BRAIN RESEARCH

Research Report

Establishing reference in language comprehension: An electrophysiological perspective

Jos J.A. Van Berkum[a,b,], Arnout W. Koornneef[a,c], Marte Otten[a], Mante S. Nieuwland[a]*

[a]*Department of Psychology, University of Amsterdam, The Netherlands*
[b]*F.C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands*
[c]*Utrecht Institute of Linguistics OTS, Utrecht, The Netherlands*

ARTICLE INFO

ABSTRACT

The electrophysiology of language comprehension has long been dominated by research on syntactic and semantic integration. However, to understand expressions like "he did it" or "the little girl", combining word meanings in accordance with semantic and syntactic constraints is not enough—readers and listeners also need to work out *what or who* is being referred to. We review our event-related brain potential research on the processes involved in establishing reference, and present a new experiment in which we examine when and how the implicit causality associated with specific interpersonal verbs affects the interpretation of a referentially ambiguous pronoun. The evidence suggests that upon encountering a singular noun or pronoun, readers and listeners immediately inspect their situation model for a suitable discourse entity, such that they can discriminate between having too many, too few, or exactly the right number of referents within at most half a second. Furthermore, our implicit causality findings indicate that a fragment like "David praised Linda because…" can immediately foreground a particular referent, to the extent that a subsequent "he" is at least initially construed as a syntactic error. In all, our brain potential findings suggest that referential processing is highly incremental, and not necessarily contingent upon the syntax. In addition, they demonstrate that we can use ERPs to relatively selectively keep track of how readers and listeners establish reference.

© 2006 Elsevier B.V. All rights reserved.

1. Introduction

Words reliably mean things. In fact, something about word meaning is invariant enough to be listed in a dictionary. Here, for example, is what the *Cambridge Advanced Learners Dictionary* lists for the word "girl":

girl (noun): **1.** a female child or young woman, especially one still at school: "*Two girls showed us round the classrooms.*"

Of course, we all know that when a word is placed in the context of other words, shades of meaning can emerge. For "girl", some of the most familiar ones are listed in the dictionary too:

girl (noun): ……… **2.** a daughter: "*We have two girls.*", "*My little girl is five.*" **3.** [usually plural] a woman worker, especially when seen as one of a group: "*shop/office girls*" **4.** [always plural] one of a group of female friends: "*I'm going out with the girls tonight.*", "*The girls at work gave it to me*"

---

* *Corresponding author.* University of Amsterdam, Department of Psychology (PN), Roetersstraat 15, 1018 WB Amsterdam, The Netherlands. Fax: +31 20 639 1656.
E-mail addresses: berkum@psy.uva.nl, J.J.A.vanBerkum@uva.nl (J.J.A. Van Berkum).

However, the flexibility of language allows us to go far beyond this. For example, as revealed by a brief Internet search, speakers can use "girl" for their dog ("This is my little girl Cassie…she's much bigger and has those cute protruding bulldog teeth"), their favorite boat ("This girl can do 24 mph if she has to"), or a recently restored World War II Sherman tank ("The museum felt that the old girl was historically unique"). Such examples reveal that for nouns, it is often not enough to just retrieve their *sense*, i.e., some definitional meaning, from our mental dictionaries.

It has been known for a long time (Frege, 1892) that there is another side to meaning, involving *reference*. To find out what "the girl" means in context, for example, we also need to work out to what or whom it refers. Is "the old girl" the tank they recently dug up and restored, or perhaps the one standing next to it? Does his powerboat or his truck do 24 mph? And which of her three daughters happens to be "my little girl"? The importance of establishing reference is even more obvious when it comes to understanding pronouns, such as "she", "they", or "those". Pronouns are amongst the most frequently encountered words in the language, and they play a vital role in talking about who did what to whom. You might be a walking dictionary when it comes to word sense, but if you're unable to work out who's who in language, you are lost.

Because of their central role in language comprehension, the processes involved in establishing reference have been studied intensively by cognitive scientists, using a wide range of behavioral methods (e.g., Garnham, 2001; Trueswell and Tanenhaus, 2005). However, so far, surprisingly few *neuro-imaging* studies on the topic exist. The dozens of EEG studies on sentence comprehension conducted over the last 25 years, for example, have predominantly addressed issues of syntax, semantics, and the interplay between the two (see Brown et al., 2000, Kutas et al., 2000; Osterhout et al., 2006, for reviews). A comparably selective focus can be seen in more recent fMRI research. In part, the lack of reference-oriented neuro-imaging research can be traced back to the fact that referential issues often require a mini-text for each trial, a requirement that is relatively difficult to combine with fMRI or EEG (Van Berkum, 2004). Furthermore, EEG research on language emerged in the sentence processing research community, which has long been dominated by issues of syntactic modularity. Finally, the two most famous and most robust ERP phenomena relevant to language comprehension, the *P600* and *N400* effects, happen to be strongly associated with syntactic and semantic unification processes respectively. Hence, for researchers looking for a feasible first ERP study, not reference, but syntax and semantics was the way to go.

In the present article, we review our ERP research on when and how readers and listeners establish reference, and we describe one new experiment in full detail. The research program at hand serendipitously took off with the discovery of a sustained frontal negative shift to referentially ambiguous nouns (Van Berkum et al., 1999a), in a study aimed at context effects in syntactic parsing. Initial follow-up work examined the replicability and generality of this referentially induced brain response (Van Berkum et al., 2003a; Van Berkum et al., 2004). In more recent work, we have begun to relate the referential negativity to models of text comprehension

(Nieuwland et al., in press) and individual differences (Nieuwland and Van Berkum, submitted for publication).

Although we will argue that the referential negativity allows us to selectively keep track of certain referential aspects of language comprehension, an important general point will be that what happens at a certain level of the comprehension system can sometimes be observed via the consequences for some other level of the system, and the ERP effect(s) typically associated with the latter. We will illustrate this in our review, as well as by means of the new experiment reported here. In the latter, we examine when and how a relatively subtle factor, the implicit causality associated with specific interpersonal verbs, affects the interpretation of a referentially ambiguous pronoun.

Note that several other psycholinguists have also begun to examine reference-related issues by means of EEG (e.g., Anderson and Holcomb, 2005; Camblin et al., 2007; Hammer et al., 2005; Magne et al., 2005; Rösler et al., 1998; Schmitt et al., 2002; Swaab et al., 2004); we refer to Camblin et al., 2007, and Van Berkum, in press, for broader surveys.

2. Signs of referential processing

2.1. The referentially induced frontal negativity (Nref effect)

To study the impact of discourse-level referential factors on incremental syntactic parsing, we conducted an ERP experiment in which we manipulated the number of candidate referents for a singular definite noun phrase (Van Berkum et al., 1999a). In the experiment, participants were asked to read several mini-stories, such as the one below (translated from Dutch, boldface added):

(1a) David had asked the boy and the girl to clean up their room before lunchtime. But the boy had stayed in bed all morning, and the girl had been on the phone all the time. David told **the girl** that had been on the phone to hang up.

In (1a), the discourse context provided a *single* unique referent for the critical noun phrase "the girl". However, to lure the syntactic parser into expecting a relative clause right after the noun, sometimes the wider discourse provided *two* equally eligible referents, as in (1b):

(1b) David had asked the two girls to clean up their room before lunchtime. But one of the girls had stayed in bed all morning, and the other had been on the phone all the time. David told **the girl** that had been on the phone to hang up.

This simple referential manipulation turned out to have a clear effect in ERPs elicited by the critical noun: relative to its referentially successful counterpart, a referentially ambiguous noun elicited a widely distributed negative deflection, emerging at about 300 ms after noun onset. The effect, which for ease of reference we dub the *Nref* effect here, was largest at anterior sites, where it was also particularly sustained (see Fig. 1, left). These features made the effect very different
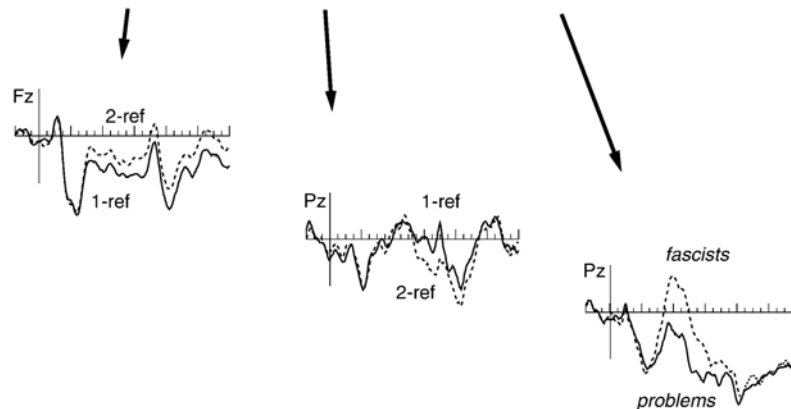
Fig. 1 – Three types of discourse-induced ERP effects during written language comprehension. From left to right: a sustained frontal negative shift (Nref effect) to a discourse-induced referential problem ("friend" is referentially ambiguous in the 2-referent context), a P600 effect to a discourse-induced syntactic problem ("there" rules out the provisional relative-clause analysis pursued at "that" in the 2-referent context), and an N400 effect to a discourse-induced semantic problem ("fascists" does not fit the wider story context). The example item is shown here in several variants (1- and 2-referent contexts, coherent/anomalous ending), but any one participant saw just a single variant. Data were obtained in a single ERP experiment, of which the referential and syntactic aspects were described in Van Berkum et al. (1999a,b), and the semantic aspects in Van Berkum et al. (1999c). Negative voltage is up, and time ranges from 150 ms before until 1200 ms after critical word onset.

from a syntax-related P600 effect (Fig. 1, middle) and a semantics-related N400 effect (Fig. 1, right), obtained for the same participants in the same written stories.

