Who are You Talking About? Tracking Discourse-level Referential Processing with Event-related Brain Potentials

Mante S. Nieuwland[1], Marte Otten[1], and Jos J. A. Van Berkum[1,2]

Abstract

■ In this event-related brain potentials (ERPs) study, we explored the possibility to selectively track referential ambiguity during spoken discourse comprehension. Earlier ERP research has shown that referentially ambiguous nouns (e.g., ''the girl'' in a two-girl context) elicit a frontal, sustained negative shift relative to unambiguous control words. In the current study, we examined whether this ERP effect reflects ''deep'' situation model ambiguity or ''superficial'' textbase ambiguity. We contrasted these different interpretations by investigating whether a discourse-level semantic manipulation that prevents referential ambiguity also averts the elicitation of a referentially induced ERP effect. We compared ERPs elicited by nouns that were referentially nonambiguous but were associated with two discourse entities (e.g., ''the girl'' with two girls introduced in the context, but one of which has died or left the scene), with referentially ambiguous and nonambiguous control words. Although temporally referentially ambiguous nouns elicited a frontal negative shift compared to control words, the ''double bound'' but referentially nonambiguous nouns did not. These results suggest that it is possible to selectively track referential ambiguity with ERPs at the level that is most relevant to discourse comprehension, the situation model. ■

INTRODUCTION

As there is only a limited number of different names around, some people probably recognize the following situation: You and a friend have two acquaintances with the same forename (e.g., ''John''). At some point, your friend tells you that he saw ''John'' the other day. But you don't know which ''John'' your friend is actually talking about, and you will have to ask for further specification. Until further notice, ''John'' remains referentially ambiguous.

Of course, referential ambiguity is not restricted to proper names, but arises whenever a reader or listener is unable to select a single referent for a name, pronoun, or noun phrase (NP) out of multiple candidates. For example, ''the girl'' is also referentially ambiguous when embedded in a two-girl story context that does not allow you to make a fair choice. This ambiguity illustrates that to get at the fullest meaning possible, one must establish both the meaning of a word (*sense* or *semantic analysis*) and its reference to a certain entity (*referential analysis* or *anaphoric processing*).

Establishing reference is of particular importance during discourse comprehension or dialogue. As conversations and stories unfold, entities are referred to over and over again. If we were unable to establish reference for repeated referring expressions, our understanding of discourse would be next to nothing. For this reason, researchers studying discourse comprehension have spent a great deal of attention on referential processing (see Graesser, Zwaan, & Millis, 1997, for a review), mainly with behavioral measures such as self-paced reading or cross-modal probing. However, unlike other linguistic subprocesses such as syntax and semantics, referential processing has received relatively little attention from researchers who study sentence or discourse comprehension with event-related brain potentials (ERPs).

Researchers have used ERPs for decades now to track the temporal dynamics of different linguistic subprocesses without a confounding task (that may induce task-specific strategies), and the multidimensional characteristics of ERPs can provide cues to the identity of the cognitive event at hand (Van Berkum, 2004). For example, a number of ERP components have been identified as to distinctively tap into syntactic or semantic aspects of language comprehension, and have subsequently been exploited to study syntactic and semantic analysis in different linguistic environments (see Brown, Hagoort, & Kutas, 2000, for a review). But until recently, no one had looked for a distinct ERP signature of referential analysis in language comprehension.

Evidence that the brain responds in qualitatively different ways to problems with reference than to problems

---

[1]University of Amsterdam, [2]F.C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

with either syntax or semantics came from a recent series of ERP studies by Van Berkum, Brown, and Hagoort (1999) and Van Berkum, Zwitserlood, Hagoort, and Brown (2003). Although semantic and syntactic processing are predominantly associated with the N400 and P600/(E)LAN, respectively (e.g., Friederici, 2002; Hagoort, Brown, & Osterhout, 1999), referentially ambiguous nouns (e.g., ''the girl'' in a two-girl context) and referentially ambiguous pronouns (e.g., ''David noticed John when he stood up.'') elicit a frontally dominant and sustained negative shift (Van Berkum, Zwitserlood, Bastiaansen, Brown, & Hagoort, 2004; Van Berkum et al., 1999, 2003; see Van Berkum, Koornneef, Otten, & Nieuwland, in press, for a review).

In the current study, we further examined the utility of this ERP signature as a tool to selectively track referential ambiguity during discourse comprehension. The rationale for this study is that the referentially induced ERP effect reported by Van Berkum and colleagues may be taken to reflect ambiguity on two different levels of language processing. First, it may reflect ''deep'' situation model ambiguity from having two referential candidates for a single anaphor. Alternatively, it may reflect a more ''superficial'' ambiguity at the level of memory for text (''textbase''; e.g., Kintsch, 1998) that arises because a single lexical concept (''the girl'') has been featured twice to denote two different entities in the discourse.

