BRAIN RESEARCH 1311 (2010) 93–109

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

The neural underpinnings of semantic ambiguity and anaphora

*Veena D. Dwivedi**, *Natalie A. Phillips*, *Stephanie Einagel*, *Shari R. Baum*

Department of Applied Linguistics, Brock University, 500 Glenridge Avenue Street, Catharines, Ontario, Canada L2S 3A1

ARTICLE INFO

Article history:
Accepted 26 September 2009
Available online 23 October 2009

Keywords:
ERPs
Slow negative shift
P600
N400
Quantifier ambiguity
Anaphora

ABSTRACT

We used event-related brain potentials (ERPs) in order to investigate how definite NP anaphors are integrated into semantically ambiguous contexts. Although sentences such as *Every kid climbed a tree* lack any syntactic or lexical ambiguity, these structures exhibit two possible meanings, where either many trees or only one tree was climbed. This semantic ambiguity is the result of quantifier scope ambiguity. Previous behavioural studies have shown that a plural definite NP continuation is preferred (as reflected in a continuation sentence, e.g., *The trees were in the park*) over singular NPs (e.g., *The tree was in the park*). This study aimed to identify the neurophysiological pattern associated with the integration of the continuation sentences, as well as the time course of this process. We examined ERPs elicited by the noun and verb in continuation sentences following ambiguous and unambiguous context sentences. A sustained negative shift was most evident at the Verb position in sentences exhibiting scope ambiguity. Furthermore, this waveform did not differentiate itself until 900 ms after the presentation of the Noun, suggesting that the parser waits to assign meaning in contexts exhibiting quantifier scope ambiguity, such that such contexts are left as underspecified representations.

1. Introduction

An important task in communication is the ability to keep track of what is being talked about. For example, if I utter the sentence, "*The tree in my backyard is beautiful*," the listener would need to know whether I am referring to the willow tree or the Japanese maple tree discussed earlier in our conversation. Presumably, the structure of the previous discourse would have an impact on the listener's first guess (Haviland and Clark, 1974). Thus, understanding how perceivers interpret definite noun phrases (e.g., *the tree*, *the girl*, etc.) in context is a key ingredient to building a model of on-line discourse comprehension.

The goal of the present study was to investigate the interpretation of definite NPs in contexts that are semantically ambiguous. Unlike previous works that have examined semantic ambiguity from a lexical perspective (e.g., whether *bank* is a financial institution or the side of a river) the present work defined semantic ambiguity using a grammatical construct, that of scope ambiguity. 'Scope ambiguity' results for sentences that contain more than one quantifier, such as *all*, *every*, *one*, *a*, and *some*. For example, sentences such as *Every kid climbed a tree* are ambiguous, despite the fact that they lack any syntactic or lexical ambiguity. The different meanings are the result of different logical orders in which the quantifiers are interpreted. On one interpretation, it is the case that for *every* (∀) child, a (∃) tree was climbed, which results in an inference that several trees were climbed. This reading is called the 'surface scope' reading, since the order of interpre-

* *Corresponding author*. Fax: +1 905 688 2360.
E-mail address: veena.dwivedi@brocku.ca (V.D. Dwivedi).

0006-8993/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2009.09.102

tation of the quantifiers matches the surface linear order of the quantifiers in the sentence. On another reading, called the inverse scope reading, the interpretation is that it is the case that there is a (∃) tree, such that every (∀) kid climbed it. The inverse scope reading results in a meaning where just one tree was climbed.

The first published work to investigate the comprehension of scope ambiguous sentences was by Kurtzman and MacDonald (1993), who conducted a computer-controlled acceptability judgment task. Participants read whole sentences containing quantifier noun phrases (NPs) such as "*Every kid climbed a tree*" which were then followed by another possible continuation sentence, either "*The trees were full of apples*" (consistent with the surface scope meaning) or "*The tree was full of apples*" (consistent with the inverse scope meaning). At the end of either continuation sentence, participants were asked to judge whether it formed a good continuation of the first (ambiguous) context sentence or not. Furthermore, the experiment included unambiguous control context sentences such as "*Every kid climbed a different tree. The trees were…*" and "*Every kid climbed the same tree. The tree was…*" One of the many findings of that seminal work was that, overall, participants preferred a plural continuation for ambiguous sentences (77% of the time; the corresponding plural and singular continuations that followed the unambiguous control contexts were judged at rates of approximately 85% each; see also Kemtes and Kemper, 1999). Thus, there was a facilitation effect for plural NP anaphors after such scope ambiguous sentences.

We decided to further investigate interpretive processes at the syntax–semantics interface by building on the findings of Kurtzman and MacDonald (1993). Whereas their results regarding the preferred interpretation are clear (and expected on theoretical grounds, see Dwivedi, 1996), we note that the task required participants' conscious judgments measured at the end of the continuation sentences. Although such judgments are useful as a guiding measure to understanding the processing of such sentences, this method has two drawbacks. First, it calls upon a meta-linguistic assessment of the stimuli and as such is not reflective of real-time unconscious processing. Second, these judgments were taken only after the entire sentence had been read; in other words, it is unclear whether there were earlier decisions regarding parsing that were entertained and then discarded, or whether the ultimate judgment made was the only grammatical choice considered. In addition, in that study, reaction time measures were not included at all. Furthermore, recent behavioural findings indicate that these effects have not been fully replicated (see Tunstall, 1998; Filik et al., 2004; Paterson et al., 2008; as well as Anderson, 2004). One potential reason why findings have been equivocal is that the above-mentioned studies examined several linguistic factors simultaneously—e.g., type of verb phrase, type of verb, type of quantifier, and order of quantifiers.

The present work seeks to address these issues in the following way: first, conscious judgments regarding the sentences in question were not required; in this way we hoped to understand the natural processing of such stimuli in real-time. Furthermore, design of the stimuli was limited to one syntactic structure, using a specific order of two quantifiers, as well as using a particular kind of verb (see below). We hoped that this more constrained design would yield less equivocal findings. Furthermore, our goal was to chart the time course of interpretation in the second disambiguating sentence, using the time resolution afforded by event-related brain potentials (ERPs). ERPs reflect voltage changes in the electrical brain activity associated with cognitive processing. This methodology is particularly useful for our purposes because it allows us to examine the processing of language stimuli on-line with very high temporal resolution (on the order of milliseconds) and adequate spatial resolution (through scalp distribution). More importantly, there are several ERP components (reviewed below) that are specifically associated with distinct aspects of lexical-semantic and syntactic processing. Thus, the nature of the ERP components elicited might provide a qualitative understanding of the nature of the linguistic processing undertaken—that is, whether semantic, syntactic or other processing mechanisms are recruited during the comprehension of the stimuli.

At present, there are at least two possibilities afforded by Kurtzman and MacDonald's findings: either they are correct regarding the surface scope preference for "every" resulting in a preference for a plural definite NP continuation, or not. If it is the former, then the corollary would be (i) the singular NP continuation should not be preferred. If, on the other hand, they are incorrect, then (ii) a preference for the singular NP continuation is a possibility instead (see Filik et al., 2004; Paterson et al., 2008). A third possible prediction would be that (iii) there might be no preference for either the plural or singular continuation.

In order to assess predictions (i)–(iii), we created two-sentence discourses where the first (context) sentence displayed quantifier scope ambiguity, and the continuation sentence began with either a plural or singular definite NP (note again that the plural reading is consistent with the surface scope reading whereas the singular marking corresponds to the inverse scope reading of the context sentence). A control condition, exactly analogous to Kurtzman and MacDonald's paradigm, was constructed in order to ensure that the effects obtained were indeed due to context, and not due to the fact that two different kinds of nouns (plural vs. singular) were being compared. Continuation sentences were preceded by two different kinds of contexts: Ambiguous and Control. Thus, this within-subjects study was defined by two independent variables: type of context (Ambiguous (A) or Control (C)) and type of continuation sentence (Plural (P) or Singular (S)), and measurements occurred at the Noun (N), Verb (V), and Verb + 1 (V+1) position. Table 1 lists the four conditions explicitly.

Regarding prediction (i), if the findings of Kurtzman and MacDonald (1993) are on track, then the preferred interpretation of the ambiguous context sentence should be the plural continuation sentence. Thus, there should be no empirical difference between this condition (Ambiguous–Plural; AP) and the Control–Plural (CP) condition. Given that the singular continuation is hypothesized to be the non-preferred condition, then the Ambiguous–Singular (AS) condition should differ empirically from the Control–Singular (CS). If prediction (ii) is correct, then the reverse situation should hold: there should be no empirical difference between AS and CS but instead between AP and CP. If instead there is no preference for either continuation (prediction (iii)), then conditions AS and AP should pattern together.

Table 1 – Experimental sample stimuli.

Type of context sentence	Type of continuation sentence
Ambiguous:	Plural
Every kid climbed a tree.	The trees were in the park.
	N V V+1 *
Control	Plural
Every kid climbed a different tree.	The trees were in the park.
	N V V+1
Ambiguous:	Singular
Every kid climbed a tree.	The tree was in the park.
	N V V+1
Control	Singular
Every kid climbed the same tree.	The tree was in the park.
	N V V+1

* ERPs were recorded at the Continuation sentences only, at the following positions: N = Noun position; V = Verb position; and V+1 = after Verb position.

