BRAIN RESEARCH 1118 (2006) 155–167

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Individual differences and contextual bias in pronoun resolution: Evidence from ERPs

Mante S. Nieuwland[a,*], *Jos J.A. Van Berkum*[a,b]

[a]*University of Amsterdam, Department of Psychology (PN), Roetersstraat 15, 1018 WB Amsterdam, The Netherlands*
[b]*F.C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands*

ARTICLE INFO

Article history:
Accepted 7 August 2006
Available online 7 September 2006

Keywords:
Pronoun resolution
Anaphoric processing
Individual difference
Reading Span
Referential ambiguity
ERPs

ABSTRACT

Although we usually have no trouble finding the right antecedent for a pronoun, the co-reference relations between pronouns and antecedents in everyday language are often 'formally' ambiguous. But a pronoun is only really ambiguous if a reader or listener indeed perceives it to be ambiguous. Whether this is the case may depend on at least two factors: the language processing skills of an individual reader, and the contextual bias towards one particular referential interpretation. In the current study, we used event related brain potentials (ERPs) to explore how both these factors affect the resolution of referentially ambiguous pronouns. We compared ERPs elicited by formally ambiguous and non-ambiguous pronouns that were embedded in simple sentences (e.g., "Jennifer Lopez told Madonna that she had too much money."). Individual differences in language processing skills were assessed with the Reading Span task, while the contextual bias of each sentence (up to the critical pronoun) had been assessed in a referential cloze pretest. In line with earlier research, ambiguous pronouns elicited a sustained, frontal negative shift relative to non-ambiguous pronouns at the group-level. The size of this effect was correlated with Reading Span score, as well as with contextual bias. These results suggest that whether a reader perceives a formally ambiguous pronoun to be ambiguous is subtly co-determined by both individual language processing skills and contextual bias.

1. Introduction

Almost no conversation goes by without somebody using a pronoun to refer to some object, action or individual. Pronouns are our linguistic 'short-cuts' for maintaining reference to topics that are in the focus of our conversation. In fact, comprehending pronouns is such common practice, we usually feel as if we understand their antecedents immediately (e.g., Clark and Sengul, 1979). Interestingly, we generally do not even notice that the co-reference relations between pronouns and their antecedents in everyday language are often 'formally' ambiguous, that is, when the linguistic features of a pronoun (e.g., male/female, singular/plural) by itself do not warrant the retrieval of a unique antecedent. For instance, given the sentence-fragment "Bruce Willis hated Al Pacino because he…", most people will automatically take 'Al Pacino' as the antecedent of 'he' (as the most likely continuation of this sentence will describe the characteristics of Al Pacino which he was hated for), even though 'he' might, in principle, also refer to 'Bruce Willis'. In

* *Corresponding author*. Fax: +31 20 6391656.
E-mail address: m.s.nieuwland@uva.nl (M.S. Nieuwland).

0006-8993/$ – see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2006.08.022

the current study, we used event-related brain potentials (ERPs) to explore how the resolution of such formally ambiguous pronouns is subtly influenced by both individual differences in language processing skills and sentence-level contextual bias.

How listeners and readers comprehend pronouns has been extensively studied for many years now (e.g., Garnham, 2001, for review). An important topic in this line of research has been whether pronouns and antecedents are immediately matched on syntactic features like number and gender (e.g., Arnold et al., 2000; Gernsbacher, 1989; McDonald and MacWhinney, 1995; Sanford and Garrod, 1989; see also Osterhout and Mobley, 1995; Schmitt et al., 2002; Streb et al., 1999, for examples of ERP studies addressing pronoun resolution). Another thoroughly investigated issue comprises the strategies that readers and listeners opt for when no referent matches sufficiently or when more than one referent matches equally well, i.e. when readers may engage in optional strategic processing to identify the referent of an ambiguous pronoun (e.g., Greene et al., 1992; Jarvikivi et al., 2005; MacDonald and MacWhinney, 1990; Tyler and Marslen Wilson, 1982). For example, many researchers have tried to unravel whether readers display certain preferences or strategies when resolving ambiguous pronouns (e.g., Crawley et al., 1990, Gernsbacher and Hargreaves, 1988; Jarvikivi et al., 2005). However, although individual differences in other aspects of language comprehension have long been established (e.g., Daneman and Carpenter, 1980), potential differences in how people resolve formally ambiguous pronouns has remained relatively unexplored territory.

It is not hard to see how people may differ in processing a pronoun like ‘he’ in “Bruce Willis hated Al Pacino because he…”, because the formal ambiguity it conveys is a rather subtle linguistic phenomenon that can easily be overlooked. Though ‘he’ can, in principle, be taken to refer to either celebrity, readers might just as well immediately take on the – to them, and at that moment – most obvious referential commitment and completely miss out on this formal ambiguity. Alternatively, some readers may indeed take into account both possible referential interpretations, and engage in additional inferencing to identify the correct antecedent. In short, a potential difference in how individuals process formally ambiguous pronouns may lie in whether they do or do not actually perceive anything ambiguous about these pronouns. Of course, the question then remains why some people do and others do not notice anything ambiguous.

Traditionally, a particular explanatory role for individual variation in measures of language comprehension has been assigned to differences in working memory capacity (e.g., Daneman and Carpenter, 1980; Just and Carpenter, 1992). Although the concept and measurement of working memory capacity remains a topic of debate (e.g., Caplan and Waters, 1999; Engle, 2002; MacDonald and Christiansen, 2002; Waters and Caplan, 1996), there is a wealth of data suggesting that individuals who perform better on a typical verbal working memory task like the Reading Span task also perform better on both off-line and on-line measures of language comprehension (e.g., Just and Carpenter, 1992). One of the general findings on the relationship between Reading Span score and language comprehension is that high span readers are more likely to elaborate their discourse models with optional, knowledge-based inferences than low span readers (e.g., Calvo, 2001; Linderholm, 2002; St. George et al., 1997; Whitney et al., 1991). Furthermore, in the case of lexical-semantic ambiguity during sentence processing, high span readers seem to have both possible interpretations more readily available than low span readers (e.g., Miyake et al., 1994).

The abovementioned types of evidence seem to provide convergent support for the hypothesis that high span readers will generally be more ‘sensitive’ to the different referential interpretations that are conveyed by formally ambiguous pronouns. Because high span readers seem more sensitive to the subtle constraints that guide language comprehension (e.g., MacDonald and Christiansen, 2002), they are more likely to temporarily take into account that either referential interpretation may ultimately be correct. Low span readers, on the other hand, may tend to be ‘insensitive’ to formal ambiguity and immediately take on the first referential commitment that comes to mind (for example, they may automatically rely on heuristic strategies or contextual bias to a greater extent than high span readers). Of course, the likelihood that a reader will immediately take on one particular referential interpretation upon encountering a formally ambiguous pronoun is not just based on his or her language processing skills, it is also greatly influenced by the degree of contextual bias (the syntactic and semantic/pragmatic constraints from the preceding context) towards one or the other interpretation. Contextual bias involves the complex interplay of different types of informational cues, for instance, whether a referent was the first-mentioned entity (e.g., Gernsbacher and Hargreaves, 1988) or the grammatical subject of a sentence (e.g., Crawley et al., 1990), but more importantly, it also involves world knowledge plausibility (e.g., McKoon et al., 1993). For example, one reason why it may be hard to immediately determine who is meant by ‘he’ in a sentence like “Bruce Willis told Al Pacino that he…” is that, in the real world, it makes approximately equal sense that somebody tells something about himself (e.g., “that he was sorry.”) as to tell something about another person (e.g., “that he was an arrogant drunk.”). But this relative, contextual ‘neutrality’ can be subtly altered in either direction: in the sentence “Bruce Willis *secretively whispered* to Al Pacino that he…”, the more likely antecedent is ‘Bruce Willis’ because people tend to be more secretive about themselves than about others. Alternatively, ‘Al Pacino’ is the more likely antecedent in “Bruce Willis *reproachfully told* Al Pacino that he…”, because people are generally more reproachful because of someone else’s actions than their own (see Koornneef and Van Berkum, 2006; Van Berkum et al., in press, for evidence in support of an immediate influence of verb-based plausibility on pronoun resolution). Furthermore, connectives that precede pronouns may also play an important role by signaling the coherence relations between the surrounding clauses. For example, while ‘because’ signals that the second clause will describe the cause of what is described in the first clause, ‘so’ signals

Table 1 – Example items (approximate translation from Dutch) with increasing contextual bias

Sentence	% NP1	% NP2	Contextual bias
Anton forgave Michael the problem because his … … car was a wreck.	50	50	0
The chemist hit the historian while he … … was laughing hard.	53	47	6
Bruce Willis told Al Pacino that he … … was a bit promiscuous.	37	63	26
Linda invited Anna when her … … irritating brother wasn't around.	70	30	40
The businessman called the dealer just as he … … left the trendy club.	20	80	60

The truncated sentences were used in the referential cloze task, while the complete sentences were used in the ERP experiment. %NP1 and %NP2 are the percentages of people in the referential cloze task who encircled the first- and second-mentioned entity respectively. Contextual bias was computed as the absolute difference between %NP1 and %NP2.

that the second clause will describe its consequence (e.g., Stevenson et al., 2000). In effect, people will generally take 'he' in "Bruce Willis hated Al Pacino *because* he…" to refer to Al Pacino, as Al Pacino is the most likely cause of Bruce Willis 's hatred, but most people will take 'he' in "Bruce Willis hated Al Pacino *so* he…" to refer to Bruce Willis, because Bruce Willis is the person most likely to act upon his own hatred.