In the former case, this referentially induced ERP can actually inform us about the waxing and waning of referential accessibility, that is, the changes in focus on story characters that accompany their comings and goings. For instance, although readers and listeners have a default focus on story protagonists (e.g., Zwaan & Radvansky, 1998), the accessibility of secondary characters (or objects) is partly dependent on whether they are part of the same scenario as the protagonist (e.g., Glenberg, Meyer, & Lindem, 1987). If they are, their accessibility will be greater, perhaps reflecting the need to keep them readily available as a future referent (see also Horton & Rapp, 2003, for an account in terms of perceptual perspective of story protagonists). Given that readers construct and continuously update a mental model that allows focused characters to be more easily accessed, they can, in principle, also use this model to avoid referential ambiguity. For example, when two girls were originally mentioned in a story, but one girl has passed away early in the story (and is therefore out of focus), ''girl'' in ''when they met a year later, he told the girl'' is not referentially ambiguous. Pursuing this logic, if the frontal negative shift reported by Van Berkum et al. (1999, 2003) reflects situation model referential ambiguity, it should also prove to be sensitive to this waxing and waning of referential accessibility.

In contrast, the ERP effect reported by Van Berkum et al. may also reflect a more superficial, low-level ambiguity. Such an account might go along the lines of what are called bottom-up, memory-based accounts of text processing (e.g., Kintsch, 1998; Myers & O'Brien, 1998). According to such accounts, a referring expression will automatically activate all antecedents that share features in common with that anaphor (e.g., when different entities are referred to by the same lexical concept), be it by a process of spreading activation (e.g., Kintsch, 1998) or resonance (e.g., Myers & O'Brien, 1998). As a consequence, these antecedents will initially be returned to working memory, independent of whether this will ultimately facilitate further processing, and irrespective of whether they are suitable candidates. Although it would be interesting if the ERP effect reported by Van Berkum et al. reflects such a ''superficial'' ambiguity, this effect then could not inform us on how readers and listeners ultimately establish reference during discourse comprehension.

In short, the abovementioned competing accounts illustrate that the functional significance of this ERP effect still needs to be established. Whether it can actually be used to selectively track discourse-level referential processing hinges on whether it is sensitive to discourse-level changes in referential accessibility of story characters. In the experiment reported below, we addressed this issue by examining whether the elicitation of this ''referentially induced'' ERP effect is averted by a discourse-level semantic manipulation that prevents referential ambiguity.

Subjects listened to short stories describing a protagonist and two secondary characters engaged in conversation. Three story types were created by manipulating the denotation of the secondary characters and/or the story content (see Table 1), but every story contained two NPs that referred to one of the secondary characters (the precritical and critical NP, respectively). In the 1–1 referent condition, both the precritical and critical NPs had only one possible referent, hence, were referentially unambiguous. In the 2–2 referent condition, both NPs had two possible referents, hence, were referentially ambiguous. But in the 2–1 referent condition, the precritical NP was referentially ambiguous, whereas the critical NP was referentially unambiguous.

We measured ERP responses to precritical and critical NPs. In the 2–2 referent condition, both ambiguous NPs should elicit a frontal, sustained negative shift (as reported by Van Berkum et al., 1999, 2003) relative to the corresponding NPs in the 1–1 referent condition. The precritical NP in the 2–1 referent condition was also referentially ambiguous, and should thus also elicit this ERP effect. Our crucial comparison, however, was at the critical NP in the 2–1 referent condition (when one secondary character had already left the scene). If this unambiguous NP still elicits a frontal negative shift, the ERP effect reported by Van Berkum et al. can be taken to reflect superficial, ''textbase'' ambiguity. Alternatively, if the unambiguous NP does not elicit a frontal negative shift but an ERP that resembles the 1–1 referent condition

Table 1. Example Story (Approximate Translation from Dutch)

1–1 Referent

At the family get-together, Jim had been talking to his nephew who was very much into politics and his uncle who was really into history. But Jim himself was only interested in sports, cars, girls etc. The **uncle** who was into history left early, but the nephew who was into politics kept rambling on. Jim didn't understand one bit and got rather bored. He told the **nephew** who was into politics that politicians should not systematically neglect delightful and important subjects like sports and girls.

2–2 Referent

At the family get-together, Jim had been talking to one nephew who was very much into politics and another one who was really into history. But Jim himself was only interested in sports, cars, girls etc. The **nephew** who was into history kept telling boring stories, and the other one also kept rambling on. Jim didn't understand one bit and got rather bored. He told the **nephew** who was into politics that politicians should not systematically neglect delightful and important subjects like sports and girls.

2–1 Referent

At the family get-together, Jim had been talking to one nephew who was very much into politics and another one who was really into history. But Jim himself was only interested in sports, cars, girls etc. The **nephew** who was into history *left early*, but the nephew who was into politics kept rambling on. Jim didn't understand one bit and got rather bored. He told the **nephew** who was into politics that politicians should not systematically neglect delightful and important subjects like sports and girls.

Precritical and critical NPs are printed in boldface, and the crucial event in the 2–1 referent condition is printed in italics.

instead, the referentially induced ERP can be taken to reflect ''deep'' discourse-level referential ambiguity.

METHODS

Thirty-one right-handed, healthy college students (14 men, mean age = 21.8 years, native speakers of Dutch) listened to 90 naturally spoken Dutch stories. Each story consisted of five sentences and contained one protagonist and two secondary characters. The protagonist was always denoted with either a proper name (e.g., ''Jim'') or a simple definite NP (e.g., ''the police-officer''). The two secondary characters were always denoted with an NP followed by a relative clause of at least four words (e.g., ''the nephew who was into history''). In the 1–1 referent condition, the secondary characters were denoted with different NPs (e.g., one uncle who... and one nephew who...). In the 2–2 and 2–1 referent condition, identical NPs were implied by using the substitution ''another'' to circumvent any differential effect of lexical repetition of the critical NPs across conditions. The subsequent relative clauses consisted of at least three words and differed only in the final word (e.g., ''who was into history'' and ''who was into politics'').

