The syntax-semantics interface during sentence processing: ERP evidence on relative clause attachment ambiguity in Greek*

Konstantina Olioumtsevits¹ & Srđan Popov²

¹Aristotle University of Thessaloniki, ²University of Groningen ogkonsta@lit.auth.gr, s.popov@rug.nl

Περίληψη

Η παρούσα μελέτη διερευνά τον ρόλο της σημασιολογίας στην προσάρτηση αμφίσημων αναφορικών προτάσεων. Ένα χρονομετρικό πείραμα διεξήχθη με ενήλικες φυσικούς ομιλητές της ελληνικής, κατά το οποίο καταγράφηκαν ηλεκτροφυσιολογικά σήματα του ανθρώπινου εγκεφάλου. Στα πειραματικά ερεθίσματα, οι δύο πιθανοί κόμβοι σύνδεσης σχημάτιζαν δομές γενικής κτητικής ή προθετικής με-φράσης. Η απουσία ενός P600 σε όλες τις πειραματικές συνθήκες δείχνει ότι, λαμβάνοντας άμεσα υπόψη τις διαθέσιμες σημασιολογικές πληροφορίες, οι δομικές προτιμήσεις των ομιλητών εξαλείφονται. Ωστόσο, ένα N400 προκλήθηκε μόνο στις προτάσεις με δομές γενικής κτητικής, υποδηλώνοντας ότι σημασιολογικά χαρακτηριστικά των δεδομένων δομών μπορούν να δυσκολέψουν τη διαδικασία σημασιολογικής ενσωμάτωσης.

Λέζεις-κλειδιά: αναφορική πρόταση, αμφισημία, σύνταζη, σημασιολογία, Προκλητά Δυναμικά (ERPs)

1 Introduction

1.1 Behavioural evidence on relative clause attachment and the role of semantics

Literature on relative clause attachment has mainly revolved around the question: How are sentences with two potential antecedents processed? In sentences such as in (1), the relative clause (RC) can syntactically modify either of the two noun phrases that precede it (NP1 or NP2). Both attachment options lead to a grammatically correct sentence, rendering the utterance ambiguous. In the absence of disambiguating information, speakers' preference drives the resolution, with the choice of the first noun phrase leading to high attachment, whereas that of the second noun phrase to low attachment (Frazier 1978).

(1) The journalist interviewed [the daughter]_{NP1} of [the colonel]_{NP2} [who had had the accident]_{RC}.

The resolution of relative clause attachment ambiguity tells an intricate story of how the human comprehension mechanism (hereafter, the parser) operates, and how different types of linguistic information interact in this process. Since the late 1970s, attachment preferences have been an attractive research topic. Numerous studies using various behavioural methods have categorized languages into two classes: high- and low-attaching (e.g., Cuetos and Mitchell 1988, Frazier, and Clifton 1996). A low

_

^{*} While writing the paper, the first author received funding from the European Union's Horizon2020 research and innovation programme under the Marie Skłodowska Curie grant agreement No 765556, as part of the MultiMind project.

attachment preference has been explained by Frazier's (1978) model, according to which, new phrases are integrated to the most recent constituent – favouring a syntactically minimal interpretation. A structurally driven model has been proposed by Gibson and Pearlmutter (1998) to explain a high attachment preference, according to which, attachment closer to the predicate is preferred in languages with relatively free word order. However, studies with differing methods have yielded contradictory results (Fernández 2002), implying that disambiguation is a multi-factorial phenomenon.

The preposition conjoining two noun phrases has also been found to be a major aspect of relative clause interpretation (Gilboy, Sopena, Clifton, and Frazier 1995). In particular, the preposition with seems to trigger low attachment, even in languages in which high attachment is otherwise preferred in sentences with an of-phrase. The theory by Frazier and Clifton (1996) offers an explanation based on theta domains imposed by different preposition types, that is, functional prepositions, such as of, versus lexical prepositions, such as with. However, in contrast to the preposition with, other lexical prepositions, such as near, have been found to trigger high attachment (Gibson, Pearlmutter, Canseco-Gonzalez, and Hickock 1996). Findings on processing dissimilarities among lexical prepositions provide further evidence that disambiguation cannot be guided solely by syntactic cues. The prepositions with and near are structurally similar, yet they are defined by different semantic properties.