It is generally thought that contextual bias makes us prefer one referential interpretation over another by determining the relative amounts of focus on entities in the discourse model, and that entities that are more focused are more easily accessible for subsequent reference (e.g., Gordon et al., 1993; Greene et al., 1992; Sanford and Garrod, 1989; Stevenson et al., 2000). Hence, contextual bias may also determine the online resolution of a pronoun that, in principle, can refer to two antecedents. In biased sentences, readers are more likely to immediately settle on the character that is in greater focus and less likely to engage in additional processing than if there is no bias and the two characters are equally focused candidates. In fact, if a sentence has a very strong contextual bias, one particular referent may be focused so much that readers immediately take this referent as the antecedent and not even notice anything ambiguous about a 'formally' ambiguous pronoun. For example, it might be just as easy to find the antecedent for 'he' in a two male referent context like "Bruce Willis reproachfully told Al Pacino that he…", as in a 1 male referent context like "Jennifer Lopez reproachfully told Al Pacino that he…".

In the present study, we examined how both language processing skills (as reflected by Reading Span performance) and contextual bias influence the processing of formally ambiguous pronouns. The use of event-related brain potentials (ERPs) enabled us to study on-line language comprehension without the need of an artificial probe task. In particular, we were able to exploit a recently discovered ERP effect that can be used to selectively track referential ambiguity during language comprehension (Nieuwland et al., in press; Van Berkum et al., 1999, 2003, 2004). This ERP effect, a sustained, frontal negative shift, is elicited by referentially ambiguous nouns (e.g., 'girl' in a two-girl context; Van Berkum et al., 1999, 2003) relative to non-ambiguous nouns. Furthermore, and of particular relevance to the current study, results from a spoken sentence-study (see Van Berkum et al., 2004) showed that this ERP effect is also elicited by formally referentially ambiguous pronouns, that is, singular, male pronouns in a context containing two males (e.g., 'he' in "John shot David as he…").

In a follow up study, Nieuwland et al. (in press) showed that this referentially induced ERP effect is elicited by ambiguity from actually having two equally eligible referents in the discourse model, and not by 'superficial' ambiguity from two similar entities having been mentioned in the discourse. For example, if two girls had originally been mentioned, but one of them had left the scene (e.g., had left the room or died), critical nouns did not elicit the frontal negative shift any longer. Taken together, these findings suggest that referential ambiguity can be selectively tracked with ERPs at the level that is most relevant to language comprehension, the situation model.

The abovementioned, referentially induced ERP effect allowed us, in the present study, to examine both the influence of individual differences in language processing skills and of contextual bias on pronoun resolution. Individual differences in language processing skills were first assessed with a new Dutch version of the Reading Span Task (Van den Noort et al., 2006), while contextual bias for each sentence was established in a separate referential cloze task (see Table 1). Subjects read sentences that were presented on a screen word-by-word, while we measured ERP-responses to the singular pronouns that were embedded in the sentences. The sentences contained two males, two females or one male and one female. Hence, the male pronouns in a 2-male referents context and the female pronouns in a 2-female context were 'formally' ambiguous pronouns. All pronouns in the 1-male and 1-female referent context were non-ambiguous.

We expected 'formally' ambiguous pronouns to elicit a sustained, frontal negative shift relative to non-ambiguous pronouns, replicating the results from the spoken-sentence study reported by Van Berkum et al. (2004). Crucially, we expected the size of this referentially induced ERP effect to be correlated with Reading Span score, as individual language processing skills should increase the reader's 'sensitivity' to formal referential ambiguity. Furthermore, similar to how a discourse context that prevents referential ambiguity also averts the elicitation of the referentially induced ERP effect (Nieuwland et al., in press), sentence-level differences in contextual bias may become visible in amplitude differences of this effect. Therefore, we expected the referential ambiguity effect to be dependent on the degree of contextual bias towards one particular referential interpretation.

Finally, to provide a background against which to evaluate the effects of referential ambiguity, we also included two conditions involving highly salient anomalies: lexical-semantic anomalies (e.g., "Al Pacino told Madonna told that she wasn't a very friendly *sausage* to be around.")

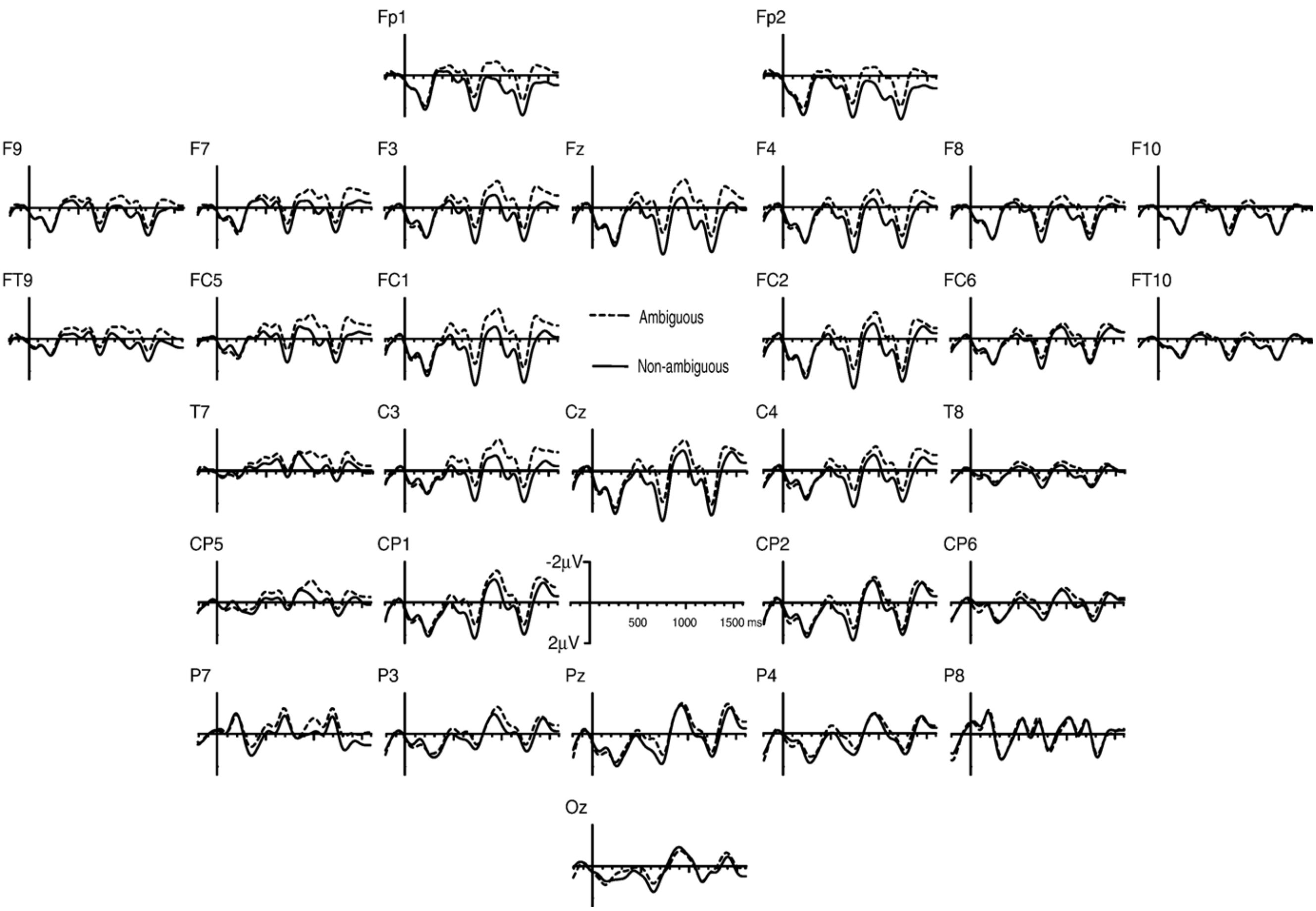

Fig. 1 – Grand averages elicited by formally ambiguous and non-ambiguous pronouns. In this and all following figures, negativity is plotted upwards and waveforms are filtered (5 Hz high cut-off, 48 dB/oct) for presentation purpose only.