The first sentence of each story introduced all three characters. The second sentence always made explicit reference only to the protagonist. The third sentence described the actions of both the secondary characters: In the 1–1 and 2–1 referent condition, one secondary character left the scene (e.g., ''died'' or ''left the room''). In the 2–2 referent condition, both characters stayed in the current scene. The fourth sentence was used to again foreground the protagonist. The fifth sentence was identical across conditions and followed the template ⟨protagonist⟩ ⟨matrix-verb⟩ ⟨secondary character⟩ ⟨remainder⟩. This final sentence involved a secondary character that was still present in the scene. Whether this secondary character was first-mentioned in either Sentence 1 or 3 was balanced over the set of stories. To justify the use of the complete relative clause in the final sentence, the theme of the final sentence was always closely related to the content of the relative clause (e.g., following ''who was into politics...,'' the remainder of the sentence would subsequently describe political affairs).

Three different versions of each story were recorded by the same female speaker with a normal speaking rate and intonation. To control for systematical, prosodic differences between referentially ambiguous and nonambiguous speech segments, we had 20 different subjects participate in a forced-choice decision task. After subjects were fully informed about possible prosodic differences in case of referential ambiguity, their task was to classify all speech segments (truncated before the disambiguating word, e.g., ''the nephew who was into'') as taken from a sentence in which this segment was referentially ambiguous or not. The results from this pretest showed that they were not able to classify segments above chance level.

Three trial lists were used for the ERP experiment. For the first list, 30 stories of each condition were pseudorandomly mixed with 60 filler stories (materials from Nieuwland & Van Berkum, 2006b, Experiment 1), such that neither condition occurred more than two times consecutively, whereas trials of each type were matched on average list position. The other two lists were derived from the first by rotating the different trial types. The total of 150 stories was divided in 10 blocks, separated by a pause. Each trial was separated from the next by a 5-sec silence and was preceded by a short warning tone. Total time-on-task was approximately 80 min. Participants were seated in front of two loudspeakers and were informed that they would be listening to short stories. They were instructed to listen for comprehension and minimize movement. No additional task demands were imposed.

The electroencephalogram (EEG) was recorded from 30 standard scalp locations, amplified (band-pass filtered at 0.03–100 Hz), digitized at 500 Hz, and re-referenced to the mean of left and right mastoids. Ocular and muscular artifacts were corrected by means of a procedure based on Independent Component Analysis (e.g., Makeig, Jung, Ghahremani, Bell, & Sejnowski, 1997). Then, epochs that ranged from −500 msec to +2000 msec relative to critical word onset were extracted and normalized (by subtraction) to a 150-msec preonset baseline. Subsequently, segments with potentials exceeding ±75 μV were rejected, and the remainder was screened for drift artifacts. If the total rejection rate exceeded 40%, data of the participant were excluded. Six participants were excluded, leaving an average segment loss of 18% across the remaining 25 participants.

RESULTS

As shown in Figure 1, the referentially ambiguous, precritical NPs from both the 2–2 and 2–1 referent condition elicited a sustained, frontal negative shift relative to unambiguous NPs in the 1–1 referent condition. The corresponding difference waveforms and their scalp distributions are displayed in Figure 2. These results are very comparable to, and thus, replicate the findings by Van Berkum et al. (2003).

The results from our crucial comparison are displayed in Figure 3. The critical NPs in the 2–2 referent condition again elicited a frontal negative shift relative to critical NPs in the 1–1 referent condition, which had a distribution similar to the effect elicited by precritical NPs. However, this effect did not reappear in the 2–1 referent condition where one character had left the scene. The corresponding difference waveforms and their scalp distributions are displayed in Figure 4. These results suggest that a discourse-level semantic manipulation that prevents referential ambiguity also averts the elicitation of the ERP effect reported by Van Berkum et al. (1999, 2003, 2004).

The effects of referential ambiguity at the precritical NPs in the 2–2 and 2–1 referent conditions and at the critical NPs in the 2–2 referent condition were

