of word order? *Cognitive Brain Research, 24,* 526-543.

Chwilla, D. J., Kolk, H. H., & Vissers, C. T. W. M. (2007). Immediate integration of novel meanings: N400 support for an embodied view of cognition. *Brain Research, 1183,* 109–123.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambric, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review. Special Issue: Memory strength and recency judgments, 12,* 769-786.

Coulson, S., & Kutas, M. (2001). Getting it: Human event-related brain response to jokes in good and poor comprehenders. *Neuroscience Letters, 316,* 71–74.

Coulson, S., & Van Petten, C. (2002). Conceptual integration and metaphor: An event-related potential study. *Memory & Cognition, 30,* 958–968.

Coulson, S., King, J., & Kutas, M. (1998). Expect the unexpected: Event-related brain responses to morphosyntactic violations. *Language and Cognitive Processes, 13,* 21–58.

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

De Grauwe, S., Swain, A., Holcomb, P. J., Ditman, T., & Kuperberg, G. R. (2010). Electrophysiological insights into the processing of nominal metaphors. *Neuropsychologia, 48,* 1965-1984.

De Vega, M. (2005). El procesamiento de oraciones con conectores adversativos y causales. / Processing of sentences with causal or adversative connectives. *Cognitiva, 17,* 85-108.

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2007). The contributions of lexico-semantic and discourse information to the resolution of ambiguous categorical anaphors. *Language and Cognitive Processes, 22,* 793–827.

Ditman, T., Holcomb, P. J., & Kuperberg, G. R. (2008). Time travel through language: Temporal shifts rapidly decrease information accessibility during reading. *Psychonomic Bulletin & Review, 15,* 750–756.

Donaldson, M. L., Reid, J., & Murray, C. (2005). Causal sentence production in children with language impairments. *International Journal of Language and Communication Disorders, 42,* 155-186.

French, L. A. (1988). The Development of Children’s Understanding of “Because” and “So”. *Journal of Experimental Child Psychology, 45,* 262-279.

Friederici A. D. & Weissenborn, J. (2007). Mapping sentence form onto meaning: The syntax-semantic interface. *Brain Research. Special Issue: Mysteries of meaning, 1146,* 50-58.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6,* 78-84.

Gernsbacher, M. A. (1997). Coherence cues mapping during comprehension. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships: Studies in the production and comprehension of text* (pp. 3-21). Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc.

Greenhouse, S., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24,* 95–112.

Hagoort, P. (1993). Impairments of lexical-semantic processing in aphasia: Evidence from the processing of lexical ambiguities. *Brain and Language, 45,* 189–232.

Hagoort, P. (2003). Interplay between Syntax and Semantics during Sentence Comprehension: ERP Effects of Combining Syntactic and Semantic Violations. *Journal of Cognitive Neuroscience, 15,* 883-889.

Hampton, A. & Weber-Fox, C. (2008). Non-linguistic auditory processing in stuttering: Evidence from behavior and event-related brain potentials. *Journal of Fluency Disorders, 33,* 253-273.

Haviland, S. E., & Clark, H. H. (1974). What’s new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Behavior, 13,* 512–521.

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research, 19,* 59–73.
Hoen, M. & Dominey, P. F. (2000). ERP analysis of cognitive sequencing: A left anterior negativity related to structural transformation processing. *NeuroReport: For Rapid Communication of Neuroscience Research, 11,* 3187-3191.
Holt, D. J., Lynn, S. K., & Kuperberg, G. R. (2009). Neurophysiological correlates of comprehending emotional meaning in context. *Journal of Cognitive Neuroscience, 21,* 2245–2262.
Ilan, A. B. & Polich, J. (1999). Tobacco smoking and memory scanning: Behavioral and event-related potential effects. *Nicotine & Tobacco Research, 1,* 233-240.
Jamieson, J. P. & Harkins, S. G. (2007). Mere effort and stereotype threat performance effects. *Journal of Personality and Social Psychology, 93,* 544-564.
Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99, 122*–149.
Katz, E. W. & Brent, S. B. (1968), Understanding Connectives. *Journal of Verbal Learning & Verbal Behavior, 7,* 501-509.
Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effect of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior, 23,* 115–126.
Kim, A. & Osterhout, L. (2005). The independence of combinatory semantic processing: Evidence from event-related potentials. *Journal of Memory and Language, 52,* 205-225.
Klin, C. M. (1995). Causal inferences in reading: From immediate activation to long-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 1483–1494.
Kluender, R. & Kutas, M. (1993). Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. *Journal of Cognitive Neuroscience, 5,* 196-214.
Kolk, H. H. J., & Chwilla, D. J. (2007). Late positivities in unusual situations. *Brain and Language, 100,* 257–262.
Kolk, H. H. J., Chwilla, D. J., van Herten, M., & Oor, P. J. W. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language, 85,* 1-36.
Kuperberg, G. R. (2007). Neural mechanisms of language comprehension: Challenges to syntax. *Brain Research, Special Issue: Mysteries of Meaning, 1146,* 23–49.
Kuperberg, G. R., Paczynski, M., & Ditman, T. (2011). Establishing Causal Coherence across sentences: An ERP Study. *Journal of Cognitive Neuroscience, 23,* 1230-1246.
Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science, 207,* 203–205.
Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307,* 161–163.
Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II: 1994–

