Cognitive Processes in Reading Comprehension: An ERP study on the maintenance of global coherence in discourses

Chris Noone

S1154257

Department of Education and Child Studies,

Leiden University

First Supervisor: Ms Anne Helder MSc.

Second Supervisor: Professor Paul van den Broek

Summary

Reading comprehension is a dynamic process which is influenced by a range of text characteristics and individual characteristics (Van den Broek, 1994). An important text characteristic is coherence, a measure of the interconnectedness of information within a text, which can be assessed locally and globally. This study was intended to conceptually replicate previous research using event-related potentials (ERPs) to investigate the maintenance of global coherence. ERPs were computed for critical words in each story to compare the difference in mean amplitude of the N400 component across coherent and incoherent story conditions. The N400 is thought to be inversely related to the ease with which semantic processing is carried out (Nieuwland & Van Berkum, 2006). It was hypothesised that breaks in global coherence would produce an N400 effect relative to the coherent control condition. A repeated measures ANOVA was carried out to test whether the N400 amplitude, measured as mean amplitude in the 300 – 500ms post critical word time window across 11 chosen scalp sites, was more strongly negative for critical words in the incoherent condition than for those in the coherent condition. Mean amplitude values for the N400 were more negative for critical words in the incoherent condition than those in the coherent condition across 7 out of 11 scalp sites, but the differences did not reach significance.

Introduction and theoretical framework

Reading comprehension is a dynamic process which is influenced by a range of text characteristics and individual characteristics (van den Broek, 1994). The influential text characteristics include, among others, semantic meaning and discourse context which can respectively be thought of as more important locally and globally. This study aims to conceptually replicate current research on the processing of local and global factors which affect the coherence of a discourse (Van Berkum, Zwitserlood, Hagoort, & Brown, 2003; van Berkum, Hagoort, & Brown, 1999; St George, Mannes, & Hoffman, 1994). Specifically, this study will take advantage of a specific ERP component, known as the N400, to reveal whether global coherence is maintained at a discourse level.

Reading comprehension is assumed to produce a number of representations of what is being read: a surface code, which is made up of the exact wording and syntax of the text, a text base (Johnson-Laird, 1983) representing the propositional structure and a situation model (van Dijk & Kintsch, 1983), which represents what a text is about and may include pragmatic, linguistic and world knowledge in addition to explicitly stated information. So, comprehension involves the perceptual identification of letters, the derivation of word meaning, the recognition of grammatical structure and the formation of a situation model representing what the text is about which may include information related to the text which is stored in long-term memory (i.e. prior knowledge) (St George et al., 1994). Successful comprehension is generally thought of as the construction of a coherent situation model (Nieuwland & Van Berkum, 2008). The more coherent a text is, the more the connection of ideas within the text is promoted, which results in the facilitation of new ideas being included in the ongoing situation model as these ideas are introduced (Lehman & Schraw, 2002).

Text coherence can be assessed at a local level and a global level. Local coherence involves establishing connections between the information currently being processed and the information immediately preceding this, which is held in short-term memory. High levels of local coherence enhance semantic and syntactic processing by allowing readers to construct referential and causal connections between nearby text segments. Global coherence involves establishing connections between the information currently being processed and information located much earlier in the text, which is no longer in short-term memory but is relevant to the information currently being processed, and to relevant world knowledge (Albrecht & O'Brien, 1993; Long & Chong, 2001). What connects the information currently being processed and the preceding information is a coherence relation. These are conceptual and need not be explicitly stated by linguistic markers, though they sometimes are (Sanders & Noordman, 2000).

Although the position had originally been that global coherence was not a factor in reading comprehension, or that it at least was processed much later than local coherence, now research suggests that readers readily maintain both local and global coherence (van Berkum et al., 1999; van Berkum et al., 2003; Cook, Halleran, & O'Brien, 1998; Albrecht & O'Brien, 1993). There are two classes of language models which may account for how local and global factors interact. Two-step models of language comprehension assume that people compute a local, context-independent meaning of a noun phrase before integrating it into the situation model (e.g. Fodor, 1983; Forster, 1979; Kintsch, 1988; Millis & Just, 1994). Interactive single-step models (e.g. Jackendoff, 2002; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994) allow for both local and global factors to overrule each other, depending on the strength and relevance of the local and global cues. In these models, global coherence is as important as local coherence (Nieuwland & Van Berkum, 2006). Current research indicates that there are at least two processes involved in establishing global coherence. These are activation and integration. During activation, each incoming phrase in the text activates related associative nodes in the memory network. The amount of activation is influenced by the degree to which a memory trace shares features with concepts in the incoming text, how much it has been elaborated in the text, and its recentness. The integration stage selects relevant information from the array of information activated (Kintsch, 1988; Long & Chong, 2001). This facilitates comprehension because when relevant prior knowledge is activated, lower thresholds are set for nodes in the associative network and this increases the chances that related information will be activated (St George et al., 1994). So, as reading progresses global coherence facilitates comprehension by making it easier to activate knowledge (from earlier in the text or long-term memory) which is then used to interpret the forthcoming text.

Commonly used psycholinguistic methods such as reaction time, eye-tracking and other behavioural measures have difficulty showing the interaction of local and global factors conclusively (Nieuwland & Van Berkum, 2006). Electroencephalographic (EEG) techniques are better suited to exploring issues of local and global coherence in language comprehension. Event-related potentials (ERPs) are particularly useful in this context. ERPs are measured by averaging the electrical voltage fluctuations, at the scalp using electrodes, which are time-locked to specific events or stimuli. ERPs are very temporally sensitive and this allows the examination of the rapid processes involved in language comprehension. An ERP effect generally refers to the difference between the average amplitude, within a specific time window, of two experimental conditions. ERP components can also be classified according to polarity (i.e. positive or negative), timing of onset of effect or peak amplitude, morphology (peaked or slow rising) and scalp distribution and as a result ERPs give us both quantitative and qualitative data. ERPs describe language processing online and therefore there is no need for further tasks which may elicit further processing (Van Berkum, 2008; Nieuwland & Van Berkum, 2006).

