Published in final edited form as:
J Cogn Neurosci. 2011 May ; 23(5): 1230–1246. doi:10.1162/jocn.2010.21452.

Establishing causal coherence across sentences: an ERP study

Gina R. Kuperberg[*,‡], **Martin Paczynski**[*], and **Tali Ditman**[*,‡]
[*]Department of Psychology, Tufts University, Medford, MA 02155, USA

[‡]Department of Psychiatry and MGH/MIT/HMS Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, Bldg 149, 13th Street, Charlestown, Massachusetts 02129, USA

Abstract

This study examined neural activity associated with establishing causal relationships across sentences during online comprehension. ERPs were measured while participants read and judged the relatedness of three-sentence scenarios in which the final sentence was highly causally related, intermediately related and causally unrelated to its context. Lexico-semantic co-occurrence was matched across the three conditions using a Latent Semantic Analysis. Critical words in causally unrelated scenarios evoked a larger N400 than words in both highly causally related and intermediately related scenarios, regardless of whether they appeared before or at the sentence-final position. At midline sites, the N400 to intermediately related sentence-final words was attenuated to the same degree as to highly causally related words, but otherwise the N400 to intermediately related words fell in between that evoked by highly causally related and intermediately related words. No modulation of the Late Positivity/P600 component was observed across conditions. These results indicate that both simple and complex causal inferences can influence the earliest stages of semantically processing an incoming word. Further, they suggest that causal coherence, at the situation level, can influence incremental word-by-word discourse comprehension, even when semantic relationships between individual words are matched.

Keywords

causal coherence; discourse; ERP; N400; P600; inference; language; situation model

Introduction

Causal relationships play a vital role in structuring the meaning of text and discourse by establishing physical, motivational, and psychological links between expressed events, actions, and states (Fletcher & Bloom, 1988; Schank & Abelson, 1977; Trabasso, van den Broek, & Suh, 1989; Trabasso & Van den Broek, 1985; van den Broek, 1990). There is consistent behavioral evidence that, during reading, comprehenders make use of such causal relationships to establish coherence during comprehension: people are faster to read sentences that are causally related (versus causally unrelated) to their preceding sentence (e.g., Haviland & Clark, 1974; Keenan, Baillet, & Brown, 1984; see also Bloom, Fletcher, van den Broek, Reitz, & Shapiro, 1990; Myer, Shinjo, & Duffy, 1987). They are also faster to recognize (McKoon & Ratcliff, 1989), name (e.g., Klin, 1995; Potts, Keenan, & Golding, 1988) and make lexical decisions (e.g., Potts et al., 1988; but see McKoon, Ratliff, & Ward,

Correspondence should be addressed to Gina R. Kuperberg, Department of Psychology, Tufts University, 490 Boston Avenue, Medford, MA 02155, USA; Tel: 617-726-3432; kuperber@nmr.mgh.harvard.edu.

1994) on probe words that are introduced following causally related (versus causally unrelated) sentences.

Causal coherence is often established through causal inferences – information that is not explicitly stated in the text but which is required to establish causal relationships across clauses. These inferences vary in complexity: when two statements are highly causally related, they simply entail the activation and integration of relevant real-world knowledge. For example, in comprehending highly causally related sentence pairs such as, "Dorothy poured the bucket of water on the bonfire. The fire went out", readers activate their real-world knowledge that *water extinguishes fire* (Singer, 1993; Singer, Andrusiak, Reisdorf, & Black, 1992; Singer & Halldorson, 1996). However, when statements are less causally related, more complex inferences are necessary to establish necessary and sufficient causal connections between them (Bloom et al., 1990; Fletcher, Hummel, & Marsolek, 1990; Graesser & Clark, 1985; Keenan et al., 1984; Kintsch & van Dijk, 1978; Long, Golding, Graesser, & Clark, 1990; Myers & Duffy, 1990; Myers et al., 1987; Van Dijk & Kintsch, 1983). For example, to connect the sentence, "The next day his body was covered in bruises" to a preceding statement such as "Joey's brother became furiously angry with him", one must infer that Joey's brother must have been so angry that he hit Joey. While behavioral studies have clearly established that the build up of causal coherence across clauses is a prerequisite for successful comprehension, they are limited in how much information they can yield about the nature and time-course of causal inferencing. Self-paced reading studies have generally measured reading times to whole sentences, rather than each individual word (e.g., Haviland & Clark, 1974; Keenan et al., 1984; Myers & Duffy, 1990; Myers et al., 1987). And many behavioral studies require comprehenders to recognize or respond to probe words (McKoon & Ratcliff, 1986, 1989) or questions (Singer, 1993; Singer et al., 1992; Singer & Halldorson, 1996) that are presented only after a passage has been read, requiring participants to explicitly look back at that scenario to determine whether it is consistent with any knowledge retrieved. Thus, these techniques may not probe online word-by-word comprehension processes (see Keenan, Potts, Golding, & Jennings, 1990 for discussion). In an effort to circumvent these problems, some investigators have asked participants to make lexical decisions (e.g., Potts et al., 1988) or to name (e.g., Klin, 1995; Potts et al., 1988) probe words. However, lexical decisions may still entail some retrospective context checking and naming can, in theory, be carried out without lexico-semantic processing and may therefore be relatively insensitive to causal inferential processes (Keenan et al., 1990).

Because of these limitations, it remains unclear whether the build-up of causal coherence across sentences is a truly incremental process that influences the processing of each incoming word as discourse unfolds. This would be line with models of language comprehension holding that multiple sources of information, beyond lexico-semantic and syntactic information, influence word-by-word processing (Altmann & Steedman, 1988; MacDonald, Pearlmutter, & Seidenberg, 1994; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995). Alternative possibilities are that the establishment of causal coherence has a prolonged time-course that is spread over many words, or that causal coherence is only properly established at the end of a sentence when there are clear cues (e.g. a period) for a comprehender to 'wrap-up' overall meaning.

If causal coherence is established on a word-by-word basis, an important question is how quickly, after the onset of each incoming word, real-world knowledge is available, i.e. how quickly inferences are generated. One possibility is that the generation of a causal bridging inference entails a short-lived detection of incoherence followed by the retrieval and/or selection of relevant real-world knowledge to fill in the missing link. For example, Singer and colleagues have suggested that, to link two related sentences, readers retrieve and integrate real-world knowledge only *after* they have initially constructed a mental syllogism with a missing premise (Singer, 1993; Singer et al., 1992; Singer & Halldorson, 1996). This may be even more likely to occur when readers require more complex inferences to link intermediately related sentences (Keenan et al., 1984; Myers & Duffy, 1990; Myers et al., 1987). In a series of studies, Keenan et al. (1984) and Myers et al. (1990; 1987) examined reading times and recall of sentences such as “The next day his body was covered in bruises” when preceded by (1) highly causally related sentences (e.g. “Joey’s brother punched him again and again”), (2) intermediately causally related sentences (e.g. “Joey’s brother became furiously angry with him”) and (3) unrelated sentences (e.g. “Joey went to a neighbor’s house to play”). They demonstrated that reading times to the intermediately causally related sentences were in between those to the highly causally related and unrelated scenarios. Interestingly, however, participants’ showed maximum cued recall to these intermediately related sentences compared with both other types of scenario. Myers et al. (1987) explained this non-linear pattern of recall by suggesting that the generation of complex causal inferences to the intermediately related sentence pairs involved a stage of active, deep processing which later helped facilitate their cued retrieval. It may therefore be that the generation of such complex causal inferences involved a distinct stage of retrieval and/or selection of relevant information from real-world knowledge.

A second possibility is that the activation of real-world knowledge and the generation of causal inferences influence the earliest stages of semantically processing an upcoming word. There is certainly evidence that real-world knowledge immediately impacts lexico-semantic processing both within sentences (Hagoort, Hald, Bastiaansen, & Petersson, 2004; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003) and, at least in highly constrained contexts, across sentences (Otten & van Berkum, 2007; van Berkum, Hagoort, & Brown, 1999). It remains unclear, however, whether such fast activation of real-world knowledge mediates the construction of causal bridging inferences, particularly those that are complex in nature. If this were the case, it would imply that these inferences are generated *before* the onset of the critical word (a true predictive inference; Cook, Limber, & O’Brien, 2001; Klin, Guzman, & Levine, 1999), or as soon as the perceptual features of the critical word appear.

It is important to note that both the possibilities outlined above would imply that causal bridging inferencing occurs incrementally during word-by-word online processing. The difference between the two accounts is *how* quickly real-world knowledge is available to construct causal coherence: whether relevant knowledge is active and selected very quickly *after* any incoherence is detected, or whether the relevant knowledge is already available to facilitate lexico-semantic processing of incoming words.

Finally, the *level* of processing and representation at which causal coherence is established during word-by-word processing remains unclear. It has been acknowledged for some time that the representations built during discourse comprehension are multi-layered (Kintsch, 1988, 1992; Van Dijk & Kintsch, 1983), constituting a *surface code* (the precise wording and syntax of the text), a *textbase* (the propositional structure), and a *mental model* (Johnson-Laird, 1983) or *situation model* (Van Dijk & Kintsch, 1983) which includes dimensions of causal, temporal, spatial, and other information (Zwaan, Magliano, & Graesser, 1995; Zwaan & Radvansky, 1998). There is evidence that causal inferencing occurs at each of these levels of representation: an overlap between arguments in the propositional text-base and prior stored information (Cook, Halleran, & O’Brien, 1998; Kintsch, 1988, 1998; McKoon & Ratcliff, 1989; Myers & O’Brien, 1998; Myers, O’Brien, Albrecht, & Mason, 1994; Sanford & Garrod, 1981, 1998) may lead to the activation of new information through semantic priming mechanisms (Duffy, Henderson, & Morris, 1989; O’Seaghdha, 1997; Sharkey & Sharkey, 1987). In addition, situation-level meaning can lead to the generation of causal inferences, even when lexico-semantic relationships across conditions are matched (McKoon & Ratcliff, 1989). Although lexico-semantic and situation-level information usually act synergistically with one another to build discourse coherence, there are situations in which semantic associations between individual words can override an overall discourse message, either influencing the final interpretation, such as in the classic semantic illusion phenomenon ("How many animals did Moses take into the ark"; Erickson & Mattson, 1981; Kamas, Reder, & Ayers, 1996), or during the initial stages of processing an incoming word (Garrod & Terras, 2000). What remains unclear is how this dynamic interaction plays out during the establishment of causal coherence as discourse unfolds, word by word.

Event-related potentials and the present study

We aimed to address these outstanding questions using event-related potentials (ERPs) – an online neural measure that yields information about the neurocognitive processes engaged as meaning is constructed during word-by-word comprehension, in advance of any explicit behavioral response. Our focus was on the modulation of two ERP components – the N400 and the Late Positivity/P600.

