Generalized Event Knowledge Activation During Online Language Comprehension

Ross Metusalem (rmetusal@cogsci.ucsd.edu)
University of California, San Diego
9500 Gilman Dr. #0515
La Jolla, CA 92093-0515 USA

Marta Kutas (kutas@cogsci.ucsd.edu)
University of California, San Diego
9500 Gilman Dr. #0515
La Jolla, CA 92093-0515 USA

Mary Hare (mlhare@bgnet.bgsu.edu)
Bowling Green State University
Department of Psychology, BGSU
Bowling Green, OH 43403-0228 USA

Ken McRae (kenm@uwo.ca)
University of Western Ontario
Department of Psychology, UWO, Social Science Centre 7418
London, ON, Canada

Jeffrey L. Elman (elman@cogsci.ucsd.edu)
University of California, San Diego
9500 Gilman Dr. #0515
La Jolla, CA 92093-0515 USA

Abstract

Online language comprehension is guided by knowledge regarding real-world events. However, it remains unclear whether activation of event knowledge during language comprehension is constrained by the linguistic context or is generalized, including a wide variety of information associated with the event even if that information has not been mentioned previously and does not satisfy constraints imposed by the local linguistic context. The present study addresses this issue by analyzing event-related brain potentials recorded as participants read brief scenarios describing typical real-world events. The amplitude of the N400 elicited by a contextually anomalous word was reduced if that word was related to the event described. This result suggests that online language comprehension involves construction of rich event representations that include information beyond that which is relevant to the processing of the current linguistic input.

Keywords: event knowledge; online language comprehension; event-related potentials; ERP; N400

Background

Online language comprehension is a rapid and incremental process guided by a wide variety of information sources. Some researchers characterize this process as the incremental mapping of linguistic structure onto real-world event structure, mediated in part by the comprehender's prior knowledge associated with the described event (e.g., Altmann & Mirković, 2009). Recent research has highlighted the importance of event knowledge to online language comprehension. At the lexical level, priming studies have shown, for example, that verbs activate agents, patients, and instruments typically associated with the specific actions denoted by the verbs (Ferretti, McRae, & Hatherell, 2001), that agent, patient, instrument, and location nouns activate verbs denoting the events in which they typically participate (McRae, Hare, Elman, & Ferretti, 2005), and that word triplet priming with lexically unassociated primes and targets reveals rapid activation of script information (Chwilla & Kolk, 2005). Such findings suggest that processing words in isolation activates event knowledge, resulting in subsequent activation of other entities and/or actions associated with the event.

At the sentential level, self-paced reading (Bicknell, Elman, Hare, McRae, & Kutas, 2008) and eye-tracking (Kamide, Altmann, & Haywood, 2003) studies have demonstrated that comprehenders can rapidly integrate information provided by a verb in combination with its preceding agent in order to predict likely upcoming patients. For example, Kamide et al. monitored participants' eye movements around a visual scene as the participants listened to sentences such as *The [man/girl] will ride the [motorbike/carousel]*. They found more anticipatory looks to the picture of the motorbike when the agent of *ride* was *man* than when the agent was *girl* (and similarly more looks to the carousel for *girl will ride* than for *man will ride*), even though both a motorbike and carousel are equally plausible patients of the verb *ride*. This result demonstrates that thematic role assignment is not guided by the verb alone,

but crucially by knowledge associated with the event denoted by the agent-verb combination. Online language comprehension thus makes rapid use of event knowledge.

While event knowledge is clearly important to linguistic processing, the specificity of the event knowledge activated during comprehension remains an open question. Is the activated event knowledge general, containing a wide variety of salient features associated with the event? Or is activation restricted to only what is relevant to the current linguistic context? Consider the passage in (1):

(1)   A huge blizzard swept through town last night. My kids ended up getting the day off from school. They spent the whole day outside building a big snowman in the front yard.

Given the previously discussed findings, it is likely the case that at the point of reading *building a*, a comprehender's knowledge regarding "playing in the snow" events allows for *snowman* to become activated as a likely patient (as opposed to *house*, for example). Is activation of event knowledge at this point limited to this feature of the "playing in the snow" event? If this were the case, it would indicate that event knowledge activation is constrained by the linguistic context, limiting activation to those event features that are relevant to the processing of the current linguistic input. However, it is also possible that an entire body of "playing in the snow" knowledge is activated in reading this passage. This knowledge might include, for example, the fact that the children are probably jackets, hats, and mittens, even though this has not been explicitly mentioned and is not directly relevant to comprehending the passage. This study seeks to determine whether event knowledge activation during comprehension involves rich, generalized representations such as this, or if it is limited to what is currently relevant given the linguistic context.

