Sentence relevance and memory retrieval in discourse comprehension: N400 and Lr NERP effects

Simona Amenta [1] **- Sara Eralti** [2]

[1] *Department of Psychology, University of Milano-Bicocca, Milano, Italy*

[2] *Department of Psychology, Catholic University of the Sacred Heart, Milano, Italy*

simona.amenta@unimib.it

Abstract

The present study explored the effect of different degrees of relevance in discourse comprehension by using ERPs analysis. A principle of pragmatic relevance is supposed to guide inferential mechanism underlying discourse processing. Discourse level comprehension needs a system of predictions about which information is more relevant in order to process the ongoing sentence meaning. This system should construct a specific mental model, where inferences related to the present sentence are stored and maintained. Three degrees of relevance of a new information (new sentence) with respect to an old information (target sentence) were manipulated: directly relevant; indirectly relevant; not relevant. Twenty-one subjects participated to the experiment and they were asked to try to comprehend a set of two paired sentences (old-new paired sentences) based on their conceptual relevance. Two negative deflections, peaking respectively at about 410 msec post-stimulus (N400), more right anterior-centrally distributed, and at about 550 msec (late negativity, LrN), more right central localized, were found. Repeated measures ANOVA found that the amplitude of both the N400 and LrNis modulated by the degree of relevance and by the strength of the underlying associations between the two sentences. Indirect relevance resulted in increased negativities in comparison with direct relevance. Contrarily, non-relevant condition did not produce an increasing in N400 and LrNamplitude. Unrelevance of the knowledge related to the actual mental model of sentences may induce a rapid and costless discarding of non pertinent information. The conclusive inference is that a subset of neural processes responding to degree of relevance of information is separable and cortically more frontally and centrally localized. Functional differences between N400 and LrN for relevance were discussed.

Keywords: Relevance; Memory; Mental model; ERPs; N400; LrN

---

1. Introduction

The construction of a mental model that encodes information about the objects and events related to a sentence meaning is crucial for discourse comprehension (Otten & Van Berkum, 2007). New information is constantly brought into the mental model, and connections to previous information are made in order to ensure discourse coherence. Updating a mental model involves the construction of conceptual dependencies, in order to link incoming information to already available information (Balconi, in press; Ditman & Kuperberg, 2007). These dependencies are based on inferential processes related to a principle of pragmatic relevance (Wilson & Keil, 1999; Wilson & Sperber, 2003).

In this process, two factors seem to play a key role: contextual constrain and memory (Otten & Van Berlum, 2007; Gerrig & O'Brien, 2005; Kutas & Federmeier, 2000).

Discourse context provides a variety of information that contributes to the activation of relevant interpretative background. The words contained in a text (whether it is oral or written) contribute to updating a mental model in two ways: at a message-level, by providing lexico-semantic and world knowledge-related information, allowing the prediction of upcoming words within a sentence; and, at a more general discourse-level, by providing semantic constraints though the activation of related information stored in long-term memory (Van Berkum, 2008; Otten & Van Berkum, 2007).

In other words, discourse comprehension, far from being a passive elaboration process of incoming verbal stimuli, is a proactive process governed by a system of predictive inferences about what information is coming next (Van Berkum, 2010). Thus, inferential processes select the most relevant information provided by unfolding sentences, in order to update the discourse model (Kamide et al., 2003; Levinson, 2000).

1.1. Relevance, mental models and memory load

According to the communicative principle of relevance, every sentence (or other act of ostensive communication) creates an expectation of optimal relevance (Sperber & Wilson, 1995). Relevance is defined as a function of cognitive effects and the effort any new information requires in order to be processed: the greater the cognitive effect achieved by processing a specific assumption in a specific context, and the smaller the processing effort required, the greater would be the degree of relevance. Thus, we intend relevance as an index of the level of contribution of a conceptual space to the core meaning of a sentence or a discourse (Noveck & Sperber, 2007; Sartori et al., 2006). Eachcontext makes a variety of assumptions manifest, and each individual has access to a variety of contexts. Therefore, the aim of the individual is to choose the best possible combination of assumptionswithin a context, and this choice is once again guided by the search for maximal relevance. Thus, the comprehension process does not involve computing all possible interpretations of a given context (Federmeier, 2007; Myers & O'Brien, 1998): a unique feature of the relevance-guided comprehension process is that it involves a mutual adjustment of content, context and cognitive effects, in order to maximize the value of any information, and to reduce at a minimum the cognitive effort. For example, given the context in (10), the assumption in (11) would be more relevant respectively than those in (12) and (13):

(10) People, who are interested in football match, generally love all kind of sport.
(11) John, who is a football fan, loves athletics.
(12) John, who is a football fan, plays tennis.
(13) John, who is a pilot, plays guitar.

(11) is more relevant than (12) on the contextual-effect side: it has a contextual implication related to (10) which (12) lacks. (11) and (12) are more relevant than (13) also on the processing-effort side: (13) contains additional material which is more costly to process, but yields no extra effects, since it is not related to (10).

In this process, memory has a central role, in fact, as we read or hear a sentence within a context, potentially relevant new information wanes without costs, as a function of how long-term memories are reactivated by current inputs (Gerrig & O'Brien, 2005). In other words, long-term memory could predict contents that might soon be needed. This predictive mechanism would be sensitive to complex sentence-level cues, such as conceptual links and semantic relationships between concepts (Federmeier, 2007). Thus, message-level representation of discourse context can facilitate the retrieval of concepts from long-term memory, and, at the same time, guides the search for new information, which is in turn considered plausible in light of the discourse modelthat is being constructed (Kutas & Federmeier, 2000). In this respect, memory retrieval and prediction of upcoming contents are integrated processes, which contribute to the definition and re-definition of discourse-level representation.

1.1.1. ERPs and discourse relevance processing

ERP studies provided interesting evidence about the relationship between discourse comprehension, relevance and those inferential mechanism on which relevance is based. Specifically a N400 ERP effect was found to be related to contextual expectations and pragmatic relevance in discourse (Camblin, Gordon & Swaab, 2007; Van Berkum, 2004). The so-called N400 effect was supposed to reflect some aspects of how words are related to their semantic context.

In order to explain N400 modulation, two different models have been proposed. Firstly, it was supposed that N400 directly indexed a difficulty in integrating anomalous words within a context (Fishler et al., 1983; Hagoort, Hald, Bastiaansen & Petersson, 2004). The hypothesis that the amplitude of the N400 directly reflects the difficulty in semantic composition and contextual enrichment is often referred to as the "integration view" (Kutas & Federmeier, 2000; Kutas et al., 2006). An important alternative account of the N400's amplitude variation sees it depending on the difficulty of retrieving knowledge associated with context (Gerrig & McKoon, 1998). This second hypothesis is referred to as "memory retrieval account" and it find itself in a better position to explain the variations due to the degree of relevance and general "coherence" of information within a context. Relevance signal is a main factor that influences semantic retrieval. The increased effort for memory retrieval, in fact, maybe due to the effort required by the activation of a richer set of semantic features and associations from the semantic memory representation of the sentence-context (Delong, Urbach & Kutas, 2005). The retrieval of semantic information stored in long-term memory is more easier when it is guided by converging cues elicited by the semantic/pragmatic context, than when it cued only by not-relevant information (Schacter, Addis & Buckner, 2007; Van Berkum et al., 2005). More recently, it has been clarified that N400 is actually modulated by the degree of predictability a specific word has within a specific context. Therefore, N400 does not reflect semantic violation, as it was originally thought, rather the component is index of the relation between the probabilistic re-activation of certain semantic features and the degree of confirmation they find in the unfolding context (DeLong et al., 2005; Otten & Van Berkum, 2007). Its amplitude increases as a function of the difficulty to relate a semantic content tothe context (Kutas & Federmeier, 2000; Kutas, Van Petten & Kluender, 2006).