Figure 1. Grand averages elicited by precritical NPs. In this and all the following figures, negativity is plotted upward and waveforms are filtered (5 Hz high cutoff, 48 dB/oct) for presentation purpose only.
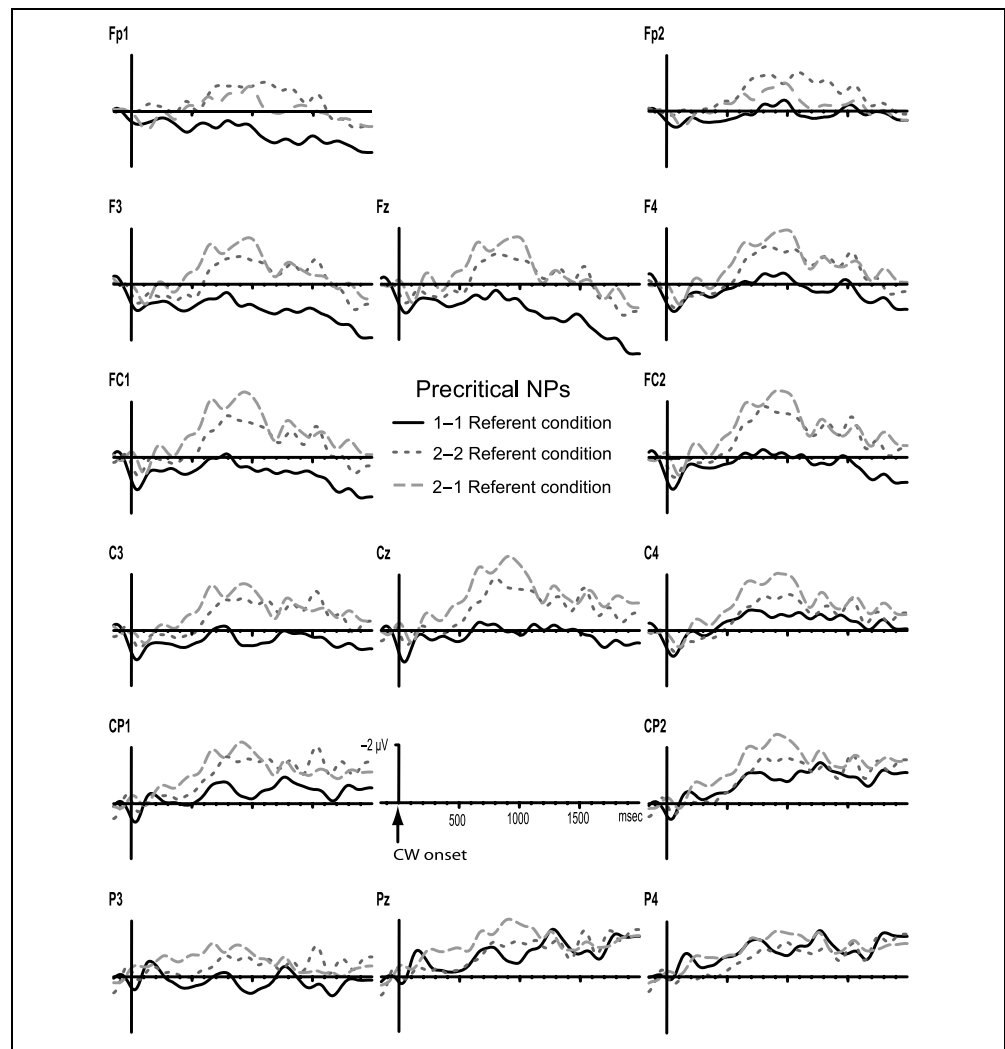
Figure 2. Difference waves at Fz for precritical NPs (2–2 minus 1–1 referent condition, and 2–1 minus 1–1 referent condition), and corresponding scalp distributions for the time windows that were used for statistical analysis.
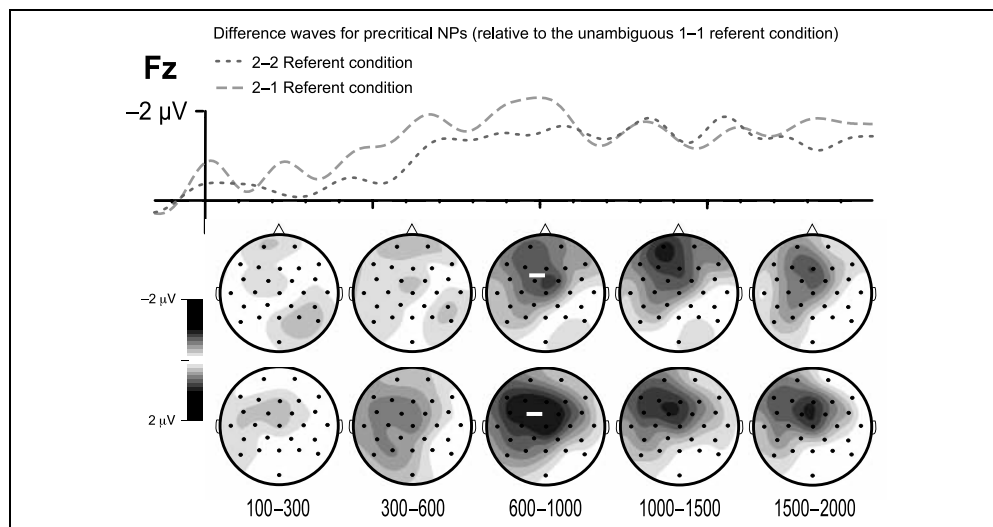
statistically significant using mean amplitude in the 500–1500 msec window at anterior channels, but not at posterior channels. Because these ambiguity effects were benchmarks for interpreting the comparison at the critical NP in the 2–1 referent condition, and we wanted to have maximal power in the statistical tests involving this critical effect, we used only anterior channels for further analysis.