Next, we discuss the possible neurophysiological form of these empirical patterns. First, the relevant ERP component associated with predictions (i) and (ii), which indicate a revision of the preferred scope assignment, would be either a P600 or an N400 effect. For example, regarding prediction (i), once the (non-preferred) singular "tree" is perceived, the processor would have to revise its initial interpretation of the scope ambiguous context sentence so that inverse scope would be assigned rather than the preferred surface scope interpretation. The ERP correlate of this revision could be a P600 effect, especially given recent conceptions of this component. Traditionally, the P600 has been conceived as a component that is elicited by structural aspects of linguistic input (Hagoort et al., 1993; Osterhout et al., 1994); however, recently it has been related to the processes of revision and repair in sentence processing. Kaan and Swaab (2003a,b) argue that the P600 actually represents a family of components distributed across the scalp (c.f., Hagoort et al., 1999; Frisch et al., 2002). P600 activity with a posterior distribution appears to index syntactic processing difficulty, whereas P600 activity with a frontal distribution is related to ambiguity resolution and/or an increase in discourse level complexity. That is, frontal P600 activity has been claimed to signal that a preferred structural analysis can no longer be maintained and must be revised. In Dwivedi et al. (2006), we found such a waveform in response to continuation sentences that were inconsistent with quantificational mood (i.e., where modal auxiliaries serve as quantifiers over events/situations; see Kratzer, 1979; Heim, 1982; Dwivedi, 1996) as defined in the previous context sentence. We interpreted this finding as possibly reflecting the cognitive process of revision of previous linguistic semantic structure. In the current experiment, if readers assign a preferential scope assignment (either predictions (i) or (ii) above) that later has to be revised, then a frontal P600 effect could be the result.[1]

[1] For a nice overview of the status of the P600, see Kuperberg (2007), where another debate surrounding positive-going waveforms is mentioned, that between the status of late positivities and the P600. Since that debate is tangential to the study at hand (see results below), we do not discuss it in detail beyond this footnote.

The N400 component has been associated with semantic congruency, where this congruency largely has to do with lexico-semantic fit (Kutas and Hillyard, 1980, 1983; St. George et al., 1997; van Berkum et al., 1999a,b; Hoeks et al., 2004, among others). To date, investigations of the N400 have largely focused on such conceptual semantic associations such as *John spread the warm bread with socks/butter*. In contrast, the present study examines a possible algorithmic computation of meaning from a compositional semantic point of view. That being said, given the fact that the purported revision would in fact be a revision in meaning, at the level of context, an N400 effect could emerge, too, either in isolation or in addition to the P600 effect (Friederici and Frisch, 2000).

Predictions (i) and (ii) rest on the assumption that perceivers assign a preferred interpretation to the initially scope ambiguous context sentences. In contrast, prediction (iii) posits no immediately preferred reading and assumes that scope is left unassigned in context sentences, thus effectively creating an underspecified context. If we think about how "the tree(s)" would be integrated into such a context, then yet another ERP waveform could be predicted, a slow negative shift. In several ERP language experiments, a slow negative shift has been interpreted as a marker of extensive use of working memory resources (e.g., Ruchkin et al., 1988; Fiebach et al., 1996, 2001; Kluender and Kutas, 1993; Kutas, 1997; Müller et al., 1997; Münte et al., 1998; Rösler et al., 1997, 1998). The present claim is that integrating a definite NP into a context that is ambiguous would result in a slow negative-going ERP component. This is due to the fact that the reference of the definite NP would be ambiguous in such a context. Recently, van Berkum et al. (1999a, 2003, 2007) have examined the question of semantic ambiguity using ERPs in both visual and auditory modalities. Those studies set up (Dutch) story contexts where the number of candidate referents for a definite NP was manipulated, such that, for example, there was either one candidate (e.g., "girl") or two mentioned in a story. As a result, reference for the definite NP was ambiguous in the two-candidate vs. one-candidate context. After the story line was established, measurements were then taken at the critical continuation sentence "David told *the girl* that…" Results indicated a frontal slow negative shift emerged in continuation sentences 280 ms after "the girl" when the previous discourse context provided two possible referents in comparison to measurements taken at that NP when the previous context provided just one referent. Thus, van Berkum et al. claim that the possibility of there being two candidate referents for the NP *the girl* was translated into the cost of either maintaining the two previously mentioned NPs in memory, or the increased search requirements for resolving the reference of *the girl*. Since that study and the present study examine the integration of referentially ambiguous definite NPs in context (prediction iii), functionally, we might expect to see a negative-going ERP component, the Nref. However, since in the present experiment, the nature of the discourse context, as well as the ambiguity at hand is completely different (such that in addition to a more complex search space, a further semantic scope computation would have to be incurred), it

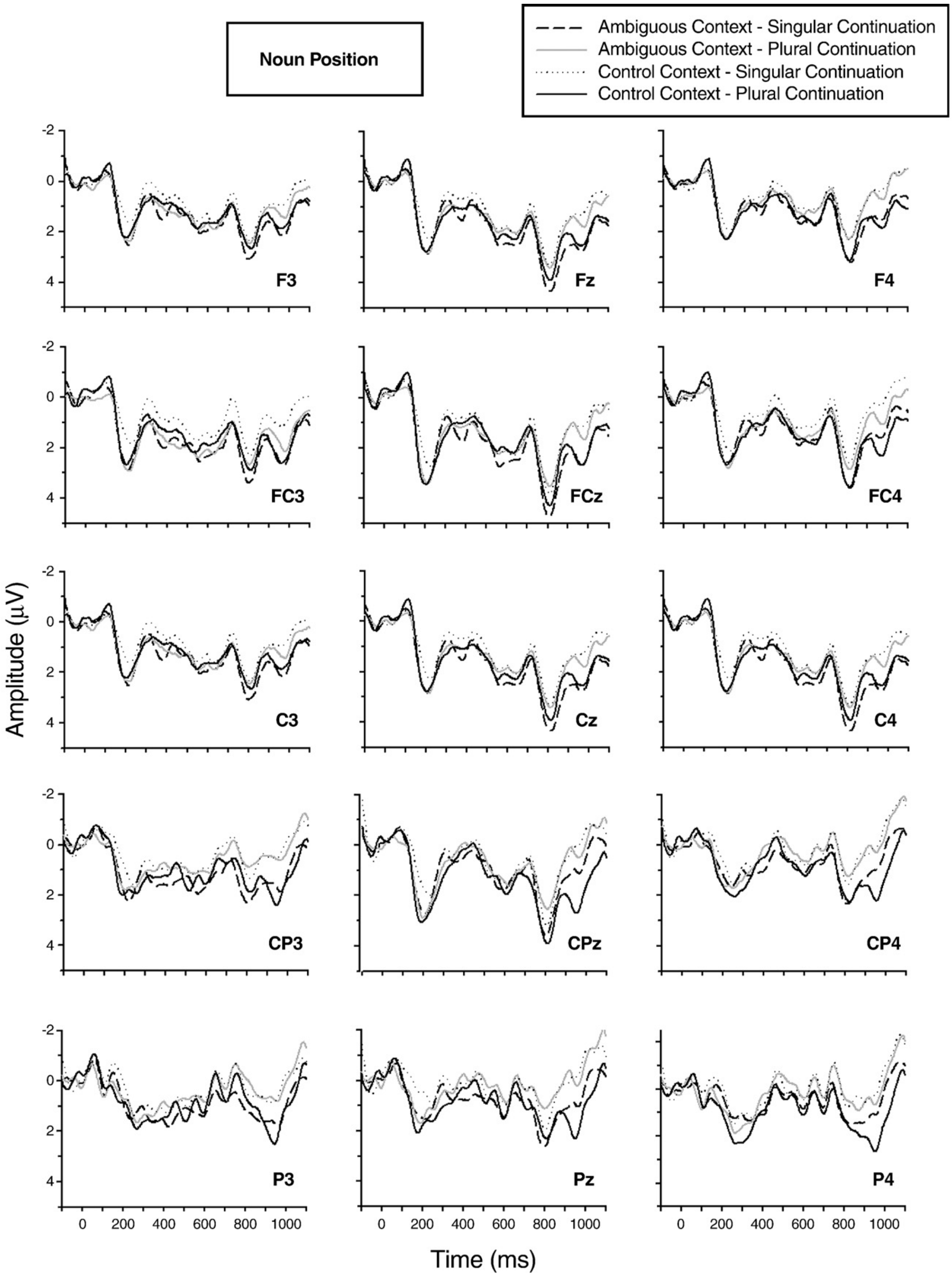

Fig. 1 – ERP recordings at Noun position.

is unclear just how similar the topography or the timing of the purported negative-going waveform might be.

Thus, for the purposes of prediction (iii), if the processor does not fully interpret the context sentence "Every kid climbed a tree," such that scope interpretation is left unresolved, an ambiguous context, especially regarding later reference, would result. This is because a unique referent for the upcoming definite NP is not provided by the context. As such, when the definite NP "the tree(s)" occurs in the critical continuation sentence, a slow negative-going wave (possibly an Nref) would be the expected ERP waveform.

In sum, this study seeks to answer at least the following two questions. First, what is the neurophysiological signature associated with integrating a definite NP into a context that is defined by scope ambiguity, and how can that inform a theory of processing such sentences? Furthermore, what is the time course of this process, that is, at what point in the sentence does integration begin to occur with respect to the previous context?

2. Results

2.1. *Electrophysiological analyses*

All statistical analyses reported below concern the ERP waveforms recorded at the onset of the Noun, Verb and Verb +1 positions. We expected that integrating the definite NP into the context would occur at the Noun or Verb position. We also analyzed effects at the Verb + 1 position in order to assess if the integration effects would be apparent near the region where integration was expected to occur.

Separate repeated measures ANOVAs were conducted for the midline and medial-lateral electrodes. The factors included were Context (2 levels: Ambiguous vs. Control), Number (2 levels: Plural vs. Singular), Time interval (4 levels: 300–500, 500–700, 700–900, 900–1100 ms), and Electrode site (5 levels: Fz, FCz, Cz, CPz, Pz) for the midline sites; for medial-lateral electrode sites, the Electrode factor was defined as anterior-to-posterior electrode sites (5 levels: F3/4, FC3/4, C3/4, CP3/4, P3/4) and Hemisphere (2 levels: left vs. right).

