ARTICLE IN PRESS
BRES-35879; No. of pages: 19; 4C:
BRAIN RESEARCH XX (2006) XXX–XXX

available at www.sciencedirect.com
ScienceDirect
www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

An electrophysiological study of mood, modal context, and anaphora

Veena D. Dwivedi[a,b,*], *Natalie A. Phillips*[b], *Maude Laguë-Beauvais*[c], *Shari R. Baum*[d]

[a]*Brock University, 500 Glenridge Ave., St. Catharines, Ontario, Canada, L2S 3A1*
[b]*Concordia University, Canada*
[c]*Université du Québec à Montréal, Canada*
[d]*McGill University, Canada*

ARTICLE INFO

Article history:
Accepted 11 July 2006

Keywords:
P600
N400
Discourse structure
Modal auxiliary
Anaphora
Event-related brain potential

ABSTRACT

We investigated whether modal information elicited empirical effects with regard to discourse processing. That is, like tense information, one of the linguistic factors shown to be relevant in organizing a discourse representation is modality, where the mood of an utterance indicates whether or not it is asserted. Event-related potentials (ERPs) were used in order to address the question of the qualitative nature of discourse processing, as well as the time course of this process. This experiment investigated pronoun resolution in two-sentence discourses, where context sentences either contained a hypothetical or actual Noun Phrase antecedent. The other factor in this 2×2 experiment was type of continuation sentence, which included or excluded a modal auxiliary (e.g., *must*, *should*) and contained a pronoun. Intuitions suggest that hypothetical antecedents followed by pronouns asserted to exist present ungrammaticality, unlike actual antecedents followed by such pronouns. Results confirmed the grammatical intuition that the former discourse displays anomaly, unlike the latter (control) discourse. That is, at the Verb position in continuation sentences, we found frontal positivity, consistent with the family of P600 components, and not an N400 effect, which suggests that the anomalous target sentences caused a revision in discourse structure. Furthermore, sentences exhibiting modal information resulted in negative-going waveforms at other points in the continuation sentence, indicating that modality affects the overall structural complexity of discourse representation.

1. Introduction

1.1. Background

During the course of discourse comprehension, information about the relative structure of events denoted by different sentences must be represented on-line. That is, different events presented in a discourse must be structured relative to other events already posited in previous context. One such relation is the temporal order in which events occur. While real-world knowledge of events and causation play a role in this structure (Kintsch, 1998; Garnham, 1981; Johnson-Laird, 1983; Webber, 1979), previous works have shown that temporal information conveyed by the grammaticalized expression of tense and tense inflection, as well as temporal adverbials such as *before*, *in an hour*, *tomorrow*, etc., affects the nature of

* *Corresponding author*. Department of Applied Linguistics, Brock University, 500 Glenridge Ave., St. Catharines, Ontario, Canada, L2S 3A1. Fax: +1 905 688 2360.
E-mail address: veena.dwivedi@brocku.ca (V.D. Dwivedi).

0006-8993/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2006.07.048

sentence processing (Trueswell and Tanenhaus, 1991; Zwaan, 1996; Münte et al., 1998a,b; Dickey, 2001).

For example, Trueswell and Tanenhaus (1991) showed that the interpretation of the ambiguous string *The student spotted*…[1] differed depending on the tense of the previous context sentence. Notably, if the previous context sentence exhibited future tense, as in, *Tomorrow the proctor will notice a student cheating on a linguistics exam*, the string *The student spotted*… was assigned a syntactic structure of a reduced relative clause, rather than the simpler, and hence preferred, structure of a main clause. In addition, Zwaan (1996) showed that temporal adverbials such as *in an hour* also affected sentence processing, such that if the information in the adverbial indicated a greater temporal distance between sentences (*in an hour* vs. *at that moment*), reading times for such sentences were longer. Furthermore, Münte et al. (1998a,b) showed that when the interpretation of the order of events presented in a sentence did not match the actual temporal order of events (where the sentences in question differed only at the adverbial *before* vs. *after*), a complexity effect ensued, indicating an increase in cognitive load.

The goal of the present work is to build on previous findings regarding tense information by showing that the closely related grammatical construct of mood must also be consulted during sentence processing. The truth-conditional compositional semantic theory[2] of Roberts (1987, 1989, 1996) argues that mood, like tense, affects the structure of events in discourse. Roberts' work builds on the theories of Heim (1982) and Kamp (1981) who posit Discourse Representation Theory as a linguistic theory of contextual update. The present work takes these linguistic theories as a starting point and investigates whether Roberts' theory of Modal Subordination (to be discussed below) exhibits any empirical effects. In Dwivedi (1996), two experiments were conducted to investigate this question, using the behavioral technique of self-paced reading, where it was shown that modal information did indeed affect sentence processing times. The present study follows up on those findings using the more fine-grained technology of Event-Related Potentials (ERPs)—this methodology allows for millisecond by millisecond time resolution of sentence processing, as well as potential information regarding the qualitative nature of linguistic processing. Before proceeding with the question at hand, a brief review of the linguistic literature is in order.

1.2. *Modality*

Jespersen (1965) introduced the sense of notional mood, where the mood of an utterance tells us whether or not it is asserted. That is, the mood of a sentence tells us something about the speaker's commitment to the truth of the proposition in the actual world. If a speaker indicates that a sentence is to be interpreted as true in the actual world, the sentence exhibits "factual mood" (Roberts, 1987, 1989). Below, (1a) is an example of this; that is, it can be interpreted as true that there is an event of novel-writing, and that John did it.

(1) a. John wrote a novel.
b. John might write a novel.
c. John must write a novel.

In contrast to (1a), sentences (1b) and (1c) exhibit "non-factual mood." In (1b), the modal auxiliary *might* introduces the notion of possibility to the claim about John and the novel. Now the speaker is not so sure about the event, hence the use of the modal auxiliary. This hypothetical meaning is expressed using "non-factual mood". There is a level of complexity added on to sentences such as (1b) vs. (1a), that is, an additional inference regarding possibility is required. Finally, in the third sentence, two possible non-factual meanings emerge. First, the epistemic reading of *must* which means, "In view of the evidence, it's the case that John must write a novel" (that is, if he wants to get his story out). This expresses the modal notion of epistemic necessity. On another deontic reading, it would mean that he is obliged to write a novel in order for him to belong to a club, for example. On both of these readings, again, an extra necessity inference regarding John writing a novel is required in contrast to sentences without a modal auxiliary.

Thus, modality in language use can be understood as introducing the notions of possibility and necessity (Kratzer, 1979, 1981; Palmer, 1990).

All languages have non-factual mood markers (Bybee and Fleischman, 1995). In English, besides the modal auxiliaries (e.g., *would*, *should*, *can*, *must*, *ought*, *may*), there are suffixes (*-able*), adverbs (e.g., *likely*, *possibly*, *maybe*), and non-factive propositional attitude verbs (e.g., *consider*, *wish*, *hope*) among others, that serve as markers of non-factual mood. That is, in *John is likely to write a novel*, or *John is considering writing a novel*, non-factual mood is expressed such that John's writing a novel is again a hypothetical scenario. We call this class of lexical items "modal operators" (cf. Heim, 1982; Asher, 1987). Linguistically, modal operators introduce an inference, either possibility or necessity.

Furthermore, when a non-factual proposition is expressed, it can be followed by other hypothetical claims. What this means is that we can temporarily add a hypothetical proposition to the Common Ground (Stalnaker, 1978), and explore its consequences, as if it were true. That is, a later clause can be interpreted as being "modally subordinate" to a previous hypothetical statement. We will explore this next.

1.3. *Discourse Representation Theory and modal subordination*

Modal subordination is a fact about natural language, which is that propositions in discourse are logically related to each other in a hierarchical structure (Roberts, 1987, 1989). Roberts compares the structure of formal proofs in logic, where hypothetical assumptions are represented as subordinate to premises, and claims that non-factual clauses are represented as subordinate to factual clauses in Discourse Representation Structure (Heim, 1982; Kamp, 1981; Kamp and Reyle, 1993).

In Discourse Representation Theory, information that is conveyed in a discourse is represented structurally by a Discourse

[1] This string can be interpreted either as a simple main clause, such as The student spotted the error or a reduced relative clause, as in The student spotted by the proctor.

[2] For an introduction to this framework, see Chierchia and McConnell-Ginet, 1990.

Representation Structure (DRS) which is graphically depicted as a box. Sentences which discourse participants assert as facts are entered one after another onto the box representing the whole discourse, called the matrix DRS. The Kamp/Heim framework uniformly analyzes definite and indefinite Noun Phrases as variables, and information in the discourse is predicated on such variables. This intuitively makes sense, given that discourses are always about something or someone, which correspond to the syntactic category NP. For example, the DRS for the simple sentence *John wrote a novel* is represented as below:

```
x      y

John   (x)
novel  (y)

x wrote y
```

DRSes are uninterpreted formal structures. To be interpreted they are 'embedded' in a truth conditional model. "Formally, an embedding is a function from discourse referents onto individuals in the model, such that the individual which a given discourse referent *r* is mapped onto displays each property corresponding to a condition on *r*" (Roberts, 1987: 15). Kamp and Heim proposed that each (anaphoric) definite NP is translated as one of the variables that is already present in the DRS. For example, the following two-sentence discourse *John wrote a novel. It was funny* is represented as below:

```
x      y      z

John   (x)
novel  (y)
x wrote y

funny (z)
z=y
```

The pronoun variable *z* is associated with the same variable *y* which was introduced for the indefinite *a novel*. The resulting DRS is true if there is an embedding function which maps *y* onto a novel that John wrote and was funny.

In DRT, the antecedent for anaphora must not only be already present in the DRS but also 'accessible' to the pronoun. Accessibility is defined structurally—an antecedent is accessible if it is located at the same or superordinate level of box structure.

Roberts (1987, 1989) extends DRT to represent factual vs. non-factual propositions. That is, propositions that are in non-factual mood, as expressed by modals, non-factive propositional attitude verbs or adverbs, are contained in an embedded or subordinate box, thereby creating different levels of discourse structure in a DRS.[3] Thus, the non-factual proposition *John might write a novel* is represented as below in Fig. 1.

The event of novel-writing "*x* write *y*" is now contained not at the top level of the box but in a subordinate box, under the scope of the possibility operator in modal logic (◊). The name *John* is not contained in the subordinate box because names

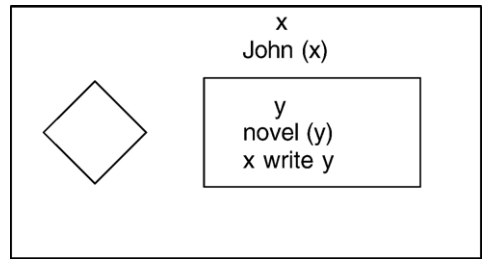

Fig. 1 – A DRS for a non-factual sentence.

are always entered at the top-level of a DRS, since their interpretation never interacts with operators. Thus, it has been minimally shown that the presence of a non-factual mood marker, such as a modal auxiliary, adds structure to a Discourse Representation. Below it will be argued that this theoretical hierarchy can account for grammatical intuitions regarding anaphoric possibilities.

For example, let us examine the two-sentence discourse below, where both the context sentence and the continuation sentence display non-factual mood.

(2) *John might write a novel. It could end quite abruptly.*

The context sentence indicates that John is not sure about whether he will be writing a novel, an abstract, or going fishing. The second sentence continues the non-factual mood and it means: "John might write a novel, and if he did, the novel could end quite abruptly (in view of what we know about John's tendencies, for example)."[4] Contrast (2) with (3), as first noted in a classic paper by Karttunen (1976).