2005. In M. Traxler & M. A. Gernsbacher (Eds.), *Handbook of psycholinguistics* (2nd ed., pp. 659–724). New York: Elsevier.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review, 104,* 211–240.

Landauer, T. K., Foltz, P. W., & Dumais, S. T. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes, 25,* 259–284.

Ledoux, K., Camblin, C. C., Swaab, T. Y., & Gordon, P. C. (2006). Reading words in discourse: The modulation of lexical priming effects by message-level context. *Behavioral and Cognitive Neuroscience Reviews, 5,* 107–127.

Lijffijt, M., Lane, S. D., Meier, S. L., Boutros, N. N., Burroughs, S., Steinberg, J. L., Gerard Moeller, F., & Swann, A. C. (2009). P50, N100, and P200 sensory gating: Relationships with behavioral inhibition, attention, and working memory. *Psychophysiology, 46,* 1059-1068.

MacDonald, M. C. & Christiansen, M. H. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review, 109,* 35-54.

MacDonald, A. W. (2008). Building a clinically relevant cognitive task: Case study of the AX paradigm. *Schizophrenia Bulletin, 34,* 619-628.

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

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109,* 163-203.

Mandler, J. M. (1985). On the comprehension of temporal order. *Language and Cognitive Processes, 1,* 309-320.

Manoach, D. S., Lindgren, K. A., Cherkasova, M. B., Goff, D. C., Halpern, E. F., Intriligator, J., & Barton, J. J. S. (2002). Scizophrenic subjects show deficient inhibition but intact task switching on saccadic tasks. *Biological Psychiatry, 51,* 816-826.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 12,* 82–91.

McKoon, G., & Ratcliff, R. (1989). Semantic associations and elaborative inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 326–338.

Missionnier, P., Leonards, U., Gold, G., Palix, J., Ibanez, V., & Giannakopoulos, P. (2003). A new electrophysiological index for working memory load in humans. *NeuroReport: Neurophysiology, Basic and Clinical, 14,* 1451-1455.

Molinaro, N., Vespignani, F., & Job, R. (2008). A deeper reanalysis of a superficial feature: An ERP study on agreement violations. *Brain Research, 1228,* 161-176.

Munte, T. F., Schiltz, K., & Kutas M. (1998). When temporal terms belie conceptual order. *Nature, 395,* 71-73.

Myers, J. L., & Duffy, S. A. (1990). Causal inferences and text memory. In A. C. Graesser & G. H. Bower (Eds.), *The psychology of learning and motivation: Inferences and text comprehension* (Vol. 25, pp. 159–173). San Diego, CA: Academic Press.

Myers, J. L., Shinjo, M., & Duffy, S. A. (1987). Degree of causal relatedness and memory. *Journal of Memory and Language,* 26, 453–465.

Nakano, H., Saron, C., & Swaab, T. Y. (2010). Speech and span: Working memory capacity impacts the use of animacy but not of world knowledge during spoken sentence comprehension. *Journal of Cognitive Neuroscience, 22,* 2886–2898.

Nieuwland, M. S., & Kuperberg, G. R. (2008). When the truth isn't too hard to handle: An event-related potential study on the pragmatics of negation. *Psychological Science, 19,* 1213–1218.

Nieuwland, M. S., & van Berkum, J. J. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research, 24,* 691–701.