In the present study, participants read brief passages describing typical real-world events. The final sentence of each passage contained either a highly expected target word (in the above example, *building a snowman*) or one of two contextually anomalous target words: one related to the event described (e.g., *building a jacket*) and one unrelated to the event described (e.g., *building a towel*). Participants' EEG was recorded as they read the passages, and the event-related brain potentials (ERPs) elicited by these three target types were contrasted. The analysis focuses on the N400, an ERP component whose amplitude is inversely proportional to the degree to which a word is expected given the preceding context (Kutas & Hillyard, 1984). It was predicted that if language comprehension involves generalized event knowledge activation, then the event-related anomalous target should become activated during the reading of the passage, while the event-unrelated anomalous target should not. This should then result in a graded N400 effect: the smallest N400 to the expected target, the largest N400 to the event-unrelated target, and an intermediate N400 to the event-related target. Such a result would indicate that although both the event-related and event-

unrelated targets violate local linguistic constraints, the event-related target becomes activated by virtue of its association with the event described. More generally, this result would support the notion that language comprehension involves generalized event knowledge activation.

Stimuli

Seventy-two experimental items (scenarios) were constructed. Each scenario consisted of three sentences and described a common real-world event. The first two sentences established the event (e.g., playing in the snow). The final sentence contained one of three sentence-medial target words: a highly expected word, a contextually anomalous word that was related to the established event (event-related anomalous target; ERA), or an equally anomalous word that was unrelated to the established event (event-unrelated anomalous target; EuRA). Each experimental item thus had three possible target words, giving the experiment three conditions.

Expected Targets

The expected targets were obtained via a cloze task in which participants read each scenario up to the word preceding the target and were asked to provide the single word most likely to come next. Participants completed the cloze task through an online form. Scenarios were presented one at a time, with all three sentences presented in paragraph format. The third sentence left off at the word preceding the target, and participants provided the most likely upcoming word in a blank text field before moving onto the next scenario. Responses could not be modified once entered.

Thirty undergraduates (twenty-three women) at the University of California, San Diego participated for course credit. All were native English speakers. Cloze probability was calculated as the percentage of participants who provided a particular response for the given scenario. The response with the highest cloze probability was chosen as the expected target word. Across the seventy-two items, mean cloze probability of the expected target was 0.81, with a standard deviation of 0.17.

Event-Related Anomalous Targets

To obtain the event-related anomalous targets (ERAs), a new group of participants completed a norming task in which they provided a list of people or things most likely to be present at each event. Participants completed this task through an online form. Scenarios were presented one at a time in paragraph format, with the expected target word obtained in the previous cloze task now filled in. Participants were instructed to read each item and to paint a mental picture of the event described. They were told that their picture would likely include prominent people or things that would participate in the event, but were not explicitly mentioned in the text. They were asked to provide up to five responses for each scenario. Responses could not be modified once entered.

Table 1: Norming results for the three rotation groups and the stimuli set overall[1]

		Group 1	Group 2	Group 3	Overall
Expected targets	Cloze probability	0.81	0.80	0.80	0.81
	Log frequency	6.95	7.01	6.88	6.95
	Orthographic length	5.58	5.71	5.75	5.68
Event-related targets	Cloze probability	0.00	0.00	0.00	0.00
	Log frequency	6.89	6.91	6.76	6.86
	Orthographic length	5.96	5.96	5.71	5.87
	Event-relatedness score	88.1	86.5	92.5	89.1
Event-unrelated targets	Cloze probability	0.00	0.00	0.00	0.00
	Event-relatedness score	0.00	0.13	0.04	0.06

Forty-five undergraduates (twenty-six women) at the University of California, San Diego participated for course credit. All were native English speakers, and none had participated in the cloze task. Each participant's responses for a particular scenario were given weighted scores based on response order (i.e., 5 for the first response, 4 for the second response, 3 for the third response, etc.). The highest scoring response that was not provided as a response in the previous cloze task (i.e., had a cloze probability of zero) was chosen as the ERA for the item. In a small number of instances, the highest scoring zero-cloze response was deemed to be a contextually sensible continuation of the scenario, despite having not been provided as a response in the cloze task. In these cases, the next highest scoring zero-cloze response was chosen. Across the seventy-two scenarios, the mean event-relatedness score of the ERA was 89.1, with a standard deviation of 34.4.