(3) *John might write a novel. #The novel/It ends quite abruptly.*

The context sentence above is in non-factual mood, whereas the continuation sentence is not—it is asserted to be true, and hence, is factual. The discourse is infelicitous (marked as "#") not because of a switch in modality – this is done in language all the time – but because of the anaphoric dependency that exists between the sentences. In the context sentence, an indefinite Noun Phrase (NP), *a novel*, is mentioned. In the continuation sentence, an anaphoric expression, here a definite NP the novel, or pronoun, *it*, refers back to that indefinite. However, the indefinite antecedent is a hypothetical object, due to the mood of that sentence, while the anaphor is asserted to actually exist. This results in an anomaly—something is asserted to exist but previous information does not support this assertion.

Roberts (1987, 1989) accounts for the contrast in (2) and (3) in the following way: first, she uses Discourse Representation Structure (DRS) and maintains that the descriptive content of non-factual clauses is always represented as subordinate to factual clauses. Furthermore, she postulates that for anaphora to be felicitous, the anaphor must either be at an equal or subordinate level compared to its antecedent. Thus, in (2), both sentences are non-factual; hence, both antecedent and pronoun are located at equivalent non-factual "levels" (or "boxes" in terms of DRS). For this reason, the anaphoric dependency is fine. In contrast, in (3), the antecedent is located in a non-factual clause and it is consequently located at a

[3] Effectively, these subordinate boxes represent the scopal domains of modal operators (cf. Kadmon, 1987).

[4] For the sake of brevity, the resulting complex DRS is not shown.

subordinate level of the Discourse Representation. However, the anaphoric definite NP or pronoun is located in a factual clause, which means that the information is entered at a higher or superordinate level compared to the antecedent. As a result, the antecedent is "inaccessible" because it is too low in the discourse structure and, thus, anaphora is infelicitous. We note again that it is not the factual/non-factual switch in mood that creates infelicity. Compare (3) with *John wrote a novel. It might have ended quite abruptly*. In the latter example, the pronoun is in a non-factual clause and its antecedent is in a factual clause. Thus, the anaphor is subordinate to the level of the antecedent in the discourse hierarchy, and results in a felicitous discourse.

The goal of this paper is to empirically study the contrast in anaphoric felicity in discourses similar to (2) and (3). It is important to investigate this phenomenon for at least two reasons. First, while child language acquisition studies have shown that modal markers such as auxiliaries are acquired late developmentally (O'Neil and Atance, 2000; Noveck and Simon, 1996; Shatz and Wilcox, 1991) to our knowledge there are no empirical studies regarding the on-line processing of modality (in isolated sentences or in discourse context) by normal adults (for an exception, see Dwivedi, 1996). Second, to date, the majority of studies on pronoun resolution (e.g., Nicol and Swinney, 1989; Badecker and Straub, 2002; Sturt, 2003; among others) have examined antecedents that happen to be names. Although these studies have yielded important findings, a host of different questions emerge when another class of NPs is examined, namely, indefinite NPs. For example, names are inherently and unambiguously referential, whereas indefinite NPs are ambiguous between referential or specific interpretations vs. non-referential or non-specific readings (Fodor and Sag, 1982). As a result, names do not interact with the scopal properties of operators (Asher, 1987), whereas indefinite NPs do. As such, a context sentence such as *John is considering calling Martha* could be followed by *She never calls back*, since the name *Martha* exists independently of John's wishes. As was illustrated above, indefinite NP objects induce anomaly in such constructions.

In the study below, we examined whether constraints established by the use of modal markers in discourse affect on-line pronoun resolution. In particular, we examined whether anomaly is detected when an anaphoric pronoun is asserted to exist, despite the fact that its antecedent is contained in a hypothetical sentence.

We investigated this phenomenon using the neurophysiological measure of event-related potentials (ERPs). ERPs reflect voltage changes in electrical brain activity associated with cognitive processing. This methodology is particularly useful for our purposes because it allowed us to examine the processing of naturalistic language stimuli on-line with very high temporal resolution (on the order of milliseconds) and adequate spatial resolution (through scalp distribution). More importantly, for our purposes, electrophysiological studies on language have provided evidence for separate neural correlates of lexical–semantic vs. syntactic processing (see Hagoort et al., 1999 for an overview). Psycholinguistic work using ERPs began in 1980 with the publication of Kutas and Hillyard's seminal paper, and was originally focused on the study of single words and isolated sentences. It is only in the last 10 years or so that this methodology has been used to study sentence processing in discourse context (e.g., Van Petten, 1995; St. George et al., 1997; van Berkum et al., 1999a,b, 2003a,b).

With regard to the lexical–semantic ERP signatures, Kutas and Hillyard (1980, 1983) first showed that a semantically anomalous or unexpected word (e.g., *He spread the warm bread with* ***socks***) elicits a large N400 component, which is a distinct negative shift in the ERP waveform with a centro-parietal scalp distribution that begins about 200 ms after the onset of the critical word and peaks at about 400 ms post-stimulus. This is in contrast to semantically acceptable or expected words (e.g., *He spread the warm bread with* ***butter***) where the amplitude of the N400 is significantly smaller. St. George et al. (1997); van Berkum et al. (1999a,b) have recently demonstrated that, in addition to sentence-internal (conceptual) semantic information, N400 amplitude is sensitive to contextual constraints across discourse.

In contrast, the P600 is a later-occurring ERP component traditionally thought to be elicited by structural or syntactic aspects of linguistic input (Hagoort et al., 1993; Osterhout and Holcomb, 1992; Osterhout et al., 1994; Friederici, 1995, 1999, 2002; Friederici et al., 2002). This component has recently been considered as representing a family of components distributed across the scalp (Hagoort et al., 1999; van Berkum et al., 1999a,b; Friederici et al., 2002). For example, some workers argue that the cognitive processing associated with structural revision and re-analysis is reflected in a frontally distributed positivity, whereas posteriorly distributed positivity has been found in response to processes associated with structural integration and repair (Friederici et al., 2001, 2002; Hagoort et al., 1999; Kaan and Swaab, 2003a;b; Kaan et al., 2000). For the purposes of the present study, it is enough to recognize that the P600 effect represents structural, rather than semantic, processing. A recent study by Hammer et al. (2005) indicates that this component can be elicited across sentences.

It is an open question as to how the brain would process the anomaly given in (3). On the one hand, (3) can be characterized as a "meaning" or semantic anomaly. That is, the meaning of the pronoun embedded in the continuation sentence (which is factual) does not match the meaning of the antecedent in the context sentence (which is hypothetical). If this kind of "meaning mismatch" characterizes the phenomenon above, then we would predict an N400 effect for those sentences, much like an effect that van Berkum et al. (1999a,b) found.

On the other hand, linguistic theory gives us reason to expect a structural effect. That is, Roberts' (1987, 1989) Discourse Representation account of the semantic anomaly is couched in structural terms. According to her formulation, the reason why the discourse is infelicitous is because the antecedent indefinite NP is located at a subordinate level compared to the pronoun. Given this structural account of modal subordination violations, one could instead predict a P600-type of effect (note that we have no a priori reason to expect either a frontally or posteriorly distributed positivity). If a P600-type effect was indeed found, then it would be evidence for the necessity of representing non-factuality in discourse structure.

Regarding the assignment of co-reference by the processor, we assume the two-step process outlined in Garrod and

Sanford (1994) and Garrod and Terras (2000). That is, the first step ("bonding") involves a search of the mental representation for a potential antecedent without forcing any commitment. In the second stage ("resolution"), a commitment is made, and the contextual information is integrated into the (psychological) discourse representation.

1.4. *The current study*

We addressed the foregoing questions in the following way: we created two-sentence discourses where the first (context) sentence displayed non-factual mood and contained a potential NP antecedent (the direct object). The second (continuation) sentence contained a pronoun and either did or did not contain a modal auxiliary—that is, either it was non-factual or factual. We predicted an empirical difference between these two continuation sentences at the Verb position, since it is at that point that the processor has enough information to assign co-reference. However, this particular comparison is confounded by the fact that the target (continuation) sentence containing a modal auxiliary is compared to one lacking a modal, e.g., *It might end quite abruptly* vs. *It ends quite abruptly*. A Control context condition was constructed in order to ensure that the effects obtained are indeed due to context, and not the fact that two different kinds of continuation sentences are being compared. Thus, each continuation sentence – both those containing modals (non-factual) and those without modals (factual) – was preceded by two different kinds of contexts: Hypothetical and Control. This within-subjects study was defined by two independent variables: type of context (Hypothetical or Control) and type of continuation sentence (Non-factual (NF), that is, containing a modal auxiliary or Factual (F), defined by a lack of a modal auxiliary),[5] and measurements occurred at the Pronoun (P), Verb (V), and the Verb + 1 (V + 1) position. Table 1 lists the 4 conditions explicitly. Note that only Condition #H-F is anomalous, at the Verb position.

This design allowed for two direct comparisons, first between the factual sentence in the Hypothetical context, which should show anomaly at the Verb position, and the factual sentence in the Control context, which should not, and second, between the two coherent discourses – H-NF vs. C-NF – where no difference in waveforms should occur. In addition, a Filler Anomalous condition was included, where real-world violations were used, in order to elicit a classic N400 response for comparison purposes.

Table 1 – Overview of the experimental conditions

Type of context sentence	Type of continuation sentence
Control (i) John is reading a novel.	Factual It ends quite abruptly. P V V+1[a]
Hypothetical (ii) John is considering writing a novel.	Factual It ends quite abruptly. P #V V+1
Control (iii) John is reading a novel.	Non-factual It might end quite abruptly. P V V+1
Hypothetical (iv) John is considering writing a novel.	Non-factual It might end quite abruptly. P V V+1

Note that at the Verb position, Condition H-F, is anomalous, marked as #.

[a] ERPs were recorded at the Continuation sentence only, at the following sentence positions: P = Pronoun position; V = Verb position; V + 1 = after Verb position.

2. Results

2.1. *Electrophysiological analyses*

The grand average ERPs, time-locked at the position of the Pronoun (It), the Verb and the Verb + 1 position, are shown in (Figs. 2, 3A, B and 6), respectively. The general condition effects persisted over a long time period, and in order to understand the temporal dynamics associated with the changes in the waveforms, we analyzed the ERP data in consecutive 200 ms windows from 300 to 1100 ms (i.e., 4 time periods). Separate repeated measures ANOVAs were conducted for midline sites (i.e., Fz, FCz, Cz, CPz, Pz), and for medial–lateral electrode sites (i.e., F3, F4, FC3, FC4, C3, C4, CP3, CP4, P3, P4) at each of the aforementioned sentence positions. The ERP analyses reported below used SPSS v.11.0 statistical software and employed the Greenhouse–Geisser (Greenhouse and Geisser, 1959) non-sphericity correction for effects with more than one degree of freedom in the numerator. Following convention, unadjusted degrees of freedom are reported, along with the Greenhouse–Geisser epsilon value (ε) and adjusted p value. Mean square error values reported are those corresponding to the Greenhouse–Geisser correction. For ease of exposition, only significant main effects involving the linguistic factors of Context and/or Modality are reported, followed by the significant interaction effects involving these factors with the non-linguistic factors of Time and/or Electrode and/or Hemisphere. Our a priori hypothesis was that there would be a significant difference between the Hypothetical Factual condition at the Verb position (i.e., H-F; the anomalous condition) and its control (i.e., C-F), and no difference between H-NF and C-NF conditions, as the latter two conditions were coherent discourses. No integration effects were predicted to occur at the Pronoun position, nor at the Verb + 1 position.

2.2. *Statistical analyses at the pronoun position*

Fig. 2 shows the grand waveforms, averaged across all participants, at the position of the Pronoun for the H-F, C-F, H-NF, and C-NF conditions. Interestingly, most notably at left parietal sites, H-NF and H-F are shown as more negative–going compared to C-NF and C-F.

Separate repeated measures ANOVAs were conducted for the midline and medial–lateral electrodes. The factors included were: Context (2 levels: Hypothetical vs. Control), Modality (2 levels: Non-factual vs. Factual), Time interval (4 levels: 300–

[5] Note that the labels used for the levels of the context sentences, Hypothetical and Control, are analogous to the labels Non-factual and Factual. For clarity's sake; however, these different names will be used in order to differentiate levels associated with type of context sentences vs. continuation sentences.