Figure 3. Grand averages elicited by critical NPs.
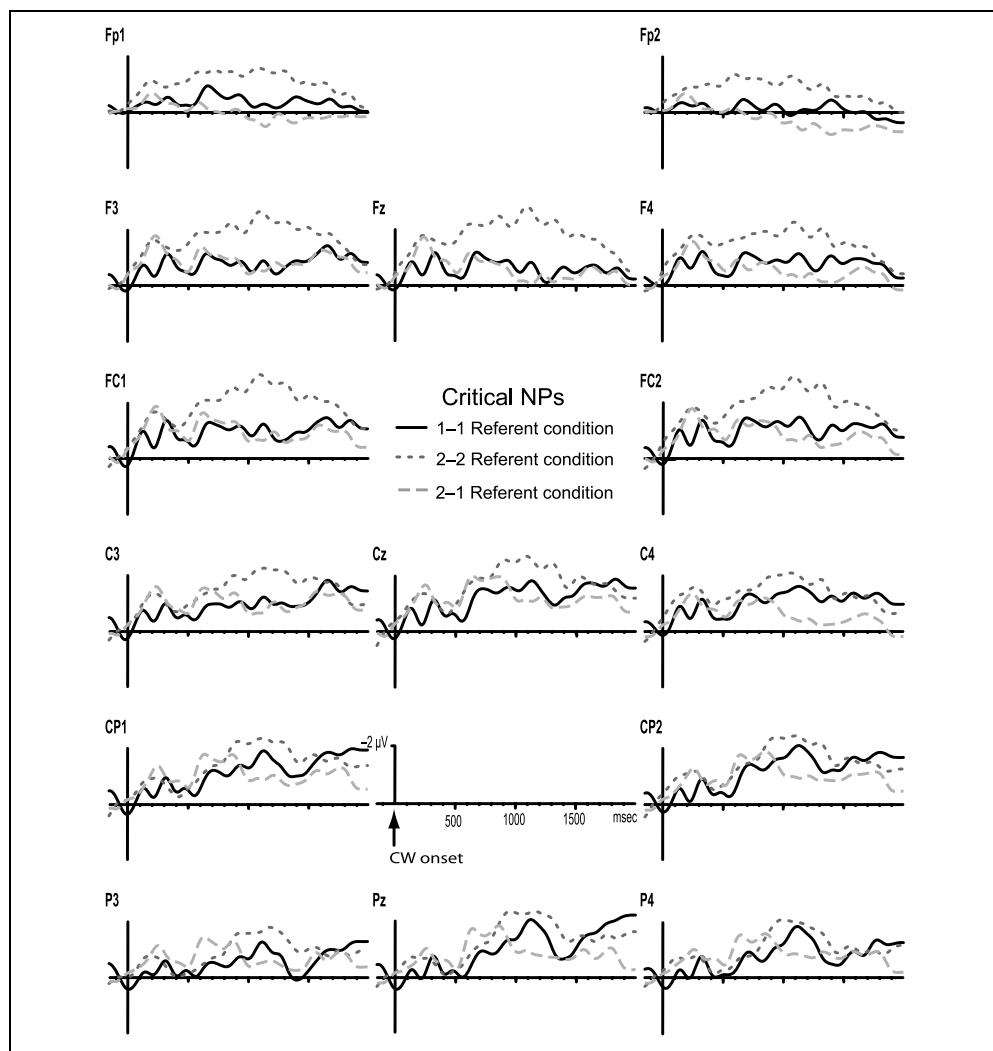
Using mean amplitude in five consecutive latency windows (100–300, 300–600, 600–1000, 1000–1500, 1500–2000 msec from word onset), the overall 2 (condition) × 15 (electrode) analysis of variance revealed significant main effects of referential ambiguity for precritical NPs in the 2–2 referent condition (600–1000 and 1000–1500 msec windows) and precritical NPs in the 2–1 referent condition (300–600, 600–1000, 1000–1500, and 1500–2000 msec windows; see Table 2).

An identical analysis for the ERPs elicited by the critical NPs revealed significant main effects of referential ambiguity for the 2–2 referent condition in the 100–300, 300–600, 600–1000, and 1000–1500 msec windows. Crucially, the 2–1 referent condition did not significantly differ from the 1–1 referent condition in any of the time windows. Subsequent comparison of the ERPs elicited by critical NPs in the 2–2 and 2–1 referent conditions revealed significant main effects of ambiguity in the 300–600 msec [$F(1,24) = 6.65$, $p = .016$], 600–1000 msec [$F(1,24) = 10.66$, $p = .003$], and 1000–1500 msec [$F(1,24) = 25.11$, $p = .000$] windows.

Finally, we analyzed ERPs elicited by referentially disambiguating words (e.g., ''politics'' and ''history'') following precritical and critical NPs (displayed in Figure 5). There were no significant differences between the 1–1 referent condition and the other two conditions in any time window (at either anterior of posterior channels), except for the 1000–1500 msec window following precritical NPs. In this time window, the 2–2 referent condition significantly differed from the 1–1 condition [$F(1,24) = 5.83$, $p = .031$] and almost significantly differed from the 2–1 condition [$F(1,24) = 5.12$, $p = .064$]. These differences are perhaps caused by differences in story content that occur after these disambiguating words: The information that both characters stay in the current scene (in the 2–2 referent condition) may be more salient or unexpected than one character leaving the scene.