The ERP analyses reported below used SPSS v.11.0 statistical software and employed the Greenhouse and Geisser (1959) non-sphericity correction for effects with more than one degree of freedom in the numerator. Following convention, unadjusted degrees of freedom are reported, along with the Greenhouse-Geisser epsilon value (ε) and adjusted p-value. Mean square error values reported are those corresponding to the Greenhouse-Geisser correction. All significant main effects are reported first, followed by the highest order interaction effects involving Context and/or Number. Unless otherwise stated, interactions were further assessed using simple effects analyses with alpha=.05. For example, a Context × Number interaction was decomposed by examining simple effects of Context at each level of Number (i.e., AP vs. AS; CP vs. CS), and then by examining the Number factor at each level of Context (i.e., AP vs. CP; AS vs. CS).

2.1.1. *Noun position: midline sites*

Fig. 1 shows the grand average waveforms, averaged across all participants, at the position of the Noun for the Ambiguous–Singular (AS), Control–Singular (CS), Ambiguous–Plural (AP) and Control–Plural (CP) conditions. The waveforms were characterised by well-defined N1–P2 components followed by sustained negative- and positive-going activity in the 300–1100 ms window, which varied across the left and right hemispheres and across the anterior-to-posterior axis of the scalp. Visually, the four conditions patterned together until about 800 ms. That is, after the Noun had been read, and the Verb was presented, another N1–P2 complex emerged at 700 ms, indicating that the next word had been presented. Just after this point, that is, at roughly 800 ms, the AP, AS, as well as CS condition become more negative-going compared to CP across midline (and medial-lateral) sites. As can be seen in Table 2, this visual observation is corroborated statistically. That is, a three-way interaction was found for Context × Number × Time (F (3, 72)=4.83, MSE=27.36, p=.005, ε=.913), where pair-wise comparisons (using Bonferroni correction) revealed that in the last time window (900–1100 ms), CS was more negative-going than CP. No other significant effects were revealed.

2.1.2. *Noun position: medial-lateral sites*

Separate repeated measure ANOVAs were conducted on the medial-lateral electrode sites which included the within factors Context (2 levels: Ambiguous vs. Control), Number (2 levels: Singular vs. Plural), Time interval (4 levels: 300–500, 500–700, 700–900, 900–1100 ms post-stimulus), anterior-to-posterior electrode sites (5 levels: F3/4, FC3/4, C3/4, CP3/4, P3/4) and Hemisphere (2 levels: left vs. right).

Table 2 shows that the medial-lateral effects mirror those found at the midline sites: there were no significant main effects of Context or Number. That is, the only significant effect that involved any of the linguistic factors was again the three-way interaction, Context × Number ×

Table 2 – F-values for ANOVA at the noun position.

Analysis	Effect (df)	F-value	Mean square error
Midline	E (4, 96)	5.873 *	328.04
	T (3, 72)	8.187 ***	287.00
	E × T (12, 288)	3.424 *	18.05
	C × N × T (3, 72)	4.829 ***	27.36
Medial-lateral	E (4, 96)	3.196 *	179.28
	E × T (12, 288)	5.123 ***	39.34
	H × T (3, 72)	4.031 **	22.16
	C × N × T (3, 72)	4.555 ***	36.14

C=Context; N=Number; T=Time; E=Electrode; H=Hemisphere.
* $p<.05$.
** $p<.01$.
*** $p<.001$.

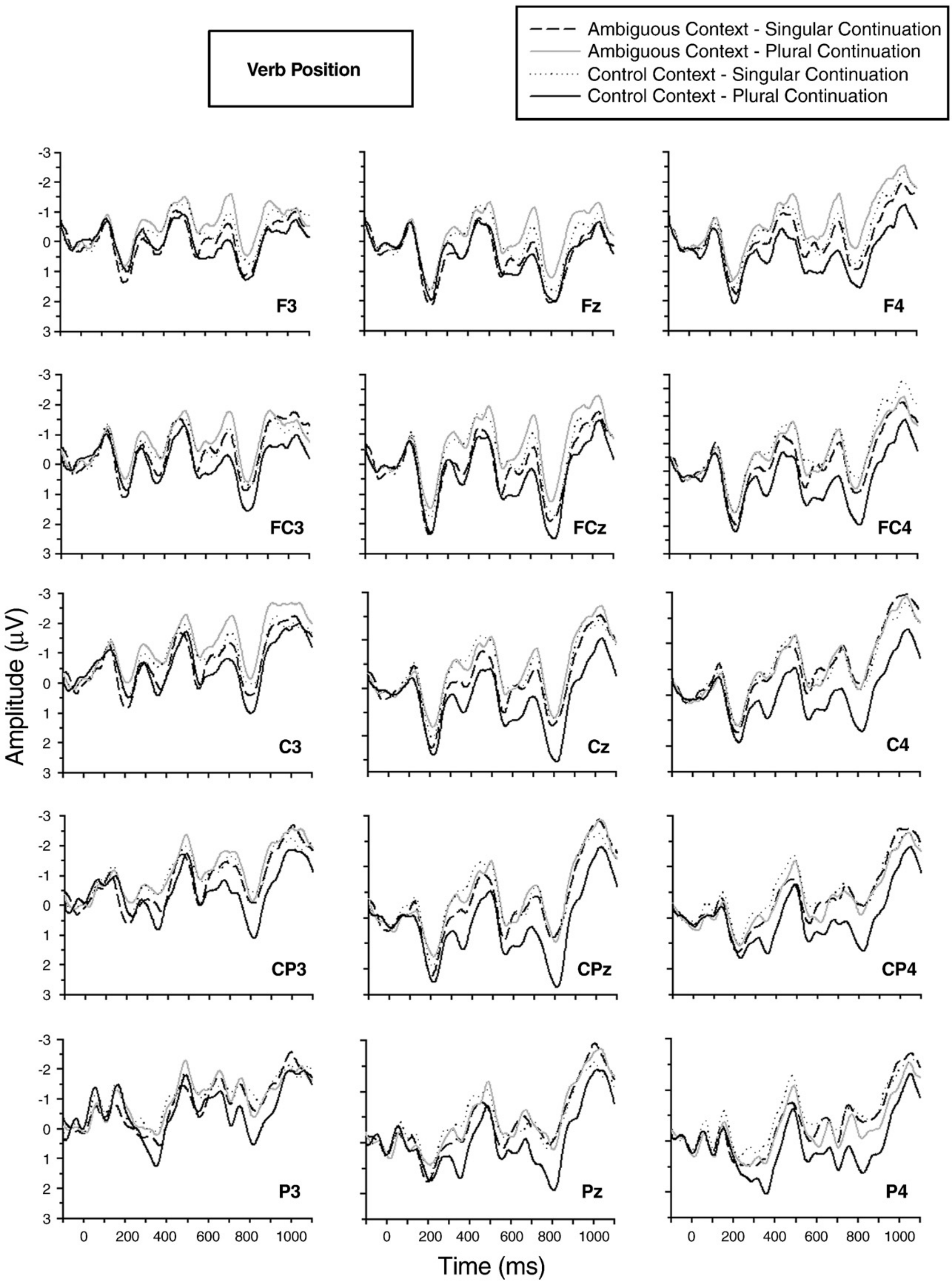

Fig. 2 – Verb position.

Table 3 – F-values for ANOVA at the verb position.

Analysis	Effect (df)	F-value	Mean square error
Midline	E (4, 96)	6.773***	213.37
	T (3, 72)	13.411***	672.28
	E × T (12, 288)	15.049***	50.02
	C × N (1, 24)	7.449*	300.01
Medial-lateral	T (3, 72)	10.930***	675.95
	E × T (12, 288)	4.558***	30.00
	C × N (1, 24)	6.667*	379.78
	N × H (1, 24)	6.521*	57.40
	H × T (3, 72)	17.348***	89.70
	C × N × E × H × T (12, 288)	2.341*	0.32

C = Context; N = Number; T = Time; E = Electrode; H = Hemisphere.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Time (F (3, 72) = 4.56, MSE = 36.14, p = .009, ε = .854). This effect revealed that CS differed significantly from CP in the last time window again (900–1100 ms, p = .03) and a trend in that same time window emerged, where AP was more negative-going than CP at p = .07 (see Fig. 1). Furthermore, AS and AP did not differ from each other in any of the time windows.

2.1.3. *Verb position: midline sites*

Fig. 2 shows the grand waveforms at the Verb position for all four conditions at both midline and medial-lateral sites. Visual inspection reveals a long-lasting negativity of AP, AS and CS at midline sites, evident at central to posterior electrodes. This long-lasting effect is furthermore apparent at medial-lateral sites in the right hemisphere, where the sustained negativity is apparent from frontal to posterior sites. Table 3 summarizes the statistical findings.

Although there were no main effects of Context or Number, a Context × Number interaction was revealed (F (1, 24) = 7.45, MSE = 40.27, p = .012). This interaction reflected the long-lasting negativity of the AP, AS and CS conditions as compared to CP. Using Bonferroni correction, pair-wise comparisons revealed that AP was significantly more negative-going than CP (p = .003) and similarly, CS was significantly more negative than CP (p = .013). Mean voltages for conditions AS and AP did not differ significantly (p = .35), nor did AS vs. CS (p = .48).[2]

2.1.4. *Verb position: medial-lateral sites*

As shown in Table 3, there was no main effect of Context or Number; however, a Context × Number interaction was again revealed (F (1, 24) = 6.67, MSE = 56.96, p = .016). Using Bonferroni correction, pair-wise comparisons indicated a pattern similar to that of midline sites. That is, AP was significantly more negative-going than its control, CP (compare mean voltages of −1.2 μV vs. −0.2 μV, respectively, where p = .001), and CS was significantly more negative than CP (compare mean voltages of −1.1 μV vs. −0.2 μV, respectively, where p = .008). Again, mean voltages for conditions AS and AP did not differ significantly (p = .57), nor did AS vs. CS (p = .55).