The N400 is a negative-going waveform that peaks at approximately 400ms and whose amplitude reflects the ease of semantically processing incoming words with respect to their preceding context and information stored in semantic memory (Kutas, Van Petten, & Kluender, 2006). The amplitude of the N400 is attenuated when there is a good (versus poor) semantic fit between the incoming word and its context – the N400 effect. This context may constitute other words in semantic priming paradigms (Bentin, McCarthy, & Wood, 1985; Rugg, 1985), sentence stems (Kutas & Hillyard, 1980, 1984), or whole discourse (van Berkum et al., 1999).

During sentence and discourse processing, the N400 can provide a sensitive index of how a situation model and pure lexico-semantic relationships interact to modulate semantic processing of each incoming word. This is because, although the amplitude of the N400 evoked by a given word can be sensitive to the situation model built up by its context (Chwilla, Kolk, & Vissers, 2007; Ditman, Holcomb, & Kuperberg, 2007, 2008; Nieuwland & Kuperberg, 2008; Otten & van Berkum, 2007; van Berkum et al., 1999), it is also modulated by pure lexico-semantic relationships between that word and its preceding content words (Ditman et al., 2007; Ledoux, Camblin, Swaab, & Gordon, 2006; Otten & van Berkum, 2007; Van Petten, 1993). Indeed, under some circumstances, close lexico-semantic relationships can temporarily dominate processing, overriding situation-level representations and leading to a so-called temporary neural semantic illusion (Nieuwland & van Berkum, 2005).

Late Positivities or P600s are a group of positive-going components which peak at a later time-point than the N400 and which can extend until approximately 900 ms after stimulus onset. They are evoked in many different situations that require or entail a second stage of processing – by words that violate the syntactic structure of their preceding context (Hagoort, 1993; Osterhout & Holcomb, 1992), by words that are highly semantically implausible/impossible with respect to their context (Kuperberg, 2007; van de Meerendonk, Kolk, Vissers, & Chwilla, In press), during the comprehension of metaphor (Coulson & Van Petten, 2002; De Grauwe, Swain, Holcomb, Ditman, & Kuperberg, Under review), jokes (Coulson & Kutas, 2001) and emotional language (Holt, Lynn, & Kuperberg, 2009; van Berkum, Holleman, Nieuwland, Otten, & Murrel, 2009). Late Positivities are likely to subsume multiple neurocognitive processes and their functional significance is debated (Coulson, King, & Kutas, 1998; Kolk & Chwilla, 2007; Kuperberg, 2007; Osterhout & Hagoort, 1999). However, there is a general consensus that such processes often involve a continued analysis of incoming word with respect to its context and information stored within long-term memory. On some accounts, Late Positivities may reflect an active attempt to integrate incoming text with information stored within long-term memory (Van Petten, Kutas, Kluender, Mitchiner, & Melsaac, 1991), possibly involving a process of selection (Federmeier, Wlotko, De Ochoa-Dewald, & Kutas, 2007), and/or the addition of new information into the discourse model (Burkhardt, 2006, 2007).

Thus far, there have been three studies that have used ERPs to examine how *causal* relationships are built up across sentences. In an early study, St. George, Mannes, & Hoffman (1997) presented participants with scenarios that encouraged the generation of a causal inference, e.g. "…She forgot about the turkey in the oven. The guests were disappointed with the ruined meal." or that did not encourage such an inference "… She put the turkey in the oven. She was disappointed when the argumentative guests ruined the meal." Participants then viewed a final sentence that confirmed any inference generated, e.g. "It was too bad the turkey was burned." The N400, averaged across all words in this final sentence, was attenuated in the inference-encouraging, relative to the control, scenarios, and the authors concluded that the earlier activation of the causal inference facilitated the semantic processing of words in the final sentence. However, because ERPs were averaged across all words in the final sentence, and because these words themselves (e.g. "burned") may have triggered an inference in both conditions, it remains unclear whether participants spontaneously generated or integrated inferences during word-by-word processing of the first two sentences.

In a more recent study, Yang, Perfetti, and Schmalhofer (2007) measured ERPs to the first word of a sentence that was either the same (referentially explicit) or semantically related (referentially paraphrased) to the final word of a preceding causally-related sentences (e.g. "…the bomb hit the ground and exploded/blew up. The <u>explosion</u>…"). "Explosion" evoked a smaller N400 when either "exploded" or "blew up" appeared in the previous sentence, than when a bridging inference had to be generated from information stored within semantic memory and integrated into the discourse model (e.g. "…the bomb hit the ground. The <u>explosion</u>…"). The authors concluded that the activation and integration of causal bridging inferences was facilitated by lexical priming across the two sentences. This study illustrates the contribution of lexico-semantic priming to the online establishment of causal relationships, as discussed above, but it does not speak to whether the situation model itself plays a role in establishing coherence online. In addition, as there was no inclusion of a condition that discouraged any inference activation entirely, it remains unclear how and when such new information was added to the discourse model during online processing.

Finally, a study by Burkhardt (2007) measured ERPs to words such as "pistol" in scenarios such as, "Yesterday a Ph.D. student was shot/killed/found dead/downtown. The press reported that the <u>pistol</u> was probably…". At the point of encountering "pistol", following "killed" and "found dead", a reader must integrate a causal inference (that someone must have been shot) into the discourse model in order to establish full discourse coherence. Interestingly, rather than evoking an effect within the N400 time window, "pistol" evoked a late positivity or P600 effect that showed an incremental increase across the three conditions with the largest positivity following "found dead". These findings were interpreted as suggesting that the addition of new information engaged later neurocognitive processes, distinct from the N400 (see also Burkhardt, 2006).

The current study used ERPs to examine the neural correlates of establishing causal coherence across three-sentence scenarios. Importantly, our design was parametric and based on the studies by Keenan et al. (1984) and Myers et al. (1990; 1987) that manipulated the degree of causal relationship across sentences, see Table 1. In highly causally related scenarios, information from world knowledge needed to establish causal relationships is easily accessible. For example, in the highly causally related scenarios in Table 1, to establish a causal relationship between the final sentence, “She had sunburn on Monday” and its preceding context, one would need access to real-world knowledge that the absence of sunscreen can lead to a sunburn if one has fair skin. In sentences that are causally completely unrelated to their context, such real-world knowledge is not readily available and comprehenders generally do not succeed in establishing causal coherence. In intermediately causally related scenarios, comprehenders will usually succeed in establishing causal coherence, but this requires the generation of a more complex causal inference that adds new information to the discourse model. For example, in the intermediately causally related scenarios in Table 1, to establish a causal relationship between the final sentence and its preceding context, one must not only gain access to real-world knowledge that the absence of sunscreen can lead to sunburn, but one must infer that, on this occasion, Jill forgot to apply the sunscreen to her skin.

We measured ERPs as readers processed these three types of three-sentence scenarios in which the final sentence unfolded word by word. Importantly, unlike the original stimuli used by Keenan et al. (1984) and Myers et al. (1990; 1987), we matched the lexico-semantic relationships between the individual content words across all three conditions using a Latent Semantic Analysis (Landauer & Dumais, 1997). This enabled us to examine the timing of establishing causal coherence at the level of the situation model, with any effects of lexico-semantic relationships being equated across the three conditions. We aimed to distinguish between several possibilities:

The first possibility was that close lexico-semantic relationships across individual words within the discourse context would dominate the semantic processing of an upcoming word between 300–500ms, overriding the situation model built up by the context. In other words, the processing of each incoming critical word would be shallow. This would predict no modulation of the N400 across the three conditions. A distinction between the three types of scenarios might, however, occur during a later stage of processing, as in a previous ERP study by Nieuwland and Van Berkum (2005) who reported no N400 effect, but a robust P600 effect to words that were highly incongruous with their discourse context but that were lexico-semantically associated with individual words in this context.

The second overall possibility was that the situation model constructed from the context, rather than simply lexico-semantic relationships within that context, would influence semantic processing of the upcoming word. On this account, critical words that were highly causally related to their context would be expected to evoke a smaller N400 than the causally unrelated critical words. This attenuation of the N400 would reflect the implicit and incremental use of real-world knowledge to facilitate processing of an incoming critical word (Hagoort et al., 2004) and would be consistent with some previous ERP studies indicating that words that are incongruous with their global discourse context evoke an N400 effect (van Berkum et al., 1999). Importantly, processing would not simply rely on matches between lexico-semantic relationships in the context and lexico-semantic associations stored within memory, but rather would be driven by an interaction between the situation model built up by the context and information stored within semantic memory (McKoon & Ratcliff, 1989; Sanford & Garrod, 2005).

Given the availability of real-world knowledge to facilitate the processing of highly causally related critical words, a key question was what would happen at the point of encountering critical words in the intermediately related scenarios that required the activation and integration of more complex inferences for the text to make sense. If the activation of such complex inferences followed an initial detection of a coherence break, then intermediately related critical words should generate just as large an N400 as the causally unrelated critical words; any inferencing and addition of new information to the discourse model that allowed the intermediately related sentences to eventually be perceived as coherent might be expected to engage a distinct, later neurocognitive process and to produce a P600 effect (Burkhardt, 2007). This would suggest that the coherence break itself triggers inferential processes that entail an active retrieval and/or selection of relevant information from long-term semantic memory.

If, on the other hand, a more complex causal inference was readily available before or at the time the incoming intermediately causally related critical word is processed, then these words should be just as easy to process as the highly related critical words, i.e. they should evoke an N400 with the same amplitude. On this account, even more complex inferencing that adds new information to the discourse model is assumed to be a relatively implicit process. Importantly, the N400 modulation on the critical word itself is assumed not to reflect the work of any inferential process, but rather to reflect the semantic processing *consequences* of earlier, more implicit inferencing (for discussion see Nieuwland, Ditman, & Kuperberg, Under review; van Berkum, In press).

Materials and Methods

Construction and development of stimuli

Two-hundred-and-thirty-one scenarios (sets of sentence triplets) were initially constructed, each with three conditions: highly causally related, intermediately related and causally unrelated. In the highly causally related scenarios, the final sentence was causally linked in meaning to the first two sentences through a simple inference requiring real-world knowledge. In the intermediately related scenarios, the same first and third sentences were used as in the highly related scenarios, but the second sentence was constructed so that the reader was required to make a more complex inference to connect the second and third sentences. The highly related and intermediately related scenarios were the same as those described in previous studies (Ditman & Kuperberg, 2007; Kuperberg, Lakshmanan, Caplan, & Holcomb, 2006). The causally unrelated scenarios were constructed by modifying either the highly related or the intermediately related scenarios so that the third sentence did not logically follow from the first two sentences. Example scenarios are presented in Table 1.