Table 2. F Values and p Values for Main Effects of Referential Ambiguity in Five Consecutive Time Windows (at Anterior Channels Only)

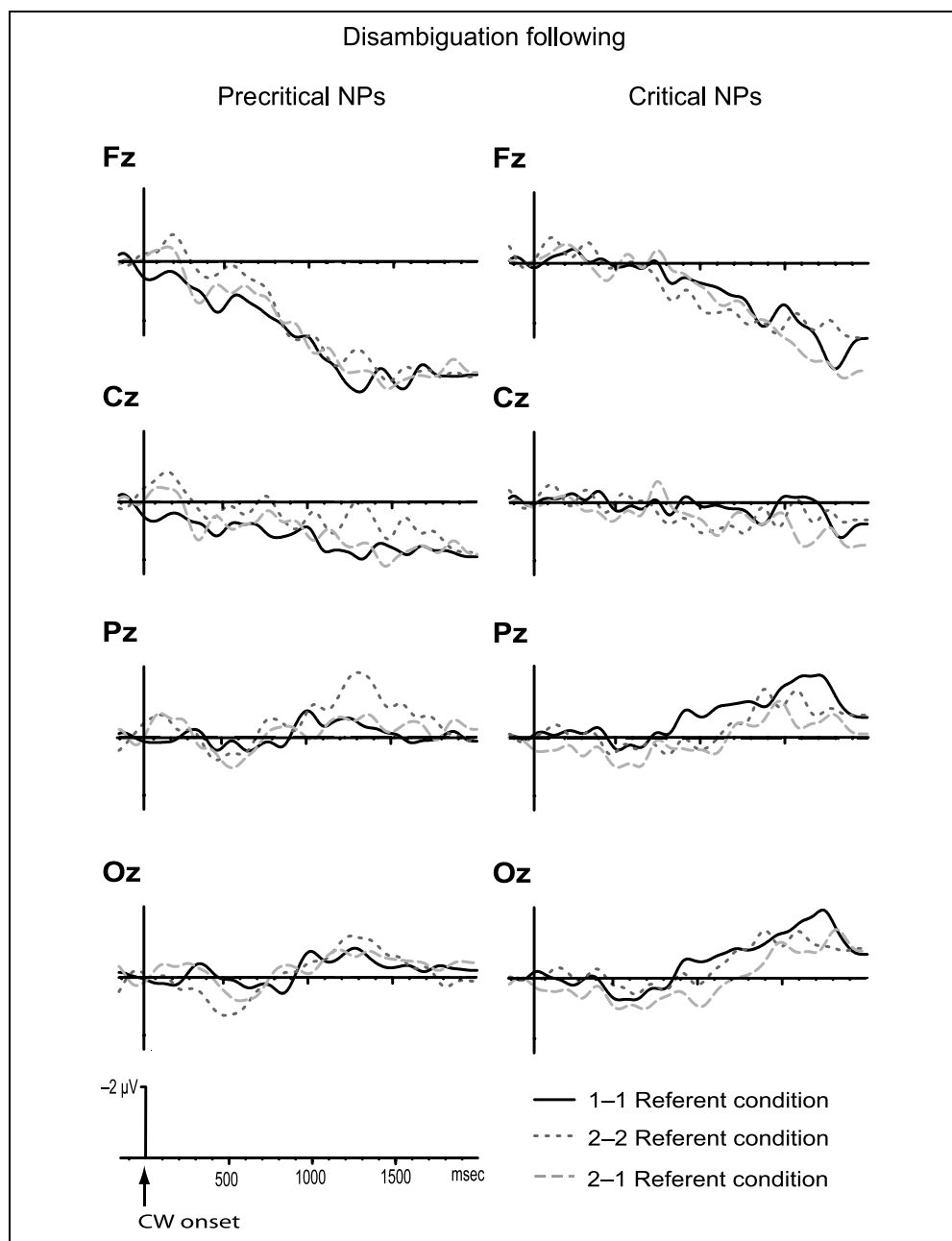|	Latency Range from Acoustic Word Onset (msec)				
	100–300	300–600	600–1000	1000–1500	1500–2000
Precritical NPs					
2–2 referent condition					
F	0.47	1.28	14.00	17.41	2.94
p	.50	.269	.001***	.000****	.099
2–1 referent condition					
F	0.80	7.56	18.20	12.86	7.06
p	.38	.01*	.000****	.001***	.014*
Critical NPs					
2–2 referent condition					
F	5.58	4.55	4.61	8.34	0.33
p	.03*	.04*	.042*	.008**	.57
2–1 referent condition					
F	0.96	0.02	0.45	1.16	0.86
p	.34	.97	.51	.292	.36

For all F tests, numerator $df = 1$, denominator $df = 24$.

*$p < .05$, **$p < .01$, ***$p < .005$, ****$p < .001$.

Importantly, the referentially induced ERP effect persisted after participants had heard the disambiguating words (the mean distance between NPs and disambiguating words was 1083 msec, min = 673, max = 1139). Together with the absence of any immediate effect of disambiguating words, this suggests that although the