and referentially failing pronouns (e.g., 'she' in "Al Pacino told Bruce Willis that *she* wasn't a very nice person to be around."[1]). ERPs elicited by the critical words in both types of anomalies were compared to ERPs elicited by their non-anomalous and referentially successful counterparts respectively (e.g. "Al Pacino told Madonna that *she* wasn't a very nice *person* to be around."). We expected lexical-semantic anomalies to elicit an N400 effect (e.g., Kutas and Hillyard, 1980), and referentially failing pronouns to elicit a P600 effect (e.g., Osterhout and Mobley, 1995; Van Berkum et al., 2004). This would allow us to examine whether the expected effect of language processing skills is specific to referential ambiguity or whether it would also affect other aspects of language comprehension (e.g., Bornkessel et al., 2004; Friederici et al., 1998; Gunter et al., 2003; Munte et al., 1998; Vos et al., 2001).

2. Results

2.1. *Main effects of referential ambiguity*

Fig. 1 displays the grand average waveforms elicited by the referentially ambiguous and non-ambiguous pronouns at all electrode locations. As visible from this figure, ambiguous pronouns elicited a sustained, slightly left-lateralized, frontal negative shift relative to non-ambiguous controls. Using mean amplitude in the 400–1500 ms latency window (based on visual inspection of the ERP waveforms), the overall Ambiguity (2) × Electrode (30) analysis of variance (ANOVA) revealed a main effect of Ambiguity ($F(1,29)=5.75$, $p=0.023$). Because simple main effects revealed a clear effect of Ambiguity at anterior electrodes only ($F(1,14)=6.74$, $p=0.015$), all subsequent analyses concerning referential ambiguity were done using mean amplitude of the 400–1500 ms window pooled over the 15 anterior electrodes.

[1] It should be noted that the present pronoun gender disagreements technically did not constitute any anomaly, because these pronouns could also be taken as referring to some third, unmentioned entity. However, under most models of pronoun resolution, readers first try to find an appropriate antecedent with the given set of discourse entities (e.g., Garnham, 2001), and such pronoun gender disagreements have been shown to elicit reliable ERP effects with both Dutch and English materials (Van Berkum et al., 2004; Osterhout and Mobley, 1995).

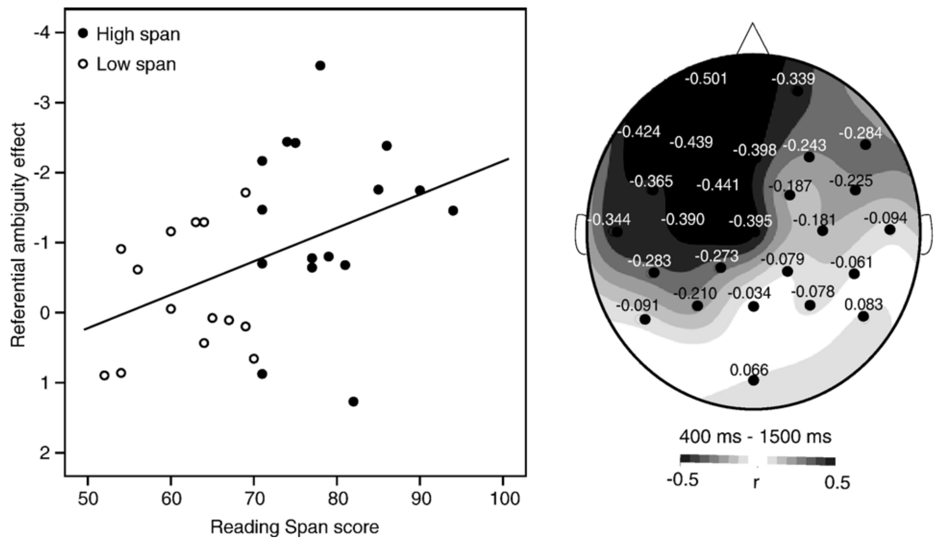

Fig. 2 – Relationship of Reading Span score and the referential ambiguity effect (left), and scalp distribution of the *r*-values corresponding to each electrode channel (right).

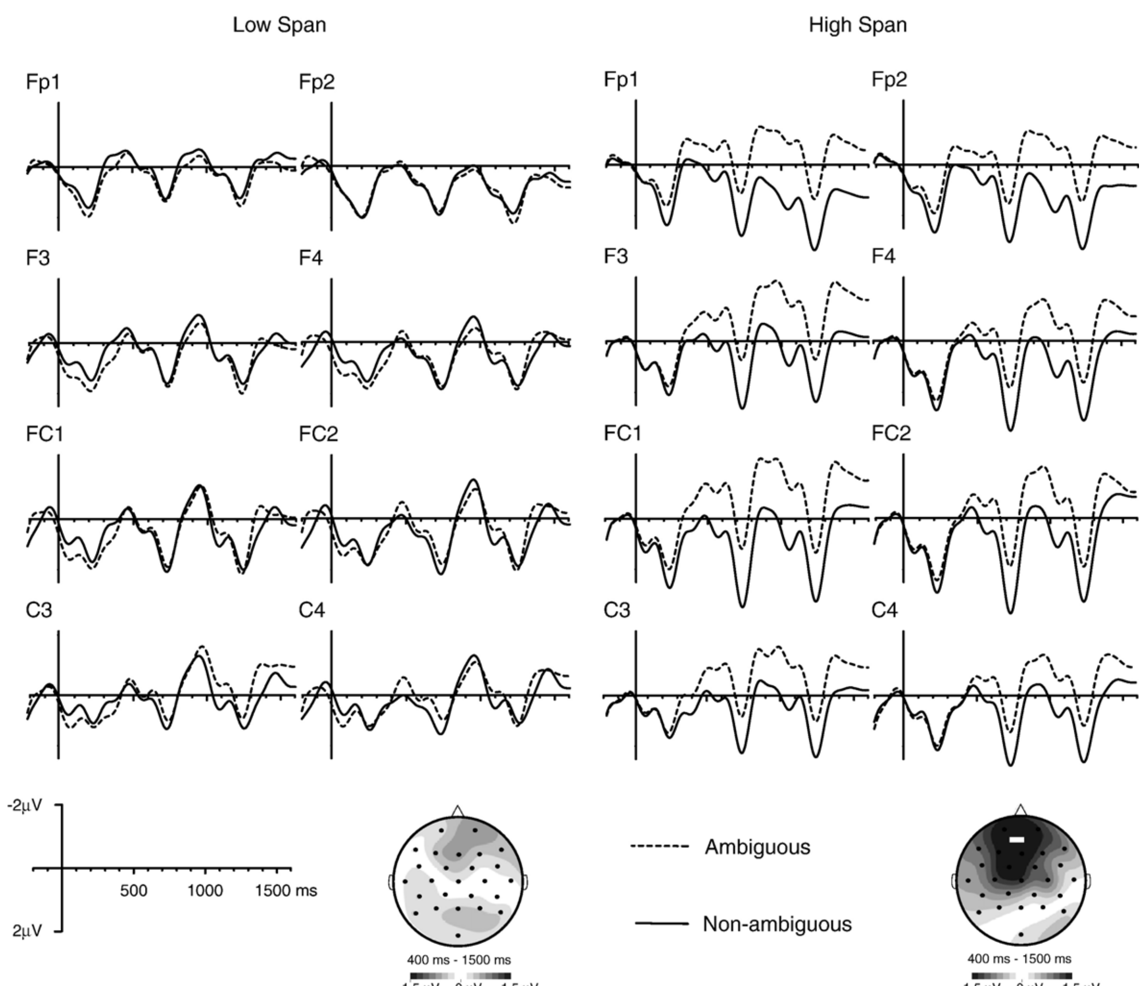

Fig. 3 – Grand averages elicited by ambiguous and non-ambiguous pronouns per Reading Span group. In this and all following figures, scalp distribution corresponding to the differential effect in the 400–1500 ms window is presented below the ERP plots.

