Processing Pronouns without Antecedents: Evidence from Event-related Brain Potentials

Ruth Filik, Anthony J. Sanford, and Hartmut Leuthold

Abstract

■ Pronouns that do not have explicit antecedents typically cause processing problems. We investigate a specific example in which this may not be the case, as in "At the interview, they asked really difficult questions," where the plural pronoun *they* has no explicit antecedent, yet is intuitively easy to process. Some unspecified but constrained set of individuals (*the interview panel* or *the company*) can be inferred as the referent, but it is not crucial to determine specifically which entities are being referred to. We propose that this contrasts with the processing of singular pronouns (*he* or *she*), for which it is necessary to determine a specific referent. We used event-related brain potentials to investigate how readers process the pronoun (*they* vs. *he/she*) in these cases. Sentences were placed in a context that either did or did not contain an explicit antecedent for the pronoun. There were two key findings. Firstly, when there was no explicit antecedent, a larger fronto-central positivity was observed 750 msec after pronoun onset for *he/she* than *they*, possibly reflecting the additional difficulty involved in establishing a referent for *he/she* than for *they* when no explicit referent is available. Secondly, there was a larger N400-like deflection evoked by *he/she* than *they*, regardless of whether there was an explicit antecedent for the pronoun. We suggest that this is due to the singular pronouns bringing about a greater integration effort than the plural pronoun. This observation adds to a growing body of research revealing fundamental differences in the way these pronouns are handled by the language processor. ■

INTRODUCTION

In order to understand a sentence, a reader or listener must identify the incoming words and assign syntactic structure to them. In addition, it is necessary to construct a coherent mental representation of what is being described. To do this, what is being read must be related to previous sentences and to general knowledge. A major device for achieving this is the use of anaphoric expressions (e.g., pronouns) that refer back to previously mentioned individuals. For example, in the sentence, "When George got home, he decided to watch TV," the pronoun *he* is an anaphor, referring back to its antecedent, *George*. The process of identifying the antecedent of an anaphoric expression is central to language comprehension, and the exploration of how it takes place has given rise to an extensive behavioral literature (cf. Garnham, 2001, for a review; Klin, Guzman, Weingarter, & Ralano, 2006; Van Gompel & Majid, 2004; Kennison, 2003, for a sample of later work). With the current research we investigate pronoun processing using event-related brain potentials (ERPs).

A growing body of evidence suggests that the different steps of language comprehension manifest in qualitatively distinct ERP components, however, this research has so far mainly concentrated on semantic and syntactic processing. For example, semantic violations are known to trigger a negative ERP deflection (N400) with a centro-parietal scalp distribution that typically peaks about 400 msec after word onset. N400 effects are widely considered to reflect semantic integration processes (see e.g., Kutas & Federmeier, 2000; Osterhout & Holcomb, 1995, for overviews), or the predictability of a word within a given context (e.g., Van Berkum, Brown, Hagoort, Zwitserlood, & Kooijman, 2005; Kutas & Hillyard, 1984). However, it should be noted that these two accounts are not mutually exclusive because higher predictability can result in easier integration. In contrast, syntactic violations tend to elicit a centro-parietal late-positive deflection 600 msec or more after word onset, thought to indicate reanalysis or repair processes (P600/"syntactic positive shift" [SPS]; Hagoort, Brown, & Groothusen, 1993), or various frontal negative deflections such as the left-lateralized anterior negativity, thought to reflect the detection of morphosyntactic errors, or phrase structure violations (see e.g., Hagoort, Brown, & Osterhout, 1999, for a review).

Despite the established importance of anaphora, there are relatively few studies that have used ERPs to investigate referential processes (e.g., Ditman, Holcomb, & Kuperberg, 2005; Hammer, Jansma, Lamers, & Münte, 2005; Swaab, Camblin, & Gordon, 2004; Schmitt, Lamers, & Münte, 2002; Streb, Rösler, & Hennighausen, 1999;

University of Glasgow, UK

King & Kutas, 1997; for an overview of earlier work, see Barber, Salillas, & Carreiras, 2004). The major focus of many previous investigations has been situations where there is gender disagreement (semantic or syntactic) between an anaphor and a potential antecedent, and findings from these studies have been mixed. For example, Osterhout and Mobley (1995, Experiment 2) compared the ERPs to personal pronouns that provided either a match or mismatch in gender to the subject noun, such as "The aunt heard that she/he had won the lottery." Mismatching pronouns elicited a P600/SPS for participants who judged the sentences as unacceptable, which was interpreted as a syntactic integration effect. Osterhout, Bersick, and McLaughlin (1997) also found P600/SPS type effects for reflexive pronouns that violated either the definitional or stereotypical gender of their antecedent. In contrast, other studies have found gender mismatching pronouns to elicit an N400 effect (King & Kutas, 1997; see also Kutas, Federmeier, Coulson, King, & Münte, 2000). Importantly, all of this work is based on the assumption that when an anaphor is encountered during reading or listening, it triggers an immediate search for an antecedent that matches in gender and number, and the effects obtained reflect the ERP response to a failure to obtain a match.