The specific ERP component of most importance to the issue of local and global coherence is the N400 effect. The N400 is a negatively deflected waveform which develops between 150ms and 250ms and peaks at approximately 400ms after onset. It is strongest in centroparietal regions. The amplitude of the N400 is inversely related to the ease with which semantic processing is carried out (Kuperberg, Paczynski, & Ditman, 2011; Nieuwland & Van Berkum, 2006; Van Berkum, 2008). The N400 can be observed in response to a range of meaningful stimuli including visual and auditory words, acronyms, gestures, symbols, sign language signs and environmental sounds. The N400 effect refers to a relative negativity which means that the waveform is at its minimum around 400ms post-stimulus. It is therefore not necessary for the amplitude of the N400 to have a negative value. A stronger N400 effect is one which has a more negative mean or peak amplitude (Kutas & Federmeier, 2009).

Evidence from the few studies that have examined this issue suggests that the amplitude of the N400 can be influenced by both local and global factors in a discourse (Nieuwland & Van Berkum, 2006). A pioneering study by St George, Mannes and Hoffmann (1994) was the first to address the influence of local and global coherence on language processing using EEG. Paragraphs were presented either with or without titles. Each of these paragraphs was locally coherent – the sentences were interconnected and made sense semantically. However, without the title there was no context and therefore a global coherence break. A strong N400 effect was found in the untitled condition compared to the titled condition. This suggests that the presence of global coherence influenced the amplitude of the N400 component. In addition, a larger P1-N1 amplitude difference component was found in the titled condition. The P1-N1 component reflects spatial attention and it was suggested that due to the presence of global coherence facilitating language processing, more resources were available to devote to attention. One drawback of this study was that the ERP component was averaged across the entire paragraph rather than at critical words. A set of experiments by Van Berkum and colleagues (1999, 2003) addressed this problem and in doing so replicated and extended the findings of St George et al. Using sentences with two possible endings (e.g. slowly/quickly), which were semantically plausible and locally coherent, and a preceding context which rendered one of these critical words anomalous allowed Van Berkum et al. to investigate whether the wider discourse context (i.e. global coherence) was taken into account in the processing of incoming critical words. On presentation of the anomalous words an N400 effect was observed. When the same sentences were heard in isolation (i.e. without any context/situation model to influence processing) the N400 effect disappeared. The results revealed that participants were maintaining global coherence and this influenced semantic processing and that the N400 effect was identical to what would be elicited by a local coherence break. A follow-up study revealed more about how local and global factors interact. This study showed that a discourse context (e.g. about an amorous peanut and his girlfriend) could overrule local semantic coherence as evidenced by sentences like “the peanut was salty” being more difficult to process (i.e. producing a larger N400) than those like “the peanut was in love”, which would in isolation produce a strong N400 effect. This provided support for the interactive single-step models of language comprehension discussed earlier (Nieuwland & Van Berkum, 2006). In summary, current research suggests that the N400 can provide a sensitive measure of how lexico-semantic relationships and the discourse model (i.e. local and global coherence) interact to influence the processing of incoming words. This research will make use of the N400 to conceptually replicate previous research on the maintenance of global coherence.

In this study, stories designed to have both local and global coherence (i.e. the control condition), or a global coherence break were used as stimuli for an ERP experiment. Two versions of each story were used, one with a context which maintains coherence with a critical word near the end of the story and one that creates a global coherence break at the critical word. These stories were presented aurally as the eye movements associated with reading cause major artifacts in EEG data. ERPs were computed for these critical words and grand averages will be compared across conditions. The null hypothesis predicted that there will be no difference in N400 amplitude between the global coherence break condition and the control condition. In this case, the close lexico-semantic relationships across individual words within the discourse context may dominate the semantic processing of an upcoming word, overruling the situation model. The alternative hypothesis predicted a significant difference in N400 amplitude between the control condition and the global coherence break conditions, with the latter being more negative in amplitude. In this case, the situation model of the text, rather than simply lexico-semantic relationships, may influence the semantic processing of upcoming words and the noticing of a global coherence break will produce the N400 effect.

Methods

Participants

11 right-handed speakers of Dutch took part in data collection. All were current university students. 9 were native speakers and both non-native participants had completed higher education through the Dutch language. The participants ranged in age from 19 to 25 (M = 22.25, SD = 2.38) and 2 were male. 8 participants provided usable data. One of the excluded participants provided data with too few clean portions. The data from the other two were unusable due to computer errors. Participants were rewarded with an entertainment voucher worth ten euro. Participants responded to campus advertising and a post on the SONA experiment management system. Exclusion criteria included colour blindness, alcohol and drug abuse, neurological and psychiatric disorders, and use of medication (except oral contraceptives). Those who did not have university level Dutch were also excluded.