Attempts were made to match the content words across all three levels of Causal Relatedness in terms of numbers of word repetitions and their semantic similarity values (SSVs) as quantified using a Latent Semantic Analysis (LSA; Landauer & Dumais, 1997; Landauer, Foltz, & Dumais, 1998).[1] The scenario triplets were counterbalanced across three lists so that, in each list, the same final sentence was not seen more than once but, across lists, all final sentences were seen in all three conditions. The scenarios were then randomized within lists.

The stimulus set was then further constrained based on a norming study that aimed to select those scenario triplets in which the three experimental conditions were most clearly distinguished in terms of their causal relatedness. Scenarios were presented to twelve undergraduate students from Tufts University (7 female, 5 male; mean age: 21.1, SD: 2.5), four for each of the three lists, who did not participate in the ERP study. The participants were asked to rate the causal connection between the third sentence in each scenario and the first two sentences on a scale from 1 (weak causal relationship) to 7 (strong causal

[1]Note that the causally unrelated scenarios used in the present study differed from the causally unrelated scenarios that we have used in the previous studies (Ditman & Kuperberg, 2007; Kuperberg et al., 2006) where LSA values of content words within these unrelated scenarios were not matched with those of the highly and intermediately related scenarios.

relationship). Scenario triplets were then eliminated if, within each triplet, any of the following criteria were met: (1) the average rating of the highly related scenario was lower then that of either the intermediately related or the causally unrelated scenario; (2) the average rating of the intermediately related scenario was the same or lower than that of the causally unrelated scenario; (3) the highly related scenario was rated below 4; (4) the intermediately related scenario was rated below 3; (5) the causally unrelated scenario was rated above 3.

After excluding scenario triplets based on these criteria, 159 scenario triplets remained. An additional pretest was carried out to verify that participants directly established a causal connection in the highly related scenarios, generated a more complex inference to the intermediately related scenarios, and failed to generate consistent inferences to the unrelated scenarios. The discourse scenarios were presented in random order to 12 Tufts undergraduate students (four for each of the three lists). At the end of each discourse scenario a “Why” question was presented, in italics. This question was constructed from the final sentence in each scenario, e.g. in the example given in Table 1, “Why did Jill have sunburn on Monday?”. Participants were asked to write a one-sentence clear answer to the question or to indicate ‘don’t know’ if there was no obvious answer. Inspection of answers indicated that, for 98% of the highly related scenarios, participants wrote responses that were very similar to the second sentence for that scenario, i.e. they repeated what they had just read. For 81% of the intermediately related scenarios, participants wrote responses that were very similar to the second sentence of the highly related condition for that scenario, even though they had not seen that sentence, i.e. they made the expected inference. For 92% of the causally unrelated scenarios, participants either indicated ‘don’t know’ or wrote responses that were very different from the second sentence of condition 1 for that scenario, i.e. they either failed to make an inference or any inferences generated were inconsistent across participants).

In all these scenarios, the first two sentences each contained between 4 and 10 words and the final sentence contained between 2 and 7 words. Critical words in the third sentence of each scenario were selected on the basis of a separate causal rating study that has previously been described (Pretest 2 ratings in Kuperberg et al., 2006). In 77 scenario triplets, the critical word was followed by one to three additional words. These fell at clause boundaries but were not accompanied by a comma. In 82 scenario triplets the critical word was the final word of the third sentence and this appeared with a period. Thus, across all participants, the same mid-sentence or sentence-final critical word appeared in all three conditions. However, across all participants, the same words did not appear in the mid-sentence position and sentence-final position.

Semantic similarity values (SSVs) were calculated using Latent Semantic Analysis (LSA; Landauer & Dumais, 1997; Landauer et al., 1998), available on the internet at http://lsa.colorado.edu. Pairwise comparisons using tasaALL space (1^{st} year college student reading level) were carried out comparing all content words in a scenario and the critical word, see Table 1. An ANOVA including all items revealed no significant differences in SSV between the three levels of Causal Relatedness, $F(2,471) = 0.140$, $p = 0.87$ and no interaction between Causal Relatedness and Position of the critical word, $F(2, 471) = 0.008$, $p = 0.99$. A main effect of Position in this ANOVA reflected the increased context (more words) preceding sentence-final critical words relative to mid-sentence critical words, $F(1, 471) = 14.4$, $p < 0.001$. In addition, an ANOVA including all items confirmed significant differences in causal ratings across the three experimental conditions, $F(2,471) = 887.8$, $p < 0.001$, with no interactions between Causal Relatedness and Position, $F(2,471) = 0.2$, $p = 0.81$, and no main effects of Position, $F(2,471) = 0.4$, $p < 0.52$, see Table 1 for means and standard deviations. Follow-ups confirmed that the mean rating of the highly related scenarios was significantly greater than that of the intermediately related scenarios, $t(316) = 16.5$, $p < 0.0001$, that was, in turn, significantly greater than that of the causally unrelated scenarios, $t(316) = 23.1$, $p < 0.0001$, see Table 1.

The final set of 159 scenario triplets was then divided into three lists so that, in each list, the same final sentence was not seen more than once but, across lists, all final sentences were seen in all three conditions. Each list thus contained 159 three-sentence scenarios, with 53 of each of the three conditions. In approximately 50% of these scenarios in each list, the critical word appeared at the sentence-final position and in approximately 50% the critical word appeared before the sentence-final position. Thus, across participants, the same mid-sentence or sentence-final critical word appeared in all three conditions. However, across all participants, the same words did not appear in the mid-sentence position and sentence-final position. Within each list, scenarios were pseudo-randomized within each list such that the same condition did not appear in more than two consecutive trials.

ERP Study

Participants—Twenty-one participants (13 female, 8 male; mean age: 20.9 years, *SD:* 1.8) were recruited from the Tufts University undergraduate population by advertising. All participants were right handed, native English speakers (no other language learned before the age of 5) who had normal or corrected-to-normal vision, were not taking any psychoactive medication, and had no history of head trauma. Written informed consent was obtained from all participants prior to participation.

Stimulus presentation and task—Participants' task was to rate how easy or difficult it was to connect the third sentence to the first two sentences in each scenario by pressing one of three buttons: easy to connect, in-between, difficult to connect. Participants were told that there were no 'right' answers to their ratings and to go with their first instincts. They were given nine practice trials after which they were given the opportunity to ask for clarification prior to beginning the main portion of the experiment.

Each participant was assigned to one of the three counterbalanced lists and was seated in a comfortable chair in a dimly lit, sound-attenuated room, separate from the experimenter. Scenarios were presented on a computer monitor using white text on a black background. In each trial, the word READY was displayed until the participant pressed any key on a response box to begin the experiment. After a fixation cross was displayed in the center of the screen for 500ms, the first two sentences were presented in their entirety for 3.4 seconds each. The third sentence was then presented word by word with each word displayed for 500ms with 100ms interstimulus interval (ISI). The mid-sentence critical words never appeared with a comma. The sentence-final critical words always appeared with a period. After each sentence, a question mark cue was displayed until the participant made his/her judgment, after which the next trial began. Participants were instructed to wait until the question mark cue before responding. In addition, at five random points in the middle of the experiment as well as after the last trial, participants were asked a comprehension question about the content of the scenario they had just read.

Electrophysiological Recording—Activity was recorded at 29 active tin electrodes placed on the scalp and held in place on the scalp by an elastic cap (Electro-Cap International, Inc., Eaton, OH), see Figure 1 for the full montage. In addition, electrodes were placed below the left eye and at the outer canthus of the right eye to monitor vertical and horizontal eye movements, and also over the left mastoid and right mastoid (recorded actively to monitor for differential mastoid activity). All EEG electrode impedances were maintained below 5 kΩ (impedance for eye electrodes was less than 10 kΩ).

The EEG signal was amplified by an Isolated Bioelectric Amplifier System Model HandW-32/BA (SA Instrumentation Co., San Diego, CA) with a bandpass of 0.01 to 40 Hz and was continuously sampled at 200 Hz by an analogue-to-digital converter. The stimuli and participants' behavioral responses were simultaneously monitored by a digitizing computer.

ERP Analysis—Analyses were first carried out on waveforms evoked by critical words in all trials, binned by the three levels of Causal Relatedness. Only trials free of ocular and muscular artifact were averaged for analysis. Analyses were conducted on mean amplitude values 325–475ms, which encompassed the N400 effect, and 500–800ms following word onset, using the 100ms of activity that preceded critical word onset as a baseline.

All sites were included in a systematic columnar "pattern of analyses" that we have applied in prior studies, (e.g., Ditman et al., 2008; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007). This approach has the advantage of yielding statistical information about the distribution of ERP effects across the whole scalp, allowing the detection of differences in the distribution of effects along the anterior-posterior (AP) axis of the scalp. In addition, at medial, lateral and peripheral electrode columns, it allows the detection of differences across the two hemispheres.

Repeated measures analyses of variance (ANOVAs) at each time window were performed at the midline, medial, lateral, and peripheral columns (see Figure 1). As we had a priori hypotheses, we proceeded straight to ANOVAs which contrasted each level of Causal Relatedness with one another (causally unrelated vs. highly related, causally unrelated vs. intermediate related, intermediately related vs. highly related). In addition to Causal Relatedness, each ANOVA included Position (mid-sentence critical word, sentence-final critical word) and AP Distribution (with the number of levels depending on the number of electrode sites in each column, see Figure 1) as within-participant factors. The three lateral electrode columns also included Hemisphere (left, right) as a within-participant factor. Follow-ups of interactions with Position were carried out by comparing Causal Relatedness for a given critical word position, rather than comparing effects of Position for given level of Causal Relatedness. The Geisser-Greenhouse correction (Greenhouse & Geisser, 1959) was applied when evaluating effects with more than one degree of freedom to protect against Type 1 errors resulting from violations of sphericity. In these cases we report original degrees of freedom and the corrected probability level.

Behavioral Analysis—Means for the 1 through 3 classifications (1: difficult to connect; 2: in between; 3: easy to connect) given to each of the three pre-determined levels of Causal Relatedness were contrasted using a repeated-measures ANOVA in which Causal Relatedness and Position were within-participant factors.