With the current article, we are interested in the processing of pronouns that do not have explicit antecedents in the text, and thus, may involve processes above and beyond the simple matching up of pronoun and antecedent. Relevant to this are recent studies on the processing of ambiguous reference (Van Berkum, Brown, Hagoort, & Zwitserlood, 2003; Van Berkum, Brown, & Hagoort, 1999). In Van Berkum et al.'s (1999) study, participants read sentences in Dutch beginning, for example, with "David told the girl that…." These sentences appeared in discourse contexts that had introduced either one possible referent for "the girl," in which case the sentence was unambiguous, or two possible referents, creating a referential ambiguity. Sentences in which *the girl* had two possible referents elicited a larger sustained frontal negativity beginning as early as 280 msec after its presentation, compared to the unambiguous condition (in which there was only one possible referent for *the girl*). A similar pattern of effects was elicited by spoken versions of the same materials (Van Berkum et al., 2003). These findings suggest that the interpretation of an anaphor seems to take place very rapidly. In addition, the observed sustained frontal negativity is clearly distinct from the syntax-related P600/SPS and the semantics-related N400. However, it is yet unclear whether a language-specific process manifests in the frontal negative ERP effect or rather an effect of differential memory load (for a discussion of this issue, see Van Berkum et al., 1999, 2003).

Perhaps most relevant to the current article is an ERP study reported by Van Berkum, Zwitserlood, Bastiaansen, Brown, and Hagoort (2004, see also Van Berkum, Kornneef, Otten, & Nieuwland, 2007). Participants heard pronouns that were either:

1. Referentially successful (one referent)
David shot at Linda as he jumped over the fence.
2. Referentially ambiguous (two potential referents)
David shot at John as he jumped over the fence.
3. Referentially failing (no suitable referents)
Anna shot at Linda as he jumped over the fence.

Van Berkum et al. observed a sustained frontal negativity for referentially ambiguous pronouns (e.g., Example 2), similar to the effects observed for referentially ambiguous nouns (Van Berkum et al., 1999, 2003). Pronouns that had no suitable referents (e.g., Example 3) elicited a standard (parietal) P600/SPS effect. This P600 effect was similar to that obtained by Osterhout et al. (1997) and Osterhout and Mobley (1995). Van Berkum et al. (2007) speculated that in the zero referent "Anna shot Linda as he…" case, syntax may get the blame because the pronoun's morphosyntactic features conflict with the possibly stronger preference for a pronoun to refer to a locally available, foregrounded referent (e.g., Garnham, 2001).

The antecedent for *he* in the "referentially failing" sentences (see Example 3) used by Van Berkum et al. is ultimately unrecoverable, as there is no obvious male character for *he* to refer to in the context. Pronouns are typically used to refer to highly salient entities in the text (e.g., Gundel, Hedberg, & Zacharski, 1993) with it generally being the case that pronouns have explicit antecedents in naturally occurring text (e.g., Erku & Gundel, 1987). However, there are important exceptions to this (Gerrig, Horton, & Stent, 2006; Greene, Gerrig, McKoon, & Ratcliff, 1994; Yule, 1982). In the present article, we are concerned with a commonplace instance of one of these exceptions, in which the plural pronoun *they* can be used felicitously without an explicit antecedent. An example of this is: "The in-flight meal I got was more impressive than usual. In fact, *they* courteously presented the food as well." In these cases, which are intuitively easy to process, some unspecified but constrained individual or set of individuals, (such as "the stewardesses" or "the airline") can be inferred as the referent, but it is not crucial to determine specifically which entity is being referred to. By substituting the pronoun *she* for *they*, the sentence becomes intuitively less acceptable. Data from a recent eye-tracking study (Sanford, Filik, Emmott, & Morrow, 2008) show that if *they* has no antecedent in cases like these, there is no disruption to eye-movement patterns, relative to the case where there is an antecedent, as in, "The in-flight meal I got from the stewardesses was more impressive than usual. In fact, *they* courteously…." However, if the singular *she* is used, there is a large disruption when there is no explicit antecedent for the pronoun, relative to the case where there is an antecedent, as in, "The in-flight meal I got [from the stewardess] was more impressive than usual. In fact, she courteously. . . ." This disruption was observed as increased reading times on the word following the pronoun.