There was also a Number × Hemisphere interaction (F (1, 24) = 6.67, MSE = 8.80, p = .017). However, pair-wise comparisons only revealed a trend (p = .09) for the singular conditions (AS, CS) to be more negative-going overall than the plural conditions (AP, CP) in the right hemisphere.

Finally, a five-way interaction of Context × Number × Electrode × Hemisphere × Time (F (12, 288) = 2.34, MSE = 0.136, p = .037, ε = .482) was revealed, which supports visual inspection where the sustained negativity exhibited by conditions AP, AS and CS compared to CP is more prevalent in later time periods in the right hemisphere. That is, pair-wise comparisons revealed that AP was significantly different from CP in roughly all time windows, with a stronger difference in the right medial-lateral sites. Furthermore, CS was also more negative than CP in right medial-lateral sites, in all electrodes except the most anterior, F4 and F3, in roughly all time windows. Again, AP and AS did not differ in any time windows. It is interesting to note that a sustained negative component was observed, rather than an N400 component. To ensure that the absence of an N400 was not due to idiosyncrasies of the experiment, responses to the Filler Anomalous and Filler Control stimuli were compared. As illustrated in Fig. 3, participants did indeed generate N400 effects in this comparison of the Filler conditions (as confirmed by ANOVA; see Table 4 where the Context effect interacts with Time and Electrode at midlines as well as medial-lateral effects; pair-wise comparisons reveal that effects were significant in the 300–500 ms time window, mostly over posterior sites and was slightly right lateralized, as expected from the literature. No other effects were significant beyond that time window, however).

2.1.5. *Verb-plus-one position: midlines*

As shown in Fig. 4, the results at this position show that the ERPs elicited in all four conditions come back together. A well-defined N1–P2 complex is revealed across all sites. This visual observation is supported by the statistical analysis as summarized in Table 5, where no significant effects emerged for the linguistic factors of Context or Number, nor did any interactions with these factors emerge.

2.1.6. *Verb-plus-one position: medial-lateral sites*

The pattern at the medial-lateral sites was similar to that of the midlines; that is, the waveforms for all four conditions come together, with a well-defined N1–P2 complex. Table 5 confirms that no significant effects for the linguistic factors of Context or Number, or any interactions involving these factors emerged.

2.1.6.1. *Results summary.* In sum, a long-lasting negative-going waveform was elicited for the experimental conditions AP and AS, and interestingly for CS. This negativity

[2] Note that simple effects analysis precludes a direct comparison between conditions AS and CP. Instead, this comparison must be inferred; i.e., if AP and CS differ reliably from CP, and AS does not differ significantly from AP and CS, then we can infer that, like AP and CS, AS also differs reliably from CP.

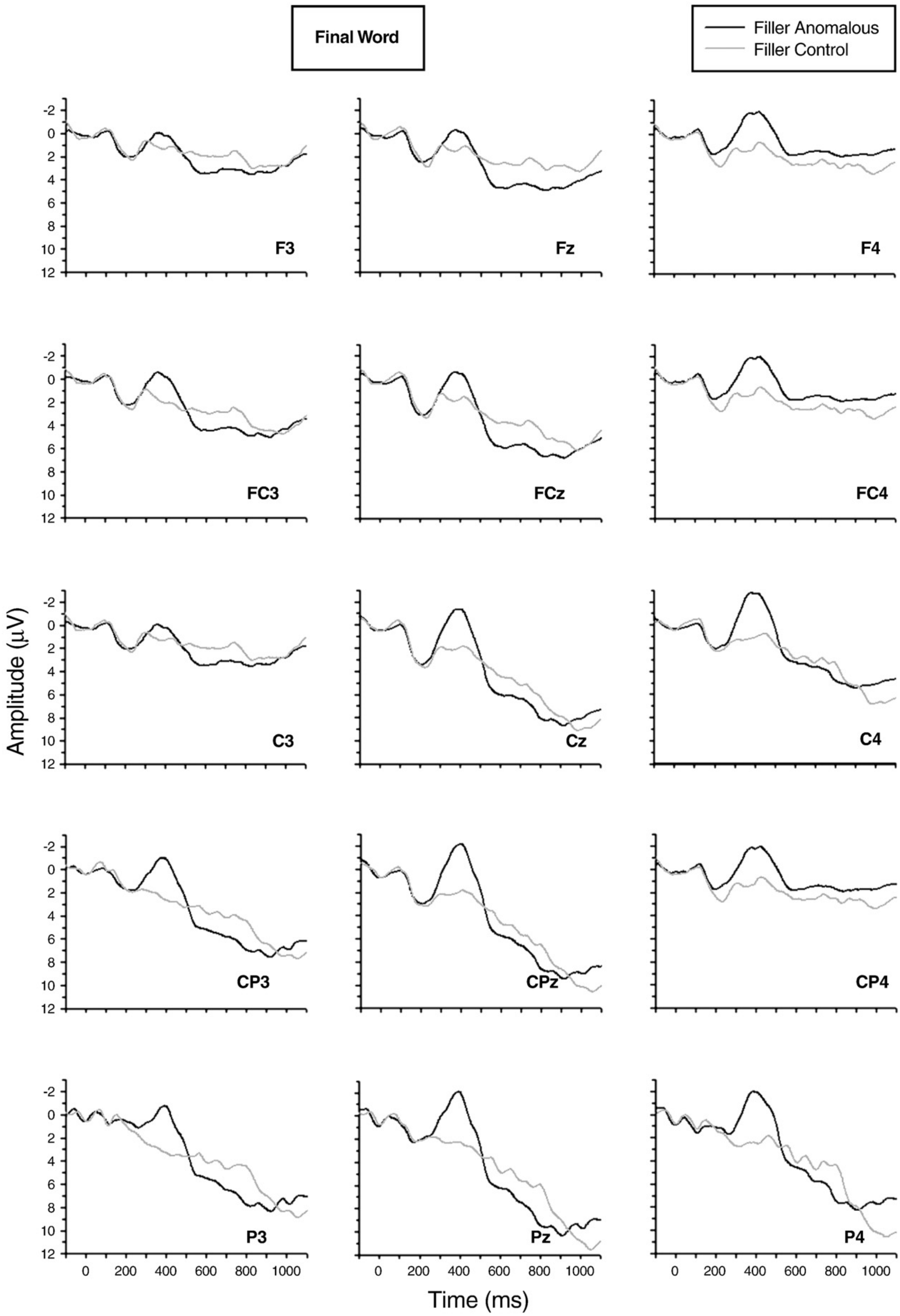

Fig. 3 – N400 fillers.

Table 4 – F-values for ANOVA at final word position for Filler anomalous vs. Filler coherent conditions.

Analysis	Effect (df)	F-value	Mean square error
Midline	E (4, 96)	10.544***	1043.411
	T (3, 72)	49.144***	3723.914
	E × T (12, 288)	36.373***	414.476
	C × T (3, 72)	12.589***	310.137
	C × E × T (12, 288)	3.365*	12.769
Medial-lateral	E (4, 96)	15.115***	1474.967
	T (3, 72)	53.107***	3737.207
	C × H (1, 24)	12.695**	217.136
	C × T (3, 72)	15.828***	487.643
	E × T (12, 288)	28.176***	357.060
	C × E × T (12, 288)	6.928***	35.590
	H × T (3, 72)	12.634***	73.884
	C × E × H × T (12, 288)	2.785*	1.110

C = Coherence; T = Time; E = Electrode; H = Hemisphere.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

was different from a classic N400 effect as depicted for filler anomalous vs. filler control items in Fig. 3. In addition, no effects reflecting a (syntactic) P600 component were elicited (but see Fig. 5 where these were elicited for the appropriate ungrammatical vs. grammatical filler conditions[3] and Table 6 which summarizes the statistical findings). The fact that AP and AS pattern together is consistent with prediction (iii), as outlined in the Introduction.

3. Discussion

The present study sought to explain how it is that the brain/parser perceives definite NP anaphors when embedded in semantically ambiguous contexts. To this end, sentences beginning with a definite NP (e.g., "*the tree(s)*") were embedded in contexts exhibiting scope ambiguity such as *Every kid climbed a tree*. Responses to continuation sentences following ambiguous contexts were compared to those following unambiguous control contexts, such as *Every kid climbed a different tree* (Control–Plural, CP condition) and *Every kid climbed the same tree* (Control–Singular, CS condition). As laid out in the Introduction, three possible empirical patterns could emerge, either (i) AS would differ from its control CS, since the singular NP is non-preferred (Kurtzman and MacDonald, 1993) or (ii) AP would differ from its control CP, if the plural NP is non-preferred. Prediction (iii) was that there would be no preference for either a singular or plural continuation, such that AS and AP would pattern together. Furthermore, the neurophysiological form associated with pattern (iii) would be a slow negative shift, possibly an Nref, due to the ambiguous nature of the context.

Predictions (i) and (ii) are ruled out since we did not find any evidence of ERP waveforms indicating a preference in interpretation. That is, no N400-like effects or P600 effects were elicited for either ambiguous condition in comparison to its control, despite the fact that participants did in fact produce these waveforms in response to our filler items (see Figs 3 and 5 for the N400 and P600 effects, respectively). Instead, we found that the Ambiguous–Plural condition and the Ambiguous-Singular condition patterned together, exhibiting sustained negativity, supporting prediction (iii). The fact that the Control–Singular condition also patterned with the Ambiguous conditions will be discussed below. Thus, the empirical finding was clearly that Ambiguous–Plural, Ambiguous-Singular, and Control–Singular conditions all exhibited a slow negative shift, along midline and right medial-lateral sites, in comparison to the Control–Plural condition, starting at 900 ms after the Noun "tree(s)" was presented and enduring over the time period of the Verb. Whereas this effect lasted throughout the presentation of the Verb "was/were" (along midline and right medial-lateral sites), no significant effects emerged after the Verb (e.g., at "in"). We discuss the significance of these effects below.