An issue raised in the literature concerns whether semantic factors guide syntactic analysis. One theoretical framework, supported by empirical evidence (Trueswell, Tanenhaus, and Garnsey 1994), is provided by constraint-based approaches: all relevant sources of information are simultaneously taken into consideration when determining syntactic assignment, discarding the idea of priority of structural cues proposed by serial models (Tanenhaus and Trueswell 1995).

Only a limited number of studies have investigated the role of semantics in relative clause interpretation. For instance, in the study by Brysbaert and Mitchell (1996) in which they classified Dutch as a high-attaching language, they used stimuli where both or only one of the potential antecedents were animate. The relative clause contained a pronoun marked for gender that agreed with only one of the antecedents. Despite the role of animacy in their paradigm and the fact that a high attachment preference was obvious mostly when only the first entity was inanimate, the authors made no analysis of the given effect. However, the relative flexibility of attachment preferences hints that lexicosemantic properties may influence structural expectations.

1.2 Event-related potentials and relative clause attachment processing

Language comprehension is a complex process that requires fast integration of information of different language levels. Electroencephalography (EEG) enables a moment-by-moment examination of sensitivity to linguistic cues, while allowing us to disentangle different types of cues through the elicitation of different event-related potential (ERP) components (Luck 2005). For instance, the P600, a positive deflection starting at around 400ms after the critical event, has been generally associated with the process of structural repair and reanalysis (Kaan 2007). In contrast, the N200 and the N400 are two negative-going waves with their onsets at around 150 ms and 300ms after the target word, respectively, reflecting a difficulty in semantic integration (Van den Brink, Brown, and Hagoort 2001, Kutas and Federmeier 2011).

ERP research on relative clause attachment processing is rather scarce, whilst also greatly limited to the examination of the effects of syntactic cues. For instance, Carreiras, Salillas, and Barber (2004) approached the phenomenon with an ERP paradigm whereby attachment ambiguities were resolved based on gender agreement. They found that a P600 effect was elicited when the speakers were presented with a non-preferred syntactic assignment, indicating an increase in syntactic processing cost. Based on the semantic aspects of grammatical gender, the authors also expected an N400 effect. However, they found no such indication, suggesting that experimental paradigms based on gender agreement are not suitable to study the interplay between syntax and semantics in relative clause attachment resolution.

A large body of ERP research has looked into the broad and rather early influence of semantics on sentence processing and expectation (e.g., Osterhout, Holcomb, and Swinney 1994, Hoeks 1997, Van den Brink, Brown, and Hagoort 2001). These findings strengthen the claim that syntactic analysis is governed by semantic, as well as structural constraints. As sentence comprehension is an incremental process, incoming words are connected with preceding words in an attempt to build up a message-level meaning. The failure of the parser to create semantic links between constituents, and the resulting N200 and/or N400 effect, indicates a semantic processing cost.

1.3 The current study

The objective of the present study is to disentangle the effects of syntax and semantics during relative clause attachment processing in Greek.

Greek is a relatively free word-order language with a rich morphological system which marks articles, nouns, adjectives and participles for gender, case and number. Of relevance to this study is how Greek expresses the possessive of-phrase and the with-phrase. The former is expressed with the genitive case, and the latter with an equivalent prepositional phrase (Tzartzanos 1991). Under the assumption that Greek speakers would show similar attachment tendencies to speakers of high-attaching languages, Papadopoulou and Clahsen (2003) conducted a study in order to answer the issue. They used stimuli with both genitive constructions and with-phrases, and the relative clause included an adjective or participle which agreed in gender with only one of the two potential antecedents. The findings confirmed their predictions: Genitive constructions showed a clear high attachment preference, whereas with-phrases showed a clear low attachment preference. All study of relative clause attachment in Greek has been behavioural as of date, and no ERP studies have been conducted on the issue.

The possessive construction and the *with*-phrase have been extensively studied, yet, given the contradictory results in the literature, there is no consensus on the phenomenon. As prepositions are also defined by semantic properties, this could play a crucial role in the syntactic disambiguation. The possessive construction denotes dependency, with the possessor being semantically more salient as the entity that establishes the given relationship. In contrast, the *with*-phrase denotes accompaniment, and the two entities involved are of equal importance on a semantic level (Langacker 1993, Lehmann and Shin 2000). Additionally, as described earlier, a research paradigm based on gender agreement blurs the investigation of semantic sources. Instead, in a sentence like *Peter asked the speech therapist of the little girl*

who was dyslexic based on the diagnosis, we are forced to establish a semantic link between dyslexic and the little girl, as we have a semantic preference about speech therapists without a language disorder. Thus, semantic properties imposed by the adjective may bias the parser towards a low attachment interpretation despite a high attachment preference.