Results

ERP data (see Figure 2)

325–475ms

Causally unrelated versus highly causally related: Critical words that were causally unrelated to the preceding context evoked a larger amplitude N400 than words that were highly causally related to their preceding context (main effect of Causal Relatedness: midline: $F(1, 20) = 12.05$, $p < .01$; medial: $F(1,20) = 10.22$, $p < .01$; lateral: $F(1, 20) = 12.00$, $p < .01$; peripheral: $F(1, 20) = 13.29$, $p < .01$). In addition, a Causal Relatedness × Position × Hemisphere interaction was observed at all non-midline columns (medial: $F (1, 20) = 11.15$, $p < .01$; lateral: $F (1, 20) = 12.39$, $p < .01$; peripheral: $F (1, 20) = 4.61$, $p < .05$). Follow-up simple effects ANOVAs at these columns revealed that the difference between causally unrelated and highly related critical words was generally left lateralized for mid-sentence critical words (main effects of Causal Relatedness, significant at left medial, left lateral, left peripheral, and right peripheral columns: all $Fs > 5.10$, all $ps < .05$), but broadly distributed across both hemispheres for sentence-final critical words (main effect of Causal Relatedness significant at all columns, all $Fs > 4.60$, all $ps < .05$). There were no other interactions with Causal Relatedness and Position (all $ps > .10$).[2]

Causally unrelated versus Intermediately related: Critical words in causally unrelated scenarios evoked a larger N400 than in the intermediately related scenarios, as evidenced by significant main effects of Causal Relatedness at midline and medial columns and marginal effects at lateral and peripheral columns (midline: $F(1, 20) = 4.72$, $p < 0.05$; medial: $F(1, 20) = 5.92$, $p < 0.05$; lateral: $F(1, 20) = 3.27$, $p < 0.10$; peripheral: $F(1, 20) = 3.04$, $p < 0.10$). No interactions with Position reached significance (all $Fs < 2.60$, all $ps > .06$).

Intermediately related versus Highly related: Unlike the contrast between the causally unrelated and highly related critical words, the N400 effect to the critical words to the intermediately related contrast was modulated by sentence position at midline sites, as reflected by a Causal Relatedness × Position × AP Distribution interaction at the midline column ($F (4, 80) = 3.60$, $p < .05$). Follow-up simple effects ANOVAs indicated that, when critical words appeared mid-sentence, the N400 to the intermediately related critical words was more negative than to the highly related words. This effect was greatest at more anterior than posterior sites, as reflected by a Causal Relatedness × AP Distribution: $F(4, 80) = 3.51$, $p < .05$. However, when critical words appeared at the sentence-final position, there were no differences between these two conditions at any electrode site (main effect of Causal Relatedness: $F < 1.00$, $p > .70$; Causal Relatedness × AP Distribution: $F < 1.00$, $p > .40$).

At non-midline columns, there were no differences in N400 evoked by intermediately and highly related (no main effects of Causal Relatedness) except at left lateralized sites where intermediately related critical words appeared to evoke a larger negativity than highly related critical words regardless of sentence position (Causal Relatedness × Hemisphere interaction at the medial column ($F(1, 20) = 5.08$, $p < .05$; marginal main effect of Causal Relatedness at the lateral ($F(1,20) = 4.14$, $p < .06$) and peripheral ($F(1,20) = 3.53$, $p < .08$) columns).

500–800ms

Causally unrelated versus Highly related: When they occurred in the mid-sentence position, critical words in the causally unrelated scenarios continued to evoke a larger negativity than highly related critical words in this time window, particularly at non-midline sites and particularly over the left hemisphere. This was reflected by significant interactions between Causal Relatedness, Hemisphere, and Position at the medial and lateral columns (medial: $F(1, 20) = 6.15$, $p < 0.05$; lateral: $F(1,20) = 6.01$, $p < 0.05$, with follow-up ANOVAs at these columns showing a significant or near-significant Causal Relatedness × Hemisphere effect only for the mid-sentence critical words, medial: $F(1, 20) = 4.65$, $p < .05$; lateral: $F(1, 20) = 3.89$, $p < 0.07$). There was also a significant main effect of Causal Relatedness at the peripheral column ($F(1, 20) = 4.78$, $p < 0.05$).

[2]For scenarios where the critical word appeared before the sentence-final position, we also examined ERPs at the sentence-final word and found no significant main effects of Causal Relatedness or interactions between Causal Relatedness and AP distribution between these two conditions (or between any other two conditions) in this time-window.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Causally unrelated versus Intermediately related: There were no significant main effects or interactions with Causal Relatedness (all Fs < 3.70, all ps > 0.06).

Intermediately related versus Highly related: There were no significant main effects or interactions with Causal Relatedness (Fs < 2.50, all ps > 0.12). There was a Causal Relatedness × Hemisphere interaction in the medial column ($F(1, 20) = 5.40$, $p < 0.05$) but analyses at left and right medial columns separately failed to reveal significant effects, ps > 0.5.

Behavioral Data

The mean classifications for each level of Causal Relatedness are given in Table 2. As expected, a repeated measures 3 (Causal Relatedness: highly related, intermediately related, causally unrelated) by 2 (Position: sentence-final critical word, mid-sentence critical word) ANOVA examining the differences in these mean ratings showed a significant main effect of Causal Relatedness ($F(2, 40) = 943.21$, $p < .001$). Pairwise comparisons indicated that the highly related scenarios were classified as more related than the intermediately related scenarios ($F(1, 20) = 280.97$, $p < .001$) that were, in turn, classified as more related than the causally unrelated scenarios ($F(1, 20) = 1005.94$, $p < .001$). There was no interaction between Causal Relatedness and Position ($F(2, 40) < 1$).

Discussion

The present study examined neural activity associated with establishing causal relationships across sentences during online comprehension. Despite being matched on lexical semantic relationships, neural modulation to critical words differentiated between the highly causally related, intermediately related and causally unrelated scenarios. Critical words in the causally unrelated scenarios evoked a larger widely distributed N400 than those in both the highly related and intermediately related scenarios. This was the case regardless of whether the critical word appeared before or at the sentence-final position. At many electrode sites, the amplitude of the N400 to critical words in the intermediately related scenarios fell in between that evoked by the highly causally related and intermediately related words with one exception: the N400 evoked by intermediately related sentence-final critical words at midline sites, was attenuated to the same degree as to the highly causally related words. No LPC/P600 effect was observed to either the causally unrelated or the intermediately related, relative to the highly related, critical words.

Overall, these findings demonstrate that causal coherence across sentences, at the situation level, can immediately influence semantic processing of incoming words, at least when participants are asked to actively judge causal relationships between sentences. Below we discuss the specific implications of these patterns of neural modulation.

N400 modulation to critical words in the causally unrelated versus the highly related scenarios

The larger N400 to critical words in the causally unrelated relative to highly related scenarios is consistent with previous studies demonstrating that readers are immediately sensitive to coherence breaks during comprehension, even when lexico-semantic relationships between individual words are close. For example, Van Berkum et al. (1999) showed that single words that were incongruous with their entire discourse context evoked a larger N400 than words that were discourse congruous, even when these words were consistent with their immediate sentence contexts. In addition, Camblin, Gordon, and Swaab (2007) demonstrated a clear N400 effect to discourse incongruous relative to congruous words that were related to their preceding content words (e.g. to “legs” in the scenario

NIH-PA Author Manuscript

NIH-PA Author Manuscript

"Lynn's wool sweater was uncomfortable and itchy. She fidgeted as the rough material irritated her skin. Lynn couldn't stop scratching her arms and legs.").[3]

In these previous studies, a variety of different types of discourse incongruities were introduced, the critical word always fell at the sentence-final position, and contextual constraint was fairly high. In the present study, congruity was always mediated by causal relatedness and in nearly 50% scenarios, the critical word did not appear with a period, falling before the end of the sentence. Thus, these findings clearly indicate that, at least when explicitly asked to make coherence links, readers' build-up of *causal* coherence is established at the situation level and influences the earliest stages of semantically processing incoming words. Moreover, they show that such coherence is built even when critical words appeared before the sentence-final word (although, in all cases, words appeared at clause boundaries). Indeed, when critical words occurred before the sentence-final position, the modulation of the N400 extended into the later 500–800ms window, suggesting that readers may have continued to establish causal coherence as more information became available.

We interpret the attenuation of the N400 to the highly related, relative to the causally unrelated, critical words as reflecting facilitation due to a match between the meaning of the incoming lexical item and an interaction between the context and the reader's stored semantic and real-world knowledge (Hagoort et al., 2004; Kutas & Federmeier, 2000; Kutas et al., 2006).

Modulation of the ERPs to critical words in the intermediately related scenarios

Of most interest was how ERPs were modulated to intermediately related critical words, relative to highly causally related and causally unrelated words. Such intermediately related critical words were congruous with their context, but required the activation and integration of complex causal inferences (Keenan et al., 1984; Myers & Duffy, 1990; Myers et al., 1987). Once again, we observed modulation of the N400 component, but no modulation of the later P600. These findings are consistent with previous studies examining ERP correlates of establishing causal coherence (St George et al., 1997; Yang et al., 2007). They do, however, contrast with those of Burkhardt (2007) who found that scenarios requiring complex causal bridging inferences were associated with P600 modulation, interpreting this as reflecting the cost of adding new information to the discourse model. Of note, however, at least some of the scenarios employed in Burkhardt's study may have required participants not only to generate a causal inference but also to temporally reorder the described events, potentially leading to later reanalysis processes. Such temporal reordering is known to engender processing costs (Claus & Kelter, 2006; Raisig, Welke, Hagendorf, & van der Meer, 2007) and may have been a factor driving the modulation of the P600 rather than causal inferencing per se. In order to further examine this hypothesis, we are currently manipulating causal relatedness and temporal order within the same paradigm.

In the Introduction we outlined two possibilities for how the N400 might be modulated in intermediately related (vs. highly causally related and unrelated) scenarios. If no causal inference was activated by the time the critical word was presented, then the intermediately related critical words should evoke an N400 that was equal in amplitude to that of the causally unrelated words. If, however, a full inference was activated, then the intermediately related critical words should generate a smaller N400 that was equal to that of the highly related critical words. As noted in the Introduction, both these predictions hold that the process of activating new information from semantic memory (i.e. the initial stage of

---

[3]In Camblin et al. (2007) associated incongruous words evoked a larger anterior positivity and larger posterior negativity than non-associated incongruous words within a later time window.

inference generation) is relatively automatic and implicit (McKoon & Ratcliff, 1989; van den Broek, 1994), and that the amplitude of the N400 to an incoming lexical item reflects the *consequence* of such activation rather than the process of inference generation itself (see Nieuwland et al., Under review; van Berkum, In press).

What we observed was an N400 to the intermediately related critical words that was consistently smaller than that to the causally unrelated words, suggesting that at least some inferencing had taken place to facilitate their processing. This finding is consistent with the early study by St. George et al. (1997) who also contrasted incongruous scenarios and scenarios that encouraged the activation of a bridging inference. In that study, however, the inference word itself was explicitly presented in a final probe sentence and the authors' measure of inferencing was the N400 averaged across all words in this final sentence. The current findings suggest that any inferencing served to facilitate the build-up of causal coherence during incremental, word-by-word processing.