3.1. *Underspecified representations*

The fact that conditions AS and AP patterned together is an important finding as it makes clear that the brain/parser does not immediately assign a logical meaning to scope ambiguous sentences. In other words, at very early stages of comprehension, the brain/parser treats scope ambiguous sentences as underspecified representations. Recent work in language processing (Christianson et al., 2001; Sanford and Sturt, 2002; Ferreira et al., 2002; Swets et al., 2008) suggests that interpretive processes are often incomplete and shallow, such that comprehenders do not commit to a particular meaning during a parse (see volume 42, number 2 of *Discourse Processes* for further articles devoted to this topic). Thus, possibly in the interest of conserving time, attention, and working memory resources, the processor leaves certain ambiguities as unresolved. The present results support the notion of a "good enough" parsing strategy, where scope ambiguous sentences are left unresolved until further disambiguating information arrives, such as an anaphor and inflected verb.

If we interpret the effect in terms of a model of anaphoric processing as assumed by Garrod and Sanford (1994) and Garrod and Terras (2000), then we can understand the findings in the following way: at stage 1 ("bonding"), the search for an antecedent to "the tree(s)" begins (i.e., in conditions AS and AP). The processor searches the earlier discourse to find an underspecified representation. Thus, at stage 2 ("resolution"), the representation must be disambiguated, such that it is consistent with the interpretation of the definite NP in the later sentence. In other words, once the processor has perceived "the trees were…" it assigns surface scope to the context sentence, or once it has perceived "the tree was…" it assigns inverse scope. The semantic computation for both interpretations is equivalent, since there was no initial preference.

3.2. *Slow negative shift*

The complex nature of the search as well as the required semantic computation accounts for the lateness, as well as the

[3] Interestingly, the ungrammatical vs. grammatical filler comparison also elicited a trend for an N400 effect, especially at medial-lateral sites. While discussion of this effect in filler stimuli is beyond the scope of this paper, we refer the reader to Kuperberg (2007) for an excellent review of the specificity of N400 and P600 effects.

extended time course, of this ERP effect. That is, the timing of this component occurs only once the Verb has been presented, that is, 900 ms after the Noun. This is in contrast to the near immediate effect noted with the Nref, at 280 ms found by van Berkum et al. (1999a). The timing difference makes sense given the differing nature of the contexts (see also van Berkum et al., 2003, where this is discussed). The present study examined contexts that were ambiguous due to underspecification. An

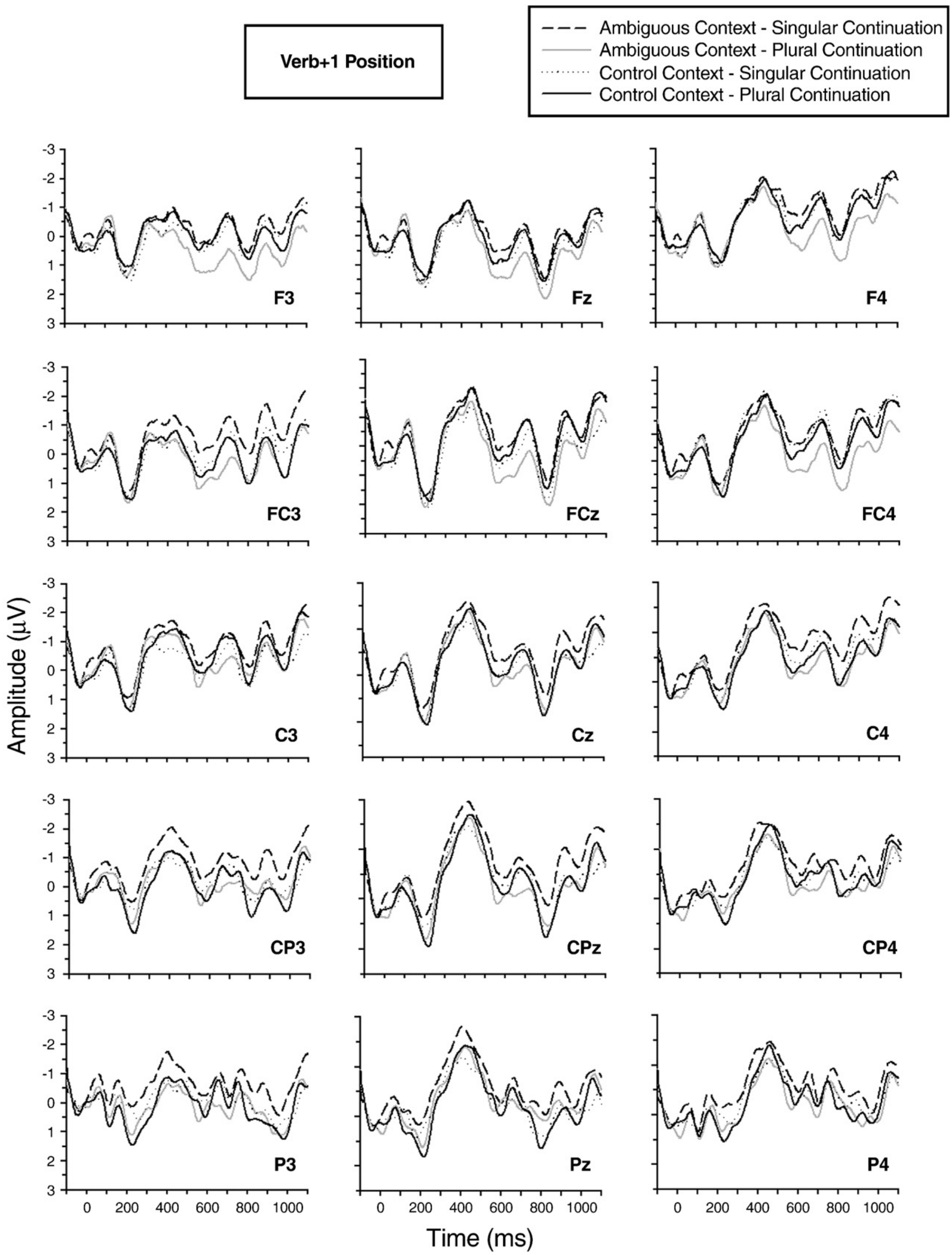

Fig. 4 – Verb+1 position.

Table 5 – *F*-values for ANOVA at the verb-plus-one (V+1) position.

Analysis	Effect (df)	*F*-value	Mean square error
Midline	E (4, 96)	9.669***	266.84
	T (3, 72)	10.269***	523.66
	E × T (12, 288)	7.679***	27.48
Medial-lateral	H (1, 24)	10.969***	759.62
	T (3, 72)	3.632*	252.97
	E × T (12, 288)	4.863***	33.25

C = Context; N = Number; T = Time; E = Electrode; H = Hemisphere.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

antecedent for the definite NP would only be available after the interpretation of the previous context sentence was complete. In contrast, contexts used in the van Berkum et al. studies were unambiguous—these offered explicit mention of more than one relevant character that could serve as a referent for the definite NP. Furthermore, later Nref effects are found in van Berkum et al. (2003) and Nieuwland and van Berkum (2008). Interestingly, the lateness of this effect, as well as its long duration, is reminiscent of the Late (negative) Slow Wave, as discussed in a study by Ruchkin et al. (1988) (see also Rösler et al., 1997). In their attempt to functionally characterize the nature of slow waves, they argue that the late negative slow wave became apparent in the 900–1200 ms latency range, which is analogous to the waveform at hand. Furthermore, they argued that this ERP waveform was sensitive to processing that is conceptually difficult. We argue that assigning scope interpretation in underspecified representations and linking that meaning with a later occurring definite NP is a complex task, requiring increased allocation of cognitive resources. Furthermore, the topography of the Late Slow Wave was broadly distributed over the centro-posterior scalp, as is the present waveform. At present, it is unclear why it is also slightly right lateralized. Perhaps the reliance on computation for semantic meaning or context, arguably a right hemisphere function, requires a stronger recruitment of right hemisphere neural circuitry (Joanette et al., 1990; Brownell et al., 1995).

Thus, although functionally the ERP component has much in common with the Nref (determining the reference of an ambiguous definite NP anaphor), the topography and timing of the waveform exhibited patterns like the Late Slow Wave, which has been argued to be elicited after performing a conceptually difficult task. Whether the waveform observed in the present study is a different version of the Nref, or another version of the several slow negativities observed that tax working memory resources due to difficulty, is left open for further research.

In sum, we believe the reason why the slow negative wave was found for conditions AP and AS is that scope ambiguous sentences, in the absence of previous context, are truly ambiguous—the brain/parser does not assign an immediate interpretation. In contrast, the Control–Plural condition, *Every kid climbed a different tree*, is clearly unambiguous. As such, integrating a definite NP into an ambiguous context elicits a negative-going waveform as compared to an unambiguous context, CP.

At this point, we must address why the Control–Singular condition, as in *Every kid climbed the same tree* patterned with AP and AS. Given the claim that the negativity associated with conditions AS and AP is the result of ambiguity, the same theoretical claim would need to be made for CS. Results from an off-line pretest (details below), where participants were asked to circle whether the singular or plural continuation sentence fit better with the context sentences, give us a clue that the "control" singular condition is unlike the control plural. The latter condition produced very strong results for a plural interpretation, 95% of the time. However, the Control–Singular condition, while clearly biased for a singular interpretation at 85%, differed significantly from the Control–Plural condition ($p < .001$). Thus, even when participants are under no time pressure to interpret such sentences, they are not doing so in a uniform manner across the control conditions. The question, of course, is from where does the increased ambiguity for Control–Singular derive?