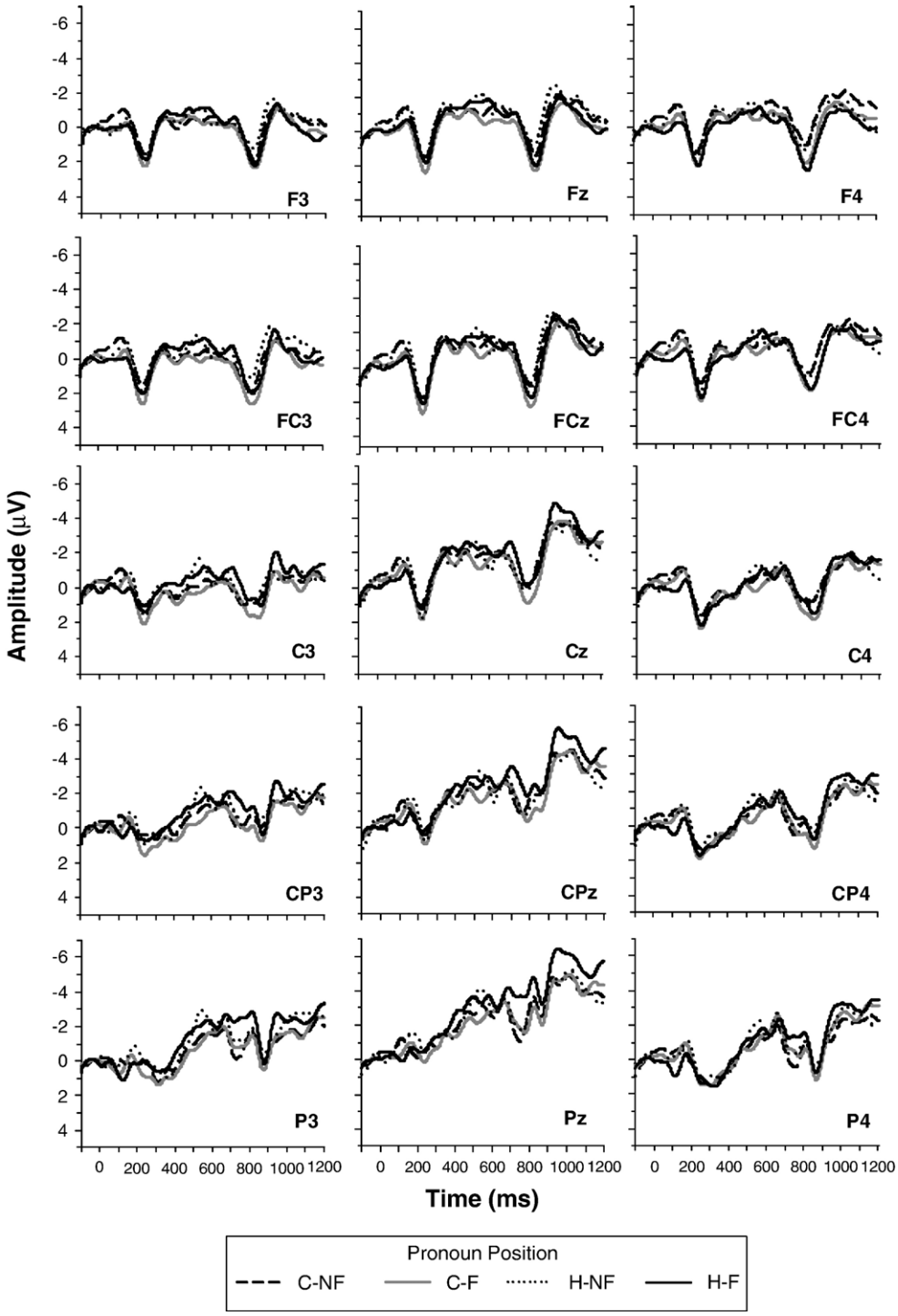

Fig. 2 – Grand average ERP waveforms measured at the Pronoun position for conditions C-NF (Control Non-factual), C-F (Control Factual), H-NF (Hypothetical Non-factual), and H-F (Hypothetical Factual). C-NF: dashed black line; C-F: solid grey line; H-NF: dotted black line; H-F: solid black line.

500 ms, 500–700 ms, 700–900 ms, 900–1100 ms), and Electrode site (5 levels: Fz, FCz, Cz, CPz, Pz) for the midline sites; for medial–lateral electrode sites, the Electrode factor was defined as anterior-to-posterior electrode sites (5 levels: F3/4, FC3/4, C3/4, CP3/4, P3/4) and Hemisphere (2 levels: Left vs. Right).

Midline analyses at the Pronoun position revealed no significant effects for the linguistic factors. Nevertheless, 2 trends were observed: first, a trend towards a main effect of Context was revealed ($F(1,22)=4.06$, $MSE=24.03$, $p=0.056$), indicating that the Hypothetical conditions (i.e., H-NF and H-F) were more negative-going than the Control conditions (i.e., C-NF and C-F). In addition, a trend towards an interaction of Modality × Electrode × Time was observed ($F(2,57)=2.79$, $MSE=4.59$, $p=0.057$, $\varepsilon=0.172$), which is consistent with visual inspection indicating that modal/non-factual conditions are more negative-going at frontal midline sites.

Medial–lateral analyses indicated no main effects of Context or Modality. A Context × Hemisphere interaction was observed ($F(1,22)=4.32$, MSE = 8.04, $p=0.049$), which pairwise comparisons revealed as Hypothetical conditions differing significantly ($p=0.03$) from Control conditions over the left hemisphere only. This showed that the H-NF and H-F conditions were more negative-going as compared to the C-NF and C-F conditions. In addition, a Context × Electrode × Hemisphere × Time interaction ($F(5,105)=4.32$, MSE = 0.17, $p=0.04$, $\varepsilon=0.396$) was observed, such that pairwise comparisons revealed a reliable difference for Hypothetical vs. Control context conditions, where the former are more negative-going for left centro-parietal sites

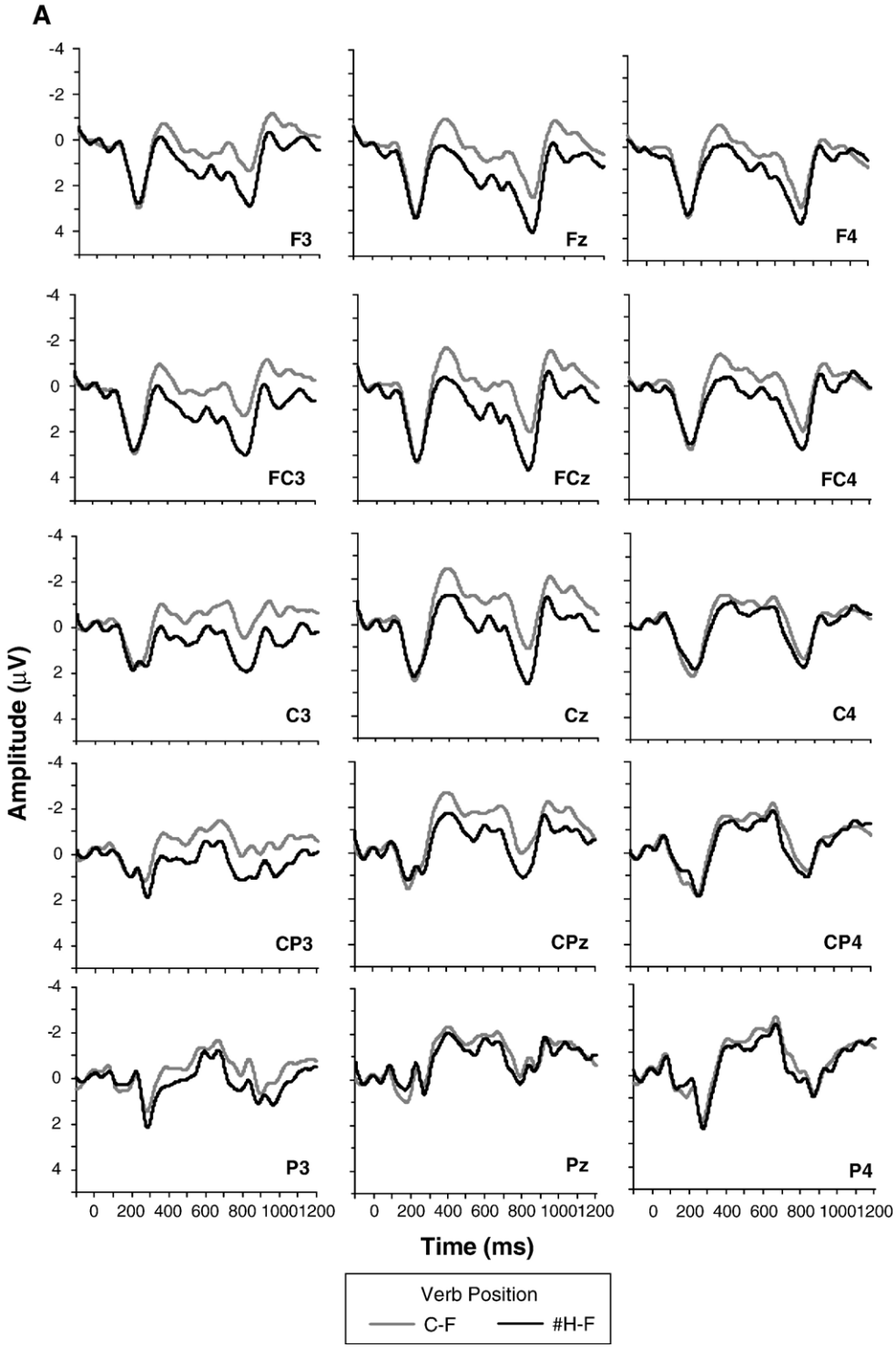

Fig. 3 – (A) Grand average ERP waveforms measured at the Verb position for conditions #H-F (#Hypothetical Factual) and C-F (Control Factual). H-F: solid black line; C-F: solid grey line. (B) Grand average ERP waveforms measured at the Verb position for conditions H-NF (Hypothetical Non-factual) and C-NF (Control Non-factual). H-NF: solid black line; C-NF: solid grey line.