At midline electrode sites, intermediately related critical words that appeared in the sentence-final position evoked an N400 that was attenuated to the same degree as to the highly related critical words. This suggests that the end of the sentence itself (indicated with by a period) may have acted as an explicit cue to encourage the establishment of a full causal inference. When the intermediately related critical words appeared before the sentence-final position, even though they were at clause boundaries, the N400 was generally not attenuated to the same degree as to the highly related critical words.[4] This partial attenuation of the N400 is consistent with the idea that causal inferencing is unlikely to be an all-or-none process (see Just & Carpenter, 1992; Lassonde & O'Brien, 2009; McKoon & Ratcliff, 1990; Sanford & Garrod, 2005; van den Broek, 1994). In other words, rather than constituting the activation/access of a specific inference (a specific lexical item), as may occur in highly-constrained contexts, inferential activation processes may be more general, facilitating the activation and processing of multiple lexical items sharing related features (Cook et al., 2001; Klin et al., 1999; Sanford & Garrod, 2005; van den Broek, 1994). (Note that the N400 evoked by a discourse-incongruous word is modulated by the degree to which this word shares semantic features with an expected word (Ditman et al., 2007; Federmeier & Kutas, 1999; Kutas & Hillyard, 1984)).

Relationship between findings and models of discourse processing

Discourse processing models fall into two major classes. Memory-based models place most emphasis on the implicit activation of information from long-term memory during the build-up of coherence (Kintsch, 1988; McKoon & Ratcliff, 1992; Myers & O'Brien, 1998; Sanford & Garrod, 1998). For example, according to the resonance model (Albrecht & O'Brien, 1993; Myers & O'Brien, 1998), incoming material continually and autonomously interacts with stored information and 'resonates' as a function of the degree of match. Similar ideas that fast 'pattern matching' facilitates the processing of upcoming material are presented in Kintsch's (1988) Construction-Integration model, in Sanford and Garrod's scenario-mapping and focus theory (Sanford, 1990; Sanford & Garrod, 1981, 1998), and in McKoon and Ratcliff's (1992) minimalist hypothesis. These memory-based models emphasize that implicit resonance with incoming material is often sufficient for inferencing. They have been relatively silent on how such material is subsequently selected or integrated into the discourse structure. Constructionalist models, on the other hand, have emphasized the idea that comprehenders actively engage inferential processes in a 'search for meaning' (Graesser, Singer, & Trabasso, 1994). These models acknowledge a phase of initial pattern

[4]Interestingly, the distribution of the effect to the intermediately-related vs. highly related scenarios was more anterior than the classic N400 effect. It is possible that this was because it reflected, in part, a prefrontally-mediated more directed search that activated relevant information, although this is speculative.

matching but place more emphasis on what information is selected to build up coherence during the integration phase of processing (Long & Lea, 2005). These two classes of models are distinguished mainly by their different emphases on the two stages of coherence building (Long & Lea, 2005), and, indeed, aspects of both models have been incorporated into hybrid models of text comprehension such as the Landscape model (Rapp & van den Broek, 2005; van den Broek, Rapp, & Kendeou, 2005).

The N400 is believed to be a memory-based ERP component: it is highly sensitive to information stored within semantic memory at various grains of representation, which impacts directly on both sentence and discourse comprehension (Kutas & Federmeier, 2000; Kutas et al., 2006; van Berkum, In press). During sentence processing, the N400 is not necessarily sensitive to the demands of selecting words from competing activated alternatives: for example, the amplitude of the N400 to unexpected words in highly constrained contexts (where selection demands are relatively high) is the same as to unexpected words in less constrained contexts (Federmeier et al., 2007). Late positivites, on the other hand, may better reflect selection processes that resolve conflict between competing lexical items (Federmeier et al., 2007, where the positivity has a more frontal distribution), or between competing sentence-level representations (Kolk & Chwilla, 2007; Kuperberg, 2007, where the positivity has a more posterior distribution). In the current dataset, ERP modulation was observed in the N400 component and not the late positivity/ P600. We interpret this as broadly supporting memory-based models of discourse processing.

On the other hand, as they stand, memory-based models cannot easily account for the pattern of N400 modulation observed across the three levels of causal relatedness in this study. In most such models, resonance is conceptualized as a function of the match between lexico-semantic relationships within the input and lexico-semantic relationships stored within long-term memory. Lexico-semantic relationships have been viewed as argument overlap between propositions (Kintsch & van Dijk, 1978), matches between semantic and contextual features (Cook et al., 1998; Myers & O'Brien, 1998) and, more recently, in terms of lexico-semantic co-occurrence, as indexed using the Latent Semantic Analysis (Kintsch, 2001; Landauer & Dumais, 1997; Landauer et al., 1998). In many situations, lexico-semantic relationships play an important role in establishing coherence: they influence ratings of text coherence, processing times of sentence pairs, and free recall of natural texts (Foltz, Kintsch, & Landauer, 1998; Wolfe, Magliano, & Larsen, 2005). In ERP studies of coherent discourse, words that are semantically related to the set of individual words in their context evoke a smaller negativity than words that are not semantically related to their context (Ditman et al., 2007; Otten & van Berkum, 2007). And, in a recent study, Yang et al. (2007) showed that lexico-semantic relationships directly influence the amplitude of the N400 during the establishment of causal coherence across clause boundaries. In the present study, however, simple lexico-semantic co-occurrence, as indexed using LSA, *was* matched across the three types of scenario. This suggests that any resonance between the input and stored information that influenced the immediate build-up of causal coherence also occurred at the situation level. Exactly what features resonated at this level, however, is unclear. We know from previous ERP studies that dimensions of the situation model including temporal variables (Ditman et al., 2008) and their linguistic codes (Ferretti, Kutas, & McRae, 2007) can immediately impact the N400. Pragmatic knowledge and relationships also play an important role (Nieuwland et al., Under review; Nieuwland & Kuperberg, 2008; van Berkum, In press). Regardless of what the precise driving factors were in the current study, the implication is that long-term memory is multilayered, with information represented at different levels, and that resonance can occur at any or all these levels.

Open questions

If resonance does indeed occur at multiple levels of semantic representation, this raises the question of how activity at each of these levels interacts during discourse processing. Under what situations do lexico-semantic relationships in the text-base override the situation model, and vice versa? In the current study, we showed that the situation model can influence the earliest stages semantically processing incoming words. However, it is clear from other studies that there are situations in which close lexico-semantic relationships within a context override situation-level meaning, leading either to a shallow final interpretation (a semantic illusion; Erickson & Mattson, 1981; Ferreira, 2003; Kamas et al., 1996; Sanford, 2002) or to no early N400 modulation, but later processing costs within the late positivity/P600 time window (Nieuwland & van Berkum, 2005; see also Hoeks, Stow, & Doedens, 2004; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). The interplay between these levels of representation may be influenced by several factors including the rate of language presentation (Camblin, Ledoux, Boudewyn, Gordon, & Swaab, 2007), linguistic focus (Sturt, Sanford, Stewart, & Dawydiak, 2004), the presence of explicit coherence markers (connectives) in the text (Millis & Just, 1994), and an individual's working memory capacity (Nakano, Saron, & Swaab, In press). In the present study, participants task was to explicitly judge causal coherence, i.e. they had a relatively high standard of coherence (a factor that has been discussed more in relation to traditional constructionalist frameworks than memory models; van den Broek, Risden, & Husebye-Hartman, 1995). In addition, the presentation rate was relatively slow (the SOA was 600ms) and participants had a long time to read the first two sentences (3.4 sentences each). It may well be that, if participants were asked to read these types of scenarios more passively, and/or the presentation of individual sentences and words was faster (Camblin, Ledoux et al., 2007), resonance would have occurred mainly at the level of lexico-semantic relationships, with the effects of the situation model influencing later stages of selection or integration, in keeping with constructionalist models. These hypotheses will be tested in future studies. What the current study does show is that, at least when comprehenders are focusing on causal relationships, they *can* activate the necessary information for complex causal inferencing fast enough to influence the earliest stages of lexico-semantic processing, and that this cannot be accounted for by lexico-semantic relationships alone.

A second open question is precisely *when* did causal inferencing take place in the present study. One possibility is that the implicit retrieval of relevant stored information occurred *before* the onset of the critical word (a true predictive inference; Cook et al., 2001; Klin et al., 1999), and that this led to 'preactivation' of its semantic features, facilitating semantic access. This may have been particularly likely in the present study as participants were given a fairly long time to read the first two sentences which may have encouraged them to predict the content of the third sentence. Another possibility is that the retrieval of this information was initiated only once the critical word was presented (a bridging inference; Singer & Ferreira, 1983). This experiment cannot distinguish these possibilities because it remains unclear whether the amplitude of N400 itself always reflects the consequences of truly predictive semantic processing (Federmeier, 2007) or whether it reflects a three-way mapping between context, stored information and lexico-semantic features of an incoming word that is initiated only once that word is presented (Hagoort, 2005; Holcomb, 1993). In other types of paradigms using highly constrained sentence and discourse contexts, however, there is evidence that the N400 can be modulated prior to the onset of a critical word by semantic predictions of that word (DeLong, Urbach, & Kutas, 2005; Federmeier, 2007; van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005). Future studies will determine whether this is also the case during the establishment of causal coherence when the context is less highly constrained.

A third set of questions asks how the current pattern of ERP data can be reconciled with the non-linear pattern of memory retrieval originally described by (Keenan et al., 1990; Myers et al., 1987), as well as the non-linear pattern of hemodynamic response reported in fMRI studies (Kuperberg et al., 2006; Mason & Just, 2004) across similar types of three-sentence scenarios. As discussed in the Introduction, participants' recall of intermediately related scenarios is superior than their recall of both highly causally related and unrelated scenarios (Keenan et al., 1990; Myers et al., 1987). And, in an fMRI study using similar stimuli to those used here, we reported that participants took longest to judge the overall coherence of intermediately related scenarios, and recruited a large network of regions distributed across right and left temporal-prefrontal cortices to these scenarios (Kuperberg et al., 2006).[5] Although the reasons for this discrepancy between the ERP and fMRI findings are unclear, one possibility is that these two techniques each indexed different stages of neurocognitive processing. ERPs measure neural response to individual events, and they therefore provided a sensitive measure of fast, online processes that were time-locked to the onset of the critical words. In contrast, fMRI indexes hemodynamic activity spread out over many words and during decision-making, and may have better captured more 'offline' processes of evaluating the full meaning of the intermediately related scenarios (including the specific inferences added to the discourse model). The increased neural activity and reaction times associated with evaluating the meaning of the intermediately related scenarios may, in turn, have reflected their active consolidation within episodic memory, leading to their improved recall on later memory testing.