Materials

Sixty globally coherent stories were constructed in Dutch, each six lines in length and aurally presented to participants. For each story, an inconsistent version was created by changing part of the second sentence. So 120 stories were used, half of which were consistent and half of which were inconsistent. Individual participants only heard one version of each story and only heard each story once. This modified the global coherence of the critical word near the end of the story. Each story began with the introduction of a character. The second line described the character or the situation. There were then 3 filler lines. Following this a sentence relating to the second sentence was presented, all of which were locally coherent. However, in one version of each story the critical word contradicted the details in the second sentence, causing a global coherence break. See Table 1 for an example of a story used in the experiment. These critical words were never used previously in each story or as critical words in any other stories. The critical words ranged in length from 191ms to 975ms (M = 496.53ms, SD = 170.86ms). Critical word onset ranged from 14.95s to 25.99s (M = 20.47s, SD = 2.32). This variation in onset eliminated the possibility of expectation effects. All of these stories were recorded by the same female speaker at a regular speaking rate and intonation. To minimise differences in rate and intonation, the recordings were edited so that the critical word was the only difference between the two conditions of a story.

Table 1.

Example of a story with incoherency modification in brackets and critical word underlined.

(S1) Oscar and Mike are ten-year-old twins who do almost everything together.
(C) Because there has been a lot of snow in the past few days, they don't have to go to school today.
(I) Because the weather has been very hot in the past few days, they don't have to go to school today.
(f) Their mother works for a large company and is working from home today.
(f) Today she is very busy with her work, she has to finish a report.
(f) To not disturb her, Oscar and Mike are playing outside.
(T) They decide to build a <u>snowman</u> in the backyard.

Procedure

The experiment session began with the participants being introduced to the lab and having the electrode net application procedure explained to them. Having read an information leaflet regarding the study, participants signed a letter of informed consent. Following the application of the electrode net, it was explained to the participants that they would listen to 60 stories and answer a comprehension question following each story. Participants were instructed to move as little as possible and were informed that there were plenty of chances to blink between stories and to therefore refrain from blinking during the presentation of stories. To increase comfort and reduce movements, the height of the screen was adjusted so the centre was at eye-level for each participant. The experiment was presented on a 15 inch LCD monitor approximately 55cm from where the participants were seated. The instructions and fixation crosses were presented in black on a light grey background. This was intended to reduce eye blink artifacts. The stories were presented through speakers and the participants were instructed to listen carefully and to focus their gaze on the fixation cross during each story. First, two practice stories were presented. Then the electroencephalographic (EEG) activity of the participants was recorded as the participants listened to 60 stories which varied in length from 17s to 30s (M = 22.83s, SD = 2.41s). Four story lists were pseudo-randomly created so that no participant heard the same story in a different condition and so that no more than 2 stories from one condition were ever presented in a row. After each story, a comprehension question with a yes or no answer was presented and responded to using the keyboard. Participants answered an average of all of these questions correctly. Before each story and each question, participants had the opportunity to blink before pressing a button to continue. After the 20^{th} and 40^{th} stories, longer breaks were allowed. The experimental task was executed by commercial experimental control software, namely E-prime (Psychology Software Distribution, Pittsburgh, Pennsylvania). E-prime controls the presentation of the stimuli, the recording of trial information and the sending of critical word event information to Netstation, the EEG data recording package. Following the completion of the task, the electrode net was removed and participants were given a chance to freshen up. Two supplementary tasks were then completed, the CBM Maze Task and the IRI Perspective Taking questionnaire which are not relevant to this study. Participants took part in the experiment individually and each session took approximately 90 minutes from when they arrived to when they left.