Figure 4. Difference waves for critical NPs.
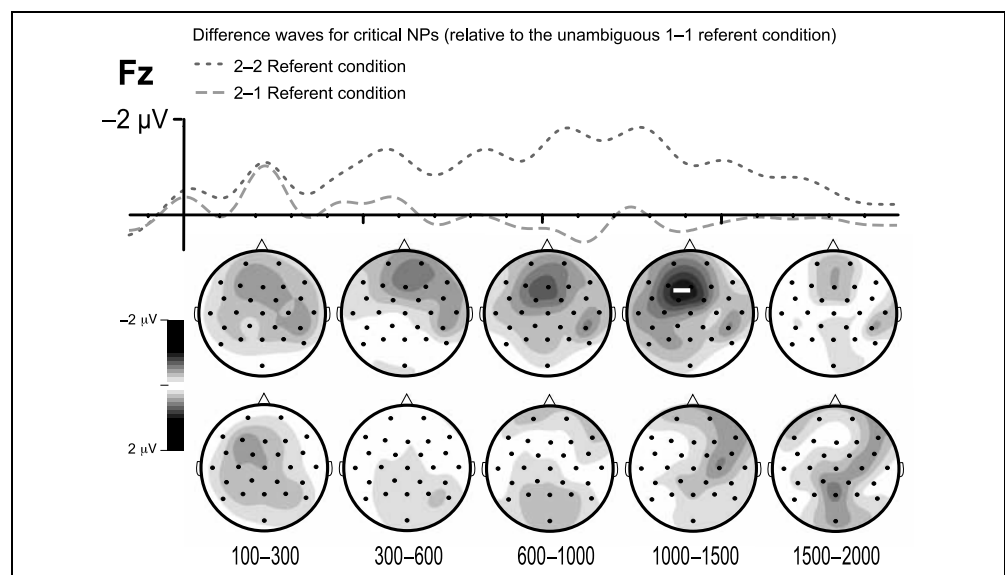
Figure 5. ERPs elicited by disambiguating words following precritical (left) and critical (right) NPs at midline electrodes.

Disambiguation following

Precritical NPs   Critical NPs

Fz   Fz

Cz   Cz

Pz   Pz

Oz   Oz

−2 μV

500   1000   1500   msec

CW onset

—— 1–1 Referent condition
······ 2–2 Referent condition
– – – 2–1 Referent condition

referentially induced ERP effect is immediately triggered by referentially ambiguous words, it ultimately outlasts the resolution of this ambiguity.

DISCUSSION

Our ERP findings confirm earlier evidence for rapid referential analysis in discourse comprehension (Van Berkum et al., 1999, 2003; see Tanenhaus, Spivey Knowlton, Eberhard, & Sedivy, 1995, for comparable results from visual world eye-tracking experiments) and are consistent with delays observed in the reading of referentially ambiguous nouns (e.g., Garnham, 2001; Myers & O'Brien, 1998). More importantly, our results suggest that

the frontal, sustained negative shift reported by Van Berkum et al. (1999, 2003) reflects genuine referential ambiguity from having two eligible referents in the current model of the discourse, and not superficial ambiguity from two entities having been mentioned in the discourse. As such, it has proven sensitive to changes in referential accessibility that typically accompany the comings and goings of story characters, and can thus be brought into action as a tool to selectively track referential processing during discourse comprehension.

Upon encountering a referentially ambiguous word, a listener is temporally unable to construct a coherent mental model of what is described in the discourse, and has to take into account that both eligible candidates may

turn out to be the intended referent. Note that these candidates do not need to be semantically equivalent (e.g., "two girls"), as a similar ERP effect is also elicited by lexically ambiguous words that can refer to equally likely, though semantically different entities (e.g., "a pitcher" in "The pitcher fell down..."; Hagoort & Brown, 1994). In this latter case, one could argue that lexically ambiguous words (when presented in a sufficiently neutral context) force the reader to set up two competing referents. But whether the referential candidates are semantically equivalent or not, referential ambiguity may very well demand additional working memory resources, which may explain why the referentially induced ERP effect resembles ERP effects that are associated with increased working memory load in linguistic (e.g., Kutas, 1997) as well as nonlinguistic tasks (e.g., Donaldson & Rugg, 1999), and why it is correlated with reading span performance (Nieuwland & Van Berkum, 2006a).

Potential cues to the functional neuroanatomical underpinnings of the referentially induced ERP effect come from a recent functional magnetic resonance imaging study by Rodd, Davis, and Johnsrude (2005). In their study, subjects were presented lexically ambiguous words in relatively neutral sentence contexts (e.g., "He saw the hair/hare..."), thus comparable to the materials known to evoke sustained, frontal negative shifts in the ERP (Hagoort & Brown, 1994). These ambiguous sentences were found to increase activity in temporal and frontal brain regions, including the left inferior frontal gyrus (LIFG) and, to some extent, its right hemisphere counterpart. In recent years, the LIFG has been associated with a range of processes, and there is still discussion about its functional subdivision. However, there seems to be a growing consensus that at least the anterior part of the LIFG (BA 45/47) is involved in making comparisons of semantic representations held in working memory during linguistic and nonlinguistic processing (e.g., Bookheimer, 2002; Kaan & Swaab, 2002; Wagner, Pare-Blagoev, Clark, & Poldrack, 2001; Thompson-Schill, D'Esposito, Aguirre, & Farah, 1997), or, as Novick, Trueswell, and Thompson-Schill (2005) put it, detecting and resolving conflict when multiple characterizations of the linguistic or nonlinguistic input point toward incompatible interpretations. It thus seems reasonable to predict that the LIFG (or, more specifically, at least BA 45/47) is also involved in processing referential ambiguity. We are currently undertaking a functional neuroimaging study to test this prediction and to explore which areas of the human brain are differentially involved in dealing with referential ambiguity, referential failure (when no referential candidate is directly available at all), and semantic incongruity during language comprehension.

Although the present findings show that ERPs can be used to selectively track referential analysis in discourse comprehension, our results also raise new questions with regard to the resolution of referential ambiguity. The referentially induced ERP effect proved sensitive to discourse-level semantic information, but did not immediately disappear upon encountering a disambiguating word. These seemingly paradoxical results indicate that the increased working memory load due to referential ambiguity has relatively long-lasting effects on language processing (see also Kutas, 1997, for similar results using syntactic complexity manipulations). In other words, our findings may suggest that although making a demand on processing capacity has an immediate effect, letting go of this demand is a more gradual process. Further research is needed to clarify these issues.