A series of studies addressed the contribution of N400 in clarifying inferential processes at discourse-level, since the component was found to reflect inferential operations involved in understanding sentence in context (Camblin, Gordon & Swaab, 2007). Thus, context-dependent N400 marksthe degree of fitness of specific conceptual expectations within and interpretative background, with a better fit giving a smaller N400 (Balconi & Pozzoli, 2004; Kutas, Van Petten & Kluender, 2006). It has been shown that N400 is clearly sensitive to the level of contextual support from prior information (sentence or text). Text-based inferences, bridging, and elaborative inferences were tested by using N400 measure, where the degree of supportiveness from previous context (Ditman & Kupernberg, 2007), or direct/indirect relatedness (Yang, Perfetti & Schmalhofer, 2005) were modulated. An increased N400 was revealed in case of a non-supportive and not-relevant context in comparison to contexts that are directly related (referentially explicit or paraphrased). These inference-dependent N400 attenuations can be seen as the result of automatic "anticipatory" processing thatis in turn modulated by the effort sustained by the long-term memory in order to integrate not-relevant information. Burkhardt and Roehm (2007) found that the amplitude of the N400 is modulated by the degree of plausibility of a conceptual dependency and the strength of the underlying associations: indirect dependencies result in an increased negativity. More generally, strongest relations between two entities revealed the most reduced N400 also in discourse-based processes (Kutas & Federmeier, 2000).

To summarize, data collected so far provides strong evidence in favor of the interpretation of the N400's modulation as an index of the easiness or difficulty to retrieve stored conceptual knowledge associated with a meaningful stimulus (at any level of the linguistic representation), which is related to both the stored representations and the retrieval cues provided by the preceding context.

Another index concerned with discourse comprehension and information retrieval from long-term memories retrieval is a late negative component, named LrN, which is consistently found in experimental manipulation inducing the re-analysis and re-establishment of discourse coherence (e.g. Arzouan, Goldstein & Faust, 2007; Otten et al., 2007). A memory-based mechanism is thought to intervene in order to complete the integration process of semantic level, after the mismatch between semantic cues and discourse context is detected (Friederici, Steinhauer & Frisch, 1999; Mecklinger et al., 1995). More recently a similar ERP effect was found in response to mismatching condition in discourse comprehension (Otten & Van Berkum, 2008).

In the present paper we explored the effect of degree of relevance on N400 modulation. In particular, we investigated the influence that relevance has on the inferential process activated when new sentences must be related to previous sentences. To our knowledge, no previous studies have widely explored the direct link between the modulation of degrees of relevance and N400 variations. We expect the N400 amplitude to reflect the computational resources used in retrieving the expected/unexpected meaning from long-term memory, and made salient (relevant) by the contextual information. Secondly, we expect the late negativity (LN) to vary as a function of lateprocesses activated by the lack of direct conceptual links between. ERPs and behavioral data (relevance judgment) will be compared in order to directly test the convergence between subjective evaluation on relevance and cortical response to relevance modulation.

2. Experiment

2.1. Method

2.1.1. Subjects

Fifteen students (12 women, 9 men) enrolled in the psychology faculty of the Catholic University of Milan took part in the experiment. The mean age was 22.7 years (SD = 1.46, range: 21-25). All subjects gave informed written consent for participating in the study. The participants were native Italian speakers with normal or corrected-to-normal vision, and they were all right-handed. Exclusion criteria were history of psychopathology or neurological disorders for the subjects or immediate family.

2.1.2. Stimuli

The stimulus materials consisted of sentences, organized in different set of four sentences each one. Each set was composed by a target sentence followed by three new sentences that differed in the degree of relevance with respect to the target, respectively: (a) relevant sentences, (b) indirectly relevant sentences, and (c) not-relevant sentences. Each sentence was composed by the following syntactic structure: subject-verb-past participle. The last final word (participle) was considered as the critical word, since it allowed the subject to establish the direct/indirect relevance of the sentences for the target (Balconi & Pozzoli, 2004; 2005). The frequency of word occurrence and the word length were checked because of the potential effect of those two factors for the N400 component (Osterhout & Nicol, 1999). We conducted a successive item analysis to eliminate possible effect owing to the cross-condition variability of the critical word.

Three different degrees of relevance of the information contained in the new sentence in relation to the information presented in the old sentence (target sentence) were predisposed: directly relevant; indirectly relevant; not-relevant. We defined the degree of relevance of the new sentence for its target as a function of the fact the degree to which the two sentences could be intuitively related. In the "relevant" condition, the information contained in the new sentence was directly relevant for the conceptual representation of the target. In this case, a low cognitive effort was required to establish a conceptual link between the two sentences. The "indirectly relevant" condition supposed that the conceptual instances induced by the new sentences were relevant for the target, yet the associative relation between old and new information was not directly manifest; thus a greater cognitive effort would be required to link the two conceptual domains produced by the activated mental model. Finally, in the "not-relevant", the two segments were unrelated, even if potentially meaningful in some of the conceptual contexts accessible to the subject; in this condition, limited contextual cognitive effects were expected to be produced.

The sentences were tested in a pre-experiment phase, in which 11 participants, who did not take part to the experimental phase, were asked to evaluate the perceived link between two-paired sentences (target-new). Those participants were similar to the experimental participants for mean age (23.04 years; SD = 1.35) and were all native Italian speakers, had normal vision, and were right-handed. The relevance was tested by using two different scales: conceptual proximity of the two sentences in terms of inferential process ("How much do you consider the content reported in the new sentence to be conceptually related to the content of the target sentence?"); and automaticity of the conceptual link ("How much did you consider the conceptual relationship between the two sentences automatic?").