Eye-movement data can indicate when a reader experiences processing difficulty, but does not yield much information concerning the nature (e.g., semantic or syntactic) of the processes involved, whereas different ERP components can distinguish different underlying language processes. Thus, although the eye-tracking data reported in Sanford et al. (2008) suggest that readers experienced processing difficulty for *he* and *she* without an explicit antecedent, but not for *they*, the results do not inform us about the nature of the underlying processes involved in locating or inferring a suitable antecedent for the pronoun. Hence, using the same experimental design as Sanford et al., we aim to use ERPs to investigate the neural mechanisms underlying pronoun resolution in these cases. Our first question is whether singular and plural pronouns with missing antecedents elicit differential ERP effects. One possibility is that although an antecedentless *they* results in no apparent processing difficulty in behavioral measures, its resolution may rely on the same underlying neural mechanism as an antecedentless *he* or *she*, resulting in the same pattern of ERP effects for *he/she* and *they* when there is no explicit antecedent for the pronoun. Another possibility is that ERP data may correspond to the eye-tracking results, with differential ERPs for *he/she* and *they* when there is no explicit antecedent for the pronoun.

It is necessary to consider how these effects may manifest themselves in the ERPs, and specifically, whether our manipulation will produce an effect that maps onto any of the established ERP markers of language processing. There are a number of possibilities. Firstly, because a pronoun without an explicit antecedent may be more difficult to integrate into a reader's semantic representation than a referentially successful one, differences between *he/she* and *they* could be observed in the N400 range. Secondly, we may expect differences in the parietal P600/SPS, as found previously (Van Berkum et al., 2004; Osterhout et al., 1997; Osterhout & Mobley, 1995), if processes involved in inferring an antecedent are similar to those involved in detecting a mismatch between pronoun and antecedent, because in the latter case, the mismatch may force inferring what the intended antecedent might be. Thirdly, these differences may be observed in a sustained frontal negativity, such as that associated with referentially ambiguous nouns and pronouns (e.g., Van Berkum et al., 1999, 2003, 2004), which in our study may reflect the relative difficulty in establishing a suitable antecedent for the pronoun. Finally, to examine whether *he/she* versus *they* and the presence or absence of an antecedent influence information processing already at an early stage of lexical access, we determined N1 amplitudes as the latter have been consistently shown to sensitively indicate influences on lexical processing (e.g., Hauk & Pulvermüller, 2004; Sereno, Brewer, & O'Donnell, 2003; Sereno, Posner, & Rayner, 1998; for a review, see Sereno & Rayner, 2003). We also determined P1 amplitudes as a control for possible word length effects (Hauk & Pulvermüller, 2004), that is, an absence of a P1 effect would rule out the possibility that later occurring effects are merely a reflection of an initial effect of shorter word length for *he/she* than *they* sentences.

If it is, indeed, the case that *he/she* and *they* are processed differently when they have missing antecedents, this would raise the more general issue of whether there may be processing differences between the two classes of pronoun irrespective of whether an antecedent is present. That is, whether there are fundamental differences in the types of antecedent search procedures set up by the two types of pronoun. This is an important issue which, to our knowledge, has not been studied using ERPs. There is, however, considerable evidence from eye tracking that antecedent matching may differ for *they* and *he/she*, such that the search process for singular pronouns draws on more processing resources than that for *they* (Sanford et al., 2008; Sanford & Filik, 2007; Moxey, Sanford, Sturt, & Morrow, 2004). Thus, our second question is whether there will be any other ERP differences observed for *he/she* versus *they*, regardless of whether or not there is an antecedent for the pronoun. However, it is not yet clear in which way this difference may emerge in the ERP.

METHODS

Participants

Twenty-one right-handed native English speakers (11 men, 10 women) from the Glasgow University community were paid to take part.

Materials and Design

We constructed 160 sentence pairs (available from the first author) based on those used by Sanford et al. (2008). Each trial consisted of a context sentence followed by a critical target sentence (see Table 1 for an example).

The experiment used a 2 (antecedent/no antecedent) × 2 (plural/singular pronoun) design. The first sentence of each item depicted a simple event in which an antecedent for the pronoun in the target sentence was either introduced explicitly, or left implicit. In the singular version, the explicit antecedent was a singular agent (e.g., *stewardess*), and in the plural version, it was a plural agent (e.g., *stewardesses*). The second sentence included either a singular (*he/she*) or a plural (*they*) pronoun, which was intended to co-refer with the agent

Table 1. Example Material

Antecedent He/she	The in-flight meal I got from the stewardess was more impressive than usual. In fact, she courteously presented the food as well.
No antecedent He/she	The in-flight meal I got was more impressive than usual. In fact, she courteously presented the food as well.
Antecedent They	The in-flight meal I got from the stewardesses was more impressive than usual. In fact, they courteously presented the food as well.
No antecedent They	The in-flight meal I got was more impressive than usual. In fact, they courteously presented the food as well.

in the first sentence. The pronoun always appeared in subject position, and was always the third word of the second sentence. In the case of singular pronouns, the choice of *he* or *she* was determined by stereotypical gender biases for the intended agent. The remainder of the target sentence following the pronoun was identical across experimental conditions.