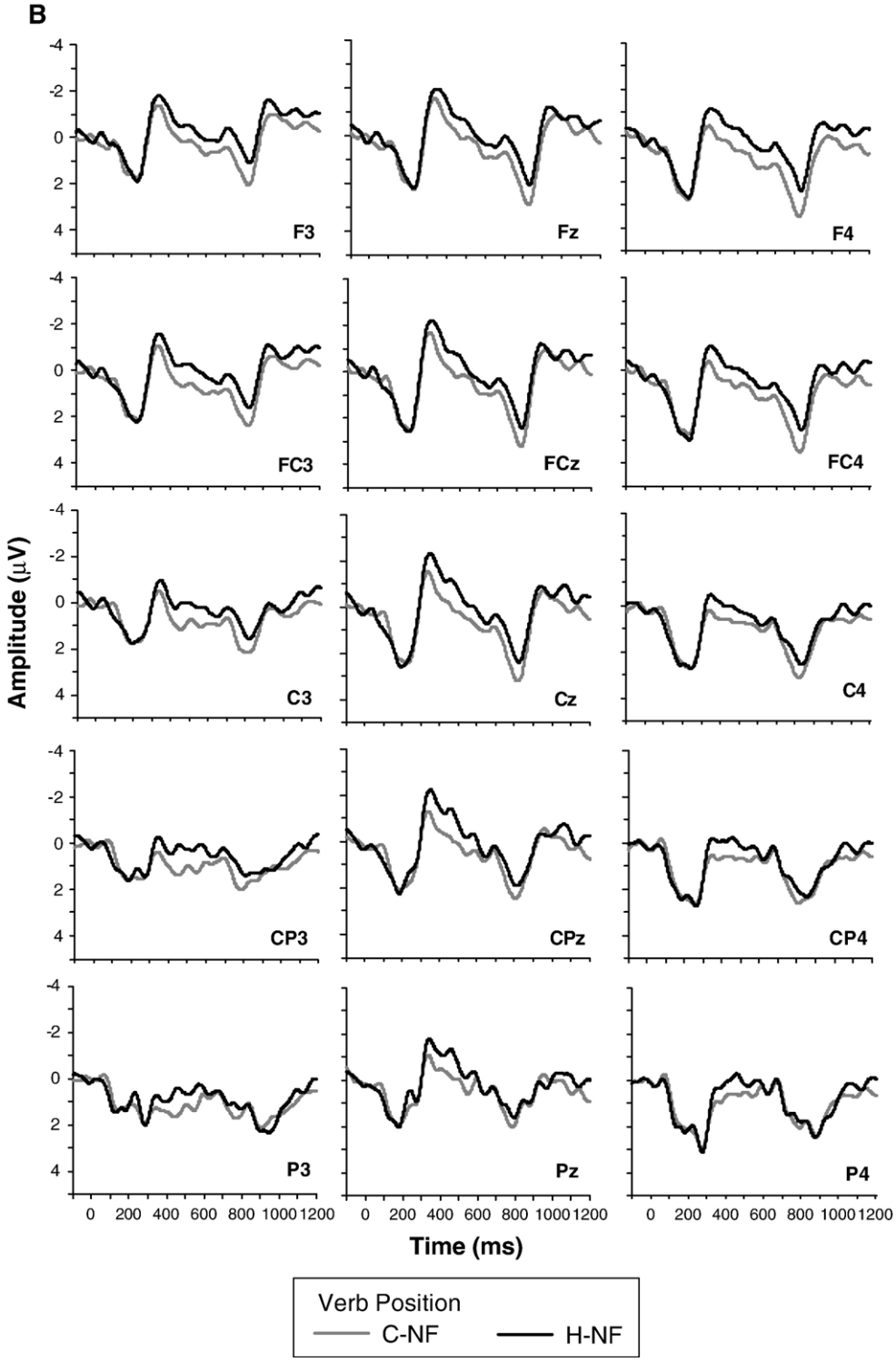

Fig. 3 (*continued*).

(i.e., C3, CP3, P3), starting at 500 ms, and maintaining a significant difference or strong trend until 1100 ms.

2.3. *Statistical analyses at the verb position*

2.3.1. *#H-F vs. C-F*

Figs. 3A and B show the grand waveforms, averaged across all participants, at the position of the Verb contrasting the #H-F and C-F conditions and the H-NF and C-NF conditions, respectively. The waveforms were characterized by well-defined N100-P200 components followed by sustained negative- and positive-going activity in the 300–1100 ms window, which varied across the left and right hemispheres and across the anterior-to-posterior axis of the scalp. Evoked responses to the word following the Verb are also visible beginning around 700 ms. As will be seen later, the comparison between the two

Non-factual conditions (i.e., H-NF, C-NF) did not differ reliably. However, central to the focus of the study, the anomalous condition (#H-F) was characterized by a slow positivity which appeared to be prominent at left and midline frontal locations (e.g., Fz, FCz, Cz, CPz, F3, F7), compared to the control condition (i.e., C-F).

2.3.1.1. Midline analyses. Separate repeated measure ANOVAs were conducted, including the factors Context (2 levels: Hypothetical vs. Control), Time interval (4 levels: 300–500 ms, 500–700 ms, 700–900 ms, 900–1100 ms), and Electrode site (5 levels: Fz, FCz, Cz, CPz, Pz), which compared conditions #H-F vs. C-F, and H-NF vs. C-NF, respectively.[6]

The #H-F vs. C-F comparison yielded a significant main effect of Context ($F(1,22)=5.02$, $MSE=35.50$, $p=0.036$) and a significant higher order interaction of Context × Electrode ($F(5,110)=4.68$, $MSE=6.14$, $p=0.014$, $\varepsilon=0.406$).

The Context × Electrode interaction was decomposed examining the simple effect of Context for each midline electrode (using Bonferroni correction) and is reported in Table 2. These analyses revealed that the #H-F condition was more positive than the C-F condition at all midline electrode sites except Pz. In order to better understand the time course of the difference between #H-F and C-F, the mean amplitude of the waveforms was computed in 50 ms windows ranging from 300 to 1100 ms for the sites at which the conditions differed (Fz, FCz, Cz, and CPz). These data were analyzed in a Context × Electrode × Time ANOVA which showed that the anomalous condition (i.e., #H-F) differed significantly as early as the second time window, 350 ms–400 ms, for all sites except CPz, where it was almost significant ($p=0.07$). Furthermore, this long-lasting positivity was significant over later time windows as well, most notably those typically associated with the P600 component (i.e., between 600 and 800 ms).

2.3.1.2. Medial-lateral analyses. Separate repeated measures ANOVAs were conducted on the medial–lateral electrode sites (i.e., F3, F4, FC3, FC4, C3, C4, CP3, CP4, P3, P4) which included the within factors Context (2 levels: Hypothetical and Control), Time interval (4 levels: 300–500 ms, 500–700 ms, 700–900 ms, 900–1100 ms), anterior-to-posterior electrode sites (5 levels: F3/4, FC3/4, C3/4, CP3/4, P3/4), and Hemisphere (2 levels: Left vs. Right).

Similar to the midline analysis, the #H-F vs. C-F analysis also yielded a main effect of Context ($F(1,22)=6.81$, $MSE=37.95$, $p=0.016$). A Context × Hemisphere interaction approached significance ($F(1,22)=3.92$, $MSE=8.61$, $p=0.06$) and reflected a trend towards a difference between #H-F and C-F over the left hemisphere compared to the right hemisphere. No other higher level interactions including Context emerged.

[6] Note that an omnibus ANOVA was not conducted using Modality as a factor. This was because direct comparisons between the non-factual and factual verb conditions would not have been ideal due to the fact that an additional word was present prior to the verb in the non-factual condition (i.e., the modal auxiliary). The processing of this additional word could have resulted in a difference in the pre-stimulus baseline leading into the time-locked activity at the Verb position. As such, we ran separate ANOVAs instead, each contrasting only #H-F vs. C-F, and H-NF vs. C-NF.

Table 2 – Pairwise comparisons for conditions #H-F and C-F at midline sites; reported using Bonferroni method

Site	F [a]	*df*	P
Fz	7.94	(1,22)	0.01
FCz	7.41	(1,22)	0.01
Cz	6.78	(1,22)	0.02
CPz	4.30	(1,22)	0.05
Pz	0.48	(1,22)	0.50

[a] Value for F is Wilks' lambda

2.3.2. H-NF vs. C-NF

For the H-NF vs. C-NF comparison at the midline sites, no significant main effect of Context was found ($F(1,22)=2.07$, $MSE=30.87$, $p=0.165$), nor did any significant interactions with Context emerge (all $Fs<1.7$, all p values >0.05).

For the medial–lateral sites, the main effect of Context was again not significant ($F(1,22)=3.36$, $MSE=42.76$, $p=0.08$), nor were there any higher order interactions.

As this comparison is crucial to our hypothesis, we followed up the #H-F vs. C-F and H-NF vs. C-NF comparisons using a difference-waveform analysis, as shown in Fig. 4. Visual inspection of Fig. 4 confirms that the positivity associated with the #H-F vs. C-F comparison is most evident at midline and left frontal sites. Separate repeated measures ANOVAs were conducted on the midline and medial–lateral electrode sites which included the within factors of Context Difference (2 levels: (H-NF minus C-NF) and (#H-F minus C-F)), as well as the Time, Electrode and Hemisphere factors. At the midline sites, the difference waveforms for #H-F minus C-F are indeed more positive-going than H-NF vs. C-NF, confirmed by a main effect of Context Difference ($F(1,22)=6.50$, $MSE=78.05$, $p=0.02$); at medial–lateral sites, this main effect was highly significant ($F(1,22)=8.11$, $MSE=94.32$, $p<0.001$). In addition, a trend towards an interaction between Context Difference × Hemisphere was observed ($F(1,22)=3.80$, $MSE=9.80$, $p=0.06$), which is consistent with visual inspection that the difference is left-lateralized.

2.3.3. Summary

A P600-like effect with Left frontal distribution was observed for the anomalous discourse (i.e., #H-F) at the Verb position compared to its control (i.e., C-F), whereas no significant difference was evident between the congruent (i.e., Non-factual) discourses at the Verb position.

It is interesting to note that a P600-like component was observed, rather than an N400 component. To ensure that the absence of an N400 was not due to idiosyncrasies of the experiment, responses to the Filler Anomalous and Filler Control stimuli were compared. As illustrated in Fig. 5, participants did indeed generate N400 effects in this comparison of the Filler conditions ($F(1,22)=19.91$, $MSE=38.35$, $p<.001$).

2.4. *Statistical analyses at the Verb + 1 position*

Fig. 6 shows the grand waveforms, averaged across all participants, at the position immediately after the Verb (i.e., Verb + 1 position), for all four conditions. The waveforms are characterized by an N100–P200 complex early on, and then again later when the next word was presented at 600 ms.

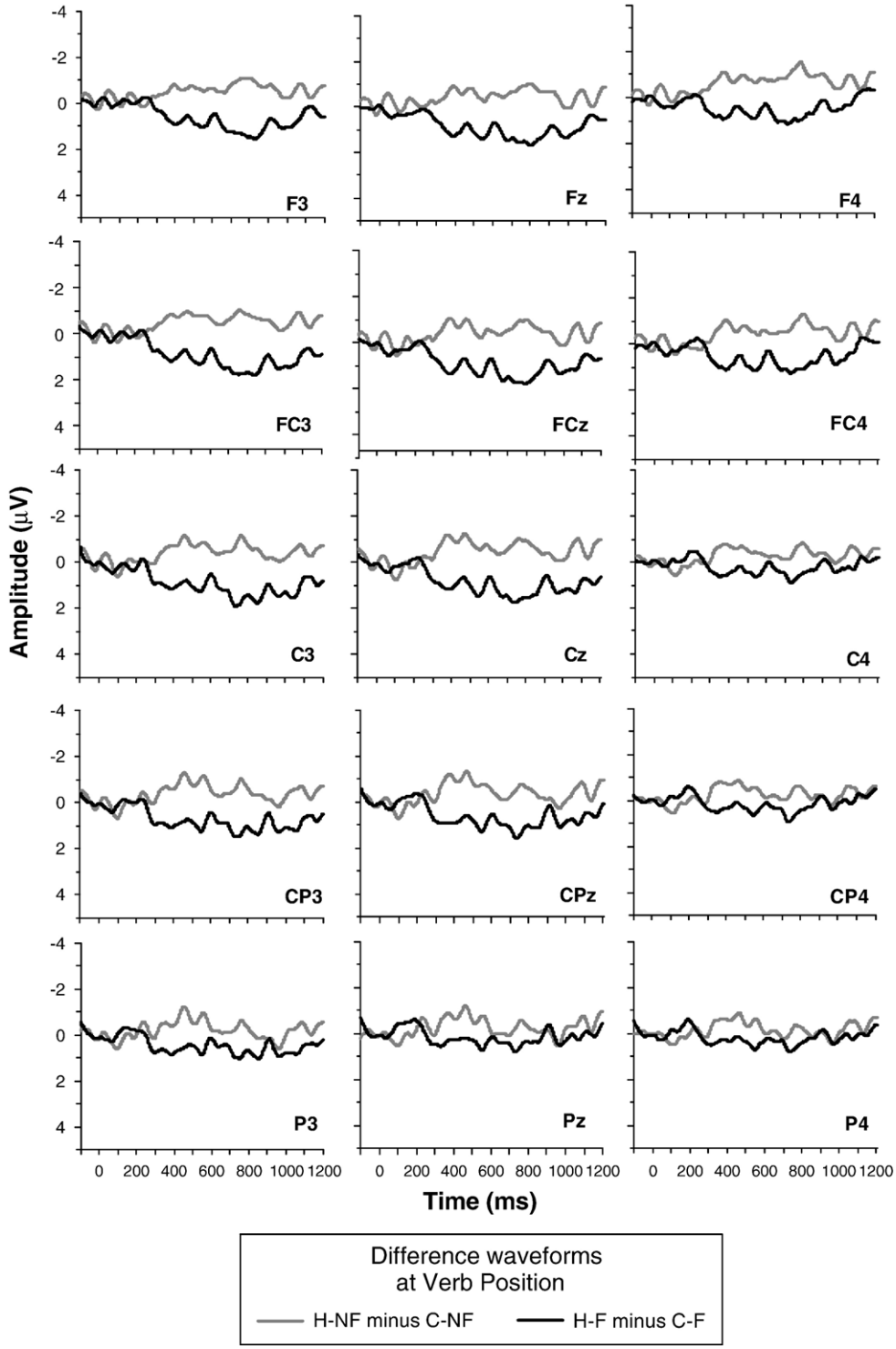

Fig. 4 – Grand average ERP difference waveforms measured at the Verb position for conditions H-NF minus C-NF (Hypothetical Non-factual minus Control Non-factual) and #H-F minus C-F (Hypothetical Factual minus Control Factual). H-NF minus C-NF: solid grey line; H-F minus C-F: solid black line.