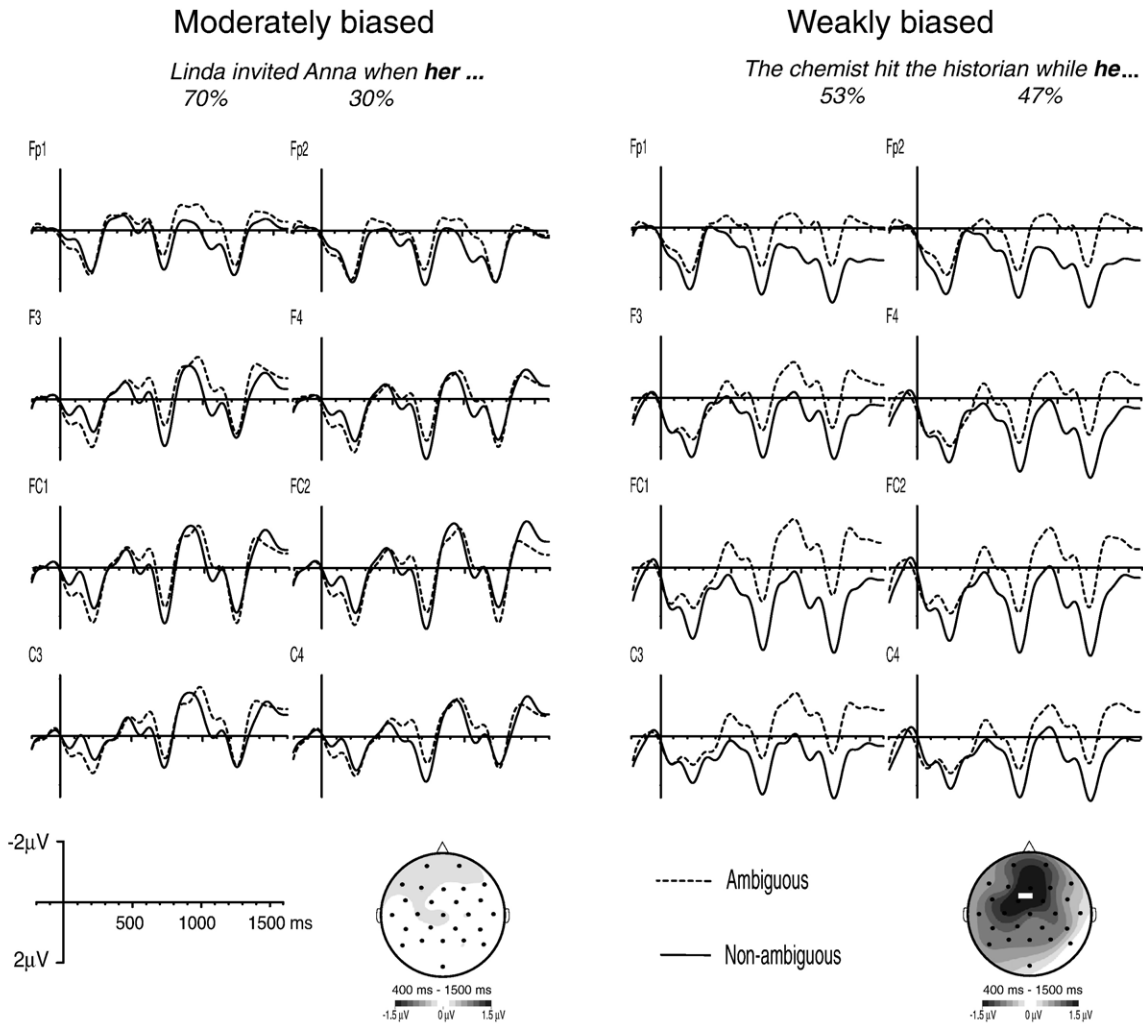

Fig. 4 – Grand averages elicited by ambiguous and non-ambiguous pronouns for moderately and weakly biased sentences. The numbers below the example sentences indicate the percentages of people who took the critical pronoun to refer to the first-mentioned entity (%NP1) or the second-mentioned entity (%NP2).

2.2. *Effects of Reading Span*

Reading Span scores ranged from 52 to 94 ($M=71$, $SD=10.8$). Importantly, this score was significantly correlated with the size of the referential ambiguity effect (Pearson's $r=0.41$, $p=0.024$; see Fig. 2).

Our thirty subjects were subsequently grouped into low span ($N=14$) and high span ($N=16$) groups using a median split. Reading Span score for the low span group ranged from 52 to 70 ($M=61.9$, $SD=6.1$, 4 males, mean age 20.8), and from 71 to 94 for the high span group ($M=78.9$, $SD=7.1$, 6 males, mean age 20.3). Grand average ERPs for the two groups are displayed in Fig. 3. The overall ANOVA revealed a significant Ambiguity(2)×Span Group(2) interaction effect ($F(1,28)=6.57$, $p=0.016$). Simple main-effect analysis within each group showed that the referential ambiguity effect was found to be significant in the high span group ($F(1,15)=11.39$, $p=0.004$), but not in the low span group. ($F(1,13)=0.003$, $p=0.956$).

2.3. *Effects of contextual bias*

Because of the small number of trials available per participant for grouping with regard to contextual bias, trials were only grouped into moderately biased and weakly biased items (mean bias 37.0% and 7.8% respectively; bias was computed as the absolute difference between the percentages of people who encircled the first- or second-mentioned entity for each sentence in the referential cloze task). Fig. 4 displays the grand average ERPs for moderately and weakly biased items and corresponding scalp distributions of the ambiguity effect. The Ambiguity(2)×Bias(2) interaction was only marginally significant ($F(1,29)=3.110$, $p=0.088$). Simple main-effect analysis within each bias type showed that the referential ambiguity effect was found to be significant in the weakly biased items ($F(1,29)=8.715$, $p=0.006$), but not in the moderately biased items ($F(1,29)=0.75$, $p=0.393$).

Although the abovementioned simple main effects are not completely covered by a significant interaction effect (perhaps

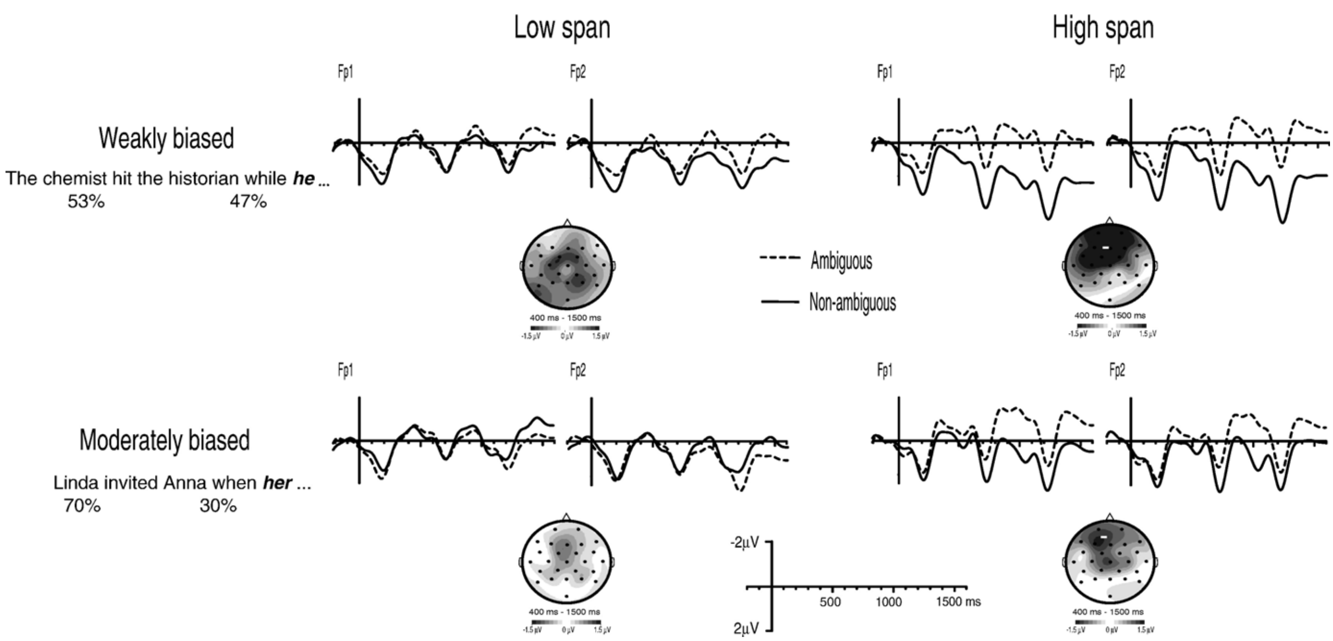

Fig. 5 – Grand averages elicited by ambiguous and non-ambiguous pronouns for moderately and weakly biased sentences for each Reading Span group at electrode positions Fp1 and Fp2.