This result had several important implications. First of all, the rapid divergence of ERP waveforms elicited by referentially ambiguous and unambiguous nouns by itself suggested that language users can very rapidly determine whether a singular definite noun has a unique referent in the earlier discourse or not, in this case within at most some 300 ms. Note that the candidate referents had been introduced non-locally, in the two sentences that preceded the critical sentence with the referring noun phrase. The speed with which our readers were able to detect that the latter was referentially ambiguous is therefore at odds with the common assumption that non-local, discourse-dependent processing is slow (see Van Berkum, in press, for extensive discussion). Secondly, because the ERP effect at hand was qualitatively very different from both the classic semantic N400 effect (Kutas and Hillyard, 1980) and the classic syntactic P600 effect (Osterhout and Holcomb, 1992; Hagoort et al., 1993), it suggested that the neural systems recruited by ambiguity-related referential processing differ from those involved in semantic and syntactic integration. Third, and illustrated by Fig. 1, our results implied that ERPs allow one to selectively keep track of at least three important aspects of sentence- and discourse-level language comprehension: the analysis of syntactic structure, the integration of word meanings, and the identification of who or what is being talked about.

The initial results, redisplayed with the associated scalp topography in Fig. 2A, had been obtained with written language presented at a relatively slow and fixed rate of one word every 600 ms. However, in two replication studies in which the same stories were presented as fully connected natural speech, referentially ambiguous nouns elicited comparable sustained frontal negativities (Van Berkum et al., 2003a; see Fig. 2B). The presence of an Nref effect therefore does not hinge on whether the language is presented in written or spoken form, and reflects something stable about how readers as well as listeners establish reference.

But what exactly? All studies discussed so far had used materials in which a referentially ambiguous noun had quite frequently been followed by a referentially disambiguating relative clause (e.g., "…that had been on the phone"). The Nref effect might thus perhaps reflect the participant's expectation for a disambiguating relative clause. To examine this possibility and to more generally determine the scope of the phenomenon at hand, Van Berkum et al. (2004) had participants listen to referentially unambiguous and ambiguous singular pronouns, as in (2a) and (2b), translated from Dutch.

**(2a)** David shot at Linda as he jumped over the fence.

(2b) David shot at John as he jumped over the fence.

Relative to their unambiguous counterparts, referentially ambiguous spoken pronouns elicited a relatively sustained frontal negativity (see Fig. 2D), much like referentially ambiguous spoken nouns had done. Because in colloquial Dutch a pronoun is never followed by a restrictive relative clause, this ruled out the possibility that the Nref effect indexed the anticipation of such a clause. Rather, it confirmed our suspicion that we were looking at a more general aspect of establishing reference. As shown in Fig. 2C, we recently also

2-ref : David had asked the two girls to clean up their room before lunchtime. But one of the girls had stayed in bed all morning, and the other had been on the phone all the time. David told the...

1-ref : David had asked the boy and the girl to clean up their room before lunchtime. But the boy had stayed in bed all morning, and the girl had been on the phone all the time. David told the...
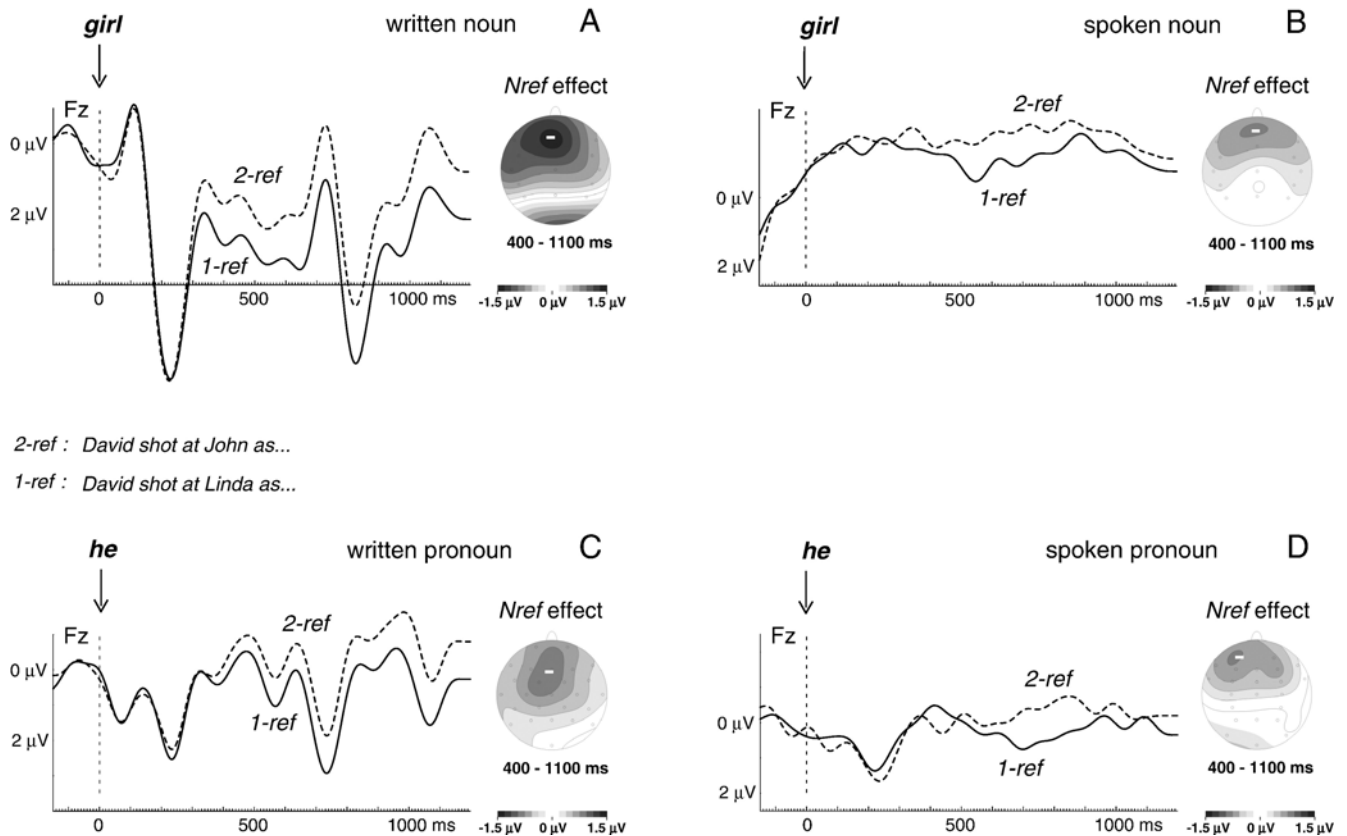
2-ref : David shot at John as...

1-ref : David shot at Linda as...

Fig. 2 – The Nref effect, a sustained frontal negative shift to referentially ambiguous written nouns (A), spoken nouns (B), written pronouns (C) and spoken pronouns (D). Data from Van Berkum et al. (1999a), Van Berkum et al. (2003a, Exp. 1, b), Nieuwland and Van Berkum (submitted for publication) and Van Berkum et al. (2004) respectively. Negative voltage is up.

obtained a sustained frontal negative shift to ambiguous pronouns in reading (Nieuwland and Van Berkum, submitted for publication), one that is very similar to the effect obtained with written nouns. In all, referential ambiguity elicits a comparable Nref effect in the ERPs to nouns and pronouns, independent of whether people read at a fixed speed of 600 ms per word or listen to fully connected natural speech.

Although the early onset of the Nref effect show that readers and listeners try to establish reference rapidly enough to that discover referential ambiguity within a mere 300–400 ms, it was as yet not clear what the relevant type of ambiguity was. Research on text comprehension (e.g., Kintsch, 1998) has shown that comprehenders not only construct a deep 'situation model' (Zwaan and Radvansky, 1998) to represent what the discourse is about, but also keep track of what has been said at a somewhat more superficial, propositional level, often referred to as the 'textbase' (Kintsch, 1998). In terms of these two levels of representation, note that "the girl" in (1b) not only sensibly refers to two *suitable* entities in the situation model, but also indexes (or 'resonates' with, Myers and O'Brien, 1998)

two recently activated memory tokens of the concept girl, regardless of their suitability. Now consider (1c):

(1c) David had asked the two girls to clean up their room before lunchtime. But one of the girls had been sunbathing in the front yard all morning, and the other had actually just driven off in his car for some serious downtown shopping. As he gazed at the empty driveway, David told the girl …

In the situation described here, "the girl" can only sensibly refer to a single person. However, in terms of recently activated memory tokens in the textbase, "the girl" is as ambiguous as before.