Visual inspection reveals that the Control Non-factual condition elicited a slow negative shift at approximately 500 ms as compared to the other 3 conditions, most notably over the left hemisphere.

The ANOVAs conducted at the Verb+1 position followed exactly the same structure as those performed for the Pronoun position. Midline analyses revealed no main effects of Context or Modality. Only a Modality × Time interaction reached significance ($F(2,53)=4.18$, MSE=2.98, $p=0.015$, $\varepsilon=0.804$). Pairwise comparisons indicated that the Non-factual conditions were more negative-going as compared to the Factual conditions in the final time window of 900–1100 ms ($p=0.02$).

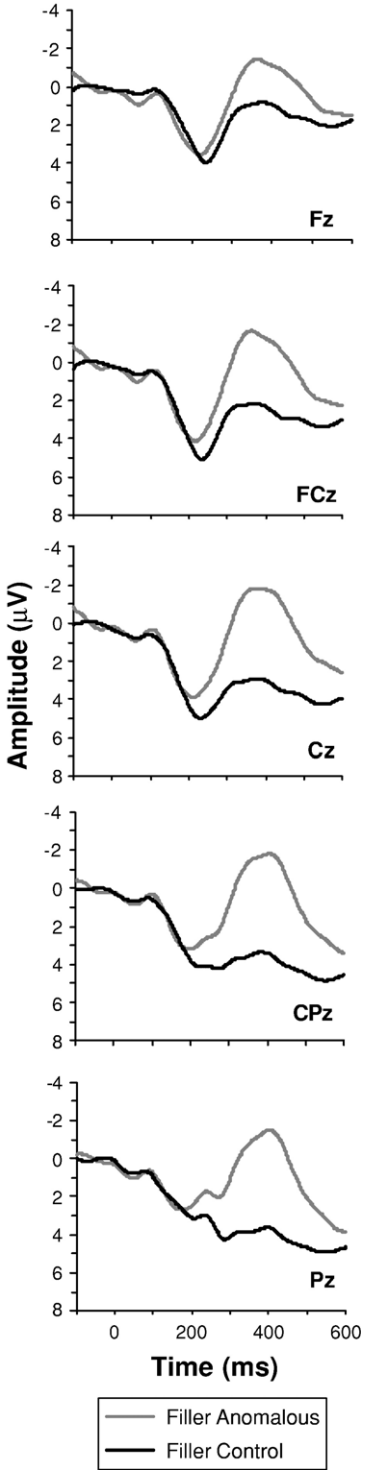

Fig. 5 – Grand average ERP waveforms measured at the final word position for the Filler Anomalous condition and the Filler Control condition. Filler Anomalous: solid grey line; Filler Control: solid black line.

At medial–lateral sites, a main effect of Context was observed ($F(1,22)=4.69$, MSE=39.53, $p=0.04$), indicating that overall, the Control conditions were more negative-going as compared to the Hypothetical conditions. In addition, a Context × Electrode × Hemisphere interaction ($F(2,38)=6.42$, MSE=2.25, $p=0.005$, $\varepsilon=0.435$) was observed. Subsequent analyses at single sites revealed that the Control conditions were more negative-going at sites C3, CP3, F4, FC4, and C4 (all p values < 0.05). A strong trend towards an interaction for Context × Modality × Electrode × Hemisphere ($F(2,45)=2.84$, MSE=2.00, $p=0.07$, $\varepsilon=0.517$) was consistent with the striking visual effect of the Control Non-factual condition differing from the other conditions. Subsequent analyses revealed that the Control Non-factual condition was significantly more negative-going than the Control Factual condition at all left centro-parietal sites; only F3 did not show the effect. In addition, a significant Modality × Hemisphere interaction was found ($F(1,22)=4.77$, MSE=12.41, $p=0.04$). Pairwise comparisons revealed that the Non-factual conditions were more negative-going over the left vs. the right hemisphere ($p<0.001$), and furthermore, over the left hemisphere, the Non-factual conditions were more negative-going than the Factual conditions ($p=0.01$). Again, visual inspection suggests that it is the Control Non-factual condition that may be driving this effect. Finally, a Modality × Time ($F(2,53)=4.09$, MSE=4.32, $p=0.02$, $\varepsilon=0.808$) interaction was observed, where subsequent analyses revealed that in the latest time window (i.e., 900–1100 ms) non-factual conditions were more negative-going than factual conditions.

3. Discussion

In the present study, ERPs were recorded while participants read well-formed vs. anomalous discourses displaying anaphora in Hypothetical and Control contexts. We were particularly interested in the ERPs elicited at the Verb position in the continuation sentences, which either contained a modal auxiliary or did not (non-factual vs. factual mood), since it was at this position that we predicted integration of contextual information to occur. The pattern of results can be summarized in the following way: as predicted, the anomalous discourse condition, #H-F, elicited a neurophysiological response at the Verb, such that the grand average waveform was more positive-going along left frontal and frontal midline sites, compared to its control (C-F). Importantly, no such difference was observed for the well-formed discourse conditions, H-NF and C-NF. Interestingly, two other positions in the continuation sentences also elicited ERP effects, of left parietal and frontal negativity. That is, at the Pronoun position (It), Hypothetical context conditions were more negative-going as compared to the Control conditions. In addition, at the position after the Verb (Verb+1), the Control Non-factual condition elicited negative-going waveforms as compared to the other three conditions.

We interpret the cognitive significance of these waveform effects in the following way: the positivity elicited at the Verb position is interpreted as a revision in terms of structural integration, whereas the negativity effects are both interpreted in terms of an increase in cognitive load, due to differing levels of complexity, to be defined below. We discuss each of these effects in turn, beginning with the predicted effect at the Verb position.

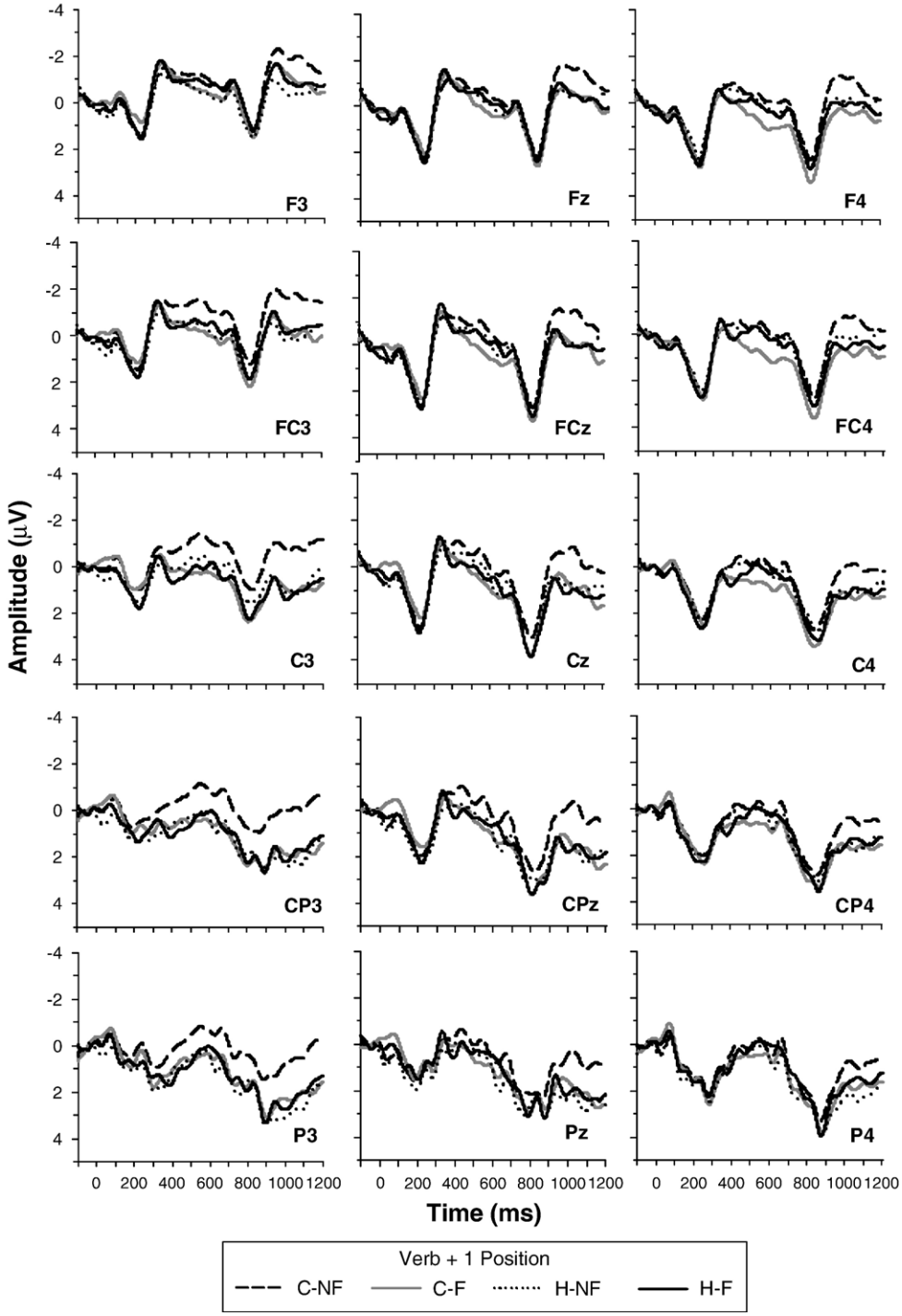

Fig. 6 – Grand average ERP waveforms measured at the Verb+1 position for conditions C-NF (Control Non-factual), C-F (Control Factual), H-NF (Hypothetical Non-factual), and H-F (Hypothetical Factual). C-NF: dashed black line; C-F: solid grey line; H-NF: dotted black line; H-F: solid black line.

3.1. *Processing discourse anomaly*

The current experiment combined ideas from recent work in linguistic theory on modality and discourse anaphora and predicted that the grammatical intuition showing anomaly in sentences that were not modally subordinate would be corroborated by neurophysiological evidence. That is, a continuation sentence in discourse context would be considered ungrammatical or anomalous because the Pronoun in the sentence was asserted to exist (since it was in a factual clause) although its antecedent, the indefinite NP object in the context sentence, was hypothetical (since it was in a non-factual clause). The neurophysiological results confirmed the grammatical intuitions, as

well as the behavioral evidence found in Dwivedi (1996), that discourses in the form of condition #H-F are anomalous, compared to condition C-F (see Experiment 1 in Dwivedi, 1996). We also examined whether the anomaly was structural or semantic/pragmatic in nature, by examining the type of ERP effect elicited. Interestingly, the response to this anomaly can be characterized as a P600-like component, with no evidence of N400 effects. That is, frontal positivity, evident at midline and left medial–lateral electrode sites, emerged at the Verb for the #H-F condition relative to the C-F. This frontal positivity is consistent with the growing literature that characterizes the P600 effect as a family of components that differ in terms of topography and underlying cognitive processes.