Conclusions

In sum, these results demonstrate unambiguously that causal coherence, driven by relationships at the situation level, can influence incremental word-by-word discourse comprehension, even when semantic relationships between individual words are matched and even when complex inferences are required to link intermediately related sentences. They further demonstrate that, at least when comprehenders are actively engaged in building causal coherence, inference processes that add new information to the discourse model can impact at the earliest stages of lexico-semantic processing.

Acknowledgments

This work was funded by National Institute of Mental Health (RO1 MH02034 to GRK), and the National Alliance for Research in Schizophrenia and Depression (NARSAD) with the Sidney J. Baer Trust (GRK). We thank Mante Nieuwland and Ellen Lau for their insightful comments on the manuscript. We are also grateful to Abigail Swain for her assistance in analyzing data and to Eric Fields for his help with manuscript preparation.

References

Albrecht JE, O'Brien EJ. Updating a mental model: Maintaining both local and global coherence. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1993; 19(5):1061–1070.

Altmann G, Steedman M. Interaction with context during human sentence processing. Cognition. 1988; 30(3):191–238. [PubMed: 3215002]

Bentin S, McCarthy G, Wood CC. Event-related potentials, lexical decision and semantic priming. Electroencephalography and Clinical Neurophysiology. 1985; 60:343–355. [PubMed: 2579801]

Bloom CP, Fletcher CR, van den Broek P, Reitz L, Shapiro BP. An on-line assessment of causal reasoning during comprehension. Memory & Cognition. 1990; 18(1):65–71.

[5]Mason & Just (2004) also showed increases in the hemodynamic response to intermediately related scenarios than highly related or causally unrelated scenarios but only within the right hemisphere. However, as discussed by Kuperberg et al. (2006), there were several methodological factors that may have limited power to detect more widespread activation in this study.

Burkhardt P. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. Brain and Language. 2006; 98(2):159–168. [PubMed: 16725188]

Burkhardt P. The P600 reflects cost of new information in discourse memory. Neuroreport. 2007; 18(17):1851–1854. [PubMed: 18090325]

Camblin CC, Gordon PC, Swaab TY. The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. Journal of Memory and Language. 2007; 56(1):103–128. [PubMed: 17218992]

Camblin CC, Ledoux K, Boudewyn M, Gordon PC, Swaab TY. Processing new and repeated names: Effects of coreference on repetition priming with speech and fast RSVP. Brain Research. 2007; 1146:172–184. [PubMed: 16904078]

Chwilla DJ, Kolk HHJ, Vissers CTWM. Immediate integration of novel meanings: N400 support for an embodied view of cognition. Brain Research. 2007; 1183:109–123. [PubMed: 17950260]

Claus B, Kelter S. Comprehending narratives containing flashbacks: evidence for temporally organized representations. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2006; 32(5):1031–1044.

Cook AE, Halleran JG, O'Brien EJ. What is readily available during reading? A memory-based view of text processing. Discourse Processes. 1998; 26(2&3):109–129.

Cook AE, Limber JE, O'Brien EJ. Situation-based context and the availability of predictive inferences. Journal of Memory and Language. 2001; 44:220–234.

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

Coulson S, Kutas M. Getting it: human event-related brain response to jokes in good and poor comprehenders. Neuroscience Letters. 2001; 316(2):71–74. [PubMed: 11742718]

Coulson S, Van Petten C. Conceptual integration and metaphor: an event-related potential study. Memory & Cognition. 2002; 30(6):958–968.

De Grauwe, S.; Swain, A.; Holcomb, PJ.; Ditman, T.; Kuperberg, GR. Electrophysiological insights into the processing of nominal metaphors. (Under review)

DeLong KA, Urbach TP, Kutas M. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nature Neuroscience. 2005; 8(8):1117–1121.

Ditman T, Holcomb PJ, Kuperberg GR. The contributions of lexico-semantic and discourse information to the resolution of ambiguous categorical anaphors. Language and Cognitive Processes. 2007; 22:793–827.

Ditman T, Holcomb PJ, Kuperberg GR. Time travel through language: Temporal shifts rapidly decrease information accessibility during reading. Psychonomic Bulletin & Review. 2008; 15:750–756. [PubMed: 18792500]

Ditman T, Kuperberg GR. The time course of building discourse coherence in schizophrenia: An ERP investigation Psychophysiology. 2007; 44:991–1001.

Duffy SA, Henderson JM, Morris RK. Semantic facilitation of lexical access during sentence processing. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1989; 15(5): 791–801.

Erickson TD, Mattson ME. From words to meaning: A semantic illusion. Journal of Verbal Learning and Verbal Behavior. 1981; 20:540–551.

Federmeier KD. Thinking ahead: The role and roots of prediction in language comprehension. Psychophysiology. 2007; 44(4):491–505. [PubMed: 17521377]

Federmeier KD, Kutas M. A rose by any other name: Long-term memory structure and sentence processing. Journal of Memory and Language. 1999; 41:469–495.

Federmeier KD, Wlotko EW, De Ochoa-Dewald E, Kutas M. Multiple effects of sentential constraint on word processing. Brain Research. 2007; 1146:75–84. [PubMed: 16901469]

Ferreira F. The misinterpretation of noncanonical sentences. Cognitive Psychology. 2003; 47:164–203. [PubMed: 12948517]

Ferretti TR, Kutas M, McRae K. Verb aspect and the activation of event knowledge. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2007; 33(1):182–196.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Fletcher CR, Bloom CP. Causal reasoning in the comprehension of simple narrative texts. Journal of Memory and Language. 1988; 27(27):235–244.

Fletcher CR, Hummel JE, Marsolek CJ. Causality and the allocation of attention during comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1990; 16(2):233–240.

Foltz PW, Kintsch W, Landauer TK. The measurement of textual coherence with Latent Semantic Analysis. Discourse Processes. 1998; 25:285–307.

Garrod S, Terras M. The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. Journal of Memory and Language. 2000; 42:526–544.

Graesser, AC.; Clark, LF. Structures and procedures of implicit knowledge. Norwood, NJ: Ablex; 1985.

Graesser AC, Singer M, Trabasso T. Constructing inferences during narrative text comprehension. Psychological Review. 1994; 101(3):371–395. [PubMed: 7938337]

Greenhouse S, Geisser S. On methods in the analysis of profile data. Psychometrika. 1959; 24(2):95–112.

Hagoort P. Impairments of lexical-semantic processing in aphasia: evidence from the processing of lexical ambiguities. Brain and Language. 1993; 45(2):189–232. [PubMed: 8358597]

Hagoort P. On Broca, brain, and binding: a new framework. Trends in Cognitive Sciences. 2005; 9(9): 416–423. [PubMed: 16054419]

Hagoort P, Hald L, Bastiaansen M, Petersson KM. Integration of word meaning and world knowledge in language comprehension. Science. 2004; 304(5669):438–441. [PubMed: 15031438]

Haviland SE, Clark HH. What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Behavior. 1974; 13:512–521.

Hoeks JCJ, Stowe LA, Doedens G. Seeing words in context: The interaction of lexical and sentence level information during reading. Cognitive Brain Research. 2004; 19:59–73. [PubMed: 14972359]

Holcomb PJ. Semantic priming and stimulus degradation: Implications for the role of the N400 in language processing. Psychophysiology. 1993; 30:47–61. [PubMed: 8416062]

Holt DJ, Lynn SK, Kuperberg GR. Neurophysiological correlates of comprehending emotional meaning in context. Journal of Cognitive Neuroscience. 2009; 21(11):2245–2262. [PubMed: 18855550]

Johnson-Laird, PN. Mental Models. Cambridge: Harvard University Press; 1983.

Just MA, Carpenter PA. A capacity theory of comprehension: Individual differences in working memory. Psychological Review. 1992; 99(1):122–149. [PubMed: 1546114]

Kamas EN, Reder LM, Ayers MS. Partial matching in the Moses illusion: response bias not sensitivity. Memory & Cognition. 1996; 24(6):687–699.

Keenan JM, Baillet SD, Brown P. The effect of causal cohesion on comprehension and memory. Journal of Verbal Learning and Verbal Behavior. 1984; 23:115–126.

Keenan, JM.; Potts, G.; Golding, J.; Jennings, T. Which elaborative inferences are drawn during reading? A question of methodologies. In: Balota, DA.; d'Arcais, GBF.; Rayner, K., editors. Comprehension processes in reading. Hillsdaly NJ: Lawrence Erlbaum Associates; 1990. p. 377-402.

Kintsch W. The role of knowledge in discourse comprehension: A construction-integration model. Psychological Review. 1988; 95:163–182. [PubMed: 3375398]

Kintsch, W. How readers construct situation models for stories: The role of syntactic cues and causal inferences. In: Healy, AF.; Kosslyn, SM.; Shiffrin, RM., editors. From learning processes to cognitive process: Essays in honor of William K Estes. Hillsdale, NJ: Lawrence Erlbaum Associates; 1992. p. 261-278.

Kintsch, W. Comprehension: A paradigm for cognition. New York: Cambridge University Press; 1998.

Kintsch W. Predication. Cognitive Science. 2001; 25:173–202.

Kintsch W, van Dijk T. Toward a model of text comprehension and production. Psychological Review. 1978; 85(5):363–394.

Klin CM. Causal inferences in reading: from immediate activation to long-term memory. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1995; 21(6):1483–1494.

Klin CM, Guzman AE, Levine WH. Prevalence and persistence of predictive inferences. Journal of Memory and Language. 1999; 40:593–604.

Kolk HHJ, Chwilla DJ. Late Positivities in unusual situations. Brain and Language. 2007; 100(3):257–262. [PubMed: 16919324]

Kuperberg GR. Neural mechanisms of language comprehension: Challenges to syntax. Brain Research, Special Issue: Mysteries of Meaning. 2007; 1146:23–49.