By using comparable critical spoken stories in which one of two persons left the scene, Nieuwland et al. (in press) have recently demonstrated that the referentially induced frontal ERP effect reflects deep, situation-model ambiguity, rather than the ambiguity associated with a more superficial resonance-type memory access mechanism. In particular, whereas stories in which a singular noun referred to two

eligible candidates in the discourse model (e.g., (1b)) elicited the Nref effect, stories such as (1c) in which one of the two candidates had been rendered an unsuitable referent did not. This shows that reference is established very rapidly at the deepest level of representation, with nouns contacting relevant situation-model discourse entities within only a few hundred milliseconds after their acoustic onset. Furthermore, these findings suggest that ERPs can be used to track certain aspects of referential processing at the level that is most relevant for comprehension, the situation model.

The fact that the Nref effect reflects something about establishing reference with respect to the situation model provides a clear constraint on its functional interpretation. Furthermore, and as discussed in the next section, we know that the effect is not a simple reflection of *any* problem with establishing reference. However, what we do not yet know is what is happening exactly in the brains of our listeners and readers when they generate this effect. First, we cannot exclude the possibility that the Nref effect is some neural correlate of *noticing* that a referring expression is ambiguous (see Nieuwland and Van Berkum, submitted for publication, for discussion). A second and related possibility is that the Nref effect indexes controlled processing as people are trying to resolve the ambiguity (e.g., by additional inferencing, and/or an additional search for cues in episodic memory of the discourse; Myers and O'Brien, 1998). A third possibility is that it reflects the additional neuronal activity required to simultaneously keep two competing referential interpretations in working memory (Gibson, 1998). The latter would explain why the Nref effect resembles the frontal ERP effects – sometimes referred to as *LAN* effects – elicited by various other expressions that tax working memory, such as (a) object-relative clauses like "The fireman *who the cop speedily rescued* sued the city over working conditions" (King and Kutas, 1995; Kutas, 1997; Müller et al., 1997) in which the reader or listener must deal with "the cop" without yet knowing what to do with "the fireman", (b) temporal expressions like "*Before* the psychologist submitted the article, the journal changed its policy" (Münte et al., 1998) in which the information supplied in the first phrase does not describe what actually happened first, or (c) expressions like "The *pitcher* fell down and broke/cursed" that contain a lexically ambiguous word (Hagoort and Brown, 1994). Sorting out whether the simultaneous pursuit of two alternative analyses, the controlled processing associated with an attempt to resolve the ambiguity, or perhaps simply the act of noticing is responsible for the Nref effect is an important issue for future research.[1]

One caveat remains to be made. There is nothing in the data obtained so far to suggest that the frontal negativity that we refer to here as the Nref effect is an exclusively reference-specific or even language-specific effect. Note, in particular, that conditions of controlled processing or increased working memory load often elicit similar frontally distributed negati-

vities, regardless of the specific cognitive domain at hand (e.g., Donaldson and Rugg, 1999; Rösler et al., 1993; Rugg and Allan, 2000). In this respect, the Nref effect is on a par with other language-relevant ERP effects, such as the N400 and P600 effects, which are also not necessarily confined to language. Importantly, however, this does not diminish the utility of these effects in unraveling the interplay of syntactic, semantic, and referential analyses during language comprehension.

2.2. Referentially induced P600 effects

Next to the referentially ambiguous and unambiguous pronouns exemplified in (2a) and (2b) respectively, participants in the Van Berkum et al. (2004) and Nieuwland and Van Berkum (submitted for publication) pronoun studies were also given sentences such as (2c), in which *zero* referents were provided for the pronoun.

(2c) Anna shot at Linda as he jumped over the fence.

As can be seen in Fig. 3A and B, these referentially failing pronouns did not elicit the Nref effect. Instead, relative to the referentially unproblematic control (2a) they elicited a P600 effect, the effect most commonly associated with syntactic problems (Hagoort et al., 1999).

These results resemble those of Osterhout et al. (1997; see also Osterhout and Mobley, 1995), who observed a P600 effect to reflexives in anomalous sentences like "The man prepared herself for the operation", as well as in more subtle gender-stereotyped sentences such as "The doctor prepared herself for the operation". However, in contrast to a 'bound' reflexive pronoun such as "herself", which must find its antecedent within the same sentence, a 'free' pronoun such as "he" or "she" can in principle refer to an as yet unmentioned additional person. The fact that such pronouns elicit a P600 effect here suggests that our readers and listeners did not make use of that possibility and simply looked for a suitable referent in the immediate context. Moreover, it suggests that when having found only female referents for "he", or only male referents for "she", people initially take the problem to be a *syntactic* one. Note that the P600 effect is not only reliably elicited by outright syntactic violations, but also by syntactic *preference* violations, i.e., formulations that are syntactically well formed but whose syntactic properties do not fit the analysis currently pursued (Hagoort et al., 1999; see Fig. 1 and Van Berkum et al., 1999a,b, for an example). If in this current analysis only locally supplied female referents are being considered, "he" can thus initially be perceived as a syntactic dead end.

It may be helpful to briefly consider why this particular construal – a syntactic problem – is available in the first place. After all, in Dutch, as in English, the choice between "he" and "she" is not controlled by the syntactic gender of the noun that introduced the referent in prior discourse, but is instead controlled by semantic factors, most notably the biological gender (sex) of the referent itself. Furthermore, and in contrast to reflexive pronouns such as "himself" or "herself", these pronouns are 'free', i.e., do not fall under a structure-sensitive grammatical binding principle (such as Chomsky's *principle A*, Chomsky, 1993), Yet, although the choice is governed by conceptual factors and the pronoun is not under structure-

---

[1] One potential complication for all three of these possibilities is that in the one study where we were able to look at it (Nieuwland et al., in press), referential *disambiguation* did not immediately cause the Nref effect to disappear. This suggests that the processing consequences of referential ambiguity are somewhat persistent.

1-ref: *David shot at Linda as...*
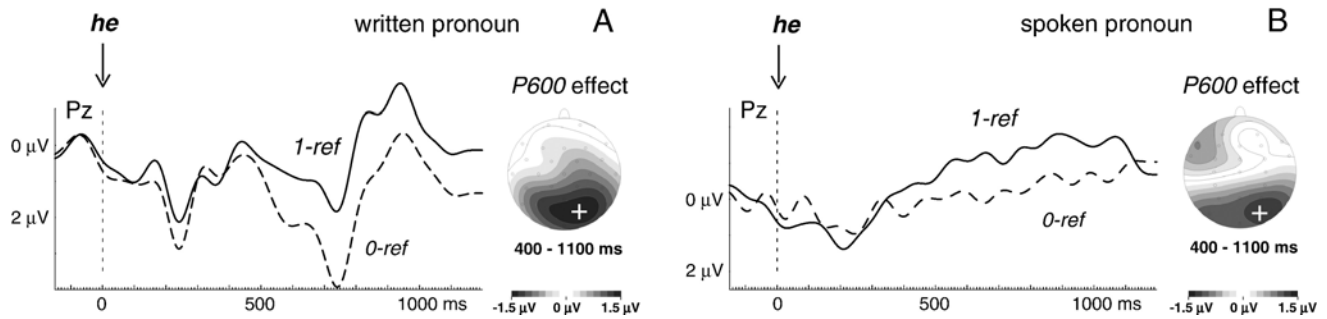
0-ref: *Anna shot at Linda as...*

Fig. 3 – Referentially induced P600 effects, elicited by written (A) and spoken (B) gender-marked pronouns for which no suitable antecedent had been introduced. Data from Nieuwland and Van Berkum (submitted for publication) and Van Berkum et al. (2004) respectively. Negative voltage is up.

sensitive grammatical control, the inflectional variants of the pronoun (for singular and plural 1st, 2nd, and 3rd person) are *grammaticized*, i.e., are part of the syntax of Dutch. Therefore, and in contrast to referring to a woman by "the man", referring to a woman by "he" can be construed as syntactically ill-formed.

It is a nontrivial observation that people indeed take it this way, because things could easily have been different. As already discussed, having too few (i.e., zero) suitable referents for a singular pronoun poses a *referential* problem that is in some ways not unlike having too many (two), and might as such also have elicited an Nref effect. The fact that it doesn't suggests that the processes that might be indexed by the Nref effect (e.g., noticing a referential problem, additional controlled processing, increased working memory load) are not triggered by *any* referential problem, and might be specific to referential ambiguity. Furthermore, like "the man", "he" also *semantically* denotes a male entity. The use of "he" for a woman could therefore be perceived as a semantic integration problem, i.e., elicit an N400 effect. Our P600 results suggest that readers and listeners initially do *not* see things this way, and at least momentarily blame the syntax (i.e., in "Anna shot at Linda as he jumped over the fence", masculine "he" should have been feminine "she").