In particular, we tested the degree of relevance, conceptual proximity and automaticity of conceptual link using Likert scales 1-7. All the measures showed similar responses by the participants. The relevant new sentences were considered directly relevant for the target (M = 6.25; SD = 1.76), proximal to the concept represented by the target (M = 5.98; SD = 1.19); automatically inferable from the target (M = 6.11; SD = 0.87). Contrarily, indirect relevant and not-relevant sentences showed a more reduced degree of relevance (respectively: M = 3.33; SD 1.13; M = 1.24; SD = 0.77), they were more conceptually distant from the target (M = 3.03; SD = 1.26; M = 1.33; SD = 0.86), and less automatically inferable from the target (M = 2.93; SD = 1.11; M = 1.29; SD = 0.83). The statistical significance of the differences was tested by a set of repeated measure ANOVAs and further contrast effects. We applied the Greenhous-Geisser (1959) correction when we evaluated effects with more than one degree of freedom to protect against Type I error, to check for the violation of the statistical assumption of equal variance of differences between the conditions of within-participants factors. The degree of relevance differed as a function of sentence condition (3), ($F(2,20) = 15.21$, $P = .001$, $\eta^2 = .36$), and types of sentence differed from each other (direct vs. indirect; indirect vs. not-relevant; direct vs. not relevant, all comparisons $P < .01$). The same result was found for the conceptual proximity ($F(2,20) = 11.20$, $P = .001$, $\eta^2 = .33$) and automaticity of inference ($F(2,20) = 15.61$, $P = .001$, $\eta^2 = .36$) relative to the main effect, with all paired comparisons showing $P < .01$.

2.1.3. Procedure

The task was performed in a quiet, darkened room in which the participants were tested one at a time. The sentences were presented orally. Two loudspeakers were placed behind the participant, to the right and left at a distance of 30 cm. The volume was distinctly audible. The sentences lasted for 3 to 4 seconds. They were introduced in a word-by-word presentation with an ISI interval of 300 ms and an inter-trial interval of 500 ms. Each participant was instructed to pay attention to each stimulus. In the warm-up phase, the participants were provided with a few examples of directly relevant, indirectly relevant and not-relevant sentences. None of the sentences presented during the experiment were used as examples, and additional examples were provided if the participants required them. Immediately after the experimental session, subjects evaluated the degree of relevance of the same set of sentences of the experimental phase, on the three scales of relevance, conceptual proximity and automaticity.

2.1.4. EEG recording

The EEG was recorded with a 32-channel DC amplifier (SYNAMPS system) and acquisition software (NEUROSCAN 4.2) (for the procedure see Balconi & Pozzoli, 2004). An ElectroCap with Ag/AgCl electrodes was used to record EEG from active scalp sites referred to the earlobes (10/20 system of electrode placement). Additionally two EOG electrodes were placed at the outer canti of the eyes. All single trial waveforms were screened for eye movement artifacts, electrodes drifting, and electromyogram artifacts by placing electrodes above and below the right eye. The data were recorded using a sampling rate of 500 Hz; an online filter was used during recording (0.1 to 50 Hz), and the impedance was kept below 5 kΩ. After EOG correction and visual inspection only artefact-free trials were considered (6% of epochs were discarded due to the presence of artifacts). Fourteen electrodes were used for the analyses (four central, Fz, Cz, Pz, Oz; ten lateral, F3, F4, C3, C4, T5, T6, P3, P4, O1, O2). Epochs comprised 700ms after stimulus onset and were calculated on a baseline of the 100 ms preceding stimulus presentation.

3. Results

3.1. Behavioral data

One-way ANOVAs were applied to subjects' evaluation of the three scales (relevance; proximity; automaticity), with type of sentence (3) as an independent factor. The participants' judgments were analogous to the pre-experiment findings, with significant main effect for relevance ($F(2,20) = 27.22$, $P = .001$, $\eta^2 = .58$). As expected, "relevance" was rated higher for directly relevant sentences in comparison to indirectly relevant and not-relevant sentences. Moreover indirectly relevant and not-relevant sentences differed from each other (see Table 1) (contrast effects significant $P = .001$). Similar results were obtained for the conceptual proximity ($F(2,20) = 19.16$, $P = .001$, $\eta^2 = .51$), with an increased perceived proximity for relevant in comparison to indirectly relevant and not-relevant sentences. The latter differed also from the indirectly relevant stimuli. Finally, the automaticity of the conceptual link was higher in response to directly relevant sentences in comparison to indirectly relevant and not-relevant sentences, with significant differences also between indirectly relevant and not-relevant sentences ($F(2,20) = 15.09$, $P = .001$, $\eta^2 = .46$).

Table 1. Subjects' ratings on evaluation scales as a function of degrees of relevance

	Relevance		Conceptual proximity		Automaticity	
	M	*(SD)*	M	*(SD)*	M	*(SD)*
Direct relevance	6.16	1.40	5.91	2.17	6.17	0.52
Indirect relevance	3.27	1.30	3.09	1.08	2.80	0.54
Unrelevance	1.49	1.09	1.98	1.60	1.68	0.29

3.2. ERP data

Morphological analysis revealed, in addition to some positive deflections, the presence of two negative peaks, respectively around 400 ms and 600 ms post-stimulus. For all the experimental conditions the two deflections appeared morphologically similar. Statistical analyses considered these two ERP effects separately.

N400 effect

The N400 component was quantified as the mean voltage of the peak in a latency range of 300-480 ms post-stimulus relative to a 100 ms pre-stimulus baseline. The appearance of the critical (final) word was considered as the onset of the time-window. The second measure considered was latency, which reproduces the peak latency detected for N400. All the statistical measures reported here consisted of repeated-measures analyses of variance (ANOVAs) with amplitude or latency measures as dependent variables. The recording site (14) and the condition (3) were considered as independent variables.

Regarding the peak amplitude, N400 varied as a function of condition ($F(2,20) = 15.43$, $P = .001$, $\eta^2 = .47$), site ($F(13,20) = 15.90$, $P = .001$, $\eta^2 = .48$) and interaction condition × site ($F(26,20) = 20.06$, $P = .001$, $\eta^2 = .52$). Specifically, contrast analysis showed an increased amplitude for indirectly relevant sentences in comparison to relevant ($F(1,20) = 6.11$, $P = .001$, $\eta^2 = .29$) and not-relevant ($F(1,20) = 13.29$, $P = .001$, $\eta^2 = .39$) sentences. None of the other comparisons was statistically significant (Table 2a).

Moreover, simple effects showed an increased frontal (Fz) and central (Cz) activity in the indirect relevant condition in comparison to the direct relevant (respectively: $F(1,20) = 23.11$, $P = .001$, $\eta^2 = .50$ and $F(1,20) = 20.15$, $P = .001$, $\eta^2 = .46$) and not-relevant condition ($F(1,20) = 18.88$, $P = .001$, $\eta^2 = .45$ and $F(1,20) = 18.65$, $P = .001$, $\eta^2 = .47$) (Figure 1a-d).

Table 2a. Mean values of N400 ERP amplitude for each condition and electrode site (midline)

	Electrode sites							
	Fz		Cz		Pz		Oz	
	Amplitude[a]							
	M	*(SD)*	M	*(SD)*	M	*(SD)*	M	*(SD)*
Direct relevance	2.26	.31	2.40	.49	1.97	.41	1.02	.33
Indirect relevance	3.97	.37	3.90	.31	2.50	.80	1.53	.47
Unrelevance	2.04	.40	2.07	.34	2.13	.41	1.01	.59
Mean	2.75	.35	2.78	.38	2.20	.62	1.26	.48

Note: [a] = Measured in mVolt.