The current study is relevant, as it bridges the gap in the literature regarding the role of semantics in establishing relative clause attachment, while also contributing to experimental ERP evidence on a cross-linguistic level. The research questions raised here are the following:

- 1) Is the process of relative clause attachment independent from nonstructural information (i.e., semantics), or do semantic cues influence the syntactic analysis?
- 2) Do the different semantic properties of the possessive construction and the *with*-phrase trigger processing dissimilarities when establishing relative clause attachments?

2 Methods

2.1 Participants

Twenty-four adult native speakers of Greek (MA = 24.4, SD = 2.6; 10 males) with no history of language impairment or neurological/psychiatric illness were recruited. They were all right-handed with normal or corrected-to-normal vision. Their hand dominance was assessed with a short version of the Edinburgh Inventory (Oldfield 1971). Regarding their level of education, they were either bachelor's or master's students. All participants took part in the experiment after giving informed consent.

2.2 Materials

The paradigm of the present study was partly based on that of the study by Papadopoulou and Clahsen (2003). Novel grammatical sentences were formed in which the main clause included two potential antecedents. The main clause was followed by a relative clause introduced by the relative complementizer. Regarding the complex noun phrase, it contained either a genitive construction or a prepositional phrase.

Contrary to the study by Papadopoulou and Clahsen (2003), relative clause attachment was manipulated based on semantic cues. The relative clause included an adjective or participle as the disambiguation point, which agreed in gender with both potential antecedents but could only refer to one of them. In the sentences (2a) and (3a) below, *experienced* is most likely to modify *the speech therapist*, since *the little girl* is probably still not very experienced in something due to age. Semantic plausibility works in the opposite direction when it comes to *dyslexic* in sentences (2b) and (3b). It should be also noted that neither inanimate nor proper nouns were used as potential antecedents, while two or three words always followed the critical word.

As the following examples illustrate, we created one core sentence, from which four experimental sentences were derived. The two experimental conditions were based on the antecedent type, genitive construction or prepositional phrase, with each having two levels: high and low attachment.

2a) genitive-high attachment (preferred)

0	Πέτρος	ρώτησε	τη
the.MASC.NOM	Peter.MASC.NOM	asked	the.FEM.ACC

λογοθεραπεύτρι-α	της	μικρ-ής
speech.therapist-FEM.ACC	the.FEM.GEN	little.girl-FEM.GEN

που	ήταν	πολύπειρ-η	στις	περιπτώσεις	δυσλεξίας.
that	was	experienced-FEM.NOM	in	cases	dyslexia

^{&#}x27;Peter asked the speech therapist of the little girl who was experienced in dyslexia cases.'

2b) genitive-low attachment (non-preferred)

Ο Πέτρος ρώτησε τη λογοθεραπεύτρια της μικρής που ήταν δυσλεκτική βάσει της διάγνωσης.

'Peter asked the speech therapist of the little girl who was dyslexic based on the diagnosis.'

3a) prepositional phrase-high attachment (non-preferred)

О	Πέτρος	ρώτησε	τη
the.MASC.NOM	Peter.MASC.NOM	asked	the.FEM.ACC

λογοθεραπεύτρι-α	με	τη	μικρ-ή
speech.therapist-FEM.ACC	with	the.FEM.ACC	little.girl-FEM.ACC

που	ήταν	πολύπειρ-η	στις	περιπτώσεις	δυσλεξίας.
that	was	experienced-FEM.NOM	in	cases	dyslexia

^{&#}x27;Peter asked the speech therapist with the little girl who was experienced in dyslexia cases.'

3b) prepositional phrase-low attachment (preferred)

Ο Πέτρος ρώτησε τη λογοθεραπεύτρια με την μικρή που ήταν δυσλεκτική βάσει της διάγνωσης.

'Peter asked the speech therapist with the little girl who was dyslexic based on the diagnosis.'