Osterhout, L., & Hagoort, P. (1999). A superficial resemblance does not necessarily mean you are part of the family: Counterarguments to Coulson, King and Kutas (1998) in the P600/SPS–P300 debate. *Language and Cognitive Processes, 14,* 1–14.

Osterhout, L., & Holcomb, P. J. (1992). Event-related potentials elicited by syntactic anomaly. *Journal of Memory and Language, 31,* 785–806.

Otten, M., & van Berkum, J. J. (2007). What makes a discourse constraining? Comparing the effects of discourse message and scenario fit on the discourse-dependent N400 effect. *Brain Research, 1146,* 158–171.

Pijnacker, J., Geurts, B., van Lambalgen, M., Buitelaar, J., & Hagoort, P. (2010). Reasoning with Exceptions: An Event-related Brain Potentials Study. *Journal of Cognitive Neuroscience, 23,* 471-480.

Potts, G. R., Keenan, J. M., & Golding, J. M. (1988). Assessing the occurrence of elaborative inferences: Lexical decision versus naming. *Journal of Memory and Language, 27,* 399–415.

Prat, C. S., Mason, R. A., & Just, M. A. (2011). Individual differences in the neural basis of causal inferencing. *Brain and Language, 116,* 1-13.

Radvansky, G. A., Zwaan, R. A., Federico, T., & Franklin, N. (1998). Retrieval from temporally organized situation models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 1224-1237.

Repovs, G. (2004). The mode of response and the Stroop effect: A reaction time analysis. *Psiholoska Obzorja/Horizons of Psychology, 13,* 105-114.

Roll, M., Horne, M., & Lindgren, M. (2010). Word accents and morphology – ERPs of Swedish word processing. *Brain Research, 1330,* 114-123.

Rosler, F., Pechman, T., Streb, J., Roder, B., & Hennighausen, E. (1998). Parsing of Sentences in a Language with Varying Word Order: Word-by-Word Variations of Processing Demands are Revealed by Event-Related Brain Potentials. *Journal of Memory and Language, 38,* 150-176.

Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology, 22,* 642–647.

Segal, E. M., Duchan, J. F., & Scott, P. J. (1991). The role of interclausal connectives in narrative structuring: Evidence from adults' interpretations of simple stories. *Discourse Processes, 14,* 27-54.

Singer, M., & Halldorson, M. (1996). Constructing and validating motive bridging

inferences. *Cognitive Psychology, 30,* 1–38.
Singer, M. (1993). Causal bridging inferences: Validating consistent and inconsistent sequences. *Canadian Journal of Experimental Psychology, 47,* 340–359.
Singer, M., Andrusiak, P., Reisdorf, P., & Black, N. L. (1992). Individual differences in bridging inference processes. *Memory & Cognition, 20,* 539–548.
St. George, M., Mannes, S., & Hoffman, J. E. (1997). Individual differences in inference generation: An ERP analysis. *Journal of Cognitive Neuroscience, 9,* 776–787.
Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632–1634.
Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle R. W., (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498-505.
van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: Evidence from the N400. *Journal of Cognitive Neuroscience, 11,* 657–671.
van Berkum, J. J. A., Holleman, B., Nieuwland, M., Otten, M., & Murrel, J. (2009). Right or wrong? The brain's fast response to morally objectionable statements. *Psychological Science, 20,* 1092–1099.
van de Meerendonk, N., Kolk, H. H. J., Vissers, C. T. W. M., & Chwilla, D. J. (2010). Monitoring in language perception: Mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience, 22,* 67–82.
Van der Meer, E., Beyer, R., Heinze, B., & Badel, I. (2002). Temporal order relations in language comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 770-779.
van Herten, M., Kolk, H. H. J., & Chwilla, D. J. (2005). An ERP study of P600 effects elicited by semantic anomalies. *Cognitive Brain Research, 22,* 241-255.
Van Petten, C. (1993). A comparison of lexical and sentence-level context effects in event-related potentials. Special Issue: Event-related brain potentials in the study of language. *Language and Cognitive Processes, 8,* 485–531.
Van Petten, C., Kutas, M., Kluender, R., Mitchiner, M., & Melsaac, H. (1991). Fractionating the word repetition effect with event-related potentials. *Journal of Cognitive Neuroscience, 3,* 131–150.
Verbruggen, F., Logan, G. D., & Stevens, M. A. (2008). STOP-IT: Windows executable software for the stop-signal paradigm. *Behavior Research Methods, 40,* 479-483.
Vos, S. H. & Friederici, A. D. (2003). Intersentential context effects on comprehension: The role of working memory. *Cognitive Brain Research, 16,* 111-122.
Vos, S. H., Gunter, T. C., Kolk, H. H. J., & Mulder, G. (2001). Working memory constraints on syntactic processing: An electrophysiological investigation. *Psychophysiology, 38,* 41-63.
Yang, C. L., Perfetti, C. A., & Schmalhofer, F. (2007). Event-related potential indicators of text integration across sentence boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 55–89.
Ye, Z. & Zhou, X. (2008). Involvement of cognitive control in sentence comprehension: Evidence from ERPs. *Brain Research, 1203,* 103-115.