Kuperberg GR, Kreher DA, Sitnikova T, Caplan D, Holcomb PJ. The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. Brain and Language. 2007; 100(3):223–238. [PubMed: 16546247]

Kuperberg GR, Lakshmanan BM, Caplan DN, Holcomb PJ. Making sense of discourse: An fMRI study of causal inferencing across sentences. NeuroImage. 2006; 33(1):343–361. [PubMed: 16876436]

Kuperberg GR, Sitnikova T, Caplan D, Holcomb PJ. Electrophysiological distinctions in processing conceptual relationships within simple sentences. Cognitive Brain Research. 2003; 17(1):117–129. [PubMed: 12763198]

Kutas M, Federmeier KD. Electrophysiology reveals semantic memory use in language comprehension. Trends in Cognitive Sciences. 2000; 4(12):463–470. [PubMed: 11115760]

Kutas M, Hillyard SA. Reading senseless sentences: Brain potentials reflect semantic incongruity. Science. 1980; 207:203–205. [PubMed: 7350657]

Kutas M, Hillyard SA. Brain potentials during reading reflect word expectancy and semantic association. Nature. 1984; 307:161–163. [PubMed: 6690995]

Kutas, M.; Van Petten, C.; Kluender, R. Psycholinguistics electrified II: 1994–2005. In: Traxler, M.; Gernsbacher, MA., editors. Handbook of Psycholinguistics. 2. New York: Elsevier; 2006. p. 659-724.

Landauer TK, Dumais ST. A solution to Plato’s problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review. 1997; 104:211–240.

Landauer TK, Foltz PW, Dumais ST. Introduction to Latent Semantic Analysis. Discourse Processes. 1998; 25:259–284.

Lassonde KA, O’Brien EJ. Contextual specificity in the activation of predictive inferences. Discourse Processes. 2009; 46(5):1–13.

Ledoux K, Camblin CC, Swaab TY, Gordon PC. Reading words in discourse: the modulation of lexical priming effects by message-level context. Behavioral and Cognitive Neuroscience Reviews. 2006; 5(3):107–127. [PubMed: 16891554]

Long, DL.; Golding, JM.; Graesser, AC.; Clark, LF. Goal, event and state inferences: An investigation of inference generation during story comprehension. In: Graesser, AC.; Bower, GH., editors. The Psychology of Learning and Motivation: Inferences and Text Comprehension. San Diego: Academic Press; 1990.

Long DL, Lea RB. Have we been searching for meaning in all the wrong places? Defining the “search after meaning” principle in comprehension. Discourse Processes. 2005; 39(2–3):279–298.

MacDonald MC, Pearlmutter NJ, Seidenberg MS. The lexical nature of syntactic ambiguity resolution. Psychological Review. 1994; 101:676–703. [PubMed: 7984711]

Mason RA, Just MA. How the brain processes causal inferences in text. Psychological Science. 2004; 15(1):1–7. [PubMed: 14717824]

McKoon G, Ratcliff R. Inferences about predictable events. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1986; 12(1):82–91.

McKoon G, Ratcliff R. Semantic associations and elaborative inference. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1989; 15(2):326–338.

McKoon, G.; Ratcliff, R. Textual inferences: Models and measures. In: Balota, DA.; d’Arcais, GBF.; Rayner, K., editors. Comprehension Processes in Reading. Hillsdale, NJ: Lawrence Erlbaum Associates; 1990. p. 403-421.

McKoon G, Ratcliff R. Inference during reading. Psychological Review. 1992; 99(3):440–466. [PubMed: 1502273]

McKoon G, Ratcliff R, Ward G. Testing theories of language processing: an empirical investigation of the on-line lexical decision task. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1994; 20(5):1219–1228.

Millis KK, Just MA. The influence of connectives on sentence comprehension. Journal of Memory and Language. 1994; 33(1):128–147.

Myers, JL.; Duffy, SA. Causal inferences and text memory. In: Graesser, AC.; Bower, GH., editors. The Psychology of Learning and Motivation: Inferences and Text Comprehension. Vol. 25. San Diego: Academic Press; 1990. p. 159-173.

Myers JL, O'Brien E. Accessing the discourse representation during reading. Discourse Processes. 1998; 26(2&3):131–157.

Myers JL, O'Brien EJ, Albrecht JE, Mason RA. Maintaining global coherence during reading. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1994; 20(4):876–886.

Myers JL, Shinjo M, Duffy SA. Degree of causal relatedness and memory. Journal of Memory and Language. 1987; 26:453–465.

Nakano H, Saron C, Swaab TY. Speech and span: Working memory capacity impacts the use of animacy but not of world knowledge during spoken sentence comprehension. Journal of Cognitive Neuroscience. (In press).

Nieuwland, MS.; Ditman, T.; Kuperberg, GR. On the incrementatilty of pragmatic processing: An ERP investigation of underinformative scalar sentences. (Under review)

Nieuwland MS, Kuperberg GR. When the truth isn't too hard to handle: An event-related potential study on the pragmatics of negation. Psychological Science. 2008; 19:1213–1218. [PubMed: 19121125]

Nieuwland MS, van Berkum JJ. Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. Cognitive Brain Research. 2005; 24(3):691–701. [PubMed: 15894468]

O'Seaghdha PG. Conjoint and dissociable effects of syntactic and semantic context. Journal of Experimental Psychology: Learning, Memory and Cognition. 1997; 23:807–828.

Osterhout L, Hagoort P. A superficial resemblence 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. 1999; 14:1–14.

Osterhout L, Holcomb PJ. Event-related potentials elicited by syntactic anomaly. Journal of Memory and Language. 1992; 31:785–806.

Otten M, van Berkum JJ. What makes a discourse constraining? Comparing the effects of dicourse message and scenario fit on the discourse-dependent N400 effect. Brain Research. 2007; 1146:158–171. [PubMed: 16916496]

Potts GR, Keenan JM, Golding JM. Assessing the occurence of elaborative inferences: Lexical decision versus naming. Journal of Memory and Language. 1988; 27:399–415.

Raisig S, Welke T, Hagendorf H, van der Meer E. Investigating dimensional organization in scripts using the pupillary response. Psychophysiology. 2007; 44(6):864–873. [PubMed: 17850243]

Rapp DN, van den Broek P. Dynamic text comprehension: An integrative view of reading. Current Directions in Psychological Science. 2005; 14(5):276–279.

Rugg MD. The effects of semantic priming and word repetition on event-related potentials. Psychophysiology. 1985; 22:642–647. [PubMed: 4089090]

Sanford, AJ. On the nature of text-driven inference. In: Balota, DA.; d'Arcais, F.; Rayner, K., editors. Comprehension processes in reading. Hillsdale, NJ: Erlbaum; 1990.

Sanford AJ. Context, attention, and depth of processing during interpretation Mind Language. 2002; 17:188–206.

Sanford, AJ.; Garrod, SC. Understanding Written Language: Explorations of Comprehension Beyond the Sentence. New York: Wiley; 1981.

Sanford AJ, Garrod SC. The role of scenario mapping in text comprehension. Discourse Processes. 1998; 26:2–3.

Sanford AJ, Garrod SC. Memory-based approaches and beyond. Discourse Processes. 2005; 39:205–224.

Schank, RC.; Abelson, RP. Scripts, Plans, Goals, and Understanding: An Inquiry into Human Knowledge Structures. Hillsdale, N.J: L. Erlbaum Associates; 1977.

Sharkey NE, Sharkey AJ. What is the point of integration? The loci of knowledge-based facilitation in sentence processing. Journal of Memory and Language. 1987; 26(3):255–276.

Singer M. Causal bridging inferences: Validating consistent and inconsistent sequences. Canadian Journal of Experimental Psychology. 1993; 47:340–359.

Singer M, Andrusiak P, Reisdorf P, Black NL. Individual differences in bridging inference processes. Memory & Cognition. 1992; 20(5):539–548.

Singer M, Ferreira F. Inferring consequences in story comprehension. Journal of Verbal Learning and Verbal Behavior. 1983; 22:437–448.

Singer M, Halldorson M. Constructing and validating motive bridging inferences. Cognitive Psychology. 1996; 30(1):1–38. [PubMed: 8660781]

St George M, Mannes S, Hoffman JE. Individual differences in inference generation: An ERP analysis. Journal of Cognitive Neuroscience. 1997; 9(6):776–787.

Sturt P, Sanford AJ, Stewart A, Dawydiak E. Linguistic focus and good-enough representations: An application of the change-detection paradigm. Psychonomic Bulletin & Review. 2004; 11(5):882–888. [PubMed: 15732698]

Tanenhaus MK, Spivey-Knowlton MJ, Eberhard KM, Sedivy JC. Integration of visual and linguistic information in spoken language comprehension. Science. 1995; 268(5217):1632–1634. [PubMed: 7777863]

Trabasso T, van den Broek P, Suh SY. Logical necessity and transitivity of causal relations in stories. Discourse Processes. 1989; 12:1–25.

Trabasso T, Van den Broek PW. Causal thinking and the representation of narrative events. Journal of Memory and Language. 1985; 24:612–630.

van Berkum, JJA. The neuropragmatics of ‘simple’ utterance comprehension: An ERP review. In: Sauerland, U.; Yatsushiro, K., editors. Semantics and Pragmatics: From Experiment to Theory. Basingstoke: Palgrave; (In press)

van Berkum JJA, Brown CM, Zwitserlood P, Kooijman V, Hagoort P. Anticipating upcoming words in discourse: evidence from ERPs and reading times. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2005; 31(3):443–467.

van Berkum JJA, Hagoort P, Brown CM. Semantic integration in sentences and discourse: Evidence from the N400. Journal of Cognitive Neuroscience. 1999; 11(6):657–671. [PubMed: 10601747]

van Berkum JJA, Holleman B, Nieuwland M, Otten M, Murrel J. Right or wrong? The brain’s fast response to morally objectionable statements. Psychological Science. 2009; 20(9):1092–1099. [PubMed: 19656340]

van de Meerendonk N, Kolk HHJ, Vissers CTWM, Chwilla DJ. Monitoring language perception: Mild and strong conflicts elicit different ERP patterns. Journal of Cognitive Neuroscience. (In press).

van den Broek, PW. Causal inferences and the comprehension of narrative texts. In: Graesser, AC.; Bower, GH., editors. The Psychology of Learning and Motivation: Inferences and Text Comprehension. San Diego: Academic Press; 1990.

van den Broek, PW. Comprehension and memory of narrative texts: Inferences and coherence. In: Gernsbacher, MA., editor. Handbook of Psycholinguistics. San Diego: Academic Press; 1994. p. 539-583.

van den Broek PW, Rapp DN, Kendeou P. Integrating memory-based and constructionist processes in accounts of reading comprehension. Discourse Processes. 2005; 39(2&3):299–316.

van den Broek, PW.; Risden, K.; Husebye-Hartman, E. The role of readers’ standards for coherence in the generation of inferences during reading. In: Lorch, RF.; O’Brien, EJ., editors. Sources of coherence in reading. Mahwah, NJ: Lawrence Erlbaum Associates, Inc; 1995. p. 353-373.

Van Dijk, TA.; Kintsch, W. Strategies of Discourse Comprehension. New York: Academic Press; 1983.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Van Petten C. 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. 1993; 8:485–531.