In order to prevent participants from seeing more than one version of each core sentence, four lists were created. The final set of the experimental materials, based on a verification test conducted prior to testing, was counterbalanced across lists. Each list was composed of 80 experimental sentences and 60 fillers, thus 140 pseudorandomized stimuli in total. Concerning the structure of the fillers, they all contained a grammatical violation. The adjective or participle in the relative clause either agreed in gender with the subject of the main clause or with none of the available nouns.

2.3 Procedure

The participants were comfortably seated in a chair in front of a screen, while the electroencephalogram was recorded. Using the E-Prime software, stimuli were presented in the centre of the screen in white letters (Calibri 24 pt) on a black background. Each sentence was presented word-by-word in the following order of events: a fixation cross first appeared for 500ms indicating the beginning of the stimulus, and then a blank screen for 300ms. Each word appeared for 300ms, always followed by a blank screen for 300ms. The last word of the sentence was presented with a full stop indicating the end of the stimulus. Given the present experiment was a passive task, 20% of stimuli – half grammatical and half ungrammatical sentences – were followed by a grammaticality judgement question. Written and spoken instructions were first provided to participants. In order to familiarize them with the procedure, a training session preceded the actual experiment. No further clarifications were provided during the experimental session. The whole experiment lasted approximately 30 min.

2.4 EEG data

Using the eego MyLab system (ANT Neuro Inc, Enschede, The Netherlands), EEG was recorded from 32 Ag/AgCI scalp electrodes (WaveGuard) positioned according to the standard 10-20 system. Eye movements were monitored with an additional electrode above the left eye (EOG). The electrode impedance was kept below 10 k Ω . The sampling rate was set to 500 Hz and the recording reference was the common average reference.

After EEG recording was completed, data pre-processing was performed with the Brain Vision Analyzer 2.0.4 software (Brain Products, GmbH, Munich, Germany). The first step was to re-reference data to the average of both mastoids. A band-pass filter of 0.1-40 Hz was then applied, and an Independent Component Analysis eyeblink correction was performed. Subsequently, data were segmented into epochs of 1200ms, starting 200ms before the onset of the critical word and ending 1000ms after the trigger marker. For each epoch, the (automatic) artifact rejection (+/- 100 μV threshold) was applied. Baseline correction was applied to the part of the segment of 200 to 0ms. The last stage of data processing involved data averaging per subject and per condition, with the minimum number of retained trials being 60% in each condition.

The averaged values were analysed based on 9 regions of interest composed of 2 or 3 electrodes each. Four independent time windows were chosen based on visual data inspection: i) 200-400ms, ii) 400-600ms, iii) 600-800ms, and iv) 800-1000ms. Analysis was conducted in R (R Core Team, 2014). Data were analysed using repeated measures ANOVA with the following within factors: condition type (2 levels: genitive and prepositional construction), attachment type (2 levels: high and low), hemisphere (2 levels: right and left), and anteriority (3 levels: anterior, central, and posterior). Two separate repeated measure ANOVAs were performed for each time window, the first one on the lateral regions and the second one on the midline regions (without the factor 'hemisphere'). When the assumption of sphericity

violated, the Greenhouse-Geisser correction was applied. The significance level was set to p < .05.

3 Results

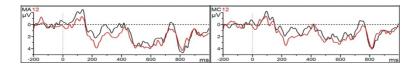
3.1 Accuracy results

The purpose of the grammaticality judgment questions was to ensure that participants paid attention throughout the entire experiment. The cut-off for including a participant in the statistical analysis was set to 60%. We could not exclude the possibility that the participants' judgments on the experimental sentences could be also affected by factors unrelated to the grammaticality of the stimuli. However, a chance-level performance would be a clear indication of a reduced alertness. According to the performance outcomes, no participants had to be excluded. Overall, they had an accuracy rate of 79.4% (SD = 9.1).

3.2 ERP results

In the first time window, neither the analysis on the lateral regions nor that on the midline revealed any significant main effects or interactions (all ps > .1). However, as the visual inspection of the topographic maps and the voltage values indicated a potential frontocentral negative effect from approximately 200ms after the onset of the critical word, we ran a post-hoc analysis on the midline regions. In the follow-up analysis, two separate repeated measure ANOVAs were performed for each condition. The effect of attachment was found to be significant for the genitive condition (F(1,23) = 4.339, p = .04), while the main effect and interaction were greatly above the significance level for the prepositional phrase condition (all ps > .1). Subsequently, paired t-tests were applied for each anteriority level in the genitive condition. The effect of attachment was found to be significant only in anterior (t(23) = -2.42, p = .02) and central regions (t(23) = -2.09, p = .04), while the voltage values demonstrated that the negative effect was caused by high attachment.