Event-Unrelated Anomalous Targets

Event-unrelated anomalous targets (EuRAs) were obtained by shuffling the ERAs across scenarios. Before this was done, the seventy-two experimental items were split into three rotation groups of twenty-four items each, allowing for three experimental lists to be constructed by rotating each group through the three conditions across the three lists. To minimize variability across the experimental lists, the rotation groups were matched on the following factors: mean cloze probability, log frequency, and orthographic length of the expected target; mean event-relatedness score, log frequency, and orthographic length of the ERA.

ERAs were shuffled across the items within each rotation group to obtain the EuRAs, thereby matching the ERAs and EuRAs for lexical factors within each group. EuRAs were chosen such that they were all zero-cloze, and in all but two of the seventy-two scenarios, EuRAs had event-relatedness scores of zero. (The two exceptions had extremely low event-relatedness scores of 1 and 3.) In addition, the shuffling was done in such a way as to match the ERAs and EuRAs within each scenario for animacy and concreteness. This was done so that if the ERA constituted an animacy or concreteness violation with respect to preceding context, the EuRA constituted the same violation. The norming results for each rotation group and the stimuli set overall are presented in Table 1.

Lexical Associations

One concern in constructing the stimuli was that the ERAs might be significantly more likely than the EuRAs to be lexically associated with the expected word. Such a confound might undermine the experiment, as the predicted graded N400 effect could then be accounted for by priming of the ERA by the expected word, as opposed to the activation of event knowledge. The University of South Florida Free Association Norms (Nelson, McEvoy, & Schreiber, 1998) were consulted to ensure that the ERAs and EuRAs were on average associated to equal degrees with their corresponding expected targets. Sixty-five of the seventy-two scenarios' expected targets appeared in the Nelson norms; mean association scores for the ERAs and EuRAs were calculated across these sixty-five items. The mean association score for the ERAs was 0.0005, and for the EuRAs was 0.0001.[2] These extremely low mean association scores and the small difference between them were deemed acceptable for the purposes of the study.

---

[1] Recall that the event-unrelated and event-unrelated targets consist of the same lexical items. Mean log frequency and orthographic length of the EuRAs is therefore equal to that of the ERAs and is not reported in the Table 1.

[2] The association score for a given word is calculated simply as the proportion of participants that provided the word (ERA or EuRA) in response to the cue word (expected target). The mean scores for ERAs and EuRAs thus correspond to one response per two thousand participants and one response per ten thousand participants, respectively.

Experiment

To examine the specificity of the event knowledge activated during online language comprehension, participants' EEG was recorded as they read the carefully constructed scenarios described in the previous section. To review, each scenario consisted of three sentences. The first two sentences established a typical real-world event. The third sentence contained one of three possible target words: a highly expected word, an anomalous but event-related word (ERA), or an anomalous and event-unrelated word (EuRA). It was hypothesized that the expected word would elicit the smallest N400, the EuRA would elicit the largest N400, and the ERA would elicit an intermediate N400. Such a finding would indicate that although both the ERA and EuRA violated local linguistic constraints, the ERA was activated through the activation of generalized event knowledge by the preceding context.

Methodology

Materials The materials consisted of seventy-two scenarios constructed according to the previously discussed criteria. Three experimental lists were created based on the grouping of the seventy-two items into the three rotation groups, such that each experimental item occurred exactly once in each condition across the three lists and exactly once in each list. In addition to the seventy-two scenarios, twenty-four fillers were included. Like the experimental items, these were three-sentence scenarios describing real-world events. None contained any anomalous words. Presentation order of experimental items and fillers was fully randomized for each participant.

Participants Thirty undergraduates (twenty-two women) at the University of California, San Diego participated for course credit. All were right-handed native monolingual English speakers with normal or corrected-to-normal vision. None reported any history of learning or reading disabilities or neurological or psychiatric disorders.