Items were arranged in four different stimulus presentation files. Each item appeared in only one of its four possible conditions in a given file, but appeared in all conditions over the four files. A given file comprised 40 materials in each of the four conditions. Thus, each participant viewed 160 experimental items, 40 in each condition. Each file also included 160 filler items of a similar length to the test materials, of which six served as practice items. An example filler item would be, ''The trains had been running late over the festive period. Many customers sent letters of complaint to the station.'' All fillers were grammatically well formed and did not contain any anomalies. Their purpose was to prevent participants from noticing regularities in the experimental materials, and thus, potentially adopting strategies while reading them, rather than reading normally. Experimental and filler items were presented in a fixed pseudorandom order, such that no more than two experimental items appeared in a row.

Participants were tested in an electrically shielded booth with ambient light kept at a low level. Word stimuli were presented in white 16-point Helvetica font on a black background at the center of a 15-in. computer monitor at a viewing distance of 90 cm.

Procedure

Participants were informed about the electroencephalogram (EEG) procedure and experimental task. After electrode application, they were seated in a booth where they read the materials from a computer screen. There were six practice trials to familiarize them with the procedure, after which the experimenter answered any questions. Then the first block of 34 experimental trials was presented followed by seven blocks of 40 trials. Blocks were separated by a break, the duration of which was determined by the participant.

The trial sequence was as follows. Each trial started with the presentation of the first sentence of each sentence pair. Participants pressed the spacebar on a computer keyboard when they had finished reading it. A blank interval of 2500 msec followed, after which a fixation cross was presented in the center of the screen for 1000 msec. Then the word-by-word presentation of the second sentence started, during which participants were asked to maintain fixation at the center of the screen. Each word was displayed centrally for 300 msec, with 200-msec blank intervals between successive word presentations. A break of 2500 msec separated each experimental trial. There was no secondary task.

Electrophysiological Measures

A BIOSEMI Active-Two amplifier system was used for continuous recording of EEG activity from 72 Ag/AgCl electrodes over midline electrodes Fpz, AFz, Fz, FCz, Cz, CPz, Pz, POz, Oz, and Iz, over the left hemisphere from electrodes IO1, Fp1, AF3, AF7, F1, F3, F5, F7, F9, FC1, FC3, FC5, FT7, C1, C3, C5, M1, T7, CP1, CP3, CP5, TP7, P1, P3, P5, P7, PO3, PO7, O1, two nonstandard positions PO9′ and O9′, which were located at 33% and 66% of the M1–Iz distance, respectively, and from the homologue electrodes over the right hemisphere. EEG and electrooculogram (EOG) recordings were sampled at 256 Hz. The on-line reference electrode was the Biosemi Common Mode Sense (CMS) electrode (see www.biosemi.com/faq/cms&drl.htm for details). Off-line, all EEG channels were recalculated to an averaged mastoid reference. Trials containing blinks were corrected using a dipole approach (BESA Version 5.1.6). Automatic artifact detection software (BESA) was run and trials with nonocular artifacts (drifts, channel blockings, EEG activity exceeding ±120 μV) were discarded. The analysis epoch started 200 msec prior to the onset of the pronoun and lasted for a total duration of 1500 msec.

Data Analysis

For artifact-free trials, the signal at each electrode site was averaged separately for each experimental condition time-locked to the onset of the pronoun. Before the measurement of ERP parameters, EEG and EOG activity was band-pass filtered (0.3–25 Hz, 6 dB/oct).[1] The ERP waveforms were aligned to a 200-msec baseline prior to the onset of the pronoun. Peak amplitudes were determined for P1 and N1 at scalp electrodes where these deflections were maximal, using a computerized peak-picking program and predefined 60-msec search intervals: P1 (O2, 90–150 msec and 590–650 msec) and N1 (PO7, 140–200 and 640–700 msec). Moreover, we measured mean ERP amplitudes in typical time intervals 200–250 and 300–500 msec for P2 and N400-like effects, respectively, and as guided by visual inspection of experimental effects in ERP waveforms, during a 720–770 msec interval relative to pronoun onset.