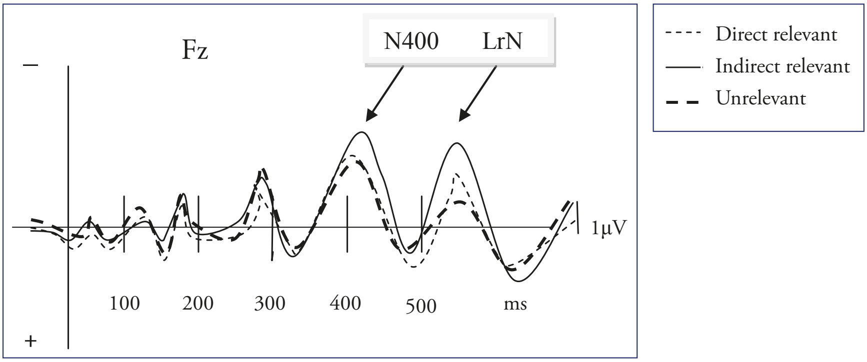

Fig. 1a

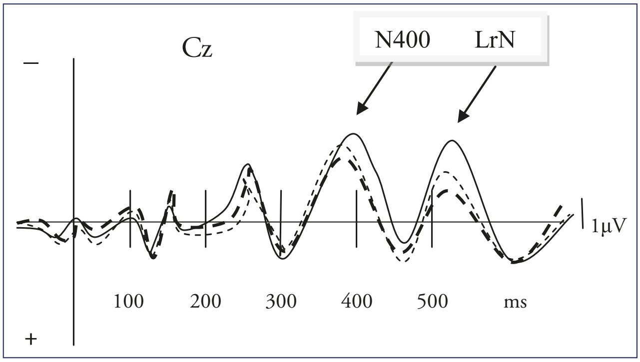

Fig. 1b

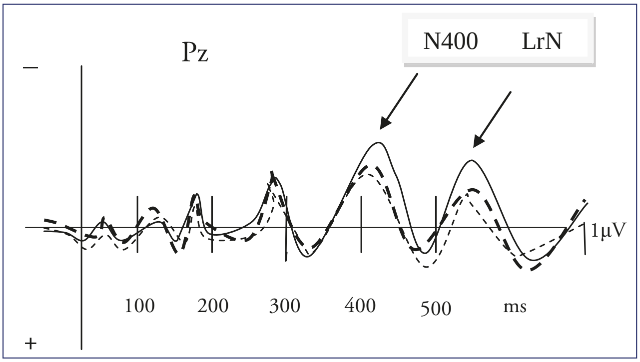

Fig. 1c

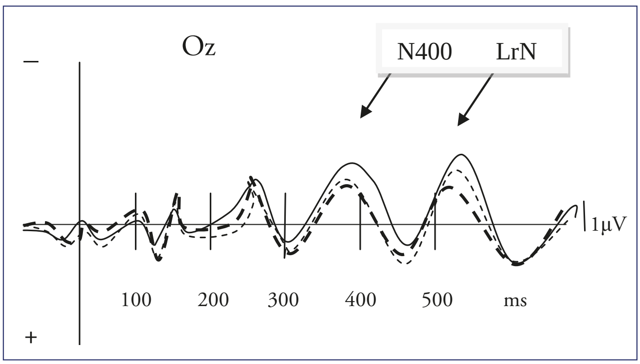

Fig. 1d

Figure 1a-d. Grand-averaged waveform (Fz) of N400 as a function of degrees of relevance (direct relevance; indirect relevance; unrelevance) and cortical sites (Fz, Cz, Pz, Oz)

In order to better assess the brain localization effect of N400, regions of interest (ROI) with eight levels were created as a function of site and lateralization: right and left frontal (F3 and F4), central (C3 and C4), parietal (P3 and P2), and temporal (T3 and T4) site. Condition ($F(2,20) = 23.43$, $P = .001$, $\eta^2 = .54$), ROI ($F(7,20) = 22.21$, $P = .001$, $\eta^2 = .50$), and condition × ROI ($F(14,20) = 20.27$, $P = .001$, $\eta^2 = .48$) were significant (Figure 2). Planned contrasts applied to main effect of ROI showed that right frontal area was more activated than right parietal ($F(1,20) = 33.33$, $P = .001$, $\eta^2 = .62$) and temporal ($F(1,20) = 49.41$, $P = .001$, $\eta^2 = .71$) ones. The same significant difference was found for right central area in comparison to right parietal ($F(1,20) = 21.21$, $P = .001$, $\eta^2 = .48$) and temporal ($F(1,20) = 14.55$, $P = .001$, $\eta^2 = .42$) areas (Figure 2).

Simple effects applied to the interaction condition × ROI showed an increased right frontal and central activity in comparison to parietal and temporal activations for indirect relevant condition (all comparisons $P = .001$).

In order to compare more directly the three conditions, a peak analysis of the condition difference wave was performed. Direct vs. indirect and not-relevant vs. indirectly relevant conditions difference waves were calculated. For the first comparison, it was shown a greater negativity for indirect relevance within Fz ($F(1,20) = 17.52$, $P = 0.001$, $\eta^2 = .48$, -2.37 μV difference) and Cz ($F(1,20) = 13.63$, $P = 0.001$, $\eta^2 = .40$, -2.31 μV difference) in comparison to the other cortical positions. No other comparison was statistically significant. The same result was found for not-relevant vs. indirectly relevant conditions comparison, with a significantly greater negativity for indirect relevance within Fz ($F(1,20) = 10.08$, $P = 0.001$, $\eta^2 = .36$, -2.14 μV difference) and Cz ($F(1,20) = 13.69$, $P = 0.001$, $\eta^2 = .40$, -2.02 μV difference).

With concerns to peaks' latencies, ANOVA showed a significant main effect for condition ($F(2,20) = 18.95$, $P = .001$, $\eta^2 = .49$), with a longer latency in indirect relevance condition compared to direct relevance ($F(1,20) = 30.18$, $P = .001$, $\eta^2 = .57$) and not-relevant condition ($F(1,20) = 30.28$, $P = .001$, $\eta^2 = .67$) (Table 2b). Moreover, the interaction condition × site was significant ($F(26,20) = 12.33$, $P = .001$, $\eta^2 = .37$). The interaction condition × ROI was showed a significant effect ($F(14,20) = 11.90$, $P = .001$, $\eta^2 = .42$). Simple effects revealed an increased delay of the peak for anterior and central sites for indirect relevance in comparison with other ROIs (all comparisons $P= .001$).