Procedure Participants sat in an electromagnetically shielded chamber and read each scenario from a computer monitor. The first two sentences of each scenario were presented in paragraph format. Once participants understood the two sentences, they pushed a button to advance to the final sentence. The final sentence was presented via rapid serial visual presentation (RSVP) with a stimulus onset asynchrony (SOA) of 350ms and a stimulus duration of 200ms. After the offset of the final word, participants answered a yes-no comprehension question before advancing to the next trial. Response hand was
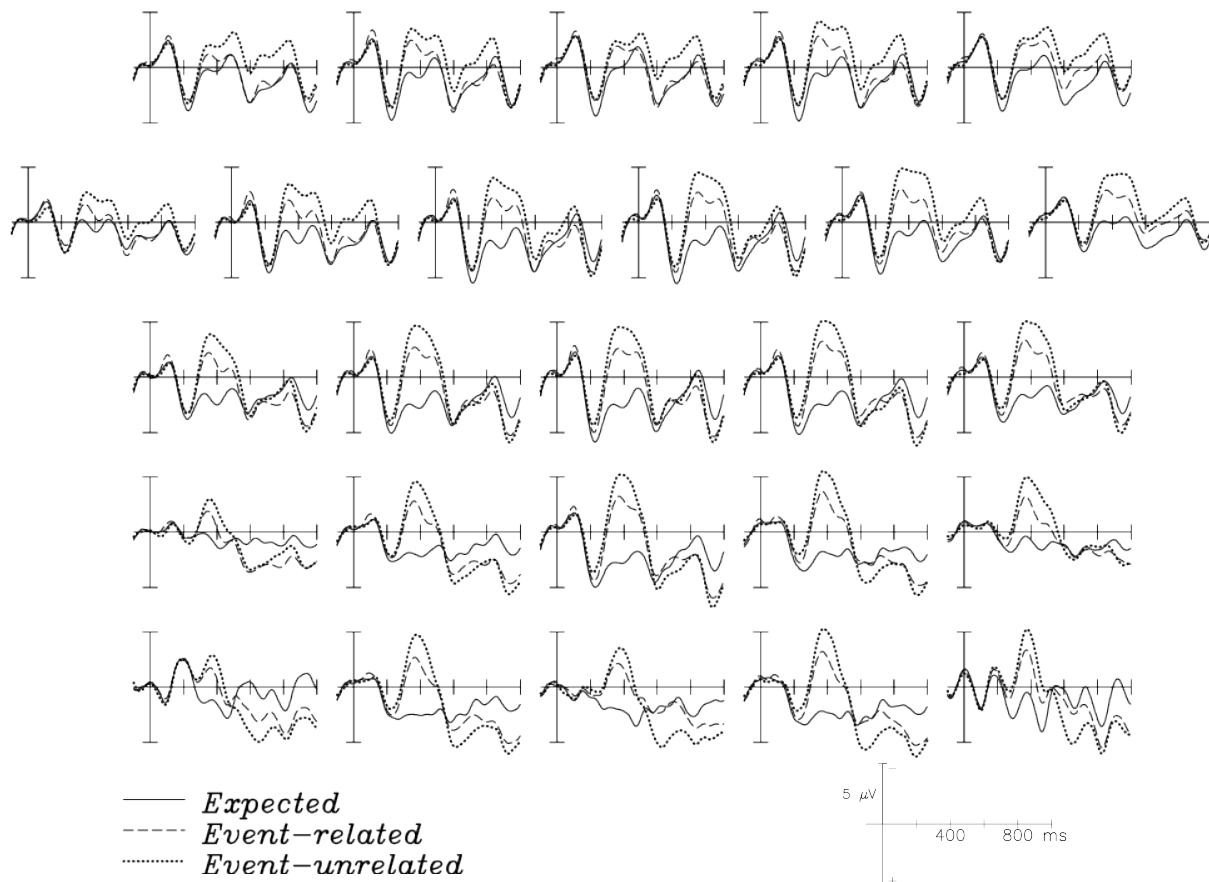

Figure 1: Grand average ERPs at all electrode sites

1061

counterbalanced across participants.

EEG was recorded from twenty-six electrodes distributed evenly over the scalp, referenced online to the left mastoid and re-referenced offline to the average of the left and right mastoids. Electrodes were placed on the outer canthus and infraorbital ridge of each eye to monitor eye movements and blinks. All electrode impedances were kept below 5KΩ. EEG was amplified with Nicolet amplifiers with a bandpass of 0.016 to 100 Hz and digitized at a rate of 250 samples per second.