Van Petten C, Kutas M, Kluender R, Mitchiner M, Melsaac H. Fractionating the word repetition effect with event-related potentials. Journal of Cognitive Neuroscience. 1991; 3:131–150.

Wolfe MBW, Magliano JP, Larsen B. Causal and semantic relatedness in discourse understanding and representation. Discourse Processes. 2005; 39(2–3):165–187.

Yang CL, Perfetti CA, Schmalhofer F. Event-related potential indicators of text integration across sentence boundaries. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2007; 33(1):55–89.

Zwaan RA, Magliano JP, Graesser AC. Dimensions of situation model construction in narrative comprehension. Journal of Experimental Psychology: Learning, Memory, and Cognition. 1995; 21:386–397.

Zwaan RA, Radvansky GA. Situation models in language comprehension and memory. Psychological Bulletin. 1998; 123(2):162–185. [PubMed: 9522683]

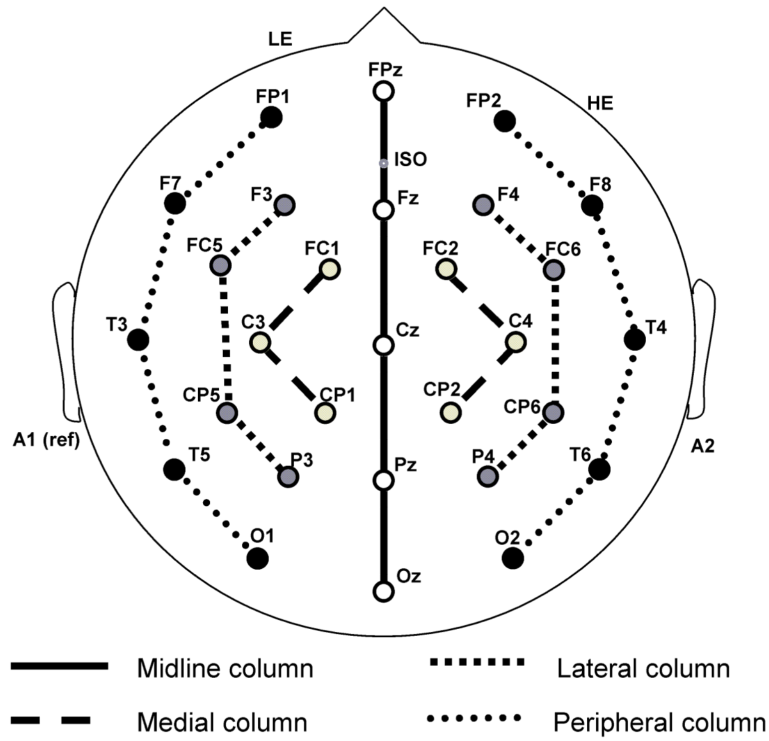

Figure 1. Electrode Montage

Electrodes placed in the standard International 10–20 System locations included five sites along the midline (FPz, Fz, Cz, Pz, and Oz) and eight lateral sites, four over each hemisphere (F3/F4, C3/C4, T3/T4, and P3/P4). Eight additional 10–20 sites were altered to form a circle around the perimeter of the scalp. These altered sites included FP1′/FP2′ (33% of the distance along the circle between T3/T4), F7′/F8′ (67% of the distance between FPz and T3/T4), T5′/T6′ (33% of the distance between T3/T4 and Oz), and O1′/O2′ (67% of the distance between T3/T4 and Oz). In addition eight extended 10–20 system sites were also used (FC1/FC2, FC5/FC6, CP1/CP2, and CP5/CP6). The dotted lines represent the four columns used in analyses (i.e., midline, medial, lateral, peripheral). In the midline ANOVA, the electrode site factor had five levels (FPz, Pz, Cz, Pz, Oz); in the medial ANOVA, it had eight levels (FC1/FC2, C2/C4, CP1/CP2); in the lateral ANOVA, it had four levels of electrode site (F3/F4, FC5/FC6, CP5/CP6, P3/P4); in the peripheral ANOVA, it had five levels (FP1′/FP2′, F7′F8′, T3/T4, T5′/T6′, O1′/O2′).

NIH-PA Author Manuscript
NIH-PA Author Manuscript

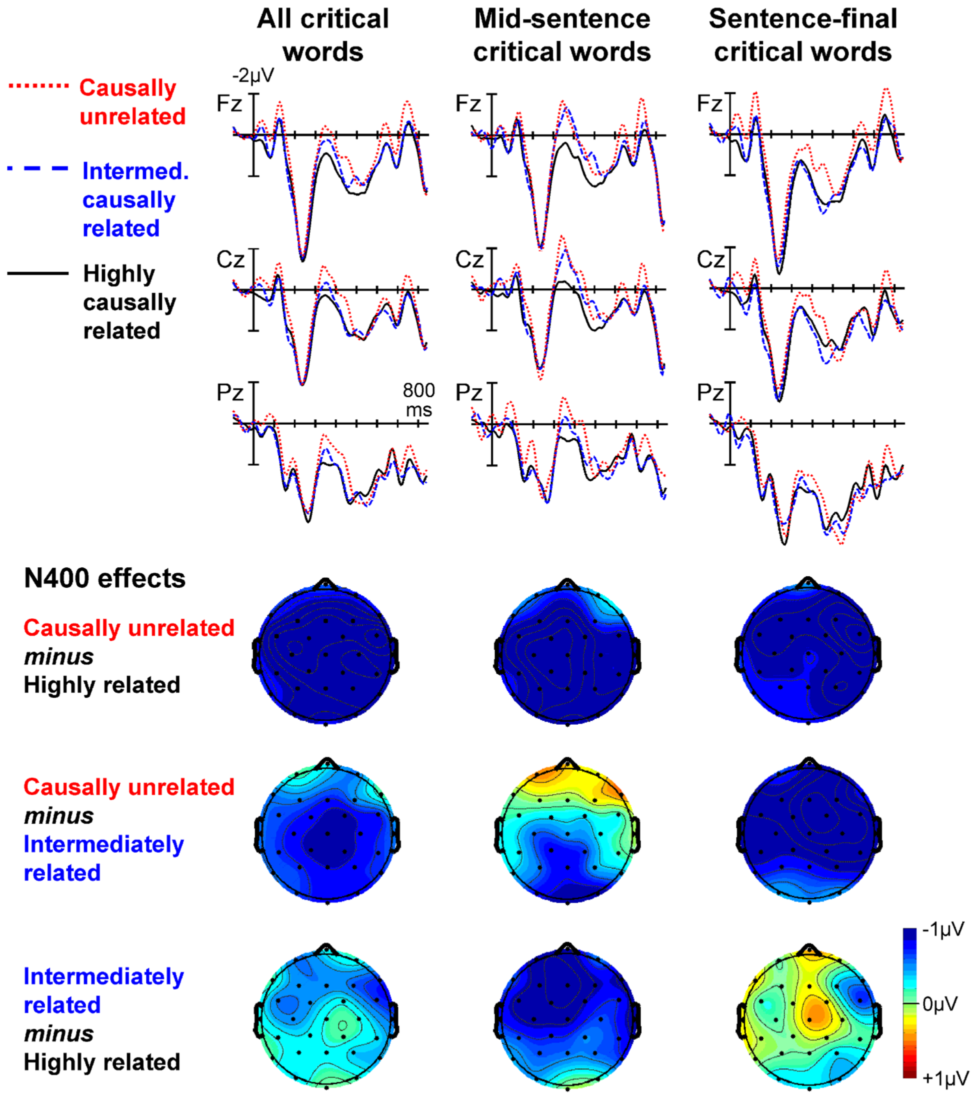

Figure 2.

Top: ERPs at midline electrode sites to critical words in highly causally related, intermediately related and causally unrelated scenarios.

Bottom: voltage maps showing mean differences across the N400 time-window (325–475ms) between critical words, comparing each level of Causal Relatedness with one another.

Table 1

Causal Relatedness	Construction and Explanation	Causal relatedness rating	LSA Semantic similarity values	Example
Highly Related	The first sentence sets up a fairly non-constraining context. The second and third sentences are highly causally related in meaning, requiring a simple inference to link them.	**All CWs:** 6.37 (0.61) **Mid-sentence CWs:** 6.36 (0.62) **Sentence-final CWs:** 6.39 (0.61)	**All CWs:** 0.16 (0.07) **Mid-sentence CWs:** 0.17 (0.07) **Sentence-final CWs:** 0.15 (0.06)	*Jill had very fair skin. She forgot to put sunscreen on. She had sunburn on Monday.* [Here the reader must infer that sunburn results from forgetting to put on sunscreen, particularly if one has fair skin.]
Intermediately Related	The same first and third sentence as in the highly related sentences. The second sentence is constructed such that the reader is required to make a complex inference to connect the second and third sentences.	**All CWs:** 4.79 (1.04) **Mid-sentence CWs:** 4.86 (1.00) **Sentence-final CWs:** 4.72 (1.08)	**All CWs:** 0.15 (0.07) **Mid-sentence CWs:** 0.17 (0.08) **Sentence-final CWs:** 0.14 (0.07)	*Jill had very fair skin. She usually remembered to wear sunscreen. She had sunburn on Monday.* [Here the reader must make the above inference but must also infer that although she usually puts on sunscreen, this time Jill forgot]
Causally Unrelated	The third sentence is the same as in both of the other conditions. In most cases the first sentence was also kept the same. The second sentence, however, is constructed in such a way as to make the third sentence not follow logically from the first two.	**All CWs:** 2.14 (0.99) **Mid-sentence CWs:** 2.11 (0.93) **Sentence-final CWs:** 2.18 (1.05)	**All CWs:** 0.16 (0.07) **Mid-sentence CWs:** 0.17 (0.07) **Sentence-final CWs:** 0.14 (0.07)	*Jill's skin always tanned well. She always put on sunscreen. She had sunburn on Monday.* [Here there is no clear inference that can be made since the second scenario establishes that it is highly unlikely that Jill would not put on sunscreen and the first scenario states that even had she forgotten, she would not be likely to sunburn]

The example given is for a scenario where the critical word appears mid-sentence. For causal relatedness ratings and LSA values, means are shown with standard deviations in parentheses.

Table 2

Causal Relatedness	Mean relatedness judgment score: All scenarios	Mean relatedness judgment score: Scenarios with mid-sentence critical words	Mean relatedness judgment score: Scenarios with sentence-final critical words
Highly related	2.88 (0.07)	2.88 (0.07)	2.87 (0.09)
Intermediately related	2.19 (0.19)	2.20 (0.20)	2.19 (0.21)
Causally unrelated	1.37 (0.17)	1.39 (0.19)	1.35 (0.21)

Mean causal relatedness judgment scores are shown. Standard deviations are indicated in parentheses.