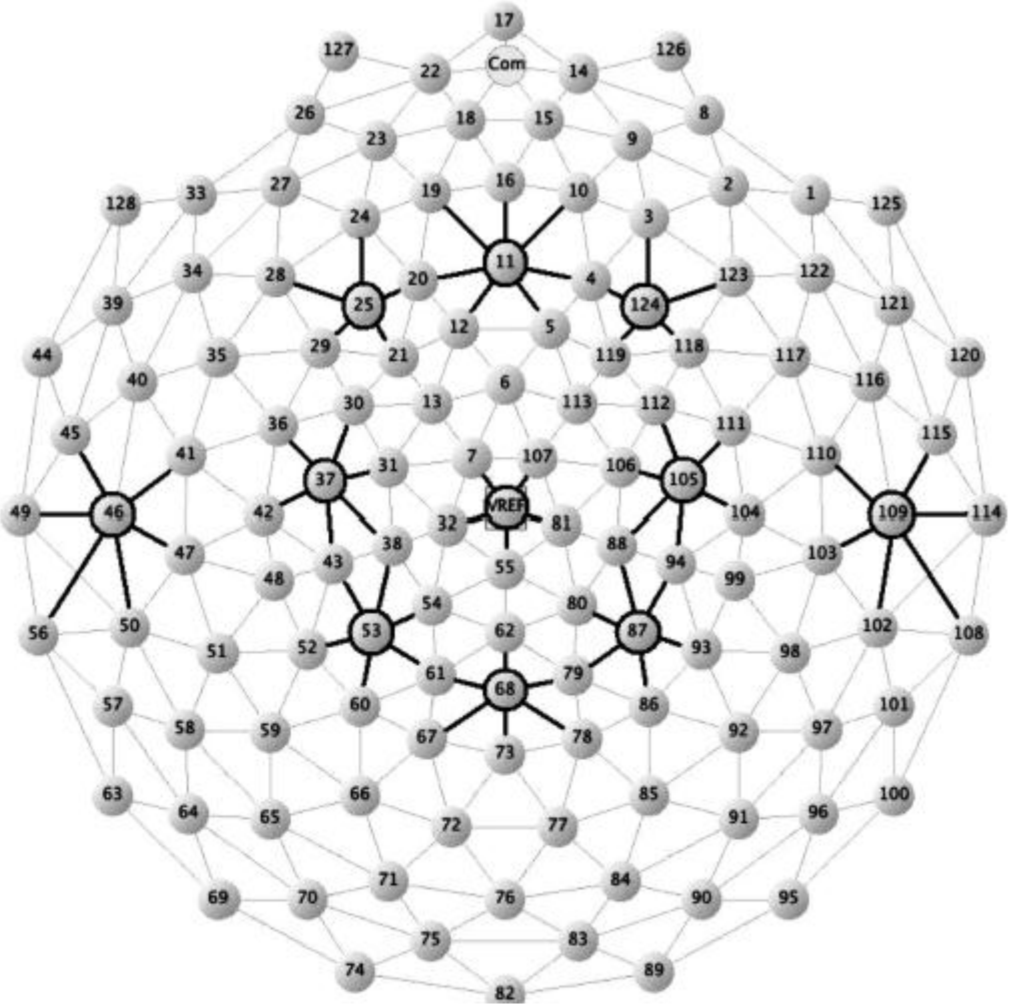
ERPs

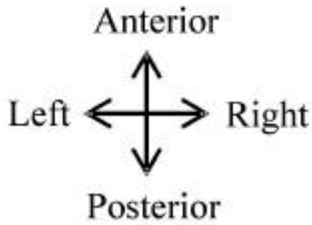
The EEG activity of the participants during the story task was acquired through 129 channel hydrocel geodesic sensor nets which were amplified using a NetAmps 300 amplifier with a digitization rate of 250Hz. Netstation, a software package designed for recording, manipulating and analysing EEG data, was used to record the EEG activity on an Apple

Macintosh,1000 MHz PowerPC G4 class computer with Mac OS 9.2.2. (Apple Computer, Cupertino, CA). Impedances were kept below 50 kΩ, an acceptable level for the system used (Ferree, Luu, Russell, & Tucker, 2001; Tucker, 1993 as cited in Yang, Perfetti & Schmalhofer, 2007). Brain Vision Analyser 2.0 (Brain Products) was used for further processing of the EEG data. First, the data was filtered using a passband range of 0.5 – 30Hz (-3dB, 48dB/Octave, high pass filter; -3dB, 12 dB/Octave, low pass filter). Then the data was rereferenced to the average of activity in all of the channels. Following this, segments beginning at 200ms before the critical word onset and ending 1000ms after the critical word onset were extracted. The removal of segments containing artifacts was then carried out. This resulted in an average of 50 segments per participant or 83% of total segments. Segments with a difference of 60µVor more in the vertical EOG channel (channel 8 – channel 126) or 40 60µV (channel 128 – channel 125) or more in the horizontal EOG channel were removed. Individual channels were removed from segments with a difference of more than 150µV between the maximum and minimum activity in the particular channel. The segments were then averaged together to create ERPs for each participant. Grand averages were then produced for each condition and from these a difference wave was created i.e. the ERP at the onset of the critical word in the consistent condition was subtracted from the ERP at the onset of the critical word in the inconsistent condition. A minimum was observed at approximately 444ms after the onset of the critical word. The mean amplitude of the averaged waves at 11 scalp regions (see Figure 1), chosen based on previous research on the N400 (e.g. Yang et al., 2007), between 300ms and 500ms was defined as the N400. It was expected that on average the amplitude for the inconsistent condition would be more negative than that of the coherent condition in the selected time window. In other words, the N400 effect would be stronger for critical words in incoherent stories.

Data Analysis

The data were statistically analysed using SPSS 19 (IBM) software. Preparatory data analyses were undertaken to assess the assumptions for the statistical test. The null hypothesis was tested using a repeated measures analysis of variance (ANOVA).





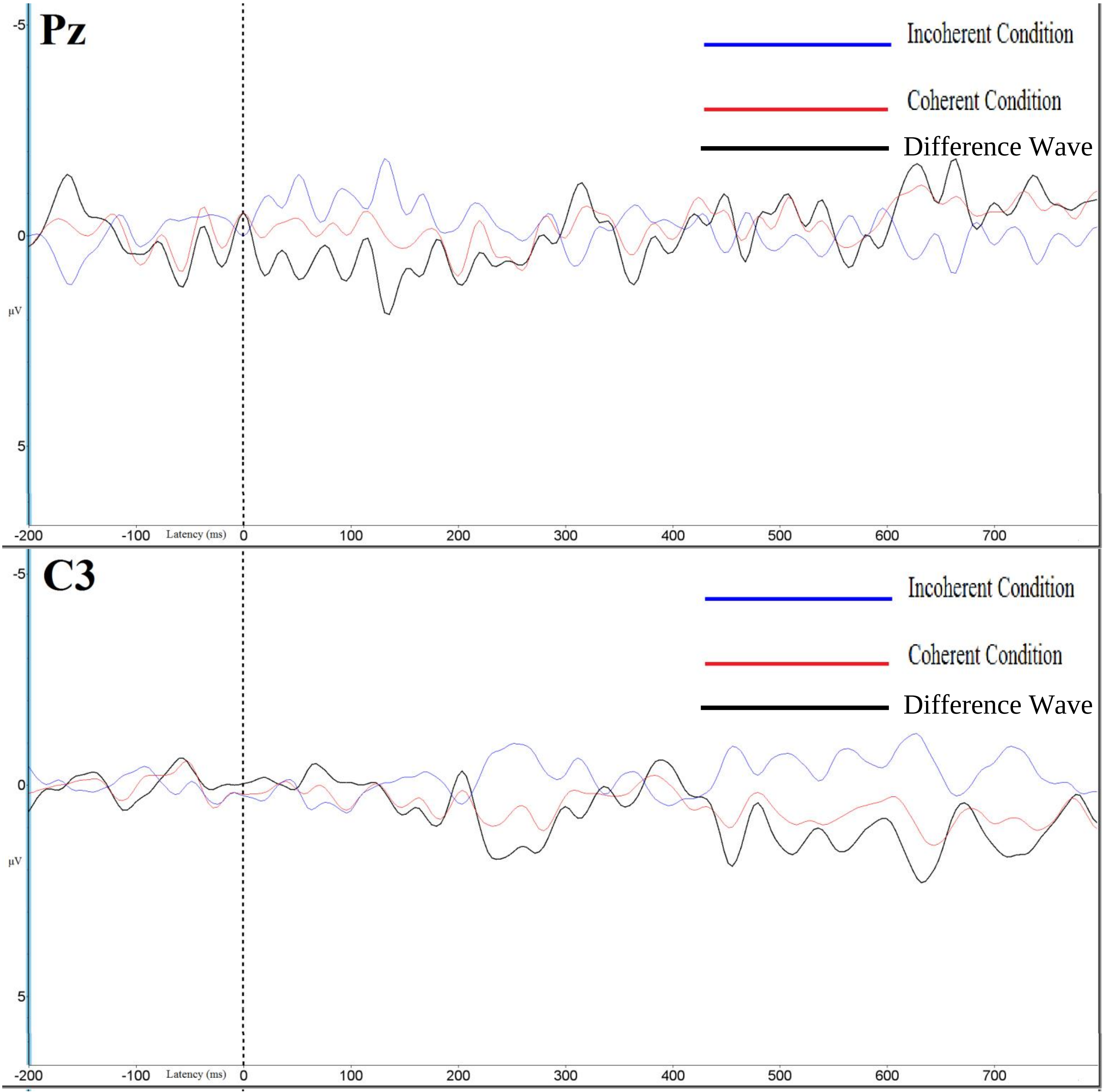
Cluster Locations	
10/20 System	**EGI 129**
F3	25
Fz	11
F4	124
C3	37
Cz	129 (VREF)
C4	105
P3	53
Pz	68
P4	87
T3	46
T4	109