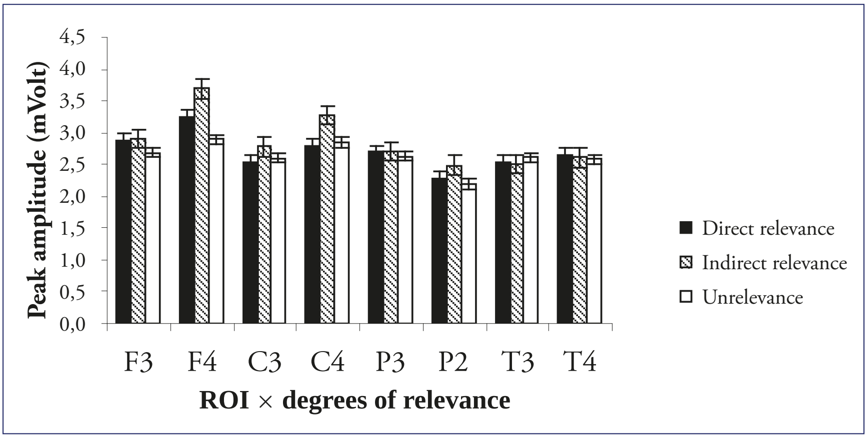

Figure 2. N400 peak amplitude as a function of ROI and degrees of relevance (error bars are standard errors)

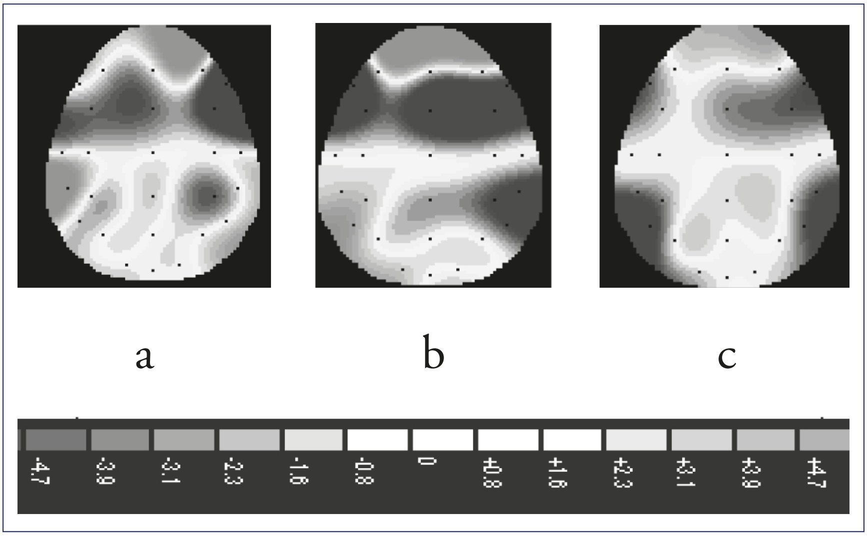

Figure 3. Topographical maps (390-420 ms) for direct relevant (a), indirect relevant (b) and unrelevat (c) conditions

Table 2b. Mean values of N400 ERP latency for each condition and electrode site (midline)

	Electrode sites							
	Fz		Cz		Pz		Oz	
	Latency[b]							
	M	(SD)	M	(SD)	M	(SD)	M	(SD)
Direct relevance	398	1.22	387	1.13	390	1.16	388	1.22
Indirect relevance	422	2.37	419	1.61	386	1.80	391	1.40
Unrelevance	397	1.13	390	1.87	387	1.09	384	1.19
Mean	405	1.58	398	1.53	387	1.35	387	1.16

Note: [b] = Measured in msec.

LN effect

The late negativity component of the ERP was quantified as the mean voltage of the peak and latency in a window range of 480-700 ms post-stimulus relative to a 100 ms pre-stimulus baseline.

Repeated measure ANOVAs were applied to the recording site (14) and to the condition (3) considered as independent variables. Regarding the peak amplitude, LN varied as a function of condition ($F(2,20) = 12.24$, $P = .001$, $\eta^2 = .43$), site ($F(13,20) = 10.98$, $P = .001$, $\eta^2 = .40$) and interaction condition × site ($F(26,20) = 18.04$, $P = .001$, $\eta^2 = .49$). Specifically, as showed by contrast analysis, it was found an increased amplitude for indirect relevance compared to direct relevance ($F(1,20) = 8.16$, $P = .001$, $\eta^2 = .32$) and the not-relevant condition ($F(1,20) = 10.20$, $P = .001$, $\eta^2 = .40$). No other comparison was statistically significant (Table 3a).

Moreover, simple effects showed an increased central (Cz) activity for indirect relevance condition in comparison to direct relevance(respectively: $F(1,20) = 7.66$, $P = .001$, $\eta^2 = .36$ and $F(1,20) = 7.15$, $P = .001$, $\eta^2 = .37$) and not-relevant condition ($F(1,20) = 18.23$, $P = .001$, $\eta^2 = .44$ and $F(1,20) = 12.76$, $P = .001$, $\eta^2 = .40$).

The analysis of ROIs showed a significant effect forcondition ($F(2,20) = 20.16$, $P = .001$, $\eta^2 = .52$), ROI ($F(7,20) = 14.13$, $P = .001$, $\eta^2 = .40$) and condition × ROI ($F(14,20) = 13.28$, $P = .001$, $\eta^2 = .41$). Planned contrasts applied to the main effect of ROI showed that right frontal-central (F4 and C4) areas were more activated than right parietal (respectively for F3 $F(1,20) = 12.36$, $P = .001$, $\eta^2 = .33$ and C3 $F(1,20) = 10.16$, $P = .001$, $\eta^2 = .31$) and temporal ($F(1,20) = 16.09$, $P = .001$, $\eta^2 = .37$ and $F(1,20) = 12.56$, $P = .001$, $\eta^2 = .31$) (Figure 4).

Table 3a. Mean values of LN ERP amplitude for each condition and electrode site (midline)

	Electrode sites							
	Fz		Cz		Pz		Oz	
	Amplitude[a]							
	M	*(SD)*	M	*(SD)*	M	*(SD)*	M	*(SD)*
Direct relevance	2.20	.35	2.44	.49	1.88	.41	1.31	.30
Dndirect relevance	2.75	.38	3.70	.37	2.32	.54	1.87	.40
Unrelevance	2.12	.40	2.05	.37	2.19	.41	1.13	.50
Mean	2.35	.38	2,73	.41	2.13	.47	1.42	.40

Note: [a] = Measured in mVolt.