Note that the referential ambiguities discussed in Section 2.2 can in principle *also* be viewed as a syntactic error (i.e., in "David shot at John as he jumped over the fence", singular "he" should actually have been plural "they"). Here, however, the absence of a P600 effect suggests that syntax doesn't get the blame. Why? We speculate that in the zero-referent "Anna shot at Linda as he..." case, syntax may get the blame because the pronoun's morphosyntactic features conflict with the very powerful preference or 'attractor' to have a pronoun refer to a locally available, foregrounded referent (i.e., either Anna, or Linda, or both; see Ariel, 2004; Garnham, 2001), a preference that may well be strong enough to win. This is reminiscent of our earlier observation (Van Berkum et al., 1999b) that discourse-induced referential ambiguity on nouns can momentarily lure the comprehension system into parsing a subsequent comple-

mentizer as a postnominal relative pronoun (heading a restrictive relative clause), even if the morphosyntax of the complementizer at hand formally prohibits this. As in the current situation, the conflict between syntactic and discourse-referential cues induced by Van Berkum et al. (1999b) elicited a P600 effect, suggesting that here too, syntax was getting the blame. In the ambiguous "David shot at John as he" case, however, the morphosyntactic features of "he" are not so much in *conflict* with a strong referential 'attractor', but simply fail to specify which of the two foregrounded referents is intended. Such temporary ambiguity might well be resolvable by additional inferencing, or by maintaining both options for a short while until more information is available, and hence does not immediately require reanalysis at the syntactic level.

Whether conflict between the morphosyntax and a strong 'referential attractor' is the right account for these referentially induced P600 effects remains to be tested. The phenomenon does bear an interesting resemblance to the fact that so-called semantic reversal anomalies, such as in "For breakfast the eggs would only eat toast and jam" (Kuperberg et al., 2003), "The cat that fled from the mice ran through the room" (Kolk et al., 2003; Van Herten et al., 2005), or "The mysterious crime had been solving" (Kim and Osterhout, 2005) also elicit a P600 effect. We agree with Osterhout and colleagues (see also Osterhout et al., 2003, pp. 285–286) that what might be relevant here is that in each case there is a powerful 'semantic attractor', an interpretation that is so tempting or compelling that the problem is seen to exist elsewhere. Such an account is similar to the attraction account we propose for explaining our *referentially* induced P600 effects, and both may be instantiations of a more general principle of resolving conflict in language comprehension.

2.3.   *Summary and implications*

The ERP studies on referential processes reviewed here demonstrate that problems with establishing reference can show up in brain potentials extremely rapidly, in at least two different ways. Before we turn to some new findings, we

discuss the main implications. First, our work has shown that readers and listeners rapidly detect whether a singular noun or pronoun successfully finds a unique referent, points ambiguously to two eligible referents, or fails to find a suitable referent in the situation model established so far. Note that in contrast to most neuro-imaging and behavioral studies in the field, our participants merely had to listen or read for comprehension, without a potentially disturbing secondary task.[2] Furthermore, and importantly, the two types of referential problems studied here are by no means unnatural. Because language unfolds in time, expressions are often temporarily ambiguous in what or whom they refer to; this is why languages provide, say, the post-nominal restrictive relative clause (e.g., "[the girl] that had been on the phone"). Also, as discussed before, there's nothing wrong with 'zero-referent' sentences like "Anna shot at Linda as he jumped over the fence". For all these reasons, we take our observations to reflect how people establish reference in everyday language comprehension.

Our ERP findings also unequivocally show that the brain responds differently to problems with having too many referents and having too few. For nouns and pronouns alike, and across spoken and written language, referential ambiguity elicits an Nref effect, a frontally dominant, sustained negative shift. A referentially failing pronoun like "he" with only women introduced, however, elicits a P600 effect (see also Osterhout and Mobley, 1995; Osterhout et al., 1997). The good news about this neuronal dissociation is that it tells us (a) that the brain handles these two referentially complex situations in a different way, recruiting partially non-overlapping neuronal ensembles, and (b) that we can use ERPs to selectively keep track of each. To unravel the workings of a system as fast, complex, and inaccessible as the linguistic brain, one thing we definitely need is time-resolved, unobtrusive, and differentially sensitive measurement techniques (Van Berkum, 2004). Although the use of EEG recording techniques has its clear drawbacks, event-related brain potentials (and quite possibly also *oscillatory* changes in the EEG, see Van Berkum et al., 2004) can in principle meet this need.

The bad news is that there is no simple one-to-one correspondence between the exact level of language representation – phonology, syntax, sense, reference – that is manipulated in an experiment and the identity of the ERP effect obtained. To be sure, the past two decades have uncovered many solid regularities, of which the best-established ones are that variations in the ease of semantic and syntactic integration relatively reliably show up in N400 and P600 effects respectively. However, the comprehension of a sentence requires an analysis at multiple levels of linguistic representations – prosodic, syntactic, semantic, referential – that are tightly linked to each other (see

Jackendoff, 2002, for a particularly helpful account). This tight linking implies that linguistic problems that *we as experimenters define* at one particular level of analysis, say, the referential 'tier', can affect the analysis at some other level (such as the syntax), and might in principle sometimes even do so to such an extent that the problem primarily *plays out* at that level. The fact that referentially failing pronouns elicit the 'syntactic' P600 effect is a case in point: "he" with only women in context is referentially as well as semantically problematic, but in the end it seems that '(morpho)syntax gets the blame'.

The fact that the comprehension system has to converge on a well-formed representation at the phonological, the syntactic, the semantic and the referential level in a way that also respects well-formedness in terms of the interfaces between levels not only complicates the use of ERPs to selectively track the various processes involved in language comprehension, but also raises a highly interesting deeper issue: if there's a problem with the input such that a stable multi-level representation cannot be arrived at, *how does the language comprehension system know where to put the blame, and/or where to try and fix things?* What our current referential findings, as well as the P600 effects to semantic reversal anomalies, point to is that we need to explore in more detail how, in cases where the concurrent analyses at different levels of language representation cannot be reconciled, the linguistic brain decides which levels of analysis can be 'clamped' as linguistically unproblematic, and which level(s) of representation should be reanalyzed.

In the pronoun studies discussed in Section 2.3 (Van Berkum et al., 2004; Nieuwland and Van Berkum, submitted for publication), the critical pronoun "he" fails to refer because no male antecedents whatsoever had been introduced in the prior text. In the new ERP experiment reported below, we exploit a specific semantic feature of certain interpersonal verbs, *implicit causality*, to manipulate the availability of suitable referents in a more subtle way. At issue is whether or not implicit causality can focus the readers' attention on particular referents quickly enough to have an impact on the referential interpretation of a gender-marked pronoun. In addition, the experiment allows us to examine whether readers also 'blame the syntax' when the set of antecedents introduced in the prior text does in principle allow for a syntactically correct reading.

3.   Establishing reference: The influence of interpersonal 'implicit causality' verbs

When asked to complete a sentence fragment such as "David praised Linda because…", readers and listeners will be inclined to continue the sentence with something about Linda, e.g., "…because she had done well". However, after "David apologized to Linda because…", people tend to continue with something about *David* instead. In "person-1 verb-ed person-2 because…" constructions, interpersonal verbs like "praise" and "apologize" thus supply information about whose behavior or state is the more likely immediate cause of the event at hand. Because this information is conveyed implicitly as part of the meaning

---

of the interpersonal verb, it is usually referred to as *implicit causality* (Garvey and Caramazza, 1974). Verbs like "praise" are said to be biased towards the person mentioned in the 2nd noun phrase, and are for that reason referred to as NP2-verbs. Verbs like "apologize", which are biased towards the person in the 1st noun phrase, are referred to as NP1-verbs.

It has been claimed that in "NP1 verb-ed NP2 because…" sentences, the information supplied by implicit causality verbs can very rapidly affect comprehension, soon after the verb. According to the *Immediate Focusing theory* (e.g., Greene and McKoon, 1995; Long and De Ley, 2000; McKoon et al., 1993), a fragment such as "David praised Linda because…" will immediately bring Linda into focus, and can as such influence the ease with which a subsequently referring pronoun is resolved. In particular, whereas a bias-consistent pronoun (in this case, "she") should be resolved more easily, the resolution of a bias-inconsistent one ("he") should be more difficult. However, the Immediate Focusing theory and the empirical data upon which it is based have been called into question by proponents of the *Clausal Integration theory* (e.g., Garnham, 2000; Garnham et al., 1996; Stewart et al., 2000). According to the latter, the information supplied by an implicit causality verb is brought to bear on comprehension towards the end of the sentence only, when people have read the main clause as well as the sub-ordinate *because*-clause and subsequently combine the causal information provided by both clauses into a single representation. Such delayed processing would be of con-siderable interest, for it would set the impact of implicit causality apart from that of a wide range of other semantic and referential factors whose impact on language compre-hension has been shown to be immediate.

We conducted a written-language ERP study to help resolve the issue. In our experiment, which was a follow-up to recent behavioral work (Koornneef and Van Berkum, 2006), native speakers of Dutch read though several coherent mini-stories, some of which contained a critical "NP1 verb-ed NP2 because…" fragment. The critical manipulation was whether the subsequent pronoun "he" was consistent or inconsistent with the verb's implicit causality bias, as illustrated by the example story below (shown in translation as well as in the original Dutch form).

(4a) NP1-biased verb, bias-consistent pronoun.

David and Linda were both driving pretty fast. At a busy intersection they crashed hard into each other. David apologized to Linda because he according to the witnesses was the one to blame.