Specifically, it has been proposed that positivities with a frontal distribution are associated with syntactic or structural revision, whereas centro-posterior positivities are associated with syntactic repair as associated with ungrammaticality, e.g., *The spoil child *throw the toys on the floor* (Osterhout and Holcomb, 1992; Hagoort et al., 1993; Frisch et al., 2002; Kaan and Swaab, 2003b; Friederici et al., 2002; among others). Given that our anomalous condition, #H-F, was associated with frontal positivity, we interpret this finding to indicate a process of revision of structural analysis. We discuss the nature of this revision next, noting that since we did not originally set out to test for a particular kind of P600 component, what follows is speculative.

The notion of structural revision vs. repair is clearly spelled out in work by Kaan and colleagues. For example, Kaan and Swaab (2003b) found frontal positivity at the auxiliary verb is in the following sentence (i) *The man is painting the house and the garage is already finished*, compared to the unambiguous (ii) *The man is painting the house but the garage is already finished*. That work claims that the processor in (i) originally assumes that *the garage* is part of a conjoined Noun Phrase, but must revise this structure to include one where the NP is instead the subject of another clause. That is, the parser must abandon a preferred structure, which is dictated by processing considerations such as Minimal Attachment (Frazier and Fodor, 1978) for one that is non-preferred. This idea can be extended to the present study in one of two ways. Either it is the case that the previous Hypothetical context sentence must be revised in order for coherence to occur at the verb position or, alternatively, the interpretation of the verb itself in the continuation sentence must be revised.

Regarding revision at the context level, hypothetical context sentences, such as *John is considering writing a novel* have two possible readings for the interpretation of the direct object. The preferred reading is where John has no idea as to the particulars of the novel in question. This is called the "narrow scope" or non-specific reading of the indefinite object, which is in construction with the modal operator *consider*. On an alternative, less preferred reading, John does in fact have a specific novel in mind. Perhaps it is one that details his past (failed) romances. Consider the following paraphrase: *John is considering writing a (particular) novel (about his past failed romances). It ends quite abruptly*! With the additional restrictive information in brackets above, anaphora is licit. The alternative, non-preferred reading is called the "wide scope" or specific reading of *a novel* which, when computed, allows the continuation sentence without the modal auxiliary to make sense (Fodor and Sag, 1982; Diesing, 1992; Kurtzman and MacDonald, 1993). In terms of Discourse Representation Theory (DRT), this would amount to "moving" the position of the object NP from a subordinate box to the top level at Logical Form (see Fig. 1). This idea complements the proposal made in Saddy et al. (2004) where the P600 component was hypothesized to indicate re-analysis regarding the scope of positive polarity items in the scope of negative operators. In the present experiment, the relevant operator is the class of modal operators.

Alternatively, a reviewer suggests that it could be the case that it is the interpretation of the continuation sentence which must be revised, ostensibly at the Verb position, in order for the sentence to fit with the previous context. All the verbs in the continuation sentences are marked with present tense inflection *-s*, e.g., *ends*, *costs*, *hurts*, etc. The interpretation of the present tense morpheme *-s* holds the key to the possible revision at stake. It has been argued that the present tense morpheme in English does not only carry a meaning of marking an event as simultaneous with the time of speech, which is the present (Reichenbach, 1947; Partee, 1973). Previous studies (Smith, 1991; Cowper, 1998; Kamp and Reyle, 1993) have claimed that this morpheme can have more than one meaning, such that it can indicate a futurate or timetable meaning, as in the following: *Charlotte teaches syntax next semester*. That is, the present tense morpheme has a non-preferred meaning of the future. Once this is assumed, the nature of the revision process at the verb becomes clear. When the parser tries to assign an antecedent to the pronoun at the Verb position, it has already computed (via the "bonding" stage, see below) that the antecedent is located in a non-factual clause—it does not exist yet. In order to structurally integrate the present sentence with the context, the parser then chooses the less preferred meaning of *-s* inflection on the verb to yield the future tense, which is inherently non-factual (Roberts, 1987, 1989, Enç, 1996, 2004; Stowell, 2004). Note the coherence of the following discourses: *John is considering writing a novel. It ends quite abruptly (once he has written it*!), also, *The artist is likely sketching a portrait. It sells easily on the street (the moment she is finished*). Thus, the revision of the meaning of the present tense morpheme in the continuation sentence could account for the positivity observed at midline and left frontal sites. In fact, if this is correct, then an alternative, broader view of the functional significance of the P600 component would be relevant. Recent works (Münte et al., 1998a,b; Schmitt et al., 2002; Kolk et al., 2003; Van Herten et al., 2005) have indicated that besides being a pure structural marker, the P600 can be manipulated by compositional semantic information involving structural re-analysis, such that it has been recently claimed to be a monitoring component that checks the veracity of the processor's interpretation. Further research is required to investigate the two possible revision analyses mentioned above, as well as the cognitive significance associated with each process.

3.2. *Negativity and increased cognitive load*

3.2.1. *Search space considerations*

Next, it is significant to note that the point at which the anomaly was detected occurred at the Verb position, not the Pronoun position. As already pointed out, this is consistent with the model of anaphoric processing as assumed by Garrod and Sanford (1994); Garrod and Terras (2000). In that model, pronouns are not fully assigned antecedents immediately. Instead, first a search process occurs for possible antecedents, which are loosely recognized, and then actual integration occurs when enough disambiguating information arrives. This especially makes sense for a Pronoun like it, which can be either an epithet as in *It is raining*…. or *It seems that*…. in addition to actually co-referring to inanimate antecedents such as books and plants. Because of the ambiguity inherent with this Pronoun, the parser waits until the Verb to assign co-reference.[7] In the present experiment, ERP effects were observed for the Hypothetical conditions only, where left centro-parietal negativity was observed, starting at about 500 ms after the onset of the Pronoun. A comparison of the waveforms in Figs. 2 and 5 indicates that this negativity does not resemble a standard N400 effect. We interpret this negativity in terms of the findings of Kluender and Kutas (1993) and others (Mecklinger et al., 1995; King and Kutas, 1995; Münte et al., 1998a,b; Friederici et al., 1998; Fiebach et al., 2002; Matzke et al., 2002; van Berkum et al., 1999a,b; van Berkum et al., 2003a,b; Hammer et al., 2005) who claim that this sort of negativity represents an increase in working memory load. For example, Fiebach et al. found negativity in response to object (vs. subject) questions with a long (vs. short) distance between the WH-question and its filler. They expected, and found, that greater working memory demands would be placed on object vs. subject WH-questions due to a violation of canonical word order expectations, and due to the increased distance over which the WH-question must be held in memory before it can be integrated with its gap. In the present study, this negativity was observed only for the Hypothetical context conditions. That is, after processing sentences such as *John is considering writing a novel*, or *The artist is likely sketching a portrait*, when the processor perceived *it* in the next sentence, increased negativity was observed vs. when the previous context sentence was a control such as *John is reading a novel*, or *The artist is displaying a portrait*. According to the Garrod and Sanford model, at the Pronoun position, only "bonding" is occurring, where the processor is searching for a potential antecedent. Clearly, Hypothetical context sentences are more complex search spaces vs. Control contexts. First, these sentences happen to be longer, as they contain extra markers of non-factuality not present in the Control context sentences. Second, the non-factual mood markers such as *considering*, *likely*, *hoping*, carry the extra inference of possibility or necessity relative to sentences without them. As such, these sentences have extra "meaning" associated with them. We note here that the notions of semantic and structural complexity are closely intertwined issues. Thus, this semantic complexity could result in an extra processing load (cf. Gennari and Poeppel, 2003), and/or, the structural complexity that ensues at the level of discourse structure once the non-factual markers are processed, complicates the search space for the processor, which would tax working memory resources. The latter explanation would follow from the linguistic theory of Roberts (1987, 1989), which assigns a higher level of structural complexity to sentences that contain non-factual mood markers, since these must be represented at a subordinate level of the Discourse Representation Structure.

3.2.2. *The complexity of narrative shift*

Unexpectedly, ERP effects emerged at the position after the Verb, or the Verb+1 position (e.g., *John is considering writing a novel. It might end* ***quite*** *abruptly*; *The artist is displaying a portrait. It would sell* ***on*** *the street*). Namely, the Control Non-factual condition was visibly negative-going over bilateral centro-parietal sites as compared to the other 3 conditions. A possible explanation for complexity effects that would result in such negativity would follow from claims by Zwaan (1996, 1999), which examined shifts in temporal information. In Zwaan (1996), it was claimed that shifts in time resulted in an increased processing load, such that participants took longer to read sentences in a text that began with *an hour later* vs. *a moment later*. Effectively, the idea is that more cognitive effort is required to integrate sentences in a text that are not temporally contiguous vs. those that are, because the default assumption that subsequent sentences in a discourse relate subsequent and contiguous events is rejected. Analogously, the claim in the present study is that sentences that shift in mood require an extra processing effort. Of the four conditions, two conditions instantiate no switch in modality: H-NF and C-F vs. the other two that do: H-F and C-NF. However, of the two switch conditions, H-F is the anomalous condition, where at the Verb position, positivity results due to structural re-analysis. This could "pull down" any possible negativities at the following word. Furthermore, the nature of the shift in that condition is to a simpler one, that is, from hypothetical to factual (although see discussion above for an alternative account). Thus, the C-NF condition is one where an event that is semantically and structurally more complex must be represented after a simpler event. The negativity associated with this condition is the cost of interpreting and building a more complex event structure. This cost is realized only after the verbal complex has been perceived, that is, after the modal auxiliary and Verb position. This delayed effect of increased reading times was also shown in Zwaan (1996) for sentences indicating time shifts. For the current study, these delays make sense if we imagine that modal and tense information, like pronouns, is anaphoric (Partee, 1973, 1984; Roberts, 1996; Schlenker, 2004). Like other anaphors, modal information is integrated in a two-step process, where resolution would only occur after the relevant event information has been perceived.

[7] A reviewer notes that a third possible reading is available, where the pronoun "it" refers to a previous event, e.g., the activity of John's writing or reading a book. While this ambiguity may indeed be present, it is unclear how the observed ERP difference between the Hypothetical and Control conditions, as elicited at the Verb position, would occur if that were indeed the case. The results at the Pronoun position, however, are indeed compatible with such an interpretation. It is clear that investigating the anaphoric possibilities of multiply ambiguous pronoun "it" is subject for further research.

Finally, results indicated that representing non-factuality has an inherent cost: overall, waveforms representing non-factual conditions in the left hemisphere were more negative-going than factual conditions at the Verb+1 position. The cost of representing modality was also observed in Dwivedi (1996); experiment 2.

In conclusion, the main contribution of the present study is to show that during sentence comprehension, a discourse structure representing modal information must be consulted. Modality affects the structural complexity of a discourse structure: the negativity apparent at the Pronoun position was an effect of the Hypothetical context that characterized the space in which its antecedent was to be found. Furthermore, frontal positive-going waveforms at the Verb position indicated that form of the modal information is likely structural in nature. In addition, negativity effects at the position following the Verb show that there is a cost to representing non-factuality in comparison to factual sentences, especially when these follow Control contexts. In summary, the finding that modality is relevant at the level of discourse structure is consistent with previous results in the literature that claim that the tense or temporal structure of a discourse shows empirical effects (see, among others Anderson et al., 1983; Carreiras et al., 1997; Trueswell and Tanenhaus, 1991; Zwaan, 1996; Dickey, 2001; Gennari, 2004). Moreover, the findings are consistent with the linguistic formulation of modal subordination and Discourse Representation Theory as conceived of by Roberts (1987, 1989, 1996); Heim (1982); Kamp (1981). Finally, the fact that a P600-like effect was elicited for a discourse anomaly that can be conceived of as a "meaning mismatch" or semantic anomaly supports the idea that this component can be elicited for (truth-conditional) compositional, rather than conceptual, semantic anomalies (Münte et al., 1998a,b; Schmitt et al., 2002; Kolk et al., 2003; Van Herten et al., 2005).