A close examination of the control conditions shows that whereas the CP condition ends with "a different tree" the CS condition ends with "the same tree." Our claim is that the presence of the definite article "the," found in CS but not CP, coupled with readings associated with "different" and "same" accounts for the effects observed. That is, the adjective "same," in *Every kid climbed the same tree* has available to it both a sentence-internal and sentence-external reading (which actually results in scope ambiguity; see footnote 4) whereas "different" (e.g., *Every kid climbed a different tree*) does not. The sentence-internal reading is one where the interpretation of "same/different" is dependent on the immediate sentence. In contrast, the sentence-external reading depends on comparisons from previous context. Thus, Carlson (1987) discussed the meaning of "same" and "different" and pointed out that these comparative adjectives usually refer to some previously mentioned element in the discourse, called the sentence-external or "deictic" reading, as below:

(1) a. The man went to the same play tonight.
b. Smith went to a different place on his vacation this year.
(Carlson, 1987, p. 531)

For example, the meaning of (1a) is that the man went to the same play tonight as compared to the one we were talking about yesterday. A similar interpretation ensues for (1b). On the other hand, Carlson noted that there are some instances where *same* and *different* do not involve a covert comparison with something previously mentioned in context. Rather, he notes that instead of a sentence-external comparison, sentence-internal comparisons are possible, as the examples below show:

(2) a. Bob and Alice attend different classes (e.g., Bob attends Biology 101 and Alice attends Philosophy 799).
b. The same salesman sold me these two magazine subscriptions (e.g., Salesman Jones sold me this subscription to Consumer Reports, and Jones, too, sold me this subscription to Cosmopolitan).
(Carlson, 1987, p. 532)

In the sentences above, "same" and "different" are now using, as their reference, elements that are found in the immediate sentence-internal context. That is, in addition to a possible sentence-external reading for (2a), where Bob and Alice attend different classes as compared to the ones that you and I like, now a comparison is also possible where they take different

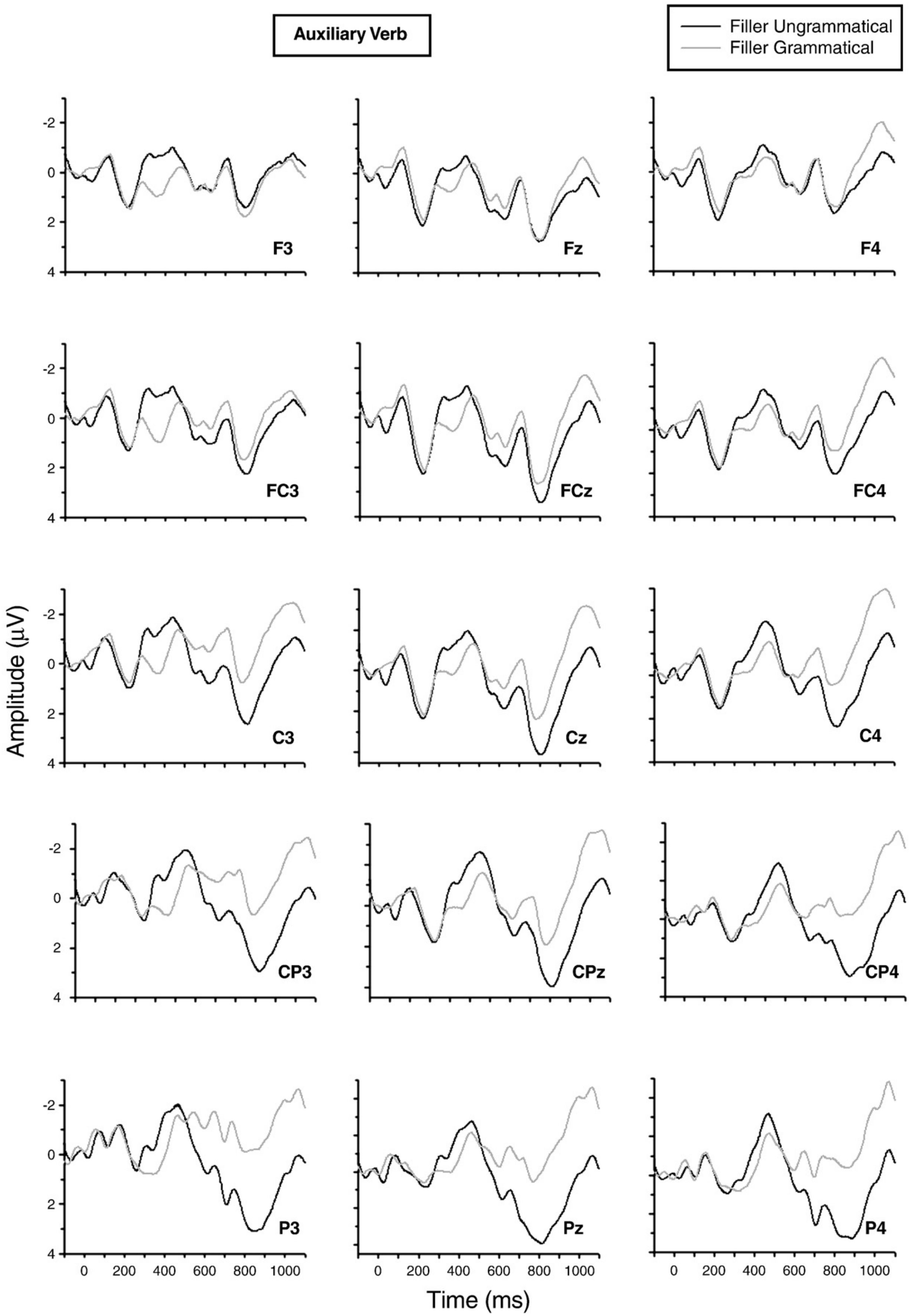

Fig. 5 – P600 fillers.

Table 6 – F-values for ANOVA at auxiliary word position for Filler grammatical vs. Filler ungrammatical conditions.

Analysis	Effect (df)	F-value	Mean square error
Midline	E (4, 96)	4.117 *	114.952
	T (3, 72)	19.121 ***	501.029
	G × E (4, 96)	6.352 **	47.645
	G × T (3, 72)	13.498 ***	138.498
	E × T (12, 288)	5.760 ***	19.177
	G × E × T (12, 288)	16.420 ***	28.441
Medial-lateral	T (3, 72)	12.315 ***	433.496
	G × E (4, 96)	10.014 ***	100.120
	G × T (3, 72)	15.369 ***	264.067
	E × T (12, 288)	4.055 **	16.353
	G × E × T (12, 288)	16.592 ***	39.676
	H × T (3, 72)	6.015 **	25.458
	G × H × T (3, 72)	9.121 ***	16.781

G = Grammar; T = Time; E = Electrode; H = Hemisphere.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

classes *as compared to each other*. For (2b), it could be the same salesman who came by my house last week who sold me these two subscriptions, or it could be the reading as made clear in (2b), where one salesman sold one person two subscriptions.

For the present experiment, *Every kid climbed the same/a different tree* was never preceded by context. Thus, it is clear that in the absence of previous context, the sentence-internal reading is the relevant reading for these sentences. Given this assumption, an explanation of the findings would be as follows: for sentences such as *Every kid climbed a different tree*, this is interpreted as every child climbing a different tree compared to every other child, and this would result in a reading where there are many trees climbed by many children. This results in an unambiguous plural interpretation of "trees."

For sentences such as *Every kid climbed the same tree*, this again, in the absence of context, would be interpreted on its sentence-internal reading, such that every child climbed the same tree as every other child. However, we submit that the presence of the definite article "the" in "the same" vs. "a" in "a different" strongly persuades the brain/parser that a sentence-external comparison might still be worthy of consideration, that is, in addition to the sentence-internal reading.

That is, perceivers are generally ready to easily accommodate information from NPs containing "the" in them, despite the fact that "the N" generally refers to old information, or previously mentioned information (Haviland and Clark, 1974; Heim, 1982; Murphy, 1984). As such, despite the overt unavailability of an antecedent, the brain/parser is still willing to entertain the sentence-external meaning associated with "the same." This is where the ambiguity arises.

Furthermore, this very ambiguity actually results in a scope ambiguity, as defined by linguistic theory (Carlson, 1987).[4]

[4] Specifically, the sentence-internal reading is consistent with the surface scope reading, where "every N" is interpreted first, and "the same" is interpreted with respect to "every" (see Barker, 2007). In contrast, the sentence-external reading would require that "the same" take scope over "every," which would be the inverse scope reading.

Thus, it seems that for this condition, too, the brain/parser waits to assign scope, since both meanings are available to it, and it does not have enough information on which to base a decision.

In sum, we have argued that in real-time, in the absence of previous context, sentences such as *Every kid climbed a tree* are truly ambiguous regarding scope possibilities. Furthermore, scope ambiguity also arises in sentences such as *Every kid climbed the same tree*. The brain/parser does not immediately resolve the semantic ambiguity; instead it leaves such constructions as underspecified. In the present case, integrating a definite NP into such a context is costly.

Thus, our ERP findings contrast with those of Kurtzman and MacDonald (1993), which showed that there was a preference for the plural interpretation, as well as our own off-line results. This could be the case because of two reasons; first, participants in previous experiments were specifically asked to choose a particular interpretation; furthermore, these studies examined processing only at the end of such sentences. In the present study, we examined processing without any meta-linguistic judgments, and we measured processing in real-time. Our results further highlight the importance of using different methodologies in investigating the nature of language processing.

In conclusion, we explored simple active sentences that exhibited scope ambiguity using the universal quantifier "every" in subject position, and the existential quantifier "a" in direct object position. The time resolution of ERPs yielded findings that were different from behavioural measures, as well as off-line measures. We found evidence that that such sentences are not disambiguated immediately; instead, the brain/parser waits to assign meaning and leaves these underspecified. However, there is a later cost to shallow processing—it must be made specific due to dependent material (definite NP) arriving later in the signal; this is reflected empirically by the long-lasting slow-negative shift. Furthermore, our findings regarding the Control–Singular condition which replaced the existential quantifier "a" in direct object position with "the same" indicate that the brain/parser is sensitive to the semantic ambiguity of this construction, which also results in a scope ambiguity.