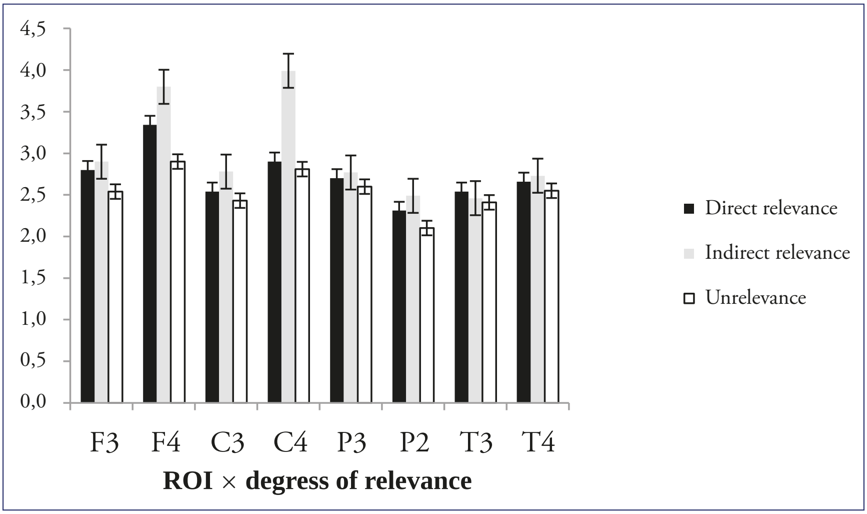

Figure 4. LrN peak amplitude as a function of ROI and degrees of relevance (error bars are standard errors)

Simple effects applied to the interaction condition × ROI showed an increased right frontal (F(1,20) = 12.70, P = .001, η^2 = .37) and central (F(1,20) = 10.77, P = .001, η^2 = .33) activity in comparison to parietal (F(1,20) = 9.12, P = .001, η^2 = .28) and temporal (F(1,20) = 12.33, P = .001, η^2 = .36) activations for indirect relevance condition (Figure 5).

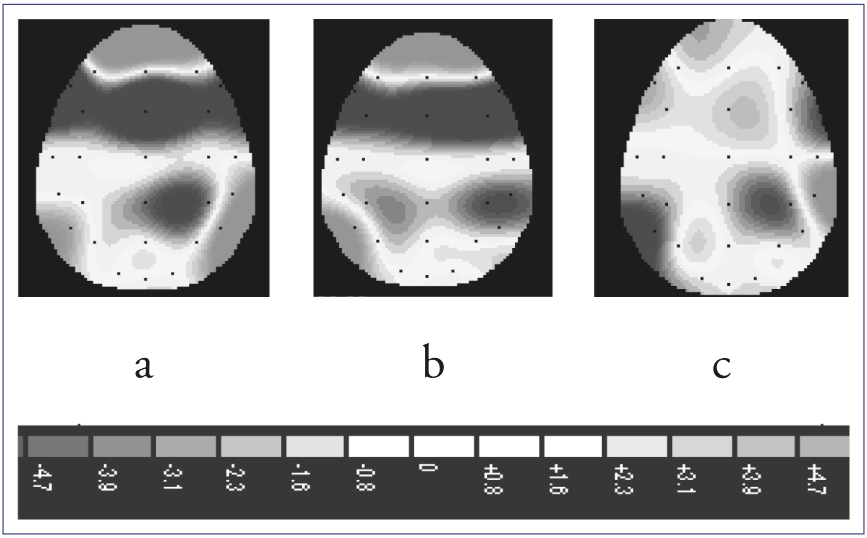

Figure 5. Topographical maps (480-520 ms) for direct relevant (a), indirect relevant (b) and unrelevat (c) conditions

Direct vs. indirect and not-relevant vs. indirect differential waves were calculated. For the first comparison, it was shown a higher negativity for indirect relevance within Cz ($F(1,20) = 12.18$, $P = 0.001$, $\eta^2 = .35$, -2.11 μV difference) than the other cortical positions. No other comparison was statistically significant. The same result was found for not-relevant vs. indirect relevance comparison, with a significant greater negativity for indirect relevance within Fz ($F(1,20) = 10.70$, $P = 0.001$, $\eta^2 = .33$, -2.50 μV difference) and Cz ($F(1,20) = 13.08$, $P = 0.001$, $\eta^2 = .39$, -1.78 μV difference).

Regarding the latency measure, ANOVA showed a significant main effect for condition ($F(2,20) = 18.95$, $P = .001$, $\eta^2 = .49$), with an increased latency for indirect relevance condition compared to direct relevance ($F(1,20) = 30.18$, $P = .001$, $\eta^2 = .57$) and not-relevant condition ($F(1,20) = 30.28$, $P = .001$, $\eta^2 = .67$) (Table 3b). No other effect was statistically significant. The condition × ROI interaction was statistically significant ($F(14,20) = 11.90$, $P = .001$, $\eta^2 = .42$). Simple effects revealed an increased delay of the peak sin response to indirect relevance for central sites in comparison to the other ROIs (all comparisons $P = .001$).

Table 3b. Mean values of LN ERP latency for each condition and electrode site (midline)

	Electrode sites							
	Fz		Cz		Pz		Oz	
	Latency[b]							
	M	*(SD)*	M	*(SD)*	M	*(SD)*	M	*(SD)*
Direct relevance	501	1.28	510	1.33	491	1.34	483	1.20
Indirect relevance	518	1.37	544	1.61	498	1.89	490	1.77
Unrelevance	488	1.18	503	1.82	490	1.67	493	1.73
Mean	502	1.23	519	1.58	493	1.63	488	1.56

Note: [b] = Measured in msec.

4. Discussion

In the present study we analyzed how relevance affects inferential processes when newly-presented information interacts with previously activated contextual assumptions. Three main results can be considered: (1) a main effect related to the degree of relevance, showed by an increased peak amplitude for both N400 and LrN in response to indirect relevant and not-relevant conditions; (2) a significant effect for N400 cortical localization which is more anteriorly right distributed and a concomitant right fronto-central distribution for LrN; (3) a similarity of direct relevant and not-relevant information processing in terms of N400 and LrN modulation.

A possible interpretation of our results allows us to suggest that the N400 is modulated as function of the degree of relevance. Secondly, we found that the brain responds quite differently to directly relevant vs. indirectly relevant vs. unrelated (not-relevant) information. These differences were supported also by a second negative deflection, the LrN, which was right fronto-centrally distributed and displayed higher amplitude in response to the indirect relevance condition in comparison to all the others. LrN supposedly indexes a late integrative semantic, elicited by coherence re-establishment processes. These main aspects are discussed.

N400 effect for relevance

The N400 ERP component could be considered as a response to semantic associative processes allowing incoming information to be integrated within the discourse mental model. As the data indicated, N400's amplitude varied as a function of the degree of relevance of new information, suggesting this component may reflect the ease of integration of new information in order to rebuild discourse coherence.

In past studies, the N400 was considered to reflect some aspect of how stimuli are related to their interpretative context (Van Berkum, 2008). In particular, it was found that N400 peaks increased when critical words or sentences were not directly related (or were not related at all) to the sentence or the discourse-context where they were embedded. It has been suggested thated that the N400 amplitude may be a signal of semantic or lexical priming between the context and the word (Otten & Van Berkum, 2008; Holcomb & Neville, 1990), or of the semantic constraints created by context (Otten et al., 2007; Kutas, 1993; Van Petten & Kutas, 1991).