Figure 1. Representation of the channel clusters which make up the most commonly used electrode scalp sites in the international 10/20 system for the repeated measures ANOVA to test differences in mean amplitude of the ERPs between coherence conditions. Reproduced from Yang, Perfetti & Schmalhofer (2007) with kind permission. [EGI (Electrical Geodesics, Inc.); VREF (vertex reference)].

Results

Data Inspection

Data inspection was first carried out to assess the assumptions for the repeated measures ANOVA. Inspection can be aided by observing what the descriptive statistics, such as mean, median, mode, spread and standardized skewness and kurtosis say about the distribution. The normality of the variables and the presence of outliers were examined through descriptive statistics, histograms and quantile-quantile plots (QQ plots). Data inspection began with a graphical analysis of the data for each condition. From the QQ plots and histograms, the distributions appeared normal. The distribution of variables can usually be inspected accurately in this manner. However, in cases of low sample size such as this, graphical analysis can be misleading. Normality can also be investigated by converting the variable to standardized scores or by using specific statistical tests. In this case, the Shapiro-Wilk test was used. This tests the null hypothesis that the sample came from a normally distributed population (Field, 2009). It is useful in small to medium sample sizes. The Shapiro-Wilk test was not significant for all but three variables (Electrode Regions 5 [p = .01],6 [p = .03] and 10 [p = .001] in the coherent condition), but on inspection of the standardised scores, the violation of normality for these variables appeared to be moderate except for the Electrode Region 10 in the Coherent condition. This violation of normality appeared to be caused by a single extreme outlier. When this was corrected by transforming it to the next closest score or by removing the outlier, the results were similar to the original test so it was decided to keep the value as it was. No action was taken for the other two variables as the ANOVA can handle such moderate violations (Luck, 2005). Sphericity was not an issue as there were only two levels of coherence.



Pz
Incoherent Condition
Coherent Condition
Difference Wave
-5
0
μV
5
-200
-100
Latency (ms)
0
100
200
300
400
500
600
700
C3
Incoherent Condition
Coherent Condition
Difference Wave
-5
0
μV
5
-200
-100
Latency (ms)
0
100
200
300
400
500
600
700