because of a relatively low power for this test), it is licensed by the results of a more fine-grained correlational analysis of the relation between contextual bias and the referential ambiguity effect. This analysis was performed by pooling trials from all subjects into 6 clusters with increasing contextual bias (each containing an approximately equally large number of trials), and calculating the mean contextual bias and referential ambiguity effect for each cluster. This analysis revealed that the size of the referential ambiguity effect was strongly and significantly correlated with contextual bias (Pearson's $r=0.94$, $p=0.006$).

2.4. Combined effects of Reading Span and contextual bias

The grand average ERP for ambiguous and non-ambiguous pronouns in moderately and weakly biased sentences for the two Reading Span groups are displayed in Fig. 5 (electrode locations Fp1 and Fp2). The Ambiguity(2) × Bias(2) × Span-group

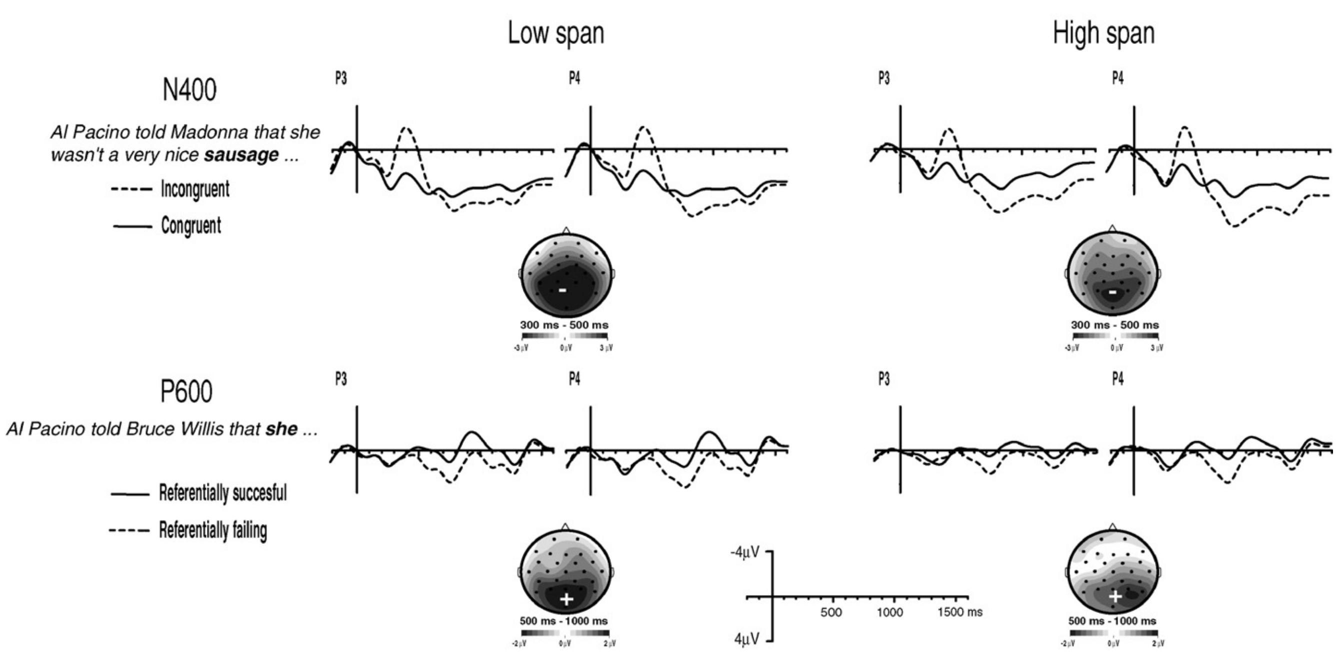

Fig. 6 – N400 effects of semantic incongruence (upper graph) and P600 effects of referential failure (lower graph) for each Reading Span group at electrode positions P3 and P4.

(2) three-way interaction was not significant ($F(1,28)=0.045$, $p=0.834$).

2.5. Reading span score and N400/P600 effects

ERP effects elicited by semantic incongruence and referential failure (i.e. semantically anomalous words and referentially failing pronouns, relative to their respective control words in the both referentially non-ambiguous and semantically congruent control condition) are displayed in Fig. 6 for both span groups (electrode locations P3 and P4). As can be seen from this figure, semantically incongruent words elicited an N400 effect relative to congruent control words in both span groups (upper graph), and the referential failing pronouns elicited a P600 effect relative to referential non-ambiguous pronouns in both groups (lower graph).

Using mean amplitude in the 300–500 ms window, the overall Congruence(2) × Electrode (30) ANOVA revealed a significant main effect of semantic congruence ($F(1,29)=68.27$, $p<0.001$). Using mean amplitude in the 500–1000 ms window, the overall Referential Failure(2) × Electrode(30) ANOVA revealed a significant main effect of referential failure ($F(1,29)=12.41$, $p=0.001$). In contrast to the ERP effect elicited by referentially ambiguous pronouns, Reading Span score did not correlate with the abovementioned N400 effect ($r=0.054$, $p=0.776$), nor with the P600 effect ($r=0.073$, $p=0.702$), taking into account only posterior channels where both effects were largest. Also, using mean amplitude difference at posterior channels, there was no significant Congruence(2) × Span Group (2) interaction ($F(1,28)=1.36$, $p=0.253$), nor a significant Referential Failure(2) × Span Group(2) interaction ($F(1,28)=0.367$, $p=0.550$).

3. Discussion

The goal of this ERP experiment was to explore how the resolution of formally ambiguous pronouns is affected by both individual differences in language processing skills and by sentence-level constraints on referential interpretation (contextual bias). Our results can be summarized as follows: (1) in line with earlier ERP studies on referential ambiguity (Van Berkum et al., 1999, 2003, 2004; Nieuwland et al., in press), formally referentially ambiguous pronouns elicited a sustained, frontal negative shift, reflecting the processing consequences of temporarily having two referential candidates for a single pronoun. (2) The amplitude of this effect was positively correlated with Reading Span score, suggesting that readers with better language processing skills are generally more sensitive to formal referential ambiguity than low skilled readers. (3) The referentially induced ERP effect was also dependent on the degree to which sentences biased readers towards one particular referential interpretation, suggesting that whether a reader experiences referential ambiguity depends on both individual reader skills and contextual bias. (4) The amplitude of two other ERP measures of language processing (N400, P600) did not correlate with Reading Span score, which could be taken to suggest that differences in language processing skills especially surface in linguistic manipulations that involve subtle constraints (like referential ambiguity) but not in manipulations that involve highly salient anomalies (in this case lexical-semantic anomalies and pronoun gender disagreements).

The present findings show that whether a reader perceives a formally ambiguous pronoun to be genuinely ambiguous depends on both reader characteristics and contextual bias. The likelihood that a reader responds differently to formally ambiguous pronouns relative to unambiguous pronouns increases with language processing skills as reflected by the Reading Span, and decreases with the strength with which sentence context biases towards one particular referential interpretation. If a reader indeed took into account both referential interpretations of formally ambiguous pronouns, the processing consequences of this perceived ambiguity are visible in the sustained, frontal negative shift as reported here.