Results

Before analysis, all epochs containing artifacts caused by blinks, eye movements, muscle tension, channel drift, or amplifier blocking were rejected offline. Participants' responses to the comprehension questions were analyzed to ensure that each participant was reading the scenarios for comprehension. Only one participant scored below 90% correct (88.9%), indicating that participants were comprehending the scenarios.

EEG was time-locked to the onset of the target words and was first averaged within participants to obtain individual participant averages for each condition. These individual participant averages were then averaged together to obtain a grand average ERP waveform for each condition. Figure 1 contains the grand average ERPs at each electrode site, arranged according to the distribution of electrodes over the scalp (i.e., frontal electrodes at the top, posterior electrodes at the bottom; midline electrodes in the middle, lateral electrodes to the side); Figure 2 presents a close-up of the grand average ERPs at the midline parietal recording site (Pz). To conduct an analysis of N400 amplitude, mean amplitudes from 200 to 500ms post-stimulus onset (relative to a 500ms pre-stimulus baseline) at each electrode for each participant were entered into a repeated measures ANOVA. A main effect of Condition was obtained [$F(2,58)=38.33$, $p<0.0001$], as was a Condition X Electrode interaction [$F(50,1450)=7.26$, $p<0.0001$]. Planned comparisons revealed the event-unrelated condition to be significantly more negative than the event-related condition [$F(1,29)=13.00$, $p<0.01$], which in turn was more negative than the expected condition [$F(1,29)=35.44$, $p<0.0001$]. This result confirms the predicted graded N400 effect: expected targets elicited the smallest N400, event-unrelated anomalous targets the largest N400, and event-related anomalous targets an intermediate N400. Analysis of the distribution of the N400 effect revealed a significant Condition X Hemisphere interaction [$F(2,58)=9.69$, $p<0.001$], a significant Condition X Laterality interaction [$F(2,58)=15.14$, $p<0.0001$], and a significant Condition X Anteriority interaction [$F(6,174)=4.96$, $p<0.01$], indicating that the N400 effect exhibited a posterior, slightly right-lateralized distribution across the scalp.

Discussion

Previous research has demonstrated the important role that event knowledge plays in online language comprehension.

However, the specificity of event knowledge activation has remained an open question. The results of the present study suggest that activated event knowledge is general, containing elements beyond what is relevant to the processing of the current linguistic input. This conclusion is supported by a reduction in N400 amplitude to a contextually anomalous word when that word is related to the event being described (and is crucially unrelated to the most expected word, as determined by consulting the South Florida Free Association Norms). This result shows that a wide range of event-relevant information is activated during online language comprehension, as opposed to only event-relevant information that meets local linguistic constraints.

It is important to note that while an event-related anomalous word elicits a reduced N400, it still elicits a larger N400 than a highly expected word. In the present study, the expected targets were more plausible patients of the preceding verbs than were the event-related targets. It is thus possible that event-related elements are activated in a gradient fashion, with those that satisfy the constraints imposed by the local linguistic context receiving greater activation. As the aforementioned study by Ferretti et al. (2001) suggests, verbs encode thematic roles in an event-specific fashion. According to this view, expected targets were closely related to the specific event denoted by the verb itself, whereas the event-related targets were related to the event described by the scenario as a whole but unrelated to the specific event denoted by the verb. This suggests that a word related to the event conveyed by the global linguistic context will receive even greater activation if it is also compatible with the event denoted by the verb.

While it is argued here that the N400 reduction for event-related targets results from activation of event knowledge, it might be argued that such a finding could arise if the event-related targets were lexically associated with the words in the preceding contexts to a greater degree than the event-unrelated targets. Associations between targets and their contexts were quantified using Latent Semantic Analysis (Landauer, Foltz, & Laham, 1998; http://lsa.colorado.edu/), a method for assessing word associations through analysis of the distribution of words in large-scale corpora. Each target received an association score between 0 and 1, with expected targets receiving a mean score of 0.276, event-related targets a mean score of 0.268, and event-unrelated targets a mean score of 0.220. A paired t-test confirmed the difference between event-related and event-unrelated targets (p=0.002). While this result suggests that the event-related targets were more strongly associated with the words in the preceding contexts than were the event-unrelated targets, this is in fact compatible the event knowledge account. Language describing real-world events will undoubtedly exhibit statistical regularities mirroring the structure of the real-world events themselves, and thus lexical co-occurrence measures calculated over large corpora should reflect the event structures encoded in event knowledge. In addition, event knowledge presumably can be derived from experience both in the real world and with language
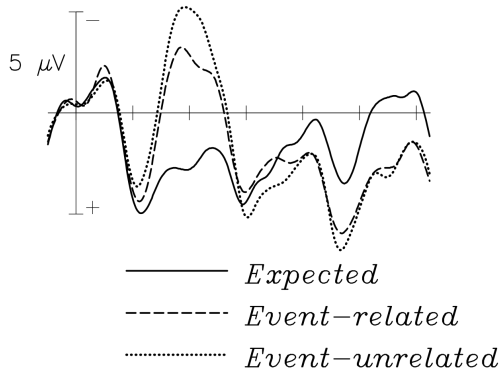
Figure 2: Grand average ERPs at Pz