ERP amplitudes at midline electrodes were analyzed separately from data recorded over lateral electrode sites. Lateral electrode sites were pooled to form regions of interest (ROIs) as recommended for the analysis of high-density electrode arrangements (cf. Dien & Santuzzi, 2005). The electrodes were divided along a left–right dimension, an anterior-to-posterior dimension, and a dorsal–ventral dimension. The six ROIs over the left hemisphere were: left anterior ventral (AF7, F7, FT7, F5, FC5), left anterior dorsal (AF3, F3, FC3, F1, FC1), left central ventral (TP7, T7, C5, CP5), left central dorsal (C3, CP3, C1, CP1), left posterior ventral (PO9′, O9′, P7, PO7, O1), and left posterior dorsal (P3, PO3, P1, P5); six homologue ROIs were defined for the right hemisphere.

Statistical analyses were performed by means of Huynh–Feldt corrected repeated measures analyses of variance (ANOVAs). For the analysis of ERP amplitude data recorded from midline electrodes, we performed an ANOVA with variables antecedent (yes, no), pronoun (*he/she*, *they*), and electrode (Fz, FCz, Cz, CPz, Pz). For the analysis of ERP deflections maximal over lateral electrode sites, we performed an ANOVA with variables antecedent (yes, no), pronoun (*he/she*, *they*), hemisphere (left, right), ant–pos (anterior, central, posterior), and verticality (ventral, dorsal).

RESULTS

Electrophysiological Data

Figure 1 depicts the grand-average ERP waveforms for the four combinations of two antecedent (antecedent vs. no antecedent) conditions by two pronoun types (*he/she*, *they*). The difference waveforms reflecting the pronoun effect (ERP[*he/she*] minus ERP[*they*]) and the antecedent effect (ERP[no antecedent] minus ERP[antecedent]) are displayed in Figures 2 and 4, respectively.

The analysis of ERP amplitudes for lateral ROIs indicated an initial positivity (P1) at about 120 msec that was larger over right than left occipito-parietal electrodes [$F(2, 40) = 7.6$, $p < 01$], but not reliably influenced by experimental conditions as a main effect (all Fs < 1), or in interaction with topographic factors (all Fs < 1.7, ps $>$ .19). The following negativity (N1) at about 170 msec was most pronounced over left parieto-occipital electrodes as indicated by the Ant–Pos × Hemisphere × Verticality interaction [$F(2, 40) = 5.1$, $p < .05$], and tended to be influenced by pronoun type [$F(1, 20) = 3.2$, $p < .10$]. Antecedent condition did not influence N1 amplitude as a main effect ($F = 2.2$, $p > .15$), nor in interaction with topographic factors (all Fs < 2.1, ps $>$ .13). A separate N1 amplitude analysis including only posterior ROIs, where N1 was maximal, indicated larger negativity for *he/she* than *they* sentences (−3.0 vs. −2.4 μV) [$F(1, 20) = 6.3$, $p < .05$]. Subsequently, a pronoun effect showed up over midline electrodes in