Our findings provide unique ERP evidence for the existence of individual differences in referential-rather than syntactic or semantic-analysis, and for individual differences in pronoun resolution in particular. [2] Following a 'processing skill' account of individual differences in language comprehension (e.g., MacDonald and Christiansen, 2002), our results suggest that high span readers (i.e. readers with better language processing skills due to both experience and biological/genetic factors) are more sensitive to the alternative interpretations that are conveyed by formally referentially ambiguous pronouns. Because our results are consistent with findings from studies on other types of ambiguity processing (e.g., lexical or syntactic ambiguity), they provide convergent evidence for the notion that high span readers are more sensitive to alternative analyses at any level of language processing (e.g., MacDonald and Christiansen, 2002).

A caveat should be made concerning our conclusion that the absence of the referential ambiguity effect in low span readers indicates that they did not notice anything ambiguous about formally ambiguous pronouns. Of course, it is logically possible that low span readers did in fact notice the referential ambiguity, but that they were unable to pursue competing referential interpretations (due to limited processing resources) and only pursued the – to them, at that moment – most plausible interpretation. Such an account is perhaps most consistent with the more traditional 'capacity'

[2] Behavioral evidence for individual differences in pronoun resolution has already been reported by Long and De Ley (2000). Their results suggest, among other things, that only skilled readers show an early effect of implicit causality on pronoun resolution (using only sentences that strongly biased towards one referential interpretation). Perhaps the present findings (that contextual bias had no differential effect on the referential ambiguity effect in high span readers) can be said to extend the Long and De Ley results, in the sense that skilled readers might only show effects of contextual bias if it is very strong, but not necessarily so if it is weak or moderate. However, the Long and De Ley study and the present study differ in important ways that complicate a direct comparison of the results: First of all, Long and De Ley grouped their subjects into skilled and less skilled readers using the Nelson–Denny Reading Test, while we used the Reading Span task, obscuring the sub-group comparison across experiments. Furthermore, Long et al. used a probe-recognition task while we gave no task other than normal comprehension.

account of individual differences in language comprehension (e.g., Daneman and Carpenter, 1980; Just and Carpenter, 1992). However, given that ERPs are generally sensitive to even very subtle perturbations of linguistic processing (e.g., Hagoort and Brown, 1994; Kutas and Hillyard, 1984), one might expect that if low span readers did notice the ambiguity, the resulting struggle to simultaneously maintain two alternative interpretations would somehow show up in the ERP (e.g., as a referentially induced ERP effect of short duration, or perhaps a qualitatively different ERP effect). As this was not the case, the most parsimonious explanation for the present results, in our view, is that whereas high span readers generally noticed the ambiguity in both weakly and moderately biased sentences, low span readers did not notice anything ambiguous about formally ambiguous pronouns as they immediately took on the – to them, and at that moment – most plausible referential interpretation.

In general, the results from the present ERP experiment are consistent with the bulk of the behavioral literature on the Reading Span task and the differences between low and high span readers on measures of online comprehension and ambiguity resolution. Perhaps the most striking example of this correspondence is that, just as high span readers are also known to be more sensitive to both interpretations of lexically ambiguous words than low span readers (e.g., Miyake et al., 1994), high span readers were more sensitive to referential ambiguity in the current study. In fact, these effects may be quite similar, as it can be argued that recognizing the ambiguity posed by two different interpretations of the same lexically ambiguous word essentially poses a referential problem. Interestingly, lexically ambiguous nouns presented in a neutral sentence context elicit a sustained, frontal negative shift (Hagoort and Brown, 1994) that resembles the ERP effect as reported here. This is consistent with the idea that a lexically ambiguous word in a sufficiently neutral context can be said to have two possible referents.

Our results are not only consistent with behavioral studies on individual differences and lexical/syntactic ambiguity resolution, but also with other ERP studies addressing individual differences in language comprehension that have reported sustained, frontal negative shifts in the ERP. For example, readers with better scores on subsequent comprehension questions have been reported to show a sustained, frontal negativity while reading syntactically complex, object relative sentences (e.g., "The reporter who the senator harshly attacked admitted the error", but poor comprehenders do not or to a lesser extent (e.g., King and Kutas, 1995; Muller et al., 1997). Furthermore, sentences that describe events in reverse chronological order (e.g., "Before the scientist…) also elicit slow negative potentials at anterior sites relative to sentences that describe events in chronological order (e.g., After the scientist…."), and the size of this ERP effect is correlated with Reading Span (Munte et al., 1998). However, the ERP effect reported here was only slightly left-lateralized, while the slow potentials from these other studies were far more asymmetrical over the left and right hemispheres (except for the Müller et al. study). These differences may be indicative of a stronger right-hemisphere involvement in processing ambiguous pronouns. Which brain structures are actually involved in processing referential ambiguity remains to be seen and is currently investigated in our lab using functional magnetic resonance imaging techniques (fMRI).

Though the present results are consistent with a large number of behavioral and ERP studies, they also raise some new questions. For example, we have taken the referentially induced ERP effect to reflect the processing consequences of readers taking into account both referential interpretations of a formally ambiguous pronoun, but the nature of these processing consequences is as yet not entirely clear. For example, the ERP effect may reflect controlled, strategic processing in order to solve the ambiguity (e.g., extensive search in episodic memory for disambiguating information, or perhaps the controlled inhibition of irrelevant information[3]), but it may also reflect the active, simultaneous maintenance of competing referential interpretations, or even a more superficial reflection of the processing consequences of the noticing 'itself'.

On the one hand, it seems likely that readers, when faced with a genuinely ambiguous pronoun, do attempt to resolve this ambiguity someway or another. In fact, most subjects indicated that when they thought a pronoun could refer to either sentential character, they *chose* the most plausible character as the correct antecedent. In this sense, the correlation between the referentially induced ERP effect and Reading Span performance might be taken to reflect the relationship between controlled processing or executive attention and working memory performance (e.g., Engle, 2002). On the other hand, results from an earlier study (Nieuwland et al., in press) suggest that there is no straightforward, temporal correspondence between the sustained nature of the referentially induced ERP effect and the referential ambiguity as conveyed by the discourse. In this earlier study, the resolution of referential ambiguity (i.e., the moment readers encounter a truly disambiguating word) was not directly observable as an immediate disappearance of the ERP effect, as one would expect if this effect reflected either controlled 'problem solving' or the active maintenance of competing interpretations (though it did ultimately

[3] Inhibition or suppression of irrelevant information is often regarded as a crucial component of comprehension (e.g., Gernsbacher and Faust, 1991), and this concept has also been used to account for ERP differences between high and low span readers. An inhibition account that is often advocated is that low span readers cannot effectively inhibit irrelevant information or dispreferred readings and therefore pursue multiple syntactic representations (e.g., Bornkessel et al., 2004; Friederici et al., 1998) or multiple word meanings (e.g., Gunter et al., 2003; but see Wagner and Gunter, 2004). However, the present results are difficult to reconcile with such an account, as low span readers were not the ones to show ERP signs of referential ambiguity. Given our results, it could be argued that low span readers were unable to inhibit the contextual cues or heuristic preferences that pointed to one particular referential interpretation, hence unable to inhibit the *preferred* reading. With respect to the functional interpretation of the referentially induced ERP effect, the possibility exists that it reflects controlled inhibition of one of the possible referential interpretations (e.g., Barrett et al., 2004), exerted after a referentially ambiguous word has actually been perceived as such.

disappear, suggesting that letting go of additional processing demands may be a more gradual process than imposing these demands; see also Kutas, 1997, for examples of comparable, sustained ERP effects that outlast syntactic complexity manipulations). Therefore, we cannot simply take the sustained nature of the present ERP effect as conclusive evidence that readers were either engaged in controlled problem-solving or pursuing competing referential interpretations during the full intervals as reported here[4], and further research is warranted to elucidate these issues.

Another interesting issue is raised by the fact that the referentially induced ERP effect was correlated with Reading Span score, while two other, well-known ERP effects (N400, P600) were not. So far, we have taken these results to suggest that differences in language processing skills especially surface in linguistic manipulations that involve subtle constraints but not in manipulations that involve highly salient anomalies. However, this does not mean that the N400 and P600 and the cognitive processes they reflect are totally insensitive to linguistic competency. For one thing, both ERPs are modulated by the amount of exposure of second-language learners to the to-be-learned language (e.g., McLaughlin et al., 2004; Osterhout et al., 2004). With regard to Reading Span performance, some studies have reported N400 or P600 amplitude differences between high and low span readers (e.g., Friederici et al., 1998; Gunter et al., 2003; St. George et al., 1997; Van Petten et al., 1997), while others have reported qualitatively different ERP responses from high and low span readers altogether (e.g., Bornkessel et al., 2004), or even differential effects of experimental filler materials on ERP responses from high and low span readers (e.g., Wagner and Gunter, 2004). These apparently intricate interactions between individual differences in linguistic competency or preferences on the one hand, and the type and complexity of linguistic materials on the other, complicate the correspondence between experimenter-defined levels of linguistic processing and the ERPs that are obtained with certain manipulations (see also Van Berkum et al., in press, for discussion). In any case, the absence of any N400 and P600 differences between high and low span readers as reported here may very well be limited to the highly salient and relatively unequivocal nature of the anomalies that were presented.