(David en Linda reden allebei behoorlijk hard. Bij een druk kruispunt botsten zij met hun auto's stevig op elkaar. David bood zijn excuses aan Linda aan omdat hij volgens de getuigen van het ongeluk alle schuld had.)

(4b) NP1-biased verb, bias-inconsistent pronoun.

David and Linda were both driving pretty fast. At a busy intersection they crashed hard into each other. Linda

apologized to David because he according to the witnesses was not the one to blame.

(David en Linda reden allebei behoorlijk hard. Bij een druk kruispunt botsten zij met hun auto's stevig op elkaar. Linda bood haar excuses aan David aan omdat hij volgens de getuigen van het ongeluk geen schuld had.)

Note that the bias of an implicit causality verb can be negated without rendering the sentence ungrammatical or incoherent. The differential processing consequence, if any, elicited by a bias-inconsistent pronoun would therefore not be the result of an overt anomaly. Instead, it would reflect something more subtle about the way in which we use various sources of information to find out what or whom is being talked about.

In this experiment, we diverged from the usual way of presenting written sentences in ERP research, word by word at a fixed word onset asynchrony of, typically, 500 or 600 ms. Instead, we used a new *Variable Serial Visual Presentation* (VSVP) procedure, in which the display time of words varies with their length and position in the sentence (cf. Otten and Van Berkum, submitted for publication; see Section 6.3 for details). Although we have no reason to doubt the validity of results obtained with fixed SVP (the findings obtained with this procedure can usually also be obtained with spoken language input; e.g., see Figs. 2 and 3), sentences presented with VSVP do feel somewhat more natural, particularly if word lengths vary considerably.

The predictions for this experiment were very straightfor-ward. If, as stated in the Clausal Integration account, readers only consider implicit causality at the end of the subordinate *because*-clause (i.e., at the end of the entire sentence), bias-inconsistent pronouns in mid-sentence should not elicit a differential ERP effect relative to bias-consistent control pronouns. However, if, as proposed in the Immediate Focusing account, verb-based implicit causality immediately foregrounds a particular referent, a bias-inconsistent pro-noun should present a problem, in that its grammaticized gender does not match the foregrounded, most salient referent. In the light of the findings reviewed in Section 2.2 (Nieuwland and Van Berkum, submitted for publication; Van Berkum et al., 2004; see also Osterhout and Mobley, 1995; Osterhout et al., 1997), we expected that such bias-incon-sistent pronouns would momentarily be perceived as having the wrong syntactic gender, and would as such elicit a P600 effect.

4. Results and discussion

The grand average ERP waveforms in Fig. 4 and the corresponding difference waveforms in Fig. 5 reveal that relative to their bias-consistent controls, bias-inconsistent pronouns elicited a positive deflection at about 400–700 ms after pronoun onset, with a maximum over centro-posterior scalp sites. Because of its polarity, timing, shape and scalp distribution, we take this to be a P600 effect. In mean-amplitude ANOVAs over all 15 posterior electrodes, the
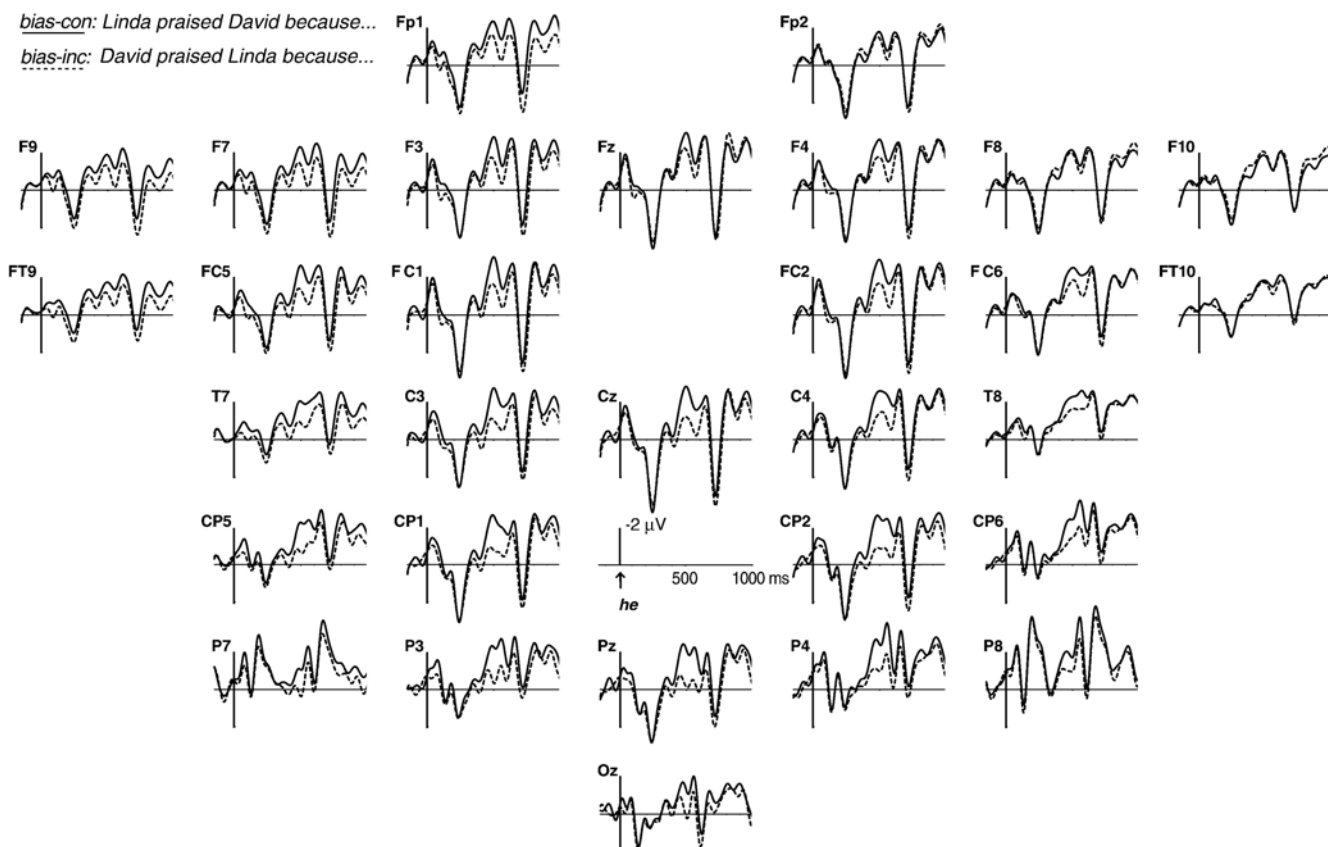
Fig. 4 – Grand average event-related brain potentials to singular pronouns whose gender-marking was consistent (solid) or inconsistent (dotted) with the implicit causality bias of a preceding verb. Negative voltage is up, and onset of the pronoun is at 0 ms.

effect was reliable in the latency ranges of 400–500 ms ($F_{1,35}=11.25$, $MSE=28.28$, $p=.002$), 500–600 ms ($F_{1,35}=9.70$, $MSE=24.98$, $p=.004$), and 600–700 ms ($F_{1,35}=4.32$, $MSE=21.54$, $p=.045$).[3]

This observation tells us a number of things. First, as in the studies reviewed before, the differential ERP effect by itself reveals that readers tried to resolve these critical pronoun rapidly enough to discover a problem in bias-inconsistent stories within at most half a second. Because the problem hinged on the implicit causality conveyed by a preceding interpersonal verb, the effect also unequivocally

demonstrates that readers make immediate use of this type of semantic information. Our ERP results therefore straight-forwardly disconfirm the *Clausal Integration theory* of how people use verb-based implicit causality information (e.g., Garnham, 2000; Garnham et al., 1996; Stewart et al., 2000), and they instead support an account in which implicit causality can immediately foreground one of two referents at the expense of the other (e.g., *Immediate Focusing theory*; Greene and McKoon, 1995; Long and De Ley, 2000; McKoon et al., 1993). The present ERP results converge with the behavioral observation that readers slow down at or immediately after bias-inconsistent pronouns (established in self-paced reading as well as eye tracking; Koornneef and Van Berkum, 2006). In all, and across three experimental paradigms with very different measures – ERP voltage, button-press reading times, and eye fixation patterns – the evidence unequivocally tells us that the implicit causality information supplied by verbs like "praise" or "apologize" can be brought to bear on language comprehension without delay, as soon as it is of relevance to the interpretation of the unfolding sentence.