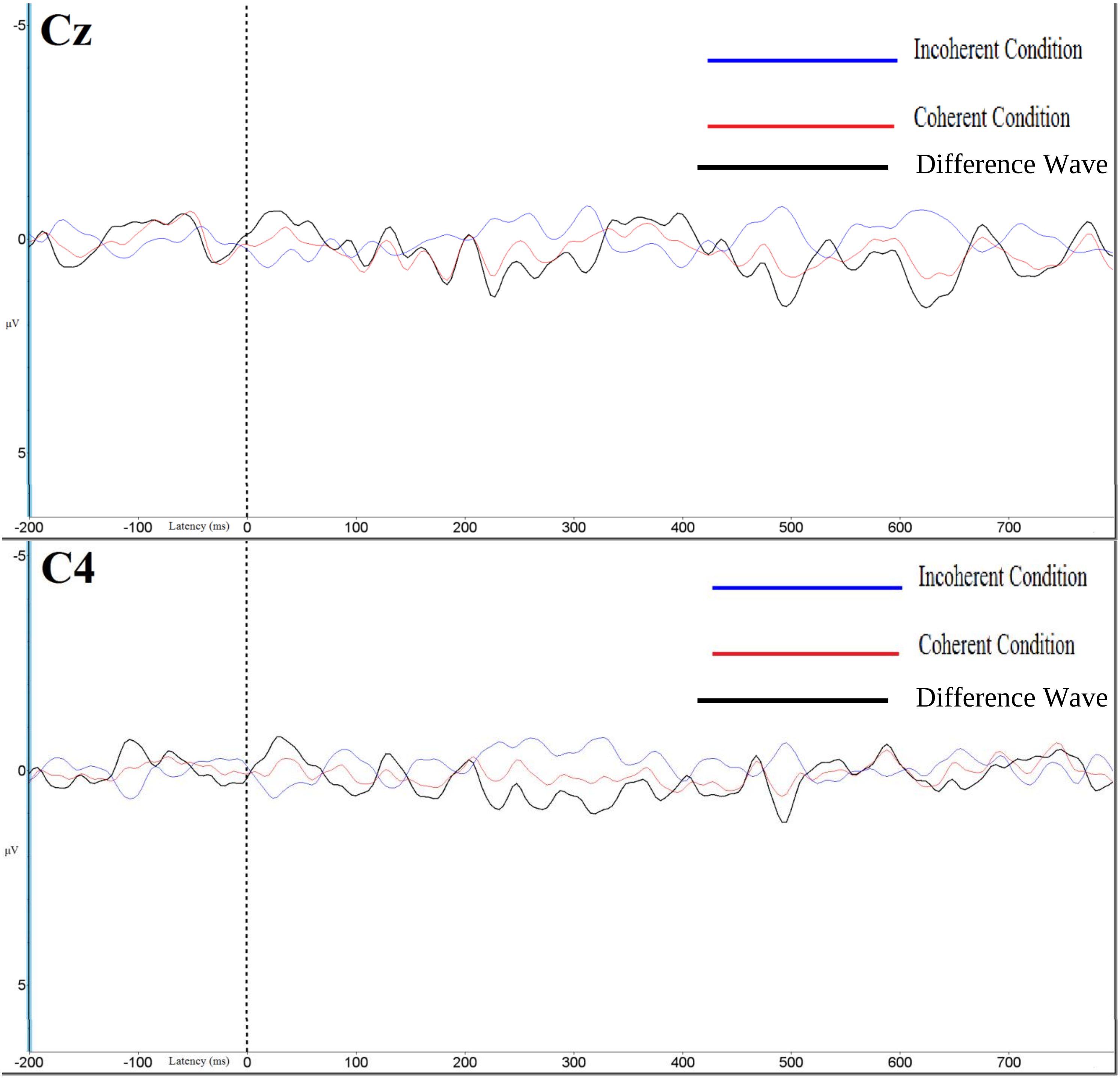


Figure 2. ERPs showing difference in amplitude between coherent and incoherent conditions at P3, C3, Cz and C4 (Red = Coherent condition, Blue = Incoherent condition, Black = Difference Wave).

Coherence Condition and N400 amplitude

Figure 2 shows the difference waves produced from the averaged waveforms in the coherent and incoherent conditions across 4 of the 11 scalp sites. A repeated measures ANOVA was carried out to test whether the N400 amplitude, measured as mean amplitude in the 300 – 500ms post critical word time window across the 11 chosen scalp sites (at F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, T3 and T4) was more strongly negative for critical words in the incoherent condition than for those in the coherent condition. Mean amplitude values for the N400 were more negative for critical words in the incoherent condition (M = .22, SD = 0.19) than those in the coherent condition (M = -.83, SD = 0.23) across all scalp sites, except those at Fz, T3 and T4, but the differences did not reach significance, $F(1,7) = .79$, $p = .41$. Therefore, the null hypothesis was accepted and the alternative hypothesis was rejected.

Table 2**.**

Mean Amplitude in microvolts as a function of story condition across 11 scalp sites.

Mean (SD)	F3	Fz	F4	C3	Cz	C4	P3	Pz	P4	T3	T4
Coherent	.06	.06	-.05	08	.29	-.02	.10	.10	22	-.40	-.18
	(1.10)	(.92)	(.66)	(.85)	(.61)	(.99)	(1.17)	(.87)	(1.02)	(1.26)	(.93)
Incoherent	.35	.06	-.24	-.14	.07	-.26	-.27	-.27	-.40	.18	-.01
	(1.45)	(1.72)	(1.16)	(.48)	(.52)	(.35)	(1.34)	(1.28)	(1.39)	(.59)	(.55)

Discussion

The aim of this study was to investigate whether breaks in global coherence at critical words in a discourse produce a larger N400 effect than critical words at the same point in a coherent discourse. Evidence of a significant difference would provide further support to previous research suggesting that global coherence breaks evoke a similar N400 component to local coherence breaks, as predicted by integrated single-step models (Jackendoff, 2002; MacDonald, Pearlmutter, & Seidenberg, 1994; Nieuwland & Van Berkum, 2006; Trueswell & Tanenhaus, 1994) The results of this study showed that on average the amplitude in the

N400 time window in response to critical words in the globally inconsistent condition was not significantly more negative than of the response to critical words in the consistent condition.

These results suggests that, in this sample, global coherence breaks did not reliably produce an N400 effect, which has been demonstrated reliably for local coherence breaks in previous research (Kutas & Federmeier, 2009). One possible explanation is that the close lexico-semantic relationships across individual words within the stories may have dominated the semantic processing of the critical words, overruling the influence of the situation model. This would contradict the predictions of integrated single-step models and the limited previous research regarding global coherence breaks and N400 amplitude. Taken as they are, the results of this study are more supportive of two-step models which predict that readers produce a local, context-independent representation of the meaning of a noun phrase prior to integrating this representation with the situation model. However, it would be unwise to make statements regarding the relative merits of these two classes of language comprehension models based on this data, because of the small sample size. Nevertheless, it should be noted that only two studies have found the N400 effect specifically in response to critical words causing global coherence breaks in discourses (Hald, Steenbeek-Planting & Hagoort, 2007; Van Berkum et al., 1999).