describing that world, meaning that co-occurrence patterns in language likely result in part from the structure of real world events and contribute to knowledge regarding those events. Lexical co-occurrence, event structure, and event knowledge are thus tightly linked, and it is unlikely that the reported results stem from lexical co-occurrence independent of event knowledge activation. This reasoning raises interesting questions regarding the interplay between knowledge of linguistic regularities and event knowledge during online language comprehension, which is an area worthy of future research.

The finding that event-related anomalous words elicit a reduced N400 stands in contrast with a previous study by Traxler, Foss, Seely, Kaup, and Morris (2000). In their study, the authors examined whether facilitated word processing during sentence comprehension might be captured by a schema-based account in which linguistic input activates a precompiled knowledge structure pertaining to the event being described. In Experiment 1, participants' eye movements were monitored as they read sentences such as *The [lumberjack/young man] chopped the axe early in the morning*. First fixation and gaze durations for the target word (*axe*) did not vary with the target's compatibility with the event denoted by the combination of the agent and verb (i.e., *axe* was read equally fast following *The young man chopped* as it was following *The lumberjack chopped*), suggesting that reading *The lumberjack chopped* did not activate a "lumberjacking" schema that contains an axe as a prototypical instrument.

Given the present finding, the result reported by Traxler et al. is quite surprising. According to the account put forth here, reading *the lumberjack chopped* should activate generalized "lumberjacking" knowledge that would likely include an axe. The source of this apparent discrepancy is unclear, although it is possible that the stimuli used here activated event knowledge more strongly than the stimuli used by Traxler et al. It is also possible that Traxler et al.'s stimuli included target words that were on average less strongly associated with the event being described, or it may simply be the case that such eye movement measures are not sensitive to the effect in question. Further examination is necessary to determine whether the findings reported by

Traxler et al. do in fact stand in contrast with those reported here, or if they are due to methodological differences.

Conclusion

Event knowledge plays an important role in online language comprehension. The present study demonstrates that activation of event knowledge is not constrained by the linguistic context, but instead is highly general, including a variety of information that is not necessarily relevant to the processing of the current linguistic input. This finding provides further support for the intimate link between language comprehension and real-world experience.

Acknowledgments

References

Altmann, G. T. M. & Mirković, J. (2009). Incrementality and prediction in human sentence processing. *Cognitive Science, 33*, 583-609.

Bicknell, K., Elman, J. L., Hare, M., McRae, K., & Kutas, M. (2008). Online expectations for verbal arguments conditional on event knowledge. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th Annual Conference of the Cognitive Science Society* (pp. 2220-2225). Austin, TX: Cognitive Science Society.

Chwilla, D. J. & Kolk, H. H. J. (2005). Accessing world knowledge: Evidence from N400 and reaction time priming. *Cognitive Brain Research, (25)*, 589-606.

Ferretti, T., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44(4)*, 516-547.

Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence of anticipatory eye movements. *Journal of Memory and Language, 49(1)*, 133-156.

Kutas, M. & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307*, 161-163.

Landuaer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to Latent Semantic Analysis. *Discourse Processes, 25*, 259-284.

McRae, K., Hare, E., Elman, J. L., & Ferretti, T. (2005). A basis for generating expectancies for verbs from nouns. *Memory & Cognition, 33(7)*, 1174-1184.

Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. http://w3.usf.edu/FreeAssociation/.

Traxler, M. J., Foss, D. J., Seely, R. E., Kaup, B., & Morris, R. K. (2000). Priming in sentence processing: Intralexical spreading activation, schemas, and situation models. *Journal of Psycholinguistic Research, 29(6)*, 581-595.