In the present experiment, final words embedded in indirectly relevant sentences elicited a larger N400 than did final words embedded in directly relevant or not-relevant sentences. Thus, we suggest that the N400 amplitude may be determined by the difficulty of restoring linguistic material at a discourse level (Van Berkum, Zwitserlood, Hagoort & Brown, 2003; Kutas & Federmeier, 2000). When sentences were matched for relevance, N400 effect was reduced and it can be suggested that subjects were able to rapidly work out how the new sentence and the discourse model were connected. Sentence-dependent N400 effect can then simply come about because the structured conceptual representations involved can act as retrieval cues to semantic long-term memory. Memory-based comprehension model highlights the contribution of the availability of plausible and relevant information (Gerrig & O'Brien, 2005). Potentially relevant additional information is processed without cost, as a function of how our content-addressable long-term memory passively "resonates" to currently active representations. In this case, inferences can be made in a sufficiently rapid way, in order to support the processing of a subsequent sentence. That is, active representations in working memory simply act as memory cues for information in long-term memory, and associated information (discourse information, general knowledge on world, etc.) is made available for further processing. Active representations are generally highly available and involve the combination of many different semantic and pragmatic cues.

Thus availability and pertinence of new information seem to be behind contextual expectations and memory retrieval accounts of context-based N400 effects (Federmeir, 2007): our content-addressable memory increases the availability of conceptually associated information in response to what we comprehend in a given context (Van Petten & Luka, 2012). This predictive mechanism supposes that speakers actively anticipate what incoming sentences may or may not say, since they anticipate relevant meaning (Otten & Van Berkum, 2009).

Moreover, the memory-based model may also explain the immediate processing of not-relevant information. In fact, we found that, when anew information is poorly relevant or completely unrelated to the previous context, the N400 amplitude is less increased in comparison to the indirect relevant condition. In this case, it is possible that, due to the low degree of relevance and consequent low level of informativeness, the new information is rapidly discarded as a contextually unrelated cue which does not fit any mental model related to the conceptual background originated by previous information. The whole independence of not-relevant information from the knowledge activated by the actual mental model may induce a rapid and cost-reduced process, since, when a stimulus is not relevant, it is rapidly discarded by the subject, without requiring any other cognitive cost.

It is therefore possible to assume that the highly unpredictable information, which do not present any relation to the previous context, are rapidly and costlessly discarded, while possibly relevant information, which are not directly related to the previous context, might be integrated within the discourse model, although requiring an effortful processing. In this last case, in fact, the contextual mental model needs to be reviewed in line with the subject's semantic and pragmatic expectations.

This explanation was also supported by the presence of a delayed peak for the N400 effect in co-occurrence with an indirect relevance condition. As our results suggest, the indirect condition produced a delayed N400, with the maximum peak amplitude at about 420 ms. The temporal delay could correspond to the increased amount of time required in order to integrate the new information through the formulation of a set of adjunctive inferences (Trueswell & Tanenhaus, 2005). Thus, a mismatching condition between expected and unexpected semantic cues may generate difficulties in updating the mental model that must be reviewed.

A second main general conclusion about N400 ERP effect contribution in discourse comprehension could support the suggestion that processes involved in discourse-dependent conceptual interpretations are similar to those involved in word or sentence-level conceptual interpretations (Otten & Van Berkum, 2008). That is, the present findings reveal that the N400 is not only sensitive to semantic manipulations and integration process, but also to the information encoded in the mental model of discourse representation. In particular, the data show that the N400 reflects processes of conceptual dependency formation between sentences, and that it can be modulated by the degree of relevance and by the accessibility of the conceptual links.

LrN

An ample negative deflection, mainly centrally distributed, was observed in the present research. The peak increased significantly in indirect relevance condition compared to both direct relevance and not-relevant conditions. From a functional point of view, whereas N400 increasing could be reflecting the initial difficulty in retrieving a direct relationship between two conceptual domains, the LrN could be the manifestation of a further attempt to gain this link by founding common conceptual domain. In previous research a sustained frontal negativity was found at sentence boundary (Friederici et al., 1996), and in some cases, it was represented as a clause ending negativity. This deflection was related to working memory load due to integration processes. A similar left-hemisphere late negativity was related to memory load in previous studies (Mecklinger et al., 1997). More generally, this late negativity has typically been reported for memory tasks with an explicit requirement to retrieve information from long-term memory. TheLrN has primarily been observed in tasks that included the retrieval of attributes from memory (i.e. a recognized item and source/context-specifying information). Nevertheless, the frontal cortical distribution of this negativity contrasts with the negative frontal effect found in the present research, suggesting that some differences between the two late deflections may exist. Further analyses should contribute to consider more deeply the cortical localization of this ERP effect. For this reason, the present conclusions about this point should be considered only speculative.

Due to its specific cortical distribution, the present negative deflection should be more directly related to a later semantic integration process supported by further inferential processes. Our findings indicate that discourse-based predictions and conceptual inferences based on discourse relevance require more than a simple semantic comparative analysis. That is, relevance-based processing needs further reconstructive process that can conciliate old conceptual predictions and actual meaning, as a consequence of the difficulty of memory retrieval.

A two steps process may be therefore hypothesized. The first phase of representation includes information arising from different levels of conceptual memories, that may converge or not converge, and this phase could be marked by the N400 modulation. In a second moment, the speaker, needing to workout a plausible relationship between parts of discourse, engage sin further integrative semantic processes able to unify the conceptual information (Otten & Van Berkum, 2008). This may be represented as a second semantic reconstructive process, where an integrated conceptual structure may take place based on conceptual inferences (Friederici, Steinhauer & Frisch, 1999). The additional effort required by this inferential and integrative mechanism presumably involves the need to restore a richer set of semantic links and associations. But, how do predictive inferences fit in? Speakers probably construct a discourse model of what they may infer and predictive inferences can be viewed as plausible representations of what may be relevant and coherent within this mental model (McDaniel, Schmalhofer & Keefe, 2001; Schmalhofer, McDaniel & Keefe, 2002).

Finally, another important aspect concerns the task-specificity of ERPs variation or, specifically, the role of the implicit vs. explicit mechanisms in the relevance processing. As emphasized by previous research, task sensitivity can be considered to be a cue of functional distinctiveness of the ERP correlates (Hagoort, Brown & Groothusen, 1993; Osterhout & Hagoort, 1999). Previous manipulation of the task-salience condition has provided clear evidence that the N400 effect is relatively independent from the degree of explicitness of the task (Balconi & Pozzoli, 2005). In other words, the salience of the task effect should not predict the size of the negative variation. In the present research we included an explicit task, since we asked the subjects to explicitly verify the degree of relevance they perceived between the paired sentences. Nevertheless, we may suppose that the recognition of the degree of relevance of new information is an automatic process that is independent from the attentional focus and immediately activated by discourse processing. Nevertheless, future research should explore in more detail this topic, in order to enlarge our knowledge on the automaticity of relevance attribution in discourse comprehension.

Acknowledgements

This project was supported by a grant from Catholic University of Milan (line D 1.1, 2009).

References

Arzouan, Y., Goldstein, A., & Faust, M. (2007). Brainwaves are stethoscopes: ERP correlates of novel metaphor comprehension. *Brain Research*, 1160, 69-81.

Balconi, M. (2010). Intentions and communication. Cognitive strategies, metacognition, and social cognition. In: Balconi, M. (ed.), *Neuropsychology of Communication*. New York: Sprinter, pp. 159-176.