The results of this study do not mean that global coherence was not maintained - it has been shown to be readily maintained. Rather it could mean that the type of top-down processing required to detect global coherence breaks may take place later. In this case, the situation model may not be available to influence the processing of incoming words until after the typical N400 time window. Evidence for this can be seen in a study by Nieuwland and Van Berkum (2005) in which no N400 effect but a strong P600 effect was found for words which were highly incongruous with the situation model but lexico-semantically related to the local context. As stated earlier, activation and integration are the processes thought to be responsible for maintaining global coherence. It is possible that the integration process is the defining difference between how local and global coherence is processed. For example, both poor and good comprehenders engage in activation successfully, but poor comprehenders have trouble with integration resulting in both groups noticing local coherence breaks but poor comprehenders failing to notice global coherence breaks (Long & Chong, 2001). Perhaps the integration of words into the situation model as it becomes available is where problems occur for poor comprehenders and this takes place in this later time window.

On the other hand, this study suffers from a very small sample size and this may have resulted in Type 2 error occurring. There simply is not enough power in this study to state in certain terms whether stories containing global coherence breaks evoke an N400 effect in relation to coherent control stories. As mentioned earlier, evidence from behavioural studies suggests that global coherence is readily maintained and some further ERP studies have suggested that global coherence breaks produce N400 waveforms with much the same onset, morphology and scalp distribution elicited by local coherence breaks (St. George et al, 1994; Van Berkum et al., 1999, 2003). Adding further support, an experiment manipulating discourse context and word association showed that incongruencies produced a larger N400 amplitude than congruent unassociated critical words (Camblin, Gordon & Swaab, 2007). Studies focusing on related but different language comprehension phenomena suggest that global context does influence the processing of upcoming words and in cases of incongruities succeed in producing N400 effects. For instance, a robust N400 effect for real-world knowledge violations has been reported, for example, stating a train in the Netherlands to be white instead of yellow as they actually are (Hagoort et al., 2004). A later study showed that world knowledge and discourse context interacted dynamically to modulate N400 amplitude so that a carefully constructed text could attenuate the N400 amplitude for world knowledge violations and increase the N400 amplitude for real world facts which do not fit the established context. The largest N400 amplitudes in this study were found when world knowledge violations interacted with an incoherent context (Hald et al., 2007). With the above evidence in mind, it seems likely that given a larger sample size, a similar significant N400 effect would have been found in this study. If such an effect can be replicated, with the methodological recommendations outlined below adopted, it would provide strong evidence for the global discourse context impacting the online processing of incoming words, as is predicted by integrated single-step models.

This study did have a number of strengths. The variation in the timing of the critical word onset negated the need for a jitter to deal with expectancy effects. Even if noticing that some stories were odd led to expecting or seeking incongruencies in the stories, the only way to notice these is through regular language comprehension (as noted by Nieuwland and Van Berkum, 2005) and this would not cause the absence of an N400 effect. The aural presentation of the story stimuli eliminated the need for reading thus reducing eye movement and blink artifacts. The presence of a comprehension question was advantageous as it ruled out the possibility that the lack of an N400 effect was due to lack of attention. Another strength was the care taken to ensure similar tone of voice and pacing of speech across the stories. There were some limitations in the study also. The most important limitation was the small sample size which led to low power and an increased chance of Type 2 error, as noted earlier. In addition to this, the minimum amount of trials per condition was used (30 per condition). Ideally, more would be included so as to improve the signal-to-noise ratio and the increase the statistical power of the ANOVA. However, it is also necessary to be mindful of participants' time commitment, their ability to pay attention for a significant length and the resources necessary to produce useful stimuli. The variation in length of the critical word, which ranged from 195ms to 975ms, may have been problematic as it could have affected the onset of the N400. The subsequent averaging of all the ERPs in the same time window may have attenuated the size of the N400 effect. Unfortunately, there were not enough participants or trials to compare the N400 for short, medium and longer length words. Future studies should keep variation in critical word latency to a minimum and increase the number of trials and participants.

This area of research warrants further attention as it is necessary to replicate the effects of global coherence breaks found in previous studies using a design which allows the comparison of ERPs in response to both local coherence breaks and global coherence break. It would also be useful to augment such replication studies with EEG coherence analysis. Electrophysiological coherence and network analysis provides a measure of changes in synchronous connections and structural coupling between distributed neural networks (Supp et al., 2004). This technique would be useful in clarifying the open question of whether processing of local coherence and global coherence are products of the same brain system or functionally separate ones (Nieuwland & Van Berkum, 2006; Supp et al., 2004). It would also be of value to perform such a study with comparing good and poor comprehenders in order to shed light on why poor comprehenders have problems with the maintenance of global coherence (Long & Chong, 2001).

In conclusion, this study sought to conceptually replicate previous studies showing a larger N400 amplitude for critical words rendered incongruous by the preceding context than for those congruent with the preceding discourse and therefore globally coherent. No such effect was found, though it cannot be conclusively ruled out. Recommendations for future research were made, including the proposal of a design to replicate and extend previous research.