The ANOVA analyses on the two following time windows did not yield any significant main effects or interactions regarding the factors condition and/or attachment (all ps > .1). Finally, in the last time window, the lateral analysis revealed a significant interaction between condition and hemisphere (F(1,23) = 5.715, p = .02). However, post-hoc paired t-tests showed that there was not a significant effect of condition in the regions of either the right or the left hemisphere (both ps > .1).



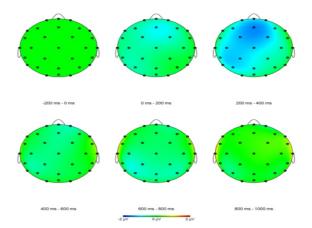


Figure 1 | Grand average ERPs for the genitive condition in two regions of interest, midline anterior (MA) and midline central (MC): the black line represents high attachment sentences (i.e., the baseline) and the red line represents low attachment sentences. The topographic maps depict differences between high and low attachment sentences.

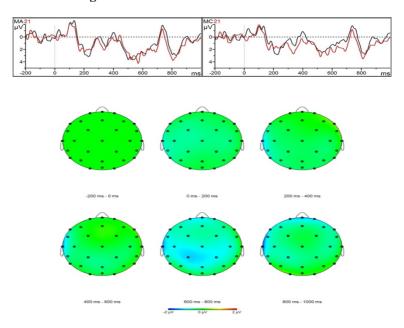


Figure 2 | Grand average ERPs for the prepositional phrase condition in two regions of interest, midline anterior (MA) and midline central (MC): the black line represents low attachment sentences (i.e., the baseline) and the red line represents high attachment sentences. The topographic maps depict differences between low and high attachment sentences.

4 Discussion

The main aim of the present study was to examine how robust semantic cues interact with structural preferences during relative clause attachment processing. The absence of any positive effect can be considered as an indication of an overall uniform performance on the syntactic level across conditions, regardless of the antecedent or attachment type. If there was an initial syntactic bias, then a P600 would be elicited to confirm that. Our findings are in support of constraint-based approaches (Tanenhaus and Trueswell 1995) and are in line with those of other studies confirming that syntax does not have *a priori* precedence over other sources of information (Trueswell,

Tanenhaus, and Garnsey 1994, Hoeks 1997). Structural preferences can be annulled by the immediate contribution of semantic information, and the parser thus avoids reanalysis (cf., Carreiras, Salillas, and Barber 2004).

The sub-objective of the present study was a more focused comparison of the distinct semantic properties of the possessive construction and the *with*-phrase. In possessive constructions, as the possessor is more semantically salient, it facilitates the processing of the possessee (Langacker 1993). The high attachment option can then only be selected if the low attachment option has been assessed and rejected. This selection processing cost is evident in the negative deflection triggered by high attachment in the genitive condition. The effect is akin to an N400 with some temporal and spatial characteristics that make it resemble the N200, also related to semantic integration (Van den Brink, Brown, and Hagoort 2001). In contrast, the two entities conjoined with the preposition *with* are of similar importance on the semantic level. Given the absence of competitive candidates, when the disambiguation point is reached, the parser selects the entity that best fits the semantic features imposed by the critical word with less effort (Lehmann and Shin 2000).

In conclusion, our data indicate that syntax is not an independent module when a structural ambiguity is encountered, providing empirical evidence on the importance of semantic cues in the extraction of a meaningful message at the sentence level. Additionally, the results provide significant insights into the processing strategies applied depending on the linguistic environment imposed by the disambiguation frame. During relative clause attachment processing, the parser focuses on semantic properties of the antecedent type when a semantic "agreement" is forced. Lastly, the findings hint at antecedent-specific features that can disrupt the process of semantic integration. These findings call for more empirical research on the role of semantics in relative clause attachment, and highlight the importance of multi-factorial study of language as a complex system of meaning-formation.