Finally, our findings shed light on models of discourse comprehension (Zwaan and Radvansky, 1998; van Dijk and Kintsch, 1983; Johnson-Laird, 1983) that take semantic ambiguity into account, since we have shown that rather than relying exclusively on a notion of ambiguity as defined by conceptual or real-world knowledge, grammatical considerations also play a role. Our results indicate that the brain/parser is sensitive to semantic ambiguity defined at a compositional semantic level, and that this sensitivity has empirical effects downstream in terms of determining the reference of definite NPs.

4. Experimental procedures

4.1. *Participants*

Twenty-five native speakers of English (15 female, mean age 21.85 years, range 18–27 years) were recruited at Concordia

University and were either paid for their participation or received partial course credit. All subjects had normal or corrected-to-normal vision and were right handed, as assessed by the Handedness Inventory (Briggs and Nebes, 1975). None of the participants reported any neurological impairments, history of neurological trauma or use of neuroleptics. Also, none of them had participated in the pilot ratings task (see below).

4.2. *Materials*

4.2.1. *Ambiguous context sentences*

Simple declarative (e.g., subject^verb^object) context sentences were constructed where the subject was a quantified NP, which was always the universal quantifier *every* paired with an animate head noun (e.g., *kid*, *tourist*, *shopper*, etc.), followed by an action verb in the past tense (e.g., *climbed*, *visited*, *squeezed*, etc.), followed by an inanimate object NP (e.g., *tree*, *statue*, *melon*, etc.) paired with the existential quantifier "a," resulting in sentences such as *Every kid climbed a tree*, *Every tourist visited a statue*, and *Every shopper squeezed a melon.*

4.2.2. *Control context sentences*

In contrast, the Control context sentences distinguished the interpretation of the indefinite objects as unambiguously singular or plural. The structure of the sentences was exactly the same as the ambiguous context sentences (e.g., subject^verb^object) except that the direct object was preceded by different adjectives. We adopted the markers used by Kurtzman and MacDonald (1993), where the Control–Singular context condition was of the form *Every kid climbed the same tree*, and the Control–Plural context condition was *Every kid climbed a different tree*. The Control–Singular (CS) context sentence was always followed by the singular continuation sentence, whereas the Control–Plural (CP) context condition was always followed by the plural continuation sentence, resulting in the following two control conditions: CS: *Every kid climbed the same tree. The tree was in the park*; and CP: *Every kid climbed a different tree. The trees were in the park.*

4.2.3. *Continuation sentences*

The form of the experimental continuation sentences was the following: the subject NP always referred back to the object NP of the context sentence. Furthermore, the subject was either plural or singular (e.g., *trees* vs. *tree*) and was followed by an auxiliary verb (e.g., *were* or *was*) and then a predicate. Half of the time the predicate was a prepositional phrase (e.g., *The tree(s)were/was* **in the park**), and the other half, it was an adjectival phrase (e.g., *The melon(s) were/was* **large and green**).

One hundred sixty scenarios were created (e.g., *kid^-climb^tree*; *shopper^squeeze^melon*; *tourist^visit^statue*, etc.) for each of the four conditions (Ambiguous–Plural, Control–Plural, Ambiguous–Singular, and Control–Singular) resulting in a total of 640 sentence pairs. In order to reduce repetition effects, the stimuli were divided into four counterbalanced lists, such that each participant saw an equal number of sentence pairs from each condition, resulting in 40 trials per experimental condition per list.

4.2.4. *Filler sentence pairs*

In addition to the experimental discourses, there were 160 filler discourses to reduce the predictability of the experimental stimuli and to reduce the chance of participants adopting particular reading strategies. Furthermore, since the predicted waveform could be of different varieties, these fillers were constructed to ensure that the participants were in fact capable of producing classic ERP effects such as the N400 and the P600 in response to semantic and syntactic ambiguity manipulations, respectively. These filler sentences are described in the section below.

4.2.4.1. *Filler anaphoric—anaphora to non-object NP antecedent*. Forty of the 160 filler discourses were coherent discourses, meaning that they were both semantically and syntactically correct. These consisted of sentences where the subject of the continuation sentence referred back to an NP which was not an object, since in the critical sentences anaphora was always to the object NP. These 40 coherent fillers were immediately followed by forced-choice comprehension questions, in order to ensure that subjects paid attention to the stimuli.[5] When the questions were presented after these stimuli, the two alternative answers were shown on the left- and right-hand sides of the computer screen, and participants had to press the corresponding button on a response pad to indicate the correct answer. The position of the correct answer was counterbalanced across trials. A sample of the filler and question appear below:

Filler (1) *On afternoons, Alice went to the babysitter.*
The afternoons were a time to relax after school.
Question: The afternoons were a time to relax after what?
Answer: SCHOOL WORK

4.2.4.2. *Incoherent N400 fillers*. Forty filler sentences used auxiliary verbs not used in the target sentences (e.g., *could*, *can*, *ought to*, *did*, and *will*) and were anomalous, but for reasons independent of grammatical constraints across sentences. Instead, these represented violations of real-world knowledge. An example of such a "Filler Anomalous" discourse is: *Celine will come to the party. She ought to bring skyscrapers*. These discourses were included in order to compare classic N400-like effects to the coherent fillers (at the final word) as described above.

4.2.4.3. *Syntactically incoherent fillers*. Eighty fillers consisted of typical P600 violations, where there was a local number agreement violation at the auxiliary verb position in the continuation sentence. For example, two possible stimuli were:

Few brothers were eating pie. They were/*was eating cake instead.

A student was feeling discouraged. The student was/*were failing the class.

[5] Questions about the Filler Anomalous sentences were not used, since the sentences did not make sense, and questions about the experimental sentences were not used in order to avoid encouraging any specific strategies for reading such sentences.

This was a very effective distractor method; in debriefing sessions post-experiment, participants often felt that the point of the experiment was to understand number (dis) agreement. Thus, we feel confident that participants were naïve as to the purpose of the experiment.

In total, each list viewed by a participant contained 320 sentence pairs: 160 target stimuli (40 from each of AS, AP, CS, CP) and 160 filler sentence pairs as described above. As noted earlier, each participant saw one list only, with sentences presented in a pseudo-random fixed sequence with the stipulation that no two trials from the same experimental condition or filler condition followed each other.

4.3. *Pretests*

In order to confirm that our stimuli yielded off-line preferences in keeping with those reported by Kurtzman and MacDonald (1993), we assessed the acceptability of all ambiguous context sentences, and half of the control contexts by evaluating these in an off-line norming study. Two semi-randomized lists were created and 32 subjects recruited at McGill University were paid $10 for their participation. None of these subjects participated in the on-line ERP experiment. In this off-line task, discourses were presented in a booklet in a pseudo-random order, with the constraint that no more than two of the same type of trial succeeded one another. In each list, 80 ambiguous context sentences were presented, as well as 80 unambiguous ones (40 Control–Singular and 40 Control–Plural). In addition, 80 fillers were used from an unrelated experiment. The participants were asked to circle the continuation sentence that fit better with the first sentence. Overall, results were consistent with those of Kurtzman and MacDonald (1993); the preferred interpretation for Ambiguous sentences (e.g., *Every kid climbed a tree*) was indeed the plural reading—73% (vs. 77% Kurtzman and MacDonald, 1993).[6] As expected, pair-wise comparisons (both by participants and by items) revealed that this differed significantly from the control conditions. That is, Ambiguous sentences judged as plural differed significantly from the Control–Plural condition, which was judged as plural 95% of the time, with very little variability ($p<.001$), as well as from Control–Singular, which was judged as singular 85% of the time ($p<.001$). Unlike the findings of Kurtzman and MacDonald, however, the CS and CP conditions in our study differed significantly *from each other* ($p<.001$). Given that these were supposed to represent ceiling scores, this was an unexpected finding. It is clear that these findings foreshadowed the on-line ERP results as discussed above.

4.4. *Procedure*

For the experimental test, participants were tested individually in one session, which lasted approximately 3 h. Short breaks were given when required. Following the application of the EEG electrodes, subjects were seated in front of a computer screen approximately 1 m away. All stimuli were presented in white text on a black background in 26 point Arial font on a Compaq Deskpro computer, with a Compaq V74 16" monitor using STIM presentation software (Compumedics Neuroscan USA, Inc., El Paso, TX). The words of the continuation sentence were presented between 0.014° and 0.089° of visual angle in the center of the computer monitor. Participants responded to the questions by using a Stim System Response Pad (Compumedics Neuroscan USA, Inc., El Paso, TX).

Each context sentence (S1) was presented in its entirety; participants pressed a button to indicate when they were ready for the continuation sentence (S2). Following an ISI of 600 ms, the continuation sentence was presented one word at a time in the centre of the screen with each word presented for 300 ms followed by an ISI of 300 ms. This presentation rate minimized eye movement artifacts in the EEG recordings and allowed for time-locking the EEG recording to the presentation of each word. Between each sentence pair there was a 3-s delay to make sure the participants read the sentences as distinct pairs. Participants were instructed to silently read the context sentence, to press a button when it had been read, and to read each individual word of the subsequent sentence. Participants were instructed not to speak, move, or blink their eyes during the presentation of the stimuli. Practice trials were included to accustom participants to the task. When required, participants responded to a comprehension question using a hand held pad. This question appeared 100 ms after the last word of certain sentence pairs, and only occurred after 25% of the filler trials. On average, participants correctly answered these questions 98.2% of the time, indicating that they were indeed paying attention. Note again that probe questions were not used on critical trials in order to ensure that participants would not develop processing strategies for these stimuli.