Going beyond the behavioral results, the fact that we observe a P600 effect to bias-inconsistent pronouns in a sentence like "David praised Linda because he…" suggests

---

[3] Fig. 5 also suggests the presence of a later short-lived positive deflection around 700–800 ms, most clearly visible at the left-anterior electrodes. This observation is supported by a reliable Consistency by Anteriority by Hemisphere interaction in a mean quadrant ANOVA in this latency range ($F_{1,35}=11.30$, $MSE=0.11$, $p=.002$), entirely due to the left-anterior quadrant (simple main effect of Consistency: $F_{1,35}=5.42$, $MSE=2.39$, $p=.026$). We are not entirely sure what to make of this short-lived effect, as it rides on a slower and possibly drift-related left-anterior negative shift that may obscure its true size (and statistical reliability). We also note that left-anterior positivities were not observed in the earlier studies with referentially failing pronouns. Pending replication, we therefore refrain from speculation.
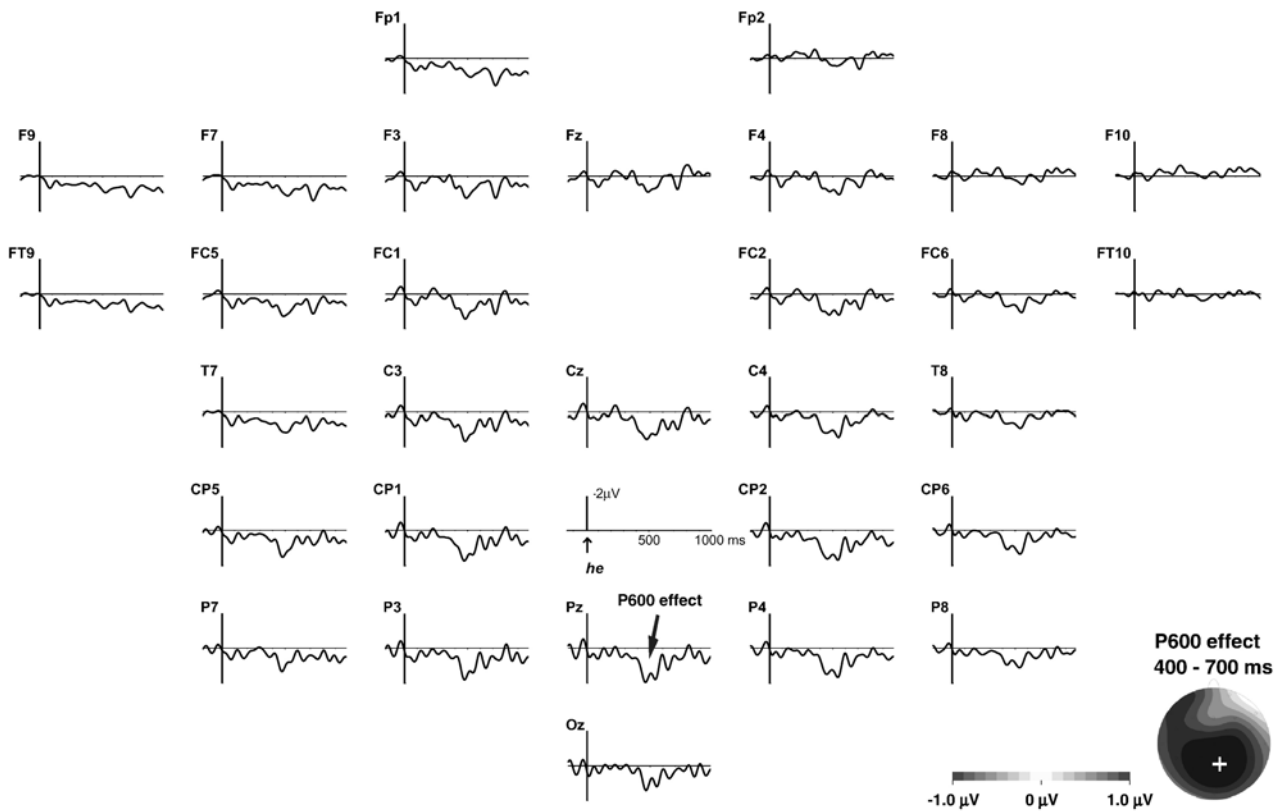
Fig. 5 – Difference waves for inconsistent–consistent pronouns, revealing a referentially induced P600 effect. Negative voltage is up, and onset of the pronoun is at 0 ms.

that even in *this* situation, readers are prepared to put the blame on syntax. As reviewed in Section 2.2, pronouns in sentences like "Anna shot at Linda as he jumped over the fence" elicited a similar P600 effect, presumably because no male antecedents had been introduced into the situation model. However, the implicit causality sentences *did* always explicitly provide a male antecedent to which "he" might refer. Our verb-induced P600 result therefore suggests that the implicit causality bias is a relatively strong one, strong enough to foreground one of the referents such that the other referent is momentarily unavailable as a candidate antecedent for the pronoun, and strong enough to also at least initially outweigh a clear morphosyntactic constraint (see Van Berkum et al., 1999b, for a comparable referential override of morphosyntax). As discussed more fully elsewhere (Koornneef and Van Berkum, 2006), the most natural account for this strong commitment is that readers use the implicit causality cue in something like "David praised Linda because…" *proactively*, and essentially *predict*, before the pronoun comes along, that the remainder of the sentence will tell us something about Linda. In terms of the attraction account discussed in Section 2.2, the consequence of such a proactive referential commitment might well be that the referential analysis is 'clamped' after "David praised Linda because" (or, perhaps, after having perceived the pronoun regardless of its gender features; i.e., "David praised Linda because <PRO>"), to such

a degree that "he" is initially construed as a syntactic problem.

The above line of reasoning does of course depend on the assumption that within the domain of language comprehension, the P600 effect specifically reflects a problem with syntactic unification. Although this is the dominant interpretation for this effect (e.g., Hagoort, 2005; Hagoort et al., 1999; Osterhout et al., 2003), it has not gone unchallenged. For reasons outlined elsewhere (Osterhout and Hagoort, 1999, see also Snijders et al., 2005), the suggestion that the P600 effect is really a *P300 oddball effect* (Coulson and Kutas, 1998) can most likely be put aside. However, P600-like effects have also been observed in paradigms where an N400 effect would have been more obvious (Kim and Osterhout, 2006; Kolk et al., 2003; Kuperberg et al., 2003; Nieuwland and Van Berkum, 2005; Van Herten et al., 2005). This could be taken to suggest that within language comprehension, the P600 effect is not exclusively associated with syntactic analysis (cf. Kolk et al., 2003; Van Herten et al., 2005). In line with our earlier remarks, however, and on grounds of parsimony, we believe that it makes more sense to consider the possibility that in these cases, syntax somehow gets the blame as well. In all, we as yet see no reason to decouple the P600 effect from syntax.

In recent ERP research (Burkhardt, 2006; Kaan et al., this volume), late posterior positivities have also been elicited by words at which the reader has to introduce new referents into

his or her model of the discourse. The P600 effects elicited in "Anna shot at Linda as he jumped over the fence" (Van Berkum et al., 2004; Nieuwland and Van Berkum, submitted for publication) might be interpreted as such. However, in "David praised Linda because he…", no new referent is required. A new-referent account would perhaps also be difficult to maintain for locally bound reflexive pronouns such as "himself", which also elicit a P600 effect if the available antecedent is female (Osterhout and Mobley, 1995). Thus, and again partly for reasons of parsimony, we prefer our current syntax-oriented interpretation. If a gender-marked pronoun fails to refer because the *preferred* referents are all of the wrong sex, its grammaticized gender feature (or 'inflection') is initially taken to be incorrect for the situation at hand, which results in a P600 effect.

As with the zero-referent P600 findings discussed in Section 2.2, note that the results could easily have been different. The problem with fragments like "David praised Linda because he…" could be construed as a semantic problem and as such elicit an N400 effect, either because an expression that semantically denotes a man ("he", "the man") is used in cases where readers focus on a woman, or because the semantic bias that is generated by the implicit causality verb is negated. Furthermore, to the extent that the need to refocus a different referent than the one expected calls for additional controlled processing, or temporarily increases working memory load, bias-inconsistent pronouns might elicit a frontal negativity. The fact that we don't see an Nref effect here again suggests that latter might, within the domain of linguistic reference, be specific to discourse-model ambiguity.

Finally, the primary implication of this experiment does not depend on assumptions about what the P600 effect might reflect, nor does it in fact depend on the exact nature of the ERP effect observed. The fact that bias-inconsistent pronouns elicit a *differential* ERP effect relative to bias-consistent ones within some 400 mss by itself reveals that the use of implicit causality information is not delayed until the end of the entire sentence. Rather, what it shows is that readers use the social cues supplied by interpersonal verbs like "praise" or "apologize" in mid-sentence, to make intelligent guesses about who might well be talked about next.

5.   Conclusions

The electrophysiology of language comprehension has long been dominated by research on syntactic and semantic integration. However, to understand expressions like "he did it" or "the little girl", combining word meanings in accordance with semantic and syntactic constraints is not enough—readers and listeners also need to work out *what or who* is being referred to. The studies reviewed here suggest that the processes involved in establishing reference can be tracked by means of ERPs relatively selectively, with high temporal precision, and without the need for an additional task. Our work has shown, first, that upon encountering a noun or pronoun, readers and listeners immediately inspect their situation model for suitable discourse entities, such that they can discriminate between having too many, too few, or

exactly the right number of referents within at most half a second. Furthermore, the new ERP results reported here demonstrate that implicit causality verbs like "praise" or "apologize" can immediately foreground particular referents, and that to the linguistic brain in action, a subsequent "he" or "she" is not as free a pronoun as linguistic theory might lead one to expect. Not only are such pronouns preferentially resolved towards available and foregrounded referents, but this on-line processing preference is such that "he" with a female referent in focus is at least initially construed as a syntactic error, regardless of the fact that a male referent has also just been introduced.