Zwaan, R. A. (1996). Processing narrative time shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22,* 1196-1207.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162–185.

Appendix A

Table 4. Fillers

Sentence type (n=45 per condition)	**Example**	**Frequency of critical word‡**	**Length of critical word**	**LSA of critical word‡‡**
Coherent	The janitor drank his coffee with cream and sugar that morning.	7.9 [1.9]	6.8 [2.2]	0.1 [0.2]
Incoherent	The janitor drank his coffee with cream and bleach in the morning.	8.1 [1.9]	6.2 [1.9]	0.1 [0.2]

The critical word in each of the example sentences is underlined here (although this was not the case in the experiment itself).

Means are shown with standard deviations in square brackets.

‡ Log transformed HAL word frequencies (Balota, et al., 2007) were taken from the English Lexicon Project database (http://elexicon.wustl.edu/). Some critical words did not exist in the HAL database and these were represented as null values in our calculations.

‡‡ Document-to-document LSA values were used to calculate semantic similarity values between the critical word and its preceding content words.

Appendix B

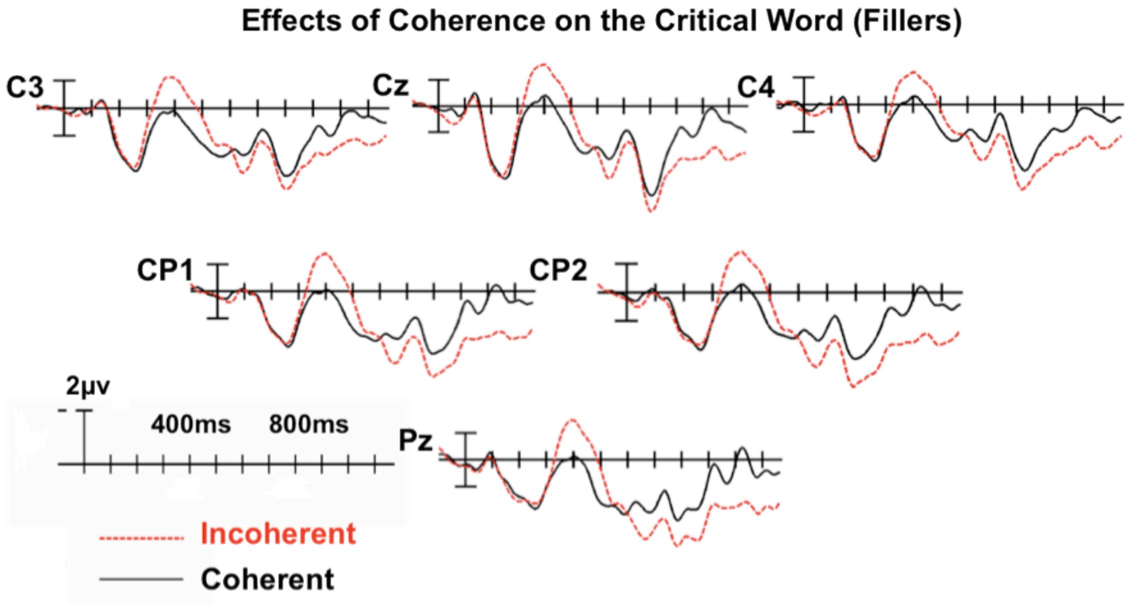


Figure 6. Grand-averaged waveforms to critical words in coherent and incoherent non-causal filler scenarios.