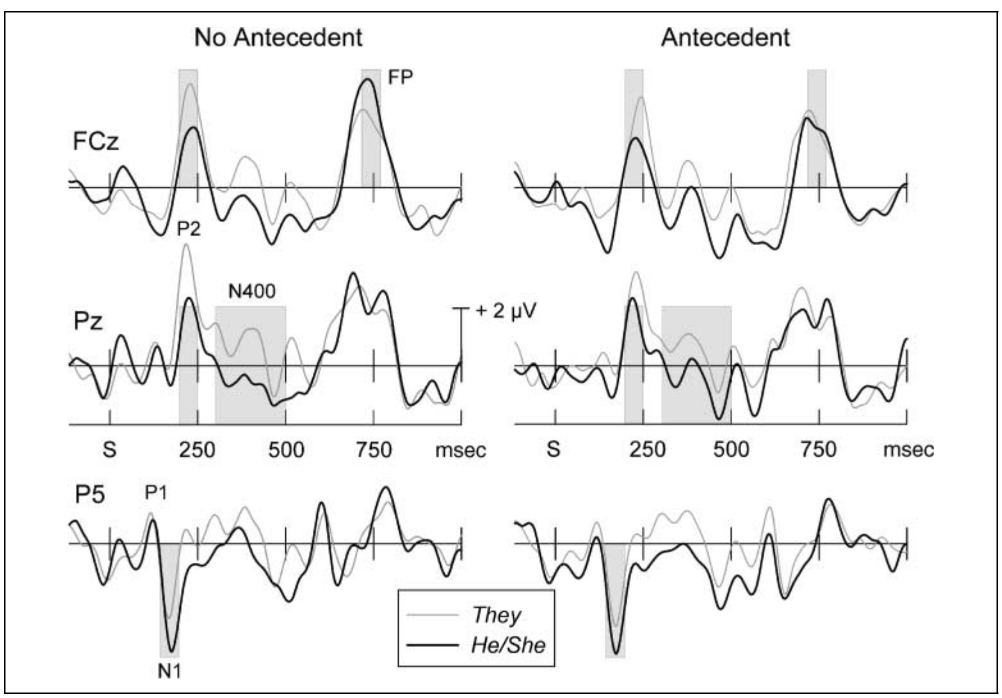


Figure 1. Grand-average ERPs at electrodes FCz, Pz, and P5 following pronoun onset (S) and onset of the word following the pronoun (500 msec) as a function of pronoun (*they* vs. *he/she*) and antecedent (left vs. right panel: no antecedent vs. antecedent). Shaded areas indicate time ranges at electrodes where N1, P2, the N400-like negativity, and the fronto-central positivity (FP) were most pronounced.

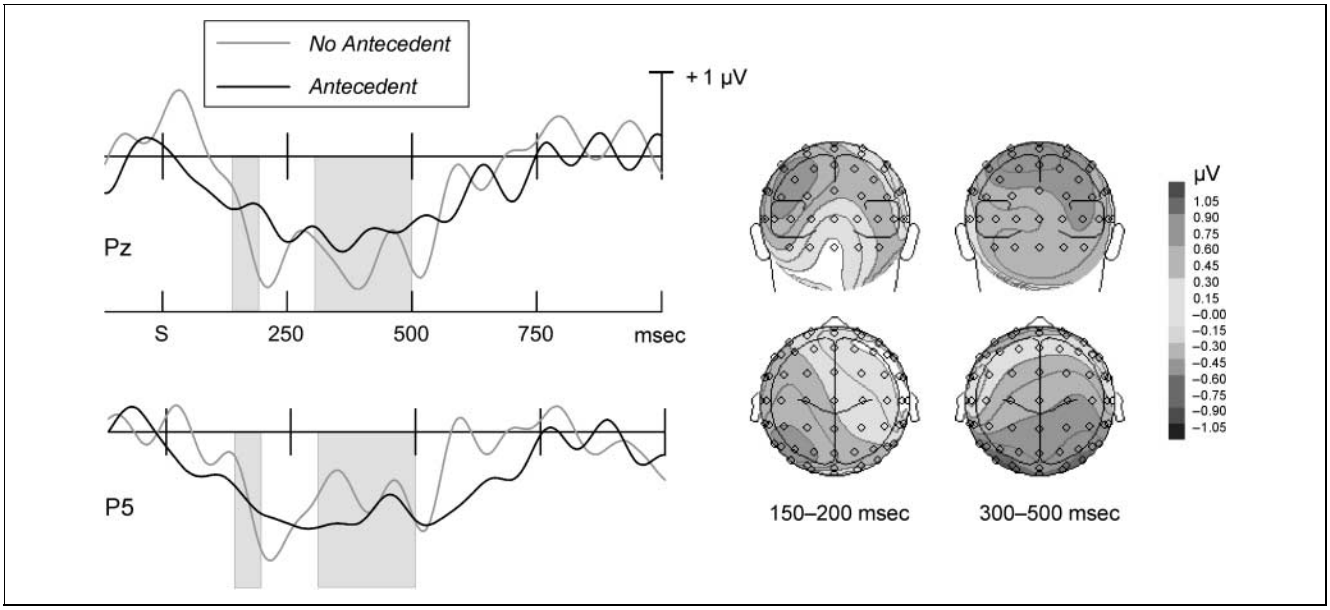


Figure 2. (Left) ERP difference waveforms reflecting the pronoun effect (*he/she* minus *they*) at electrodes Pz and P5, separately for sentences with and without antecedents. (Right) Spline-interpolated topographic maps of the pronoun effect for time intervals 150–200 and 300–500 msec after pronoun onset. Isopotential line spacing = 0.15 µV.

the centro-parietal P2 (cf. Figure 1) [$F(4, 80) = 5.4$, $p < .01$] as a smaller positivity for *he/she* than *they* sentences (2.7 vs. 1.7 µV) [$F(1, 20) = 7.2$, $p < .05$]. This P2 pronoun effect is in the opposite direction compared the one in N1, suggesting that the earlier larger negativity for *he/she* than *they* sentences extends into the P2 time range. P2 amplitude was not reliably influenced by antecedent ($F = 2.7$, $p > .11$), nor was in interaction with pronoun or electrode (all Fs < 1.1).