4.5. *Electrophysiological measures*

A commercially available nylon EEG cap containing silver/silver chloride electrodes (Quik-Cap) was used for EEG recording. The EEG was recorded from 5 midline electrode sites and 22 lateral sites. A cephalic (forehead) location was used as ground. All sites were referenced to the left ear during acquisition and re-referenced off-line to a linked ear reference. EOG was recorded from electrodes placed at the outer canthi of both eyes (horizontal EOG) and above and below the left eye (vertical EOG). EOG artifacts were corrected off-line for all subjects using a rejection criterion of $\pm 100\ \mu V$, in accordance with the procedure outlined in the Neuroscan 4.3 Edit (2004) manual. This resulted in the following artifact rejection rates: for the AP condition at N: 11.2%, at V: 8.9%, at V1: 9.4%; for the AS condition at N: 12.1%, at V: 10.3%, at V1 10.2%; for the CP condition at N: 13.5%, at V: 12.4%, at V1: 10.9%; and finally for the CS condition at N: 11.5%, at V: 10.7% and at V1: 10%.

EEG was sampled continuously with critical EEG epochs time-locked to the onset of each target word of S2: the head Noun, the auxiliary verb, and the word after the verb (i.e., Verb + 1 position; this was never the final position in the sentence). EEG data were amplified using Neuroscan Synamps in a DC-100 Hz bandwidth using a 500 Hz digitization rate. Single trial epochs were created using a −100 to

[6] This preference is the result of the fact that a majority of the participants had plural preferences for a majority of the items. That is, a combination of both participant and item factors resulted in the overall preference value of 73%. We note that while these findings are intriguing, they did not form the focus of the present study and we leave this issue for further research.

1100 ms window around the eliciting stimulus and processed off-line using Neuroscan Edit 4.3 software. For each participant, ERP averages were computed for each category of critical words in all target continuation sentences. The mean voltage amplitude of the −100 to 0 ms period of each averaged waveform was calculated and served as the 0 μV baseline for post-stimulus activity. The mean amplitude of each waveform was computed in 200 ms intervals from 300 to 1100 ms post-stimulus, yielding 4 mean amplitudes. These effects were examined across five midline electrode sites (i.e., Fz, FCz, Cz, CPz, and Pz) and medial-lateral electrode sites as defined in the Results section above.

Acknowledgments

This work was supported by a Social Sciences and Humanities Research Council of Canada award (#410-2006-1748) entitled: "An electrophysiological investigation of processing quantifiers in discourse," awarded to N. A. Phillips, V. D. Dwivedi and S. R. Baum. Part of this work was conducted while the first author was supported in part by funding from the Social Sciences and Humanities research Council of Canada (grant number 412-2003-1003) awarded to Dr. Anna Maria Di Sciullo for the Major Collaborative Research Initiative on Interface Asymmetries. This work has benefited from presentations at the Mental Lexicon Conference, the Society of Psychophysiological Research as well as the Cognitive Neuroscience Society and Canadian Linguistics Association meetings. Thanks also to Hotze Rullmann, Sid Segalowitz, and Tim Murphy for helpful discussion, as well as John Drury, Karsten Steinhauer, Paul Portner and Barbara Partee. In addition, thanks to Matt Goldhawk and Sharon Gagnon for help in preparing the manuscript.

REFERENCES

Anderson, C., 2004. The structure and real-time comprehension of quantifier scope ambiguity. Unpublished doctoral dissertation, Northwestern University, Evanston, IL.

Barker, C., 2007. Parasitic scope. Ling. Phil. 30, 407–444.

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

Brownell, H., Gardner, H., Prather, P., Martino, G., 1995. Language, communication, and the right hemisphere. In: Kirshner, H.S. (Ed.), Handbook of Neurological Speech and Language Disorders. Marcel Dekker, New York, pp. 325–349.

Carlson, G.N., 1987. Same and different: some consequences for syntax and semantics. Ling. Phil. 10, 531–565.

Christianson, K., Hollingworth, A., Halliwell, J.F., Ferreira, F., 2001. Thematic roles assigned along the garden path linger. Cogn. Psychol. 42, 368–407.

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

Dwivedi, V.D., Phillips, N.A., Laguë-Beauvais, M., Baum, S.R., 2006. An electrophysiological study of mood, modal context, and anaphora. Brain Res. 1117, 135–153.

Ferreira, F., Bailey, K., Ferraro, V., 2002. Good enough representations in language comprehension. Curr. Dir. Psychol. Sci. 11, 11–15.

Fiebach, C.J., Schlesewsky, M., Friederici, A.D., 2001. Syntactic working memory and the establishment of filler-gap dependencies: insights from ERPs and fMRI. J. Psycholinguist. Res. 30, 321–338.

Filik, R., Paterson, K.B., Liversedge, S.P., 2004. Processing doubly quantified sentences: evidence from eye movements. Psychon. B. Rev. 11, 953–959.

Friederici, A.D., Frisch, S., 2000. Verb argument structure processing: the role of verb-specific and argument-specific information. J. Mem. Lang. 43, 476–507.

Friederici, A.D., Hahne, A., Mecklinger, A., 1996. Temporal structure of syntactic parsing: early and late event-related brain potential effects. J. Exp. Psychol. Learn. 22, 1219–1248.

Frisch, S., Schlesewsky, M., Saddy, D., Alpermann, A., 2002. The P600 as an indicator of syntactic ambiguity. Cognition 85, B83–B92.

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

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

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

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

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

Haviland, S.E., Clark, H.H., 1974. What's new? Acquiring new information as a process in comprehension. J. Verbal Learn. Verbal Behav. 13, 512–521.

Heim, I., 1982. The semantics of definite and indefinite noun phrases. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Hoeks, J.C.J, Stowe, L.A., Doedens, G., 2004. Seeing words in context: the interaction of lexical and sentence level information during reading. Cogn. Brain Res. 19, 59–73.

Joanette, Y., Goulet, P., Hannequin, D., 1990. Right hemisphere and verbal communication. Springer-Verlag, New York.

Johnson-Laird, P.N., 1983. Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness. Harvard University Press, Cambridge, MA.

Kaan, E., Swaab, T.Y., 2003a. Electrophysiological evidence for serial sentence processing: a comparison between non-preferred and ungrammatical continuations. Cogn. Brain Res. 17, 621–635.

Kaan, E., Swaab, T.Y., 2003b. Repair, revision, and complexity in syntactic analysis: an electrophysiological differentiation. J. Cogn. Neurosci. 15, 98–110.

Kemtes, K.A, Kemper, S., 1999. Aging and resolution of quantifier scope effects. J. Gerontol. B-Psycol. 54, 350–360.

Kluender, R., Kutas, M., 1993. Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. J. Cogn. Neurosci. 5, 196–214.

Kratzer, A., 1979. Conditional necessity and possibility. In: Bauerle, R., Egli, U., von Stechow, A. (Eds.), Semantics from a Different Point of View. Springer, Berlin.

Kuperberg, G.R., 2007. Neural mechanisms of language comprehension: challenges to syntax. Brain Res. 1146, 23–49.

Kurtzman, H.S., MacDonald, M.C., 1993. Resolution of quantifier scope ambiguities. Cognition 48, 243–279.

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

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

Kutas, M., Hillyard, S.A., 1983. Event-related brain potentials to grammatical errors and semantic anomalies. Mem. Cogn. 11, 539–550.

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

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

Murphy, G., 1984. Establishing and accessing referents in discourse. Mem. Cogn. 12, 489–497.

Nieuwland, M.S., Van Berkum, J.J.A., 2008. The interplay between semantic and referential aspects of anaphoric noun phrase resolution: evidence from ERPs. Brain Lang. 106, 119–131.

Osterhout, L., Holcomb, P.F., Swinney, D.A., 1994. Brain potentials elicited by garden-path sentences: evidence for the application of verb information during parsing. J. Exp. Psychol. Learn. 28, 786–803.

Paterson, K.B., Filik, R., Liversedge, S.P., 2008. Competition during the processing of quantifier scope ambiguities: evidence from eye movements during reading. Q. J. Exp. Psychol. 61, 459–473.

Rösler, F., Heil, M., Röder, B., 1997. Slow negative brain potentials as reflections of specific modular resources of cognition. Biol. Psychol. 45, 109–141.

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

Ruchkin, D.S., Johnson, R., Mahaffey, D., Sutton, S., 1988. Toward a functional categorization of slow waves. Psychophysiology 25, 339–353.

Sanford, A.J., Sturt, P., 2002. Depth of processing in language comprehension: not noticing the evidence. Trends Cogn. Sci. 6, 382–386.

St. George, M., Mannes, S., Hoffman, J.E., 1997. Individual differences in inference generation: an ERP analysis. J. Cogn. Neurosci. 9, 776–787.

Swets, B., Desmet, T., Clifton, C., Ferreira, F., 2008. Underspecification of syntactic ambiguities: evidence from self-paced reading. Mem. Cogn. 36, 201–216.

Tunstall, S.L., 1998. The interpretation of quantifiers: semantics and processing. Unpublished doctoral dissertation, University of Massachusetts, Amherst.

Van Berkum, J.J.A., Brown, C.M., Hagoort, P., 1999a. Early referential context effects in sentence processing: evidence from event-related brain potentials. J. Mem. Lang. 41, 147–182.

Van Berkum, J.J.A., Hagoort, P., Brown, C.M., 1999b. Semantic integration in sentences and discourse: evidence from the N400. J. Cogn. Neurosci. 11, 657–671.

Van Berkum, J.J.A., Brown, C.M., Hagoort, P., Zwitserlood, P., 2003. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. Psychophysiology 40, 235–248.

Van Berkum, J.J.A., Koornneef, A.W., Otten, M., Nieuwland, M.S., 2007. Establishing reference in language comprehension: an electrophysiological perspective. Brain Res. 1146, 158–171.

van Dijk, T.A., Kintsch, W., 1983. Strategies in Discourse Comprehension. Academic Press, New York.

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