4. Experimental procedures

4.1. Participants

Twenty-three native speakers of English (13 female, mean age 20.9 years, range 18 to 28 years) were recruited at McGill University and paid for their participation. All subjects had normal or corrected-to-normal vision and were right handed, as assessed by the Handedness Inventory (Briggs and Nebes, 1975). None of the participants reported any neurological impairment, history of neurological trauma, or use of neuroleptics. Also, none of them had participated in the pilot ratings task (see below).

4.2. Materials

4.2.1. Hypothetical conditions

Hypothetical context sentences were constructed in order to ensure non-factuality, such that they always contained a marker of non-factual mood (such as a modal adverb *possibly*, *likely*, *perhaps*, etc. and/or a non-factive propositional attitude verb such as *consider*, *muse*, *wonder*, etc.). In addition, the context sentence also used a verb of creation (such as *paint*, *bake*, *write*) in order to further bias for a non-specific (or non-existent) reading of the indefinite NP. The continuation sentence always contained a pronoun referring back to the indefinite NP antecedent. This pronoun appeared in a sentence that either contained or did not contain one of 5 epistemic modals: *would*, *might*, *may*, *must*, *should*. In order to minimize potential confounding factors of frequency and repetition, 20 high frequency verbs were used (Francis and Kucera, 1982). Note that it was impossible to match the frequency of the verbs and modals closely, as modals are closed-class items with extremely high frequencies of occurrence. However, data suggest that ERP differences in frequency within the category of "high frequency" words are minimal (Van Petten and Kutas, 1990).

4.2.2. Control conditions

The Control context did not exhibit any non-factuality markers. Therefore, no modal adverbs or non-factive propositional attitude verbs were used in the Control contexts. Instead, sentences contained verbs of using (such as *read*, *show*, *enjoy*) which presuppose the existence of their direct objects. As such, the existence of the indefinite NP was presupposed in the control sentences; that is, it 'actually' existed. The Control context sentence was followed by the same modal and non-modal continuation sentences as in the Hypothetical condition.

4.2.3. Filler sentence pairs

In order to reduce the chance of participants adopting particular reading strategies, 100 filler sentence pairs were also included. Half of the fillers controlled for the fact that the non-factual sentences in the targets were always semantically coherent. Thus, these additional filler sentences used modal auxiliaries not used in the target sentences (e.g., *could*, *can*, *ought to*, *did*, *will*) and were anomalous, but for reasons independent of grammatical constraints across sentences. Instead, these represented violations of real-world knowledge. An example of such a "Filler Anomalous" discourse is: *Celine will come to the party. She ought to bring skyscrapers*. These discourses were included in order to compare classic N400-like effects.

The other 50 filler sentence pairs controlled for verb repetition and consisted of coherent two-sentence discourses containing anaphora that used 10 high frequency verbs, which were not used in the target sentences, repeated 5 times each. These were necessary to include since our targets contained high frequency modals, which were repeated several times. Since we predicted no anomaly in sentences with modal auxiliaries, we needed to ensure that the lack of an N400 or P600 effect here was due to the acceptability of the modals, and not due to the fact that these are repeated high frequency items. Thus, these "Filler Control" sentence pairs were constructed such that the continuation sentence contained a pronoun and a high frequency verb, e.g., *The director was deciding which scene to keep. He cut the sad scene*.

In order to ensure that subjects paid attention to the stimuli, forced-choice content questions were asked about S2; these questions, however, only followed the 50 filler control sentences. Questions about the Filler Anomalous sentences were not used, since the sentences did not make sense, and

questions about the experimental sentences were not used in order to avoid encouraging any specific strategies for reading such sentences. When the questions were presented after the Filler Control stimuli, the two alternative answers were shown on the left- and right-hand side of the computer screen, and participants had to press the corresponding button on a response pad to indicate the correct answer. The position of the correct answer was counterbalanced across trials. For example:

S1: The boy enjoys his art work.
S2: He cuts with children's scissors only.
Question: What sort of scissors does he use?
Answer: CHILDREN'S or ADULT'S

In summary, the experiment consisted of the following items: 100 modal/non-factual continuation sentences, which were preceded by a Hypothetical or Control context sentence (100 H-NF + 100 C-NF), 100 non-modal/factual sentences which were preceded by a Hypothetical or Control context sentence (100 H-F + 100 C-F) and 100 fillers (50 Filler Anomalous and 50 Filler Control). This resulted in 500 discourse pairs.

In order to reduce repetition effects, the stimuli were divided into two counterbalanced lists, such that each participant saw an equal number of sentence pairs from each condition, resulting in 50 trials per experimental condition. Each list contained 300 sentence pairs in total, and consisted of 100 filler sentence pairs and half of each of the four experimental sentence pairs. Each participant saw one list only, with sentences presented in a pseudo-random fixed sequence with the constraint that no more than two sentence pairs of the same type could follow one another.

4.2.4. *Pretests*

We evaluated the acceptability of a preliminary version of the 500 discourses by conducting a norming study. Two semi-randomized lists were created and 62 subjects recruited at McGill University were paid $5 for their participation. None of these subjects participated in the actual experiment. In this off-line task, discourses were given in a pseudo-random order, with the constraint that no more than two of the same type of trial succeeded one another. The participants were asked to rate the extent to which the second sentence "fits" with the first one on a Likert scale ranging from 0 to 4, where 0 meant that it was not a good continuation at all, and 4 meant it followed naturally. Sentences that were not strongly rated as per predictions (e.g., a Hypothetical–Factual sentence rated as highly acceptable) were then modified for use in the present experiment. In all, the ratings confirmed the intended readings: the Filler Anomalous discourses were rated close to 0, the H-F discourses were rated in the 2 range, and all other conditions were rated above 3.

4.3. *Procedure*

For the experimental test, participants were tested individually in one session, which lasted approximately 2.5 h. Short breaks were given when required. Following the application of the EEG electrodes, subjects were seated in front of a computer screen approximately one meter away. All stimuli were presented in a 26 point white Arial font in the center of a SVGA computer monitor with a black background. Each context sentence (S1) was presented in its entirety; participants pressed a button to indicate when they were ready for the continuation sentence (S2). Following an ISI of 600 ms, the continuation sentence was presented one-word-at-a-time in the center of the screen with each word presented for 300 ms followed by an ISI of 300 ms. This presentation rate minimized eye movement artifacts in the EEG recordings and allowed for time-locking the EEG recording to the presentation of each word. Between each sentence pair, there was a 3-second delay to make sure the participants read the sentences as distinct pairs. Participants were instructed to silently read the context sentence, to press a button when it had been read, and to read each individual word of the subsequent sentence. Participants were instructed not to speak, move, or blink their eyes during the presentation of the stimuli. Practice trials were included to accustom participants to the task. When required, participants responded to a comprehension question using a hand held pad. This question appeared 100 ms after the last word of certain sentence pairs, and only occurred after filler control trials. On average, participants correctly answered these questions 91.4% of the time, indicating that they were indeed paying attention. Note that probe questions were not used on critical trials in order to ensure that participants would not develop processing strategies for these stimuli.

4.4. *Electrophysiological measures*

A commercially available nylon EEG cap containing silver/silver chloride electrodes (Quik-Cap) was used for EEG recording. The EEG was recorded from five midline electrode sites and 22 lateral sites. A cephalic (forehead) location was used as ground. All sites were referenced to the left ear during acquisition and re-referenced off-line to a linked ear reference. EOG was recorded from electrodes placed at the outer canthi of both eyes (horizontal EOG) and above and below the left eye (vertical EOG). EOG artifacts were corrected off-line for all subjects using a rejection criterion of ±50 μV, in accordance with the procedure as outlined in the Neuroscan 4.3 Edit (2004) manual. EEG was sampled continuously with critical EEG epochs time-locked to the onset of each target word of S2: the Pronoun, the modal (when present), the Verb, and the word after the Verb (i.e., Verb+1 position; this was never the final position in the sentence). EEG data were amplified using Neuroscan NuAmps in a DC-100 Hz bandwidth using a 500 Hz digitization rate. Single trial epochs were created using a −100 to 1100 ms window around the eliciting stimulus and processed off-line using Neuroscan Edit 4.3 software. For each participant, ERP averages were computed for the critical words in all target continuation sentences. The mean number of trials remaining after artifact rejection for the H-F condition was 32; for the other 3 conditions (H-NF, C-NF, C-F) the mean number of trials was 38 at the Verb and Verb+1 position. At the Pronoun position, the mean number of trials was remarkably less: 26 for all 4 conditions. This could be because participants did not always have time to move their eyes to the fixation point, since the Pronoun was

always the first word to appear in the continuation sentence after the Context sentence (which was presented as a block). The mean voltage amplitude of the – 100 to 0 ms period of each averaged waveform was calculated and served as the 0 μV baseline for post-stimulus activity. The mean amplitude of each waveform was computed in 200 ms intervals from 300 to 1100 ms post-stimulus, yielding 4 mean amplitudes. These effects were examined across five midline electrode sites (i.e., Fz, FCz, Cz, CPz, Pz) and medial–lateral electrode sites as defined in Results.

Acknowledgments

The first author was previously supported in part by a Canadian Institute of Health Research grant (MOP11290) awarded to Dr. Shari Baum and a Valorisation-Recherche Québec Fellowship during the development and running of this experiment, and was later supported in part by funding from the Social Sciences and Humanities research Council of Canada (grant number 412-2003-1003) awarded to Dr. Anna Maria Di Sciullo for the Major Collaborative Research Initiative on Interface Asymmetries. Thanks to Chuck Clifton, Karsten Steinhauer, Hotze Rullmann, Paul Portner, Eric McReady and Lyn Frazier for helpful comments during the development and analysis stages of this experiment. Also, thanks to Jason McDevitt and Erika Lawrence for help in preparing stimuli, and Erin Vensel and Pablo Ruiz for help in participant testing, as well as Tsee Leng Choy and Stephanie Einagel for assistance in preparing the manuscript.

REFERENCES

Anderson, A., Garrod, S., Sanford, A., 1983. The accessibility of pronominal antecedents as function of episode shifts in narrative texts. Q. J. Exp. Psychol. A. 35, 427–440.

Asher, N., 1987. A typology for attitude verbs and their anaphoric properties. Ling. Phil. 10, 125–197.

Badecker, W., Straub, K., 2002. The processing role of structural constraints on the interpretation of pronouns and anaphors. J. Exp. Psychol: Learn. Mem. Cogn. 4, 748–769.

Briggs, G., Nebes, R., 1975. Patterns of hand preference in a student population. Cortex 11, 230–238.

Bybee, J., Fleischman, S., 1995. Modality in Grammar and Discourse. John Benjamins, Amsterdam/Philadelphia.

Carreiras, A., Carreido, N., Alonso, M., Fernández, A., 1997. The role of verb tense and verb aspect in foregrounding of information during reading. Mem. Cogn. 25, 438–446.

Chierchia, G., McConnell-Ginet, S., 1990. Meaning and Grammar: An Introduction to Semantics. MIT Press, Cambridge.

Cowper, E., 1998. The simple present tense in English: a unified treatment. Studia Linguistica 52, 1–18.

Dickey, M.W., 2001. The Processing of Tense. Kluwer, Dordrecht.

Diesing, M., 1992. Indefinites. MIT Press, Cambridge.

Dwivedi, V., 1996. Modality and discourse processing. In: Montrul, S., Kessler Robb, M. (Eds.), McGill Working Papers in Linguistics 12, Department of Linguistics, McGill University. Montreal, pp. 17–52.

Enç, M., 1996. Tense and modality. In: Lappin, S. (Ed.), The Handbook of Contemporary Semantic Theory. Blackwell Publishers, pp. 345–358.