Balconi, M., & Pozzoli, U. (2004). Elaborazione di anomalie semantiche e sintattiche con stimolazione visiva e uditiva. Un'analisi mediante correlati ERPs. *Giornale Italiano di Psicologia*, 3, 585-612.

Balconi, M., & Pozzoli, U. (2005). Comprehending semantic and grammatical violations in Italian. N400 and P600 comparison with visual and auditory stimuli. *Journal of Psycholinguistic Research*, 34, 71-98.

Burkhardt, P., & Roehm, D. (2007). Differential effects of saliency: an event-related brain potential study. *Neuroscience Letters*, 413, 115-120.

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

DeLong, K.A., Urbach, T.P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8, 1117-1121.

Ditman, T., & Kuperberg, G.R. (2007). The time course of building discourse coherence in schizophrenia: an ERP investigation. *Psychophysiology*, 44, 991-1001.

Federmeier, K.D. (2007). Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology*, 44, 491-505.

Fischler, I., Bloom, P.A., Childers, D.G., Roucos, S.E., & Perry, N.W. (1983). Brain potentials related to stages of sentence verification. *Psychophysiology*, 20, 400-409.

Friederici, A.D., Hahne, A., & Mecklinger, A. (1996). Temporal structure of syntactic parsing: early and late event-related brain potential effects. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 1219-1248.

Friederici, A.D., Steinhauer, K., & Frisch, S. (2007). Lexical integration: sequential effects of syntactic and semantic information. *Memory and Cognition*, 27, 438-453.

Gerrig, R.J., & McKoon, G. (1998). The readiness is all: the functionality of memory-based text processing. Invited article, *Discourse Processes*, 26, 67-86.

Gerrig, R.J., & O'Brien, E.J. (2005). The scope of memory-based processing. *Discourse Processes*, 39, 225-242.

Hagoort, P.J., Brown, C., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes*, 8, 439-483.

Hagoort, P.J., Hald, L., Bastiaansen, M., & Peterson, K.M. (2004). Integration of word meaning and world knowledge in language comprehension. *Science*, 304, 438-441.

Holcomb, P.J., & Neville, H.J. (1990). Auditory and visual semantic priming in lexical decision: a comparison using event-related brain potentials. *Language and Cognitive Processes*, 5, 281-312.

Kamide, Y., Altmann, G.T.M., & Haywood, S.L. (2003). Prediction and thematic information in incremental sentence processing: evidence from anticipatory eye movements. *Journal of Memory and Language*, 49, 133-156.

Kutas, M. (1993). In the company of other words: electrophysiological evidence for single word versus sentence context effect. *Language and Cognitive Processes*, 8, 533-572.

Kutas, M., & Federmeier, K.D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Science*, 4, 463-470.

Kutas, M., Van Petten, C., & Kluender, R. (2006). Psycholinguistics electrified II: 1995-2005. In: Traxler, M., & Gernsbacher, M.A. (eds.), *Handbook of Psycholinguistics*, 2nd ed. New York: Elsevier, pp. 659-724.

Levinson, S.C. (2000). *Presumptive Meanings: the Theory of Generalized Conversational Implicature*. Cambridge: MIT Press.

McDaniel, M.A., Schmalhofer, F., & Keefe, D.E. (2001). What is minimal about predictive inferences? *Psychonomic Bulletin and Review*, 8, 840-846.

Mecklinger, A., Opitz, B., & Friederici, A.D. (1997). Semantic aspects of novelty detection in humans. *Neuroscience Letters*, 235, 65-68.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A.D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: an analysis with event-related potentials. *Memory and Cognition*, 23, 477-494.

Myers, J.L., & O'Brien, E.J. (1998). Accessing the discourse representation during reading. *Discourse Processes*, 26, 131-157.

Noveck, I., & Sperber, D. (2007). The why and how of experimental pragmatics: the case of "scalar inferences". In: Burton-Roberts, N. (ed.), *Advances in Pragmatics*. Basingstoke: Palgrave Macmillan, pp. 184-212.

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

Osterhout, L., & Nicol, J. (1999). On the distinctiveness, independence, and time course of the brain responses to syntactic and semantic anomalies. *Language and Cognitive Processes*, 3, 283-317.

Otten, M., Nieuwland, M.S., & Van Berkum, J.J.A. (2007). Great expectations: specific lexical anticipation influences the processing of spoken language. *BMC Neuroscience*, 8, 89.

Otten, M., & Van Berkum, J.J.A. (2007). What makes a discourse constraining? Comparing the effects of discourse message and scenario fit on the discourse-dependent N400 effect. *Brain Research*, 1153, 166-177.

Otten, M., & Van Berkum, J.J.A. (2008). Discourse-based word anticipation during language processing: prediction or priming? *Discourse Processes*, 45, 464-496.

Otten, M., & Van Berkum, J.J.A. (2009). Does working memory capacity affect the ability to predict upcoming words in discourse? *Brain Research*, 1291, 92-101.

Sartori, G., Mameli, F., Polezzi, D., & Lombardi, L. (2006). An ERP study of low and high relevance semantic features. *Brain Research Bulletin*, 69, 182-186.

Schacter, D.L., Addis, D.R., & Buckner, R.L. (2007). Remembering the past to imagine the future: the prospective brain. *Nature Reviews Neuroscience*, 8, 657-661.

Schmalhofer, F., McDaniel, M.A., & Keefe, D.E. (2002). A unified model for predictive and bridging inferences. *Discourse Processes*, 33, 105-132.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication and Cognition*, 2nd ed. Oxford, UK: Blackwell.

Trueswell, J.C., & Tanenhaus, M.K. (2005). *Approaches to Studying World-situated Language Use: Bridging the Language-as-product and Language-action Tradition*. Cambridge, MA: MIT Press.

Van Berkum, J.J.A. (2004). Sentence comprehension in a wider discourse: can we use ERPs to keep track of things? In: Carreiras, M., & Clifton, C. (eds.), *The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond*. New York: Psychology Press, pp. 229-270.

Van Berkum, J.J.A. (2008). Understanding sentences in context: what brain waves can tell us. *Current Directions in Psychological Science*, 17, 376-380.

Van Berkum, J.J.A. (2010). The brain is a prediction machine that cares about good and bad. Any implications for neuropragmatics? *Italian Journal of Lingusitics*, 22, 181-208.

Van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 443-467.

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. *Cognitive Brain Research*, 17, 701-718.

Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open- and closed-class word. *Memory and Cognition*, 19, 95-112.

Van Petten, C., & Luka, B.J. (2012). Prediction during language comprehension: benefits, costs, and ERP components. *International Journal of Psychophysiology*, 83, 176-190.

Wilson, D., & Sperber, D. (2003). Relevance theory. In: Ward, G., & Horn, L. (eds.), *Handbook of Pragmatics. Blackwell Science*. Oxford, UK: Blackwell, pp. 607-632.

Wilson, R., & Keil, F. (1999). *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, MA: MIT Press.

Yang, C.L., Perfetti, C.A., & Schmalhofer, F. (2005). Event-related potential indicators of text integration across sentences boundaries. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 55-89.