Our ERP findings on referential processing raise two important issues for future research. One is to examine the functional and neuronal generator(s) of the frontally sustained negative shift elicited by referentially ambiguous nouns and pronouns (Nref effect). We are currently conducting an fMRI study with referentially ambiguous and unambiguous pronouns to help illuminate this issue, but we also signal the need for ERP studies examining how the Nref effect relates to other frontal negativities observed in linguistic and non-linguistic tasks. The second issue for research is a much more general one. We have seen that referential processing is not only highly incremental, but also at times highly influential. Syntax doesn't always come first. Rather, sometimes reference comes first, and if it conflicts with syntax, the latter gets the blame. What we need to know is why this happens. What is at stake is not just a matter of mapping ERP effects to levels of language analysis, or relating our present findings to, say, the 'unexpected' P600 effects with semantic reversal anomalies. Rather, what is at stake is a much deeper question. If language comprehension involves an analysis at multiple levels of linguistic representation that are linked (cf. Jackendoff, 2002), how does the system deal with conflicts between levels of analysis, such that it knows where to put the blame, and hence where to try and fix things?

6.   Experimental details for the experiment

6.1.   Participants

Participants were 36 native speakers of Dutch (27 female, mean age 21, range 18–26 years) who had not taken part in the Koornneef and Van Berkum (2006) experiments or any of its pretests.

6.2.   Materials

The critical items were 80 Dutch stories constructed around a strong implicit causality verb, as exemplified in example (4). Half of these stories were those used in the earlier self-paced reading study (Koornneef and Van Berkum, 2006, Exp. 1). These 40 stories had been constructed around twenty verbs with a strong NP1-bias and 20 verbs with a strong NP2-bias (see below for a list), selected for their strong bias on the basis of a paper-and-pencil sentence completion pretest. For the present ERP study, 40 more stories were constructed around

the same verbs, using the same procedure and criteria as for the original set (see Koornneef and Van Berkum, 2006, for details, and www.josvanberkum.nl for the complete set of Dutch items).

NP1-biased verbs		NP2-biased verbs	
Verb	*Bias*	*Verb*	*Bias*
fascineren (*fascinate*)	100%	minachten (*hold in contempt*)	100%
excuses aanbieden (*apologize*)	100%	benijden (*envy*)	100%
bekennen (*confess*)	96%	bewonderen (*admire*)	100%
teleurstellen (*disappoint*)	96%	vrezen (*fear*)	100%
vervelen (*bore*)	95%	waarderen (*appreciate*)	100%
hinderen (*bother*)	91%	bekritiseren (*criticise*)	100%
oplichten (*swindle*)	87%	prijzen (*praise*)	96%
kwellen (*torment*)	86%	complimenteren (*compliment*)	96%
storen (*disturb*)	86%	haten (*hate*)	96%
smeken (*beg*)	83%	ontslaan (*fire*)	96%
bellen (*call*)	82%	respecteren (*respect*)	95%
ergeren (*annoy*)	82%	straffen (*punish*)	95%
verbazen (*amaze*)	82%	feliciteren (*congratulate*)	95%
verontrusten (*worry*)	82%	verantwoordelijk stellen (*hold responsible*)	91%
winnen van (*win*)	82%	aanklagen (*sue*)	91%
misleiden (*mislead*)	78%	troosten (*comfort*)	91%
irriteren (*irritate*)	78%	verafschuwen (*loathe*)	86%
liegen (*lie*)	78%	aanbidden (*adore*)	86%
inspireren (*inspire*)	77%	houden van (*love*)	86%
intimideren (*intimidate*)	77%	bedanken (*thank*)	82%

The 80 critical stories were pseudo-randomly mixed with 160 coherent and highly variable filler stories (addressing a different issue, Otten and Van Berkum, 2006), such that the resulting list contained 40 stories with a bias-consistent pronoun and 40 with a bias-inconsistent pronoun, with each verb occurring once in each condition, and with the two item subsets matched on average verb bias (each 90%) and NP1:NP2 ratio (each 1:1). A second list was derived from the first by rotating each critical item across condition. Each list was presented to 18 participants.

6.3. Procedures

Participants sat in a comfortable chair in a normally lit room, and they were asked to read for comprehension, without an additional task. The stories were presented in black 36 point courier new font on a white background on a fast TFT display (Iiyama TXA 3834 MT) positioned at about 80 cm distance. Before each trial, a fixation cross was shown in the center of the screen for 2.5 s. Participants were instructed to avoid blinks and eye-movement when the words were presented on screen, and were encouraged to blink when the fixation cross was shown. To signal the start of each trial to the participant, a beep was presented 1 s before the onset of the first word. The stories were then presented word by word.

To make this presentation more natural, we used a *Variable Serial Visual Presentation* (VSVP) procedure in which the

presentation duration of each non-critical word varied with its length and position in the sentence (cf. Otten and Van Berkum, submitted for publication). Non-critical word duration consisted of a standard offset of 187 ms plus additional 27 ms per letter (with an upper bound of 10 letters for each word). In the present experiment, durations varied from 214 ms for a one-letter word to 450 ms for words consisting of ten or more letters. Between words, the screen went blank for a standard duration of 106 ms. The presentation of clause-final words preceding a comma was prolonged with an additional 200 ms. In addition, presentation time for sentence-final words was extended with an extra 293 ms, followed by a 1 s pause until the next sentence began. These various parameters were based on natural reading times (Haberlandt and Graesser, 1985; Legge et al., 1997), a subjective assessment of the naturalness of the resulting presentation, and technical constraints imposed by the video refresh rate. To avoid spurious ERP effects due to accidental differences in average word length across conditions, the critical pronoun and the three words that followed were presented with a fixed duration of 346 ms, based on the average critical word length across all 240 stories in the experiment. The word just before the critical pronoun was always the same one ("omdat", "because"), presented for 322 ms, again with the standard 106 ms interword interval. Participants did not notice the alternation between completely variable and semi-fixed word duration presentation within a single story.

6.4. EEG recording and preprocessing

The electroencephalogram was recorded from 30 silver-chloride electrodes mounted in an elastic cap (EasyCap), each referenced to the left mastoid. Blinks and vertical eye-movements were registered by placing an electrode under the left eye, also referenced to the left mastoid. Electrode impedance were kept below 5 kOhms during the experiment. The EEG was amplified with BrainAmps DC amplifiers (BrainProducts, München), filtered with a 0.03–100 Hz band-pass, sampled at a 500 Hz rate, and re-referenced off-line to the average of right and left mastoids. Blinks and eye movements were removed from the data using a procedure based on Independent Component Analysis (ICA; Jung et al., 2000; Makeig et al., 1997). After that the data were segmented in epochs from 500 ms before critical word onset until 1200 ms after critical word onset. After baseline-correcting the signals by subtracting mean amplitude in the 150 ms preceding critical word onset, we eliminated segments in which the signal exceeded ±100 μV, or which featured a linear drift of more than ±40 μV, beginning before the onset of the critical word (9% of the trials in all). For each participant the remaining signals were averaged per condition.

Acknowledgements

REFERENCES

Anderson, J.E., Holcomb, P.J., 2005. An electrophysiological investigation of the effects of coreference on word repetition and synonymy. Brain Lang. 94, 200–216.

Ariel, M., 2004. Accessibility marking: discourse functions, discourse profiles, and processing cues. Discourse Process. 37 (2), 91–116.

Brown, C.M., Hagoort, P., Kutas, M., 2000. Postlexical integration processes in language comprehension: evidence from brain-imaging research. In: Gazzaniga, M.S. (Ed.), The New Cognitive Neurosciences. MIT Press, Cambridge, MA, pp. 881–895.

Burkhardt, P., 2006. Inferential bridging relations reveal distinct neural mechanisms: Evidence from event-related brain potentials. Brain Lang. 98, 159–168.

Camblin, C.C., Ledoux, K., Boudewyn, M., Gordon, P.C., Swaab, T.Y., 2007. Processing new and repeated names: Effects of coreference on repetition priming with speech and fast RSVP. Brain Res. 1146, 172–184.

Chomsky, N., 1993. Lectures on Government and Binding: The Pisa Lectures. Mouton de Gruyter, Berlin.

Coulson, S., Kutas, M., 1998. Expect the unexpected: event-related brain responses to morphosyntactic violations. Lang. Cogn. Processes 13, 21–58.

Donaldson, D.I., Rugg, M.D., 1999. Event-related potential studies of associative recognition and recall: electrophysiological evidence for context dependent retrieval processes. Cogn. Brain Res. 8, 1–16.

G. Frege, On Sense and Nominatum. Reprinted in Martinich, A.P. (1996), The philosophy of language. 1892, pp. 199–211, Oxford: Oxford University Press.

Garvey, C., Caramazza, A., 1974. Implicit causality in verbs. Linguist. Inq. 5, 459–464.

Garnham, A., 2001. Mental Models and the Interpretation of Anaphora. Psychology Press, Hove, UK.