One other caveat that should be made here is that the elicitation of a frontal negative shift by a referentially problematic word may be restricted to referentially *ambiguous* words. That is to say, this ERP effect is not the sole "representative" of referential problems in general. ERP studies investigating other referential aspects of language comprehension (e.g., Nieuwland & Van Berkum, 2006a; Van Berkum et al., 2004, in press; Anderson & Holcomb, 2005; Swaab, Camblin, & Gordon, 2004; Harris, Wexler, & Holcomb, 2000; Streb, Rosler, & Hennighausen, 1999; Osterhout & Mobley, 1995) suggest that different referential "problems" (e.g., nonparallel discourse structures, referential failure, synonymy) may elicit ERP effects that are typically associated with semantic or syntactic problems (e.g., N400 or P600). Perhaps because establishing reference is intricately intertwined with phonological, syntactic, and semantic levels of linguistic representation, referential problems may sometimes affect the analysis at other levels and become manifest as such (see Van Berkum et al., in press, for a discussion).

In all, our results show that the ERP effect that is associated with referential ambiguity reflects genuine ambiguity from having two eligible candidates in the situation model of the discourse. This suggests that at least one type of problem with referential analysis, namely, referential ambiguity, can be selectively tracked with ERPs at the level that is most relevant to discourse comprehension, the situation model.

Acknowledgments

REFERENCES

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

Bookheimer, S. (2002). Functional MRI of language: New approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience, 25,* 151–188.

Brown, C. M., Hagoort, P., & Kutas, M. (2000). Postlexical integration processes in language comprehension: Evidence from brain-imaging research. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (2nd ed., pp. 881–895). Cambridge: MIT Press.

Donaldson, D. I., & Rugg, M. D. (1999). Event-related potential studies of associative recognition and recall: Electrophysiological evidence for context dependent retrieval processes. *Cognitive Brain Research, 8,* 1–16.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in Cognitive Sciences, 6,* 78–84.

Garnham, A. (2001). *Mental models and the interpretation of anaphora.* Philadelphia, PA: Taylor and Francis.

Glenberg, A. M., Meyer, M., & Lindem, K. (1987). Mental models contribute to foregrounding during text comprehension. *Journal of Memory and Language, 26,* 69–83.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Review of Psychology, 48,* 163–189.

Hagoort, P., & Brown, C. M. (1994). Brain responses to lexical ambiguity resolution and parsing. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing.* Hillsdale, NJ: Erlbaum.

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

Harris, T., Wexler, K., & Holcomb, P. (2000). An ERP investigation of binding and coreference. *Brain and Language, 75,* 313–346.

Horton, W. S., & Rapp, D. N. (2003). Out of sight, out of mind: Occlusion and the accessibility of information in narrative comprehension. *Psychonomic Bulletin and Review, 10,* 104–109.

Kaan, E., & Swaab, T. Y. (2002). The brain circuitry of syntactic comprehension. *Trends in Cognitive Sciences, 6,* 350–356.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition.* New York: Cambridge University Press.

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

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. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 10979–10984.

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. (2006a). Individual differences and contextual bias in pronoun resolution: Evidence from ERPs. *Brain Research, 1118,* 155–167.

Nieuwland, M. S., & Van Berkum, J. J. A. (2006b). When peanuts fall in love: N400 evidence for the power of

discourse. *Journal of Cognitive Neuroscience, 18,* 1098–1111.

Novick, J. M., Trueswell, J. C., & Thompson-Schill, S. L. (2005). Cognitive control and parsing: Reexamining the role of Broca's area in sentence comprehension. *Cognitive, Affective, & Behavioral Neuroscience, 5,* 263–281.

Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34,* 739–773.

Rodd, J. M., Davis, M. H., & Johnsrude, I. S. (2005). The neural mechanisms of speech comprehension: fMRI studies of semantic ambiguity. *Cerebral Cortex, 15,* 1261–1269.

Streb, J., Rosler, F., & Hennighausen, E. (1999). Event-related responses to pronoun and proper name anaphors in parallel and nonparallel discourse structures. *Brain and Language, 70,* 273–286.

Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Electrophysiological evidence for reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience, 16,* 715–726.

Tanenhaus, M. K., Spivey Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632–1634.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 14792–14797.

Van Berkum, J. J. A. (2004). Sentence comprehension in a wider discourse: Can we use ERPs to keep track of things? In M. Carreiras & C. Clifton Jr. (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond* (pp. 229–270). New York: Psychology Press.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language, 41,* 147–182.

Van Berkum, J. J. A., Koornneef, A., Otten, M., & Nieuwland, M. S. (in press). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research.*

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 Berkum, J. J. A., Zwitserlood, P., Hagoort, P., & Brown, C. M. (2003). When and how do listeners relate a sentence to the wider discourse? Evidence from the N400 effect. *Cognitive Brain Research, 17,* 701–718.

Wagner, A. D., Pare-Blagoev, E. J., Clark, J., & Poldrack, R. A. (2001). Recovering meaning: Left prefrontal cortex guides controlled semantic retrieval. *Neuron, 31,* 329–338.

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin, 123,* 162–185.