References

Albrecht, J. E., & O'Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(5), 1061-1070. doi:10.1037//0278-7393.19.5.1061

Camblin, C. C., Gordon, P. C., & Swaab, T. Y. (2007). The interplay of discourse congruence and lexical association during sentence processing: Evidence from ERPs and eye tracking. *Journal of Memory & Language*, *56*, 103-128. doi: 10.1016/j.jml.2006.07.005

Cook, A. E., Halleran, J. G., & O'Brien, E. J. (1998). What is readily available during reading? A memory-based view of text processing. *Discourse Processes, 26*(2-3), 109-129. doi:10.1080/01638539809545041

Field, A. (2009). *Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll).* Introducing Statistical Methods. London: SAGE

Fodor, J. A. (1983). *The modularity of mind*. Cambridge: MIT Press.

Forster, K. (1979). Levels of processing and the structure of the language processor. In: W. E. Cooper & E. C. T. Walker (Eds.), *Sentence processing: Psycholinguistic studies presented to Merrill Garrett*. Hillsdale, NJ: Erlbaum.

Hagoort, P., Hald, L. A., Bastiaansen, M. C. M., & Petersson, K. M. (2004). Integration of word meaning and world knowledge in language comprehension. Science, 304(5669), 438-441. doi:10.1126/science.1095455 Hald, L., Steenbeek-Planting, E. G., & Hagoort, P. (2007). The interaction of discourse context and world knowledge in online sentence comprehension. Evidence from the N400. *Brain research, 1146,* 210-218. doi:10.1016/j.brainres.2007.02.054

Jackendoff, R. (2002). *Foundations of language: Brain, meaning, grammar, evolution.* Oxford: Oxford University Press.

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, MA: Harvard University Press.

Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction–integration model. *Psychological Review, 95*, 163–182.

Kuperberg, G. R., Paczynski, M., & Ditman, T. (2011). Establishing Causal Coherence across Sentences: An ERP Study. *Journal of Cognitive Neuroscience*, *23*(5), 1230-1246. doi:10.1162/jocn.2010.21452

Kutas, M., & Federmeier, K. (2009) N400. *Scholarpedia*, *4*(10):7790.

Lehman, S., & Schraw, G. (2002). Effects of Coherence and Relevance on Shallow and Deep Text Processing. *Journal of Educational Psychology*, *94*(4), 738 -750. doi:10.1037//0022-0663.94.4.738

Long, D. L., & Chong, J. L. (2001). Comprehension skill and global coherence: A paradoxical picture of poor comprehenders' abilities. *Cognition*, *27*(6), 1424-1429.

Luck, S. (2005). *An Introduction to the Event-Related Potential Technique.* Massachusetts: MIT Press.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703.

McNamara, D., Kintsch, E., Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14*(1), 1-43.

Millis, K. K., & Just, M. A. (1994). The influence of connectives on sentence comprehension. *Journal of Memory and Language*, *33*(1), 128–147.

Nieuwland, M. S., & Van Berkum, J. J. A. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary change deafness in discourse comprehension. *Cognitive Brain Research, 24(3*), 691-701. doi: 10.1016/j.cogbrainres.2005.04.003.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of cognitive neuroscience*, *18*(7), 1098-1111. doi:10.1162/jocn.2006.18.7.1098

Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The interplay between semantic and referential aspects of anaphoric noun phrase resolution: Evidence from ERPs. *Brain and language*, *106*(2), 119-131. doi:10.1016/j.bandl.2008.05.001

Nieuwland, M. S., & Van Berkum, J. J. A. (2008). The Neurocognition of Referential Ambiguity in Language Comprehension. *Language and Linguistics Compass*, *2*(4), 603-630. doi:10.1111/j.1749-818X.2008.00070.x

Nieuwland, M. S., Otten, M. & Van Berkum, J. J. A. (2007). Who are you talking about? Tracking discourse-level referential processing with event-related brain potentials. *Journal of cognitive neuroscience*, *19*(2), 228-236. doi:10.1162/jocn.2007.19.2.228

O'Brien, E. J., & Albrecht, J. E. (1992). Comprehension strategies in the development of a mental model. *Journal of experimental psychology. Learning, memory, and cognition*, *18*(4), 777-84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1385615

Sanders, T. J. M., & Noordman, L. G. M. (2000). The Role of Coherence Relations and Their Linguistic Markers in Text Processing. *Discourse Studies*, *29*(1), 37-60.

St George, M., Mannes, S., & Hoffman, J. E. (1994). Global semantic expectancy and language comprehension. *Journal of Cognitive Neuroscience, 6(1),* 70-83. Nature Publishing Group. doi: 10.1162/jocn.1994.6.1.70.

Supp, G. G., Alois, C. A., Gunter, T. C., Bernard, M., Pfurtscheller, G., & Petsche, H. (2004). Lexical memory search during N400 : cortical couplings in auditory comprehension. *Medical Informatics*, *15*(7), 3-7. doi:10.1097/01.wnr.0000126219.25235.

Trueswell, J. C., Tanenhaus, M. K., & Garnsey, S. M. (1994). Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution. *Journal of Memory and Language*, *33*(3), 285–318. Doi: 10.1006/jmla.1994.1014.

Van Berkum, J. J. A., Hagoort, P., & Brown, C. M. (1999). Semantic integration in sentences and discourse: evidence from the N400. *Journal of cognitive neuroscience*, *11*(6), 657-71. doi: 10.1162/089892999563724

Van Berkum, J. J. A., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Brain research. Cognitive brain research*, *17*(3), 701-18. doi: 10.1016/S0926-6410(03)00196-4.

Van Berkum, J. J. A. (2008). Understanding Sentences in Context: What Brain Waves Can Tell Us. Current Directions in Psychological Science, *17*(6), 376-380. doi: 10.1111/j.1467-8721.2008.00609.x

Van den Broek, P. (1994). Comprehension and memory of narrative texts: Inferences and coherence. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 539 - 588). San Diego, CA: Academic Press

Van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York: Academic Press.