Garnham, A., Traxler, M., Oakhill, J., Gernsbacher, M.A., 1996. The locus of implicit causality effects in comprehension. J. Mem. Lang. 35, 517–543.

Gibson, E., 1998. Linguistic complexity: locality of syntactic dependencies. Cognition 68, 1–76.

Greene, S.B., McKoon, G., 1995. Telling something we can't know: experimental approaches to verbs exhibiting implicit causality. Psychol. Sci. 6, 262–270.

Haberlandt, K.F., Graesser, A.C., 1985. Component processes in text comprehension and some of their interactions. J. Exp. Psychol. Gen. 114 (3), 357–374.

Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9 (9), 416–423.

Hagoort, P., Brown, C.M., 1994. Brain responses to lexical-ambiguity resolution and parsing. In: Clifton, Jr., C., Frazier, L., Rayner, K. (Eds.), Perspectives on Sentence Processing. Erlbaum, Hillsdale, NJ, pp. 45–80.

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

Hagoort, P., Brown, C.M., Osterhout, L., 1999. The neurocognition of syntactic processing. In: Brown, C.M., Hagoort, P. (Eds.), The Neurocognition of Language. Oxford Univ. 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.

Jackendoff, R., 2002. Foundations of Language. Oxford Univ. Press, New York.

Jung, T.P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., Sejnowski, T.J., 2000. Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. Clin. Neurophysiol. 111, 1745–1758.

Kaan, E., Dallas, A.C., Barkley, C.M., 2007. Processing bare quantifiers in discourse. Brain Res. 1146, 199–209.

Kim, A., Osterhout, L., 2005. The independence of combinatory semantic processing: evidence from event-related potentials. J. Mem. Lang. 52, 205–225.

King, J.W., 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 (3), 376–395.

Kintsch, W., 1998. Comprehension: A Paradigm for Cognition. Cambridge Univ. Press, Cambridge.

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.

Koornneef, A.W., Van Berkum, J.J.A., 2006. On the use of verb-based implicit causality in sentence comprehension: evidence from self-paced reading and eye tracking. J. Mem. Lang. 54 (4), 445–465.

Kuperberg, G.R., Sitnikova, T., Caplan, D., Holcomb, P.J., 2003. Electrophysiological distinctions in processing conceptual relationships within simple sentences. Cogn. Brain Res. 17, 117–129.

Kutas, M., 1997. Views on how the electrical activity that the brain generates reflects the functions of different language structures. Psychophysiology 34, 383–398.

Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.

Kutas, M., Federmeier, K.D., Coulson, S., King, J.W., Münte, T.F., 2000. Language. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (Eds.), Handbook of Psychophysiology. Cambridge Univ. Press, Cambridge, pp. 576–601.

Legge, G.E., Ahn, S.J., Klitz, T.S., Luebker, A., 1997. Psychophysics of reading—XVI. The visual span in normal and low vision. Vision Res. 37 (14), 1999–2010.

Long, D.L., De Ley, L., 2000. Implicit causality and discourse focus: the interaction of text and reader characteristics in pronoun resolution. J. Mem. Lang. 42, 545–570.

Magne, C., Astésano, C., Lacheret-Dujour, A., Morel, M., Alter, K., Besson, M., 2005. On-line processing of "Pop-Out" words in spoken French dialogues. J. Cogn. Neurosci. 17, 740–756.

Makeig, S., Jung, T.P., Bell, A.J., Ghahremani, D., Sejnowski, T.J., 1997. Blind separation of auditory event-related brain responses into independent components. Proc. Natl. Acad. Sci. U. S. A. 94 (20), 10979–10984.

McKoon, G., Greene, S.B., Ratcliff, R., 1993. Discourse models, pronoun resolution, and the implicit causality of verbs. J. Exper. Psychol., Learn., Mem., Cogn. 19, 1040–1052.

Müller, H.M., King, J.W., Kutas, M., 1997. Event-related potentials elicited by spoken relative clauses. Cogn. Brain Res. 5, 193–203.

Münte, T.F., Schiltz, K., Kutas, M., 1998. When temporal terms belie conceptual order. Nature 395, 71–74.

Myers, J.L., O'Brien, E.J., 1998. Accessing the discourse representation during reading. Discourse Processes 26, 131–157.

Nieuwland, M.S., Van Berkum, J.J.A., 2005. Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. Cogn. Brain Res. 24, 691–701.

Nieuwland, M.S., Van Berkum, J.J.A., submitted for publication. Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. Manuscript.

Nieuwland, M.S., Otten, M. and Van Berkum, J.J.A., in press. Who are you talking about? Tracking discourse-level referential processing with ERPs. J. Cogn. Neurosci.

Osterhout, L., Hagoort, P.H., 1999. A superficial resemblance does not necessarily mean you are part of the family:

counterarguments to Coulson, King, and Kutas (1998) in the P600/SPS-P300 debate. Lang. Cogn. Processes 14 (1), 1–14.

Osterhout, L., Holcomb, P.J., 1992. Event-related brain potentials elicited by syntactic anomaly. J. Mem. Lang. 31, 785–806.

Osterhout, L., Kim, A., Kuperberg, G., 2006. The neurobiology of sentence comprehension. To appear in. In: Spivey, M., Joanisse, M., McRae, K. (Eds.), The Cambridge Handbook of Psycholinguistics. Cambridge Univ. Press, Cambridge.

Osterhout, L., Mobley, L.A., 1995. Event-related brain potentials elicited by failure to agree. J. Mem. Lang. 34, 739–773.

Osterhout, L., Bersick, M., McLaughlin, J., 1997. Brain potentials reflect violations of gender stereotypes. Mem. Cogn. 25 (3), 273–285.

Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., Inoue, K., 2003. Sentences in the brain: event-related potentials as real-time reflections of sentence comprehension and language learning. In: Carreiras, M., Clifton, Jr., C. (Eds.), The On-Line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond. Psychology Press, New York, pp. 271–308.

Otten, M., Van Berkum, J.J.A. What makes a discourse constraining? Comparing the effects of discourse message and scenario relevance on the discourse-dependent N400 effect (working title). Manuscript in preparation, 2006.

Otten, M., Van Berkum, J.J.A. submitted for publication. Discourse-based anticipation during language processing: Prediction or priming? Manuscript.

Rösler, F., Heil, M., Glowalla, U., 1993. Monitoring retrieval from long-term memory by slow event-related brain potentials. Psychophysiology 30, 170–182.

Rösler, F., Pechmann, T., Streb, J., Röder, B., Hennighausen, E., 1998. Parsing of sentences in a language with varying word order: word-by-word variations of processing demands are revealed by event-related brain potentials. J. Mem. Lang. 38, 150–176.

Rugg, M.D., Allan, K., 2000. Memory retrieval: an electrophysiological perspective. In: Gazzaniga, M.D. (Ed.), The New Cognitive Neurosciences. MIT Press, Cambridge, MA, pp. 805–816.

Stewart, A.J., Pickering, M.J., Sanford, A.J., 2000. The time course of the influence of implicit causality information: focusing versus integration accounts. J. Mem. Lang. 42, 423–443.

Schmitt, B.M., Lamers, M., Münte, T.F., 2002. Electrophysiological estimates of biological and syntactic gender access during pronoun processing. Cogn. Brain Res. 14, 333–346.

Snijders, T.M., Jensen, O., Hagoort, P., 2005. Neuromagnetic correlates of syntactic processing. Presented at the 11th

Annual Conference on Architectures and Mechanisms for Language Processing, Ghent, Belgium, pp. 5–7.

Swaab, T.Y., Camblin, C.C., Gordon, P.C., 2004. Reversed lexical repetition effects in language processing. J. Cogn. Neurosci. 16, 715–726.

Trueswell, J.C., Tanenhaus, M.K. (Eds.), 2005. Processing World-Situated Language: Bridging the Language-as-Action and Language-as-Product Traditions. MIT Press, Cambridge, MA.

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

Van Berkum, J.J.A. The electrophysiology of discourse and conversation. In: M. Spivey, K. McRae, M. Joanisse (Eds.), The Cambridge Handbook of Psycholinguistics. Cambridge University Press, in press.

Van Berkum, 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., Brown, C.M., Hagoort, P., 1999b. When does gender constrain parsing? Evidence from ERPs. J. Psycholinguist. Res. 28 (5), 555–571.

Van Berkum, J.J.A., Hagoort, P., Brown, C.M., 1999c. Semantic integration in sentences and discourse: evidence from the N400. J. Cogn. Neurosci. 11 (6), 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., Brown, C.M., Hagoort, P., 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 Berkum, J.J.A., Zwitserlood, P., Bastiaansen, M.C.M., Brown, C.M., Hagoort, P., 2004. So who's "he" anyway? Differential ERP and ERSP effects of referential success, ambiguity and failure during spoken language comprehension. Annual meeting of the Cognitive Neuroscience Society (CNS-2004), San Francisco, April 18–20.

Van Herten, M., Kolk, H.H.J., Chwilla, D.J., 2005. An ERP study of P600 effects elicited by semantic reversal anomalies. Cogn. Brain Res. 22, 241–255.

Zwaan, R.A., Radvansky, G.A., 1998. Situation models in language comprehension and memory. Psychol. Bull. 123, 162–185.