References

- Brysbaert, Marc, and Don C. Mitchell. 1996. "Modifier attachment in sentence parsing: Evidence from Dutch". *Quarterly Journal of Experimental Psychology*, 49, 664-695.
- Carreiras, Manuel, Elena Salillas, and Horacio Barber. 2004. "Event-related potentials elicited during parsing of ambiguous relative clauses in Spanish". *Cognitive Brain Research*, 20, 98-105.
- Cuetos, Fernando, and Don C. Mitchell. 1988. "Cross-linguistic differences in parsing: Restrictions on the use of Late Closure strategy in Spanish". *Cognition*, 30, 73–105.
- Fernández, Eva. M. 2002. "Relative clause attachment in bilinguals and monolinguals". *Advances in Psychology*, *134*, 187-215.
- Frazier, Lyn. 1978. "On comprehending sentences: Syntactic parsing strategies" (Unpublished doctoral dissertation). University of Connecticut, Storrs.
- Frazier, Lyn, and Charles Clifton. 1996. Construal. Cambridge, MA: MIT Press.
- Gibson, Edward, and Neal J. Pearlmutter. 1998. "Constraints on sentence comprehension". *Trends in Cognitive Science*, 2, 262–268.

- Gibson, Edward, Pearlmutter, Neal, Canseco-Gonzalez, Enriqueta, and Gregory Hickok. 1996. "Recency preferences in the human sentence processing mechanism". *Cognition*, 59, 23–59.
- Gilboy, Elizabeth, Sopena, Josep-M Maria, Cliftrn Jr, Charles, and Frazier, Lyn. 1995. "Argument structure and association preferences in Spanish and English complex NPs". *Cognition*, *54*, 131-167.
- Hoeks, John. 1997. "Processing coordinated structures". In M. Biemans & J. v. d. Weijer (Eds.), *Proceedings of the CLS Opening Academic Year 1997-1998*. Nijmegen, Netherlands. https://www.researchgate.net/publication/228507227_Processing_Coordinated_Structures
- Kaan, Edith. 2007. "Event-Related Potentials and language processing: A brief overview". *Language and Linguistics Compass*, 1, 571-591.
- Kutas, Marta, and Kara D. Federmeier. 2011. "Thirty years and counting: Finding meaning in the N400 component of the Event-Related Brain Potential (ERP)". *The Annual Review of Psychology*, 62, 621-647.
- Langacker, Ronald W. 1993. "Reference-point constructions". *Cognitive Linguistics*, 4, 1-38.
- Lehman, Christian, and Yong-Min Shin. 2000. *The functional domain of concomitance: A typological study of the linguistic representation of concomitant relations*. Arbeitspapier 4: Typologie des Yukatanischen. University of Erfurt.
- Luck, Steven J. 2005. An introduction to the event-related potential technique. Cambridge, MA: The MIT Press.
- Oldfield, Richard C. 1971. "The assessment and analysis of handedness: The Edinburgh inventory". *Neuropsychologia*, *9*, 97–113.
- Osterhout, Lee, Phillip J. Holcomb, and David A. Swinney. 1994. "Brain potentials elicited by garden-path sentences: Evidence of the application of verb information during parsing". *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 20, 786-803.
- Papadopoulou, Despina, and Harald Clahsen. 2003. "Parsing strategies in LI and L2 sentence processing: A study of relative clause attachment in Greek". *Studies in Second Language Acquisition*, 25, 501-528.
- Tanenhaus, Michael K., and John C. Trueswell. 1995. "Sentence comprehension". In J. L. Miller & P. D. Eimas (Eds.), *Speech, language, and communication* (pp. 217-262). San Diego: Academic Press.
- Team R Core. 2014. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.
- Trueswell, John C., Michael K. Tanenhaus, and Susan M. Garnsey. 1994. "Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution". *Journal of Memory and Language*, 33, 285-318.
- Tzartzanos, Achilleas. 1991. Νεοελληνική Σύνταζις (της κοινής Δημοτικής) [Syntax of Modern Greek]. Thessaloniki, Greece: Kyriakidi.
- Van Den Brink, Daniëlle, Colin M. Brown, and Peter Hagoort. 2001. "Electrophysiological evidence for early contextual influences during spokenword recognition: N200 versus N400 effects". *Journal of Cognitive Neuroscience*, 13, 967–985.