Enç, M., 2004. Rethinking past tense. In: Gueron, J., Lacarme, J. (Eds.), The Syntax of Time. The MIT press, Cambridge, pp. 203–215.

Fiebach, C.J., Schlesewsky, M., Friederici, A.D., 2002. Separating syntactic memory costs and syntactic integration costs during parsing: the processing of German WH-questions. J. Mem. Lang. 47, 250–272.

Fodor, J.D., Sag, I., 1982. Referential and quantificational indefinites. Ling. Phil. 5, 355–398.

Francis, W.N., Kucera, H., 1982. Frequency Analysis of English Usage. Houghton Mifflin, Boston.

Frazier, L., Fodor, J.D., 1978. The sausage machine: A new two-stage parsing model. Cognition 6, 291–326.

Friederici, A., 1995. The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. Brain Lang. 50, 259–281.

Friederici, A., 1999. The neurobiology of language comprehension. In: Friederici, A. (Ed.), Language Comprehension: A Biological Perspective. Springer, Berlin/Heidelberg/New York, pp. 263–301.

Friederici, A., 2002. Towards a neural basis of auditory language processing. Trends Cogn. Sci. 6, 78–84.

Friederici, A.D., Steinhauer, K., Mecklinger, A., Meyer, M., 1998. Working memory constraints on syntactic ambiguity resolution as revealed by electrical brain responses. Biological Psychology 47, 193–221.

Friederici, A., Mecklinger, A., Spencer, K., Steinhauer, K., Donchin, E., 2001. Syntactic parsing preferences and their on-line revisions: A spatio-temporal analysis of event-related brain potentials. Cogn. Brain Res. 11, 305–323.

Friederici, A., Hahne, A., Saddy, D., 2002. Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. J. Psycholinguist. Res. 31, 45–63.

Frisch, S., Schlesewsky, M., Saddy, D., Alpermann, A., 2002. The P600 as an indicator of syntactic ambiguity. Cognition 85, 83–92.

Garnham, A., 1981. Mental models as representations of text. Mem. Cogn. 9, 560–565.

Garrod, S., Sanford, A., 1994. Resolving sentences in a discourse context: how discourse representation affects language understanding. In: Gernsbacher, M. (Ed.), Handbook of Psycholinguistics. Academic Press, pp. 675–698.

Garrod, S., Terras, M., 2000. The contribution of lexical and situational knowledge to resolving discourse roles: bonding and resolution. J. Mem. Lang. 42, 526–544.

Gennari, S., 2004. Temporal references and temporal relations in sentence comprehension. J. Exp. Psychol. Learn., Mem. Cogn. 30, 877–890.

Gennari, S., Poeppel, D., 2003. Processing correlates of lexical semantic complexity. Cognition 89, B27–B41.

Greenhouse, S., Geisser, S., 1959. On methods in the analysis of profile data. Psychometrika 24, 95–111.

Hagoort, P., Brown, C., Groothusen, J., 1993. The syntactic positive shift as an ERP measure of syntactic processing. Lang. Cogn. Process 8, 439–483.

Hagoort, P., Brown, C., Osterhout, L., 1999. The neurocognition of syntactic processing. In: Brown, C., Hagoort, P. (Eds.), The Neurocognition of Language. Oxford University Press, Oxford, pp. 273–316.

Hammer, A., Jansma, B.M., Lamers, M., Münte, T.F., 2005. Pronominal reference in sentences about persons or things: an electrophysiological approach. J. Cogn. Neurosci. 17, 227–239.

Heim, I., 1982. The Semantics of Definite and Indefinite Noun Phrases. Unpublished Doctoral Dissertation. University of Massachusetts-Amherst.

Jespersen, O., 1965. The Philosophy of Grammar. Norton, New York.

Johnson-Laird, P., 1983. Mental Models: Towards a Cognitive

Science of Language, Inference and Consciousness. Cambridge University Press, Cambridge.
Kaan, E., Swaab, T., 2003a. Repair, revision and complexity in syntactic analysis: An electrophysiological differentiaton. J. Cogn. Neurosci. 15, 98–110.
Kaan, E., Swaab, T., 2003b. Electrophysiological evidence for serial sentence processing: A comparison between non-preferred and ungrammatical continuations. Cogn. Brain Res. 17, 621–635.
Kaan, E., Harris, A., Gibson, E., Holcomb, P., 2000. The P600 as an index of syntactic integration difficulty. Lang. Cogn. Proc. 15, 159–201.
Kadmon, N., 1987. On Unique and Non-Unique Reference and Asymmetric Quantification. Unpublished Doctoral Dissertation. University of Massachusetts-Amherst.
Kamp, H., 1981. A theory of truth and semantic interpretation In: Groenendijk, J., Janssen, T., Stokhof, M. (Eds.), Formal Methods in the Study of Language. Mathematical Centre Tracts, Amsterdam.
Kamp, H., Reyle, U., 1993. From Discourse to Logic: Introduction to Model-Theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory. Kluwer Academic, Boston.
Karttunen, L., 1976. Discourse referents. In: McCawley, J. (Ed.), Syntax and Semantics 7. Academic Press, New York, pp. 363–385.
King, J., Kutas, M., 1995. Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. J. Cogn. Neurosci. 7, 376–395.
Kintsch, W., 1998. Comprehension: A Paradigm for Cognition. Erlbaum, Hillsdale.
Kluender, R., Kutas, M., 1993. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. J. Cogn. Neurosci. 5, 196–214.
Kolk, H.H.J., Chwilla, D.J., van Herten, M., Oor, P.J.W., 2003. Structure and limited capacity in verbal working memory: a study with event-related potentials. Brain Lang. 85, 1–36.
Kratzer, A., 1979. Conditional necessity and possibility. In: Bauerle, R., Egli, U., von Stechow, A. (Eds.), Semantics from a Different Point of View. Springer, Berlin.
Kratzer, A., 1981. The notional category of modality. In: Eikmeyer, H., Rieser, H. (Eds.), Words, Worlds and Contexts: New Approaches in Word Semantics. de Gruyter, Berlin.
Kurtzman, H., MacDonald, M., 1993. Resolution of quantifier scope ambiguities. Cognition 48, 243–279.
Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.
Kutas, M., Hillyard, S.A., 1983. Event-related brain potentials to grammatical errors and semantic anomalies. Mem. Cogn. 11, 539–550.
Matzke, M., Mai, H., Nager, W., Russeler, J., Münte, T., 2002. The costs of freedom: An ERP-study of non-canonical sentences. Clinical Neurophysiology 113, 844–852.
Mecklinger, A., Schriefers, H., Steinhauer, K., Friederici, A., 1995. Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. Mem. Cogn. 23, 477–494.
Münte, T.F., Heinze, H.J., Matzke, M., Wieringa, B.M., Johannes, S., 1998a. Brain potentials and syntactic positive violations revisited: No evidence for specificity of the syntactic positive shift. Neurophychologia 36, 217–226.
Münte, T., Schlitz, K., Kutas, M., 1998b. When temporal terms belie conceptual order. Nature 395, 71–73.
Neuroscan Inc. 2004. Edit 4.3–Offline Analysis of Acquired Data. USA: Neurosoft Inc.
Nicol, J., Swinney, D., 1989. The role of structure in coreference assignment during sentence comprehension. J. Psycholing. Res. 18, 5–19.
Noveck, I., Simon, S.M., 1996. Children's understanding of epistemic modals. J. Child Lang. 23, 621–643.
O'Neil, D., Atance, C., 2000. 'Maybe my daddy give me a big piano': The development of children's use of modals to express uncertainty. First Lang. 20, 29–52.
Osterhout, L., Holcomb, P., 1992. Event-related brain potentials elicited by syntactic anomaly. J. Mem. Lang. 31, 785–804.
Osterhout, L., Holcomb, P., Swinney, D.A., 1994. Brain potentials elicited by garden-path sentences: evidence for the application of verb information during parsing. J. Exp. Psychol., Learn Mem Cogn. 28, 786–803.
Palmer, F., 1990. Modality and the English Modals. Longman, London/New York.
Partee, B.H., 1973. Some structural analyses between tenses and pronouns in English. The Journal of Philosophy 70, 601–609.
Partee, B.H., 1984. Nominal and temporal anaphora. Linguistics and Philosophy 7, 243–286.
Reichenbach, H., 1947. Elements of Symbolic Logic. University of California Press, Berkeley.
Roberts, C., 1987. Modal Subordination, Anaphora and Distributivity. Doctoral Dissertation. University of Massachusetts-Amherst.
Roberts, C., 1989. Modal subordination and pronominal anaphora in discourse. Ling. Phil. 12, 683–721.
Roberts, C., 1996. Anaphora in intensional contexts. In: Lappin, S. (Ed.), The Handbook of Contemporary Semantic Theory. Blackwell Publishers, pp. 215–246.
Saddy, D., Drenhaus, H., Frisch, S., 2004. Processing polarity items: Contrastive licensing costs. Brain Lang. 90, 495–502.
Schlenker, P., 2004. Sequence phenomena and double access readings generalized: Two remarks on tense, person, and moodIn: Gueron, Lecarme (Eds.), The Syntax of Time, pp. 555–596.
Schmitt, B.M., Lamers, M., Münte, T.F., 2002. Electrophysiological estimates of biological and syntactic gender violation during pronoun processing. Cogn. Brain Res. 14, 333–346.
Shatz, M., Wilcox, S., 1991. Constraints on the acquisition of English modals. In: Gelman, S., Byrnes, J. (Eds.), Perspectives on Language and Thought: Interrelations in Development. Cambridge University Press, Cambridge, pp. 319–353.
Smith, C., 1991. The Parameter of Aspect. Kluwer Academic Publishers, Dordrecht.
Stalnaker, R., 1978. Assertion. In: Cole, P. (Ed.), Syntax and Semantics v.9. Academic Press, New York.
St. George, M., Mannes, S., Hoffman, J., 1997. Individual differences in inference generation: An ERP analysis. J. Cogn. Neurosci. 9, 776–787.
Stowell, T., 2004. Tense and modals. In: Gueron, Lecarme (Eds.), The Syntax of Time, pp. 621–635.
Sturt, P., 2003. A new look at the syntax–discourse interface: the use of binding principles in sentence processing. J. Psycholing. Res. 32, 125–139.
Trueswell, J.C., Tanenhaus, M., 1991. Tense, temporal context and syntactic ambiguity resolution. Lang. Cogn. Processes 63, 303–338.
van Berkun, J.J.A., Brown, C.M., Hagoort, P., 1999a. Early referential context effects in sentence processing: Evidence from event-related brain potentials. J. Mem. Lang. 41, 147–182.
van Berkum, J.J.A., Hagoort, P., Brown, C.M., 1999b. Semantic integration in sentences and discourse: Evidence from the N400. J. Cogn. Neurosci. 11, 657–671.
van Berkum, J.J.A., Brown, C.M., Hagoort, P., Zwitserlood, P., 2003a. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. Psychophysiology 40, 235–248.
van Berkum, J.J.A., Zwitserlood, P., Hagoort, P., Brown, C.M., 2003b. When and how do listeners relate a sentence to the wider

discourse? Evidence from the N400 effect. Cogn. Brain Res. 17, 701–718.
Van Herten, M., Kolk, H.H.J., Chwilla, D., 2005. An ERP study of P600 effects elicited by semantic anomalies. Cogn. Brain Res. 22, 241–255.
Van Petten, C., 1995. Words and sentences: Event-related brain potential measures. Psychophysiology 32, 511–525.
Van Petten, C., Kutas, M., 1990. Interactions between sentence context and word frequency in event-related brain potentials. Mem. Cogn. 18, 380–393.
Webber, B., 1979. A Formal Approach to Discourse Anaphora. Garland, New York.
Zwaan, R., 1996. Processing narrative time shifts. J. Exp. Psychol., Learn. Mem. Cogn. 22, 1196–1207.
Zwaan, R., 1999. Situation models: the mental leap into imagined worlds. Current Directions in Psychological Science 8, 15–18.