P2 was followed by a slowly growing negative-going deflection. In the time interval 300–500 msec after pronoun onset, the widely distributed negative-going ERP waveform was more negative for *he/she* than *they* sentences (−0.6 vs. 0.2 µV) [$F(1, 20) = 16.1$, $p < .001$]. The main effect of antecedent and all interactions including this variable were not significant (all Fs < 1.5, ps $> .21$). The analysis of ERP amplitudes at lateral ROIs confirmed the main effect of pronoun [$F(1, 20) = 14.9$, $p < .001$], which was more pronounced over dorsal rather than ventral ROIs [$F(1, 20) = 16.5$, $p < .001$], and tentatively over left centro-parietal ROIs as indicated by the Pronoun × Hemisphere × Ant–Pos interaction [$F(2, 40) = 3.2$, $p = .05$; cf. Figure 2], consistent with an N400-like effect.

To examine whether the pronoun effect on the ERP waveform (N400) is present for both *he* and *she*, we analyzed ERP data separately as a function of the three pronoun types. That is, ERP data were collapsed across antecedent conditions in order to achieve a reasonable number of trials in average ERP waveform of *he* and *she* sentences. Figure 3 depicts the grand-average ERP waveforms for the three pronoun types (*they*, *he*, *she*). An ANOVA of mean ERP amplitude (300–500 msec) with variables pronoun (*they*, *he*, *she*) and electrode (Fz, FCz, Cz, CPz, Pz) revealed larger negativity for both *he* (−1.0 µV) and *she* (−0.9 µV) as compared to *they* sentences (0.2 µV) [$F(2, 40) = 4.9$, $p < .05$]. The difference in the N400-like effect between *he/she* and *they* constitutes a novel observation that we believe to be an important reflection of integration differences, which we shall discuss shortly.

As for the pronoun, the presentation of the following word ($t = 500$ msec) triggered a similar series of positive and negative ERP deflections (P1, N1, P2; cf. Figure 1). Overlap with the sustained pronoun-related negativity was indicated over lateral ROIs by less positive P1 (0.1 vs. 1.0 µV) and more negative N1 peak amplitudes (−1.5 vs. −0.7 µV) for *he/she* than *they* sentences [all Fs$(1, 20) > 15.4$, ps $< .01$], whereas the main effect of pronoun was not significant over midline electrodes for the subsequent centro-parietal positivity peaking at about 200 msec (P2) [$F(1, 20) = 2.9$, $p > .10$].

Most importantly, at about 750 msec following pronoun onset (i.e., 250 msec after the onset of the following word), a larger fronto-central positivity (FP) showed up when no explicit antecedent was provided for *he/she* sentences as compared to the other experimental conditions (cf. Figures 1 and 4). Analysis of mean ERP amplitude over midline electrodes in the time interval 720–770 msec indicated a fronto-central positivity [$F(2, 40) = 5.4$, $p < .05$] that was of marginally higher amplitude for no antecedent than antecedent conditions (2.6 vs. 2.0 µV) [$F(1, 20) = 5.3$, $p < .10$]. The Pronoun × Electrode interaction [$F(2, 40) = 5.4$, $p < .05$] and the Antecedent × Pronoun × Electrode interaction were significant [$F(2, 40) = 2.7$, $p < .05$]. As can be seen in Figure 5, the latter interaction was due to a 1-µV stronger

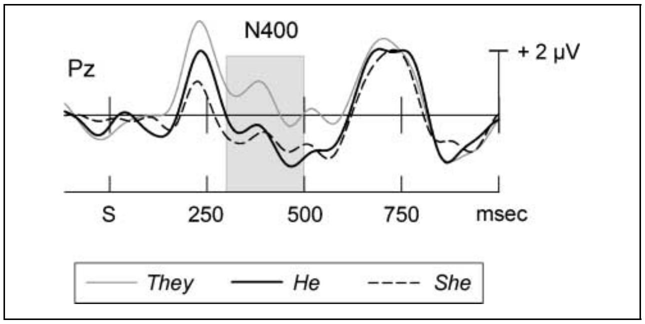


Figure 3. Grand-average ERPs at electrode Pz following pronoun onset (S) as a function of pronoun (*they* vs. *he* vs. *she*). The shaded area indicates the time range when the N400-like negativity was most pronounced.