To conclude, we have reported ERP evidence that the likelihood that a reader will take into account different referential interpretations of a formally ambiguous pronoun increases with the reader's Reading Span, and decreases with the strength of the contextual bias towards one particular referential interpretation. As such, our results show that whether a reader perceives a formally ambiguous pronoun to be genuinely referentially ambiguous is subtly co-determined by both individual language processing skills and contextual bias.

[4] Unfortunately, the present design did not allow us to examine the effects of subsequent *resolution* of referential ambiguity, because most of our sentences did not contain any conclusive cues to which referential interpretation was indeed the most correct one.

4. Experimental procedures

4.1. Subjects

After giving informed consent, thirty-two right-handed college students (11 males, mean age 20.4) participated in this study for course credit. All participants were native speakers of Dutch, without any neurological impairment, and had not participated in the referential cloze task (see below).

4.2. Reading span task

A computerized, new Dutch version of the reading span task (Van den Noort et al., 2006) was used to measure verbal working memory performance. This version consisted of five sets of 20 sentences, matched for sentence-length (number of syllables) and matched for the number of letters, number of syllables and frequency of the final word. The sentences were presented in different set sizes (2, 3, 4, 5 or 6 sentences), in random order (minimizing set-size anticipatory effects) and were read aloud by the subjects from a computer screen (minimizing experimenter effects).When a subject had finished a sentence, he/she immediately pressed the space bar triggering the onset of the next sentence. If the subject could not finish the sentence in 6.5 s, the next sentence was automatically presented. When a subject had completed all the sentences of a set, a recall-cue was presented and he/she had to recall the final words of the sentences from that set, while the experimenter registered and scored the responses. Subjects were instructed to read for comprehension with a normal pace (though encouraged to read faster if they were not able to read the sentences in 6.5 s). Reading Span score was computed as the total number of final words that were correctly recalled (note that this scoring method is less vulnerable to within-task performance variability than the original Reading Span scoring method, see also).

4.3. Materials

We created 270 Dutch sentences describing the interaction between two characters. The one-referent condition always contained one male and one female. In the two-referent condition, both characters were either male or both female. The two characters were denoted by proper names (e.g., 'John/David/Mary/Lisa'), by definite noun phrases (e.g., 'the father/son/empress/queen') or they were celebrities (e.g., 'George Bush/Bill Clinton/Madonna/Jennifer Lopez'). The two conditions were matched on frequency and length of the words that denoted the story characters (and the celebrities were matched on frequency using Google-hits on Dutch websites). Every sentence also contained a pronoun ('he/she/his/her',[5]) that could formally be taken to

[5] The Dutch equivalent of 'she' is 'zij', which can also mean the plural pronoun 'they'. To avoid this ambiguity, sentences containing the pronoun 'zij' were constructed so that a corresponding singular verb always preceded the pronoun and unambiguously constrained its interpretation to a singular 'she', e.g., "Toen Marie Lisa opzocht keek zij…").

refer to either of the two characters mentioned in the two-referent condition. We used a range of different themes for constructing the sentences to increase the content variability of our materials (see Table 1 for examples; a full set of example materials can be obtained from http://users.fmg.uva.nl/mnieuwland/NieuwlandVanBerkum_Pronoun-exp_Example_stimuli.pdf).

The 270 items were pre-tested using a written, referential cloze task: all two-referent sentences were truncated after the pronoun, pseudo-randomly mixed into six different versions. Thirty-two subjects were asked to complete every sentence with the first sentence-ending that came to mind (while keeping this continuation as simple and logical as possible), and subsequently circle the character they had in mind while completing the sentence (i.e., the antecedent of the pronoun).

The rationale for using this task was that when readers come up with the most logical sentence continuation, they have already made a referential commitment, and what they write down reflects and confirms this commitment. Note that this pretest is more natural than asking subjects to directly point out the referent (encircling without or preceding completion), and its usefulness and sensitivity have already been demonstrated in other studies (e.g., Stevenson et al., 2000; although in the Stevenson et al. study, the referential interpretation was scored by experimenters afterwards).

For each item, contextual bias was computed as the absolute percentage-difference between the percentage of people that completed it with the first mentioned character in mind and the percentage of people that had the second story character in mind.[6] Results from our completion task showed that the contextual bias of our pre-test items ranged from 0% to 100%, from which we selected 240 items (*min*=0%, max=61%, *M*=22%, SD=17%). For each item, a one-referent version was created by replacing one male character with a female character or vice versa. Another two versions were created for each item to address the differential role of language processing skills in processing different types of anomalies/ambiguities: a lexical-semantic anomaly condition by replacing one word from the one-referent version with a lexical-semantic anomalous word (e.g., "Paula always greets Josh when he enters the quiet umbrella/study hall."), and a zero-referent condition by replacing both male characters with female characters or vice versa (effectively leaving the pronoun without a directly available, eligible antecedent, e.g., "Jennifer Lopez screamed at Madonna because he …"). The lexical-semantic anomalous words were matched to their control counterparts for length and frequency.

[6] The results from this completion task showed that some subjects encircled the first and second characters equally often, while some tended to encircle the first character more often than the second character or vice versa. For example, one subject had encircled the first character in 66% of the pre-test items, while another subject did so in only 41% of the items. All other subjects scored within this range. We are aware that these differences may be related to individual differences in language comprehension skills. However, because no other task than the completion task was administered to these subjects, we were not able to investigate whether differences in sentence completion were a function of individual language comprehension skills.

Each subject was presented 60 items from each condition, evenly distributed across the contextual bias spectrum of the corresponding two-referent sentences. To counteract the transparency of presenting 240 sentences containing pronouns, we included 120 fillers sentences that also described the interaction between two characters but that did not contain any pronouns.

4.4. *Procedure*

After having completed the Reading Span task, participants were seated in front of a display monitor, and were informed that they would be reading sentences word-by-word from the middle of the screen. They were instructed to read for comprehension only, and to minimize movement. No additional task demands were imposed.

Four trial lists were used. For the first list, 60 items from each condition were pseudorandomly mixed with 120 filler sentences such that no trial type occurred more than three times consecutively and trials of each type were matched on average list position. The other lists were derived from the first by rotating the trial types. The total of 360 sentences was divided in five blocks, separated by a pause.

To parallel natural reading times, all words (except from the word that directly preceded the critical pronouns onwards) were presented using a variable serial visual presentation procedure, which was based on natural reading times (Haberlandt and Graesser, 1985; Legge et al., 1997) and a subjective assessment of the naturalness of the resulting presentation times. In principle, word duration in ms was computed as ((number of letters × 30) + 190), with a maximum of 430 ms. However, to avoid spurious ERP effects due to differences in word length, we switched to non-variable presentation from the word that preceded the critical pronoun onwards (using a word duration of 350 ms, with an exception for final words, which were presented 1000 ms). Importantly, participants did not notice this alternation between variable and fixed word duration presentation within sentences. All inter-word-intervals were 150 ms. After every final word, a blank screen was presented for 2 s, which was subsequently followed by a fixation-mark. At the onset of this fixation-mark, subjects could start the new sentence themselves using a right button-press. Total time-on-task was approximately 70 min.

4.5. *EEG recording*

The EEG was recorded from 30 standard scalp locations (Fz, Cz, Pz, Oz, Fp1/2, F3/4, F7/8, F9/10, FC1/2, FC5/6, FT9/10, C3/4, T7/8, CP1/2, CP5/6, P3/4, P7/8), amplified (band-pass filtered at 0.03 Hz–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 (see Jung et al., 2000; Makeig et al., 1997). Then, epochs that ranged from −500 ms to +1600 ms relative to critical word onset were extracted and normalized (by subtraction) to a 150 ms pre-onset 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. Two participants were excluded, leaving an average segment loss of 9% across the remaining 30 participants.

Acknowledgments

We thank Marte Otten, Petra van Alphen, Marcus Spaan, Bert Molenkamp and Geert-Jan Mertens for their help. This research was supported by an NWO Innovation Impulse Vidi grant to JVB.

REFERENCES

Arnold, J.E., Eisenband, J.G., Brown Schmidt, S., Trueswell, J.C., 2000. The rapid use of gender information: evidence of the time course for pronoun resolution from eyetracking. Cognition 76 (1), B13–B26.

Barrett, L.F., Tugade, M.M., Engle, R.W., 2004. Individual differences in working memory capacity and dual-process theories of the mind. Psychol. Bull. 130, 553–573.

Bornkessel, I.D., Fiebach, C.J., Friederici, A.D., 2004. On the cost of syntactic ambiguity in human language comprehension: an individual differences approach. Cogn. Brain Res. 21, 11–21.

Calvo, M.G., 2001. Working memory and inferences: evidence from eye fixations during reading. Memory 9 (4–6), 365–381.

Caplan, D., Waters, G.S., 1999. Verbal working memory and sentence comprehension. Behav. Brain Sci. 22 (1), 77–126.

Clark, H.H., Sengul, C.J., 1979. In search of referents for nouns and pronouns. Mem. Cogn. 7 (1), 35–41.

Crawley, R.A., Stevenson, R.J., Kleinman, D., 1990. The use of heuristic strategies in the interpretation of pronouns. J. Psycholinguist. Res. 19 (4), 245–264.

Daneman, M., Carpenter, P.A., 1980. Individual differences in working memory and reading. J. Verbal Learn. Verbal Behav. 19 (4), 450–466.

Engle, R.W., 2002. Working memory capacity as executive attention. Curr. Dir. Psychol. Sci. 11, 19–23.

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

Garnham, A., 2001. Mental Models and the Interpretation of Anaphora. Taylor and Francis, Philadelphia PA.

Gernsbacher, M.A., 1989. Mechanisms that improve referential access. Cognition 32, 99–156.

Gernsbacher, M.A., Faust, M.E., 1991. The mechanism of suppression—A component of general comprehension skill. J. Exper. Psychol., Learn., Mem., Cogn. 17, 245–262.

Gernsbacher, M.A., Hargreaves, D.J., 1988. Accessing sentence participants: the advantage of first mention. J. Mem. Lang. 27 (6), 699–717.

Gordon, P.C., Grosz, B.J., Gilliom, L.A., 1993. Pronouns, names, and the centering of attention in discourse. Cogn. Sci. 17 (3), 311–347.

Greene, S.B., McKoon, G., Ratcliff, R., 1992. Pronoun resolution and discourse models. J. Exper. Psychol., Learn., Mem., Cogn. 18 (2), 266–283.

Gunter, T.C., Wagner, S., Friederici, A.D., 2003. Working memory and lexical ambiguity resolution as revealed by ERPs: a difficult case for activation theories. J. Cogn. Neurosci. 15, 643–657.

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., Brown, C.M., 1994. Brain responses to lexical ambiguity resolution and parsing. In: Clifton, C., Frazier, L., Rayner, K. (Eds.), Perspectives on Sentence Processing. Lawrence Erlbaum Associates, Hillsdale, NJ.

Jarvikivi, J., van Gompel, R.P.G., Hyona, J., Bertram, R., 2005. Ambiguous pronoun resolution: contrasting the first-mention and subject-preference accounts. Psychol. Sci. 16 (4), 260–264.

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 (10), 1745–1758.

Just, M.A., Carpenter, P.A., 1992. A capacity theory of comprehension: individual differences in working memory. Psychol. Rev. 99 (1), 122–149.

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.

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.

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 (4427), 203–205.

Kutas, M., Hillyard, S.A., 1984. Brain potentials during reading reflect word expectancy and semantic association. Nature 307, 161–163.

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.

Linderholm, T., 2002. Predictive inference generation as a function of working memory capacity and causal text constraints. Discourse Processes 34 (3), 259–280.

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 (4), 526–570.

MacDonald, M.C., Christiansen, M.H., 2002. Reassessing working memory: comment on Just and Carpenter (1992) and Waters and Caplan (1996). Psychol. Rev. 109 (1), 35–54.

MacDonald, M.C., MacWhinney, B., 1990. Measuring inhibition and facilitation from pronouns. J. Mem. Lang. 29, 469–492.

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, 10979–10984.

McDonald, J.L., MacWhinney, B., 1995. The time-course of anaphor resolution—Effects of implicit verb causality and gender. J. Mem. Lang. 34, 543–566.

McKoon, G., Ward, G., Ratcliff, R., Sproat, R., 1993. Morphosyntactic and pragmatic factors affecting the accessibility of discourse entities. J. Mem. Lang. 32 (1), 56–75.

McLaughlin, J., Osterhout, L., Kim, A., 2004. Neural correlates of second-language word learning: minimal instruction produces rapid change. Nat. Neurosci. 7, 703–704.

Miyake, A., Just, M.A., Carpenter, P.A., 1994. Working memory constraints on the resolution of lexical ambiguity: maintaining multiple interpretations in neutral contexts. J. Mem. Lang. 33 (2), 175–202.

Muller, H.M., King, J.W., Kutas, M., 1997. Event-related potentials elicited by spoken relative clauses. Cogn. Brain Res. 5 (3), 193–203.

Munte, T.F., Schiltz, K., Kutas, M., 1998. When temporal terms belie conceptual order. Nature 395 (6697), 71–73.

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

Osterhout, L., Mobley, L.A., 1995. Event-related brain potentials elicited by failure to agree. J. Mem. Lang. 34 (6), 739–773.
Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., Inoue, K., 2004. Sentences in the brain: event-related potentials as real-time reflections of sentence comprehension and language learning. In: Carreiras, M., CliftonJr. Jr., C. (Eds.), The On-Line Study of Sentence Comprehension: Eyetracking, ERP, and beyond. Psychology Press.
Sanford, A.J., Garrod, S.C., 1989. What, when, and how? Questions of immediacy in anaphoric reference resolution. Lang. Cogn. Processes 4 (3–4), SI235–SI262.
Schmitt, B.M., Lamers, M., Munte, T.F., 2002. Electrophysiological estimates of biological and syntactic gender violation during pronoun processing. Cogn. Brain Res. 14, 333–346.
Stevenson, R., Knott, A., Oberlander, J., McDonald, S., 2000. Interpreting pronouns and connectives: interactions among focusing, thematic roles and coherence relations. Lang. Cogn. Processes 15, 225–262.
St. George, M., Mannes, S., Hoffman, J.E., 1997. Individual differences in inference generation: an ERP analysis. J. Cogn. Neurosci. 9 (6), 776–787.
Tyler, L.K., Marslen Wilson, W.D., 1982. Conjectures and refutations: a reply to Norris. Cognition 11 (1), 103–107.
Van Berkum, J.J.A., Brown, C.M., Hagoort, P., 1999. 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., Zwitserlood, P., 2003. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. Psychophysiology 40 (2), 235–248.
Van Berkum, J.J.A., Koornneef, A.W., Otten, M., Nieuwland, M.S., in press. Establishing reference in language comprehension: An electrophysiological perspective. Brain Res.
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 den Noort, M.W.M.L., Bosch, M.P.C., Haverkort, M., Hugdahl, K., 2006. Is there a decline in verbal working memory over age? A study with the new standard computerized reading span test. Eur. J. Psychol. (May).
Van Petten, C., Weckerly, J., McIsaac, H.K., Kutas, M., 1997. Working memory capacity dissociates lexical and sentential context effects. Psychol. Sci. 8, 238–242.
Vos, S.H., Gunter, T.C., Kolk, H.H., Mulder, G., 2001. Working memory constraints on syntactic processing: an electrophysiological investigation. Psychophysiology 38, 41–63.
Wagner, S., Gunter, T.C., 2004. Determining inhibition—Individual differences in the 'lexicon context' trade-off during lexical ambiguity resolution in working memory. Exp. Psychol. 51, 290–299.
Waters, G.S., Caplan, D., 1996. The measurement of verbal working memory capacity and its relation to reading comprehension. Q. J. Exp. Psychol. A Human Exp. Psychol. 1, 51–79.
Whitney, P., Ritchie, B.G., Clark, M.B., 1991. Working-memory capacity and the use of elaborative inferences in text comprehension. Discourse Processes 14 (2), 133–145.