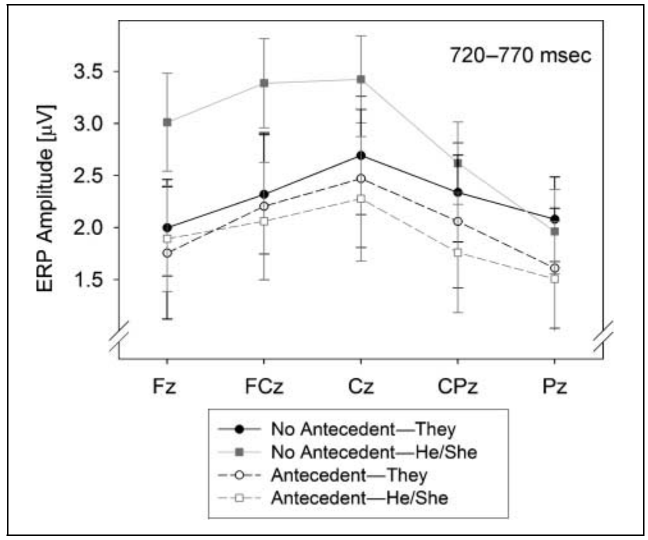


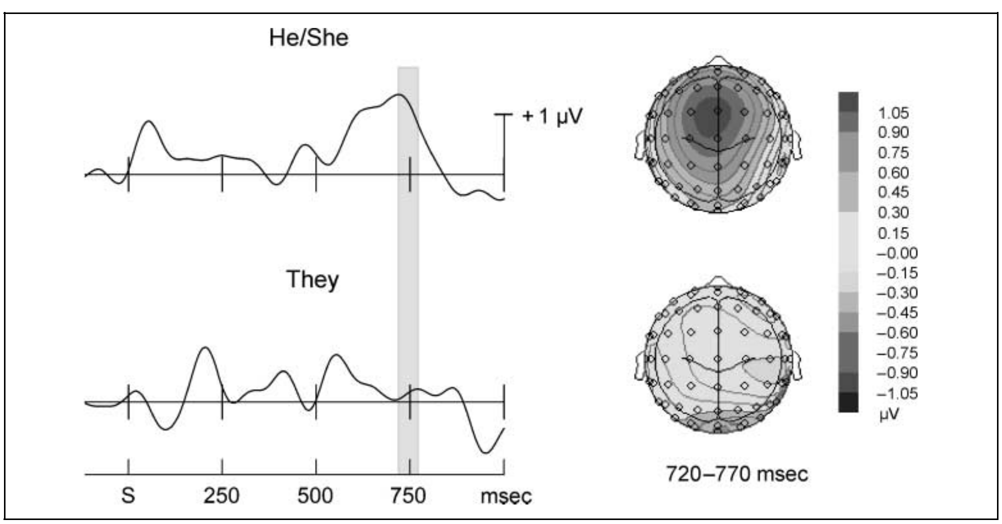
Figure 5. Mean ERP amplitudes over midline electrode sites for time range 720–770 msec after pronoun onset as a function of antecedent (no antecedent vs. antecedent) and pronoun (*they* vs. *he/she*). Error bars indicate standard errors of the mean.

positivity, mainly over fronto-central electrodes (cf. Figure 4), in *he/she* sentences with no antecedent as compared to *he/she* sentences with antecedent and *they* sentences with and without antecedent [all $Fs(2, 40) > 3.7$, $ps < .05$]. The ROI analysis confirmed this effect pattern over bilateral fronto-lateral electrodes as indicated by the significant Antecedent × Pronoun × Ant–Pos interaction [$F(2, 40) = 4.2$, $p < .05$].

DISCUSSION

The results showed two key and novel findings. First, differential effects for singular (*he/she*) versus plural (*they*) pronouns began very early after the onset of pronoun presentation, emerging as a larger negativity for *he/she* than *they* sentences. This pronoun effect was already present at the N1 deflection,[2] which suggests, in conjunction with other N1 findings (cf. Hauk & Pulvermüller, 2004), that *he/she* versus *they* differ very early during word processing, presumably at a level associated with lexical access (cf. Sereno & Rayner, 2003), by which we mean the process of initiating the procedure associated with a word. Following the stronger posterior negativity triggered by *he/she* than *they* sentences in the N1 time range (150–200 msec), the pronoun effect was most pronounced as a sustained negativity in the N400 time range, maximal over centro-parietal electrodes. We believe this pronoun-related influence on the ERP waveform to be related to the classical N400 effect. Furthermore, additional analyses revealed that the size of the negative effect evoked by *he* and *she* was the same, with both being significantly larger than the N400 evoked by *they*. Although word frequency and word length are known to influence the N400, it is unlikely that the pronoun-related ERP effects obtained here are due to these basic features. A word length effect is highly unlikely as all pronouns are fairly short in length. Moreover, whereas the P1 component has been found to sensitively reflect word length

Figure 4. (Left) ERP difference waveforms reflecting the antecedent effect (*no antecedent* minus *antecedent*) at the FCz electrode separately for *he/she* and *they* sentences. (Right) Spline-interpolated topographic maps of the antecedent effect for *he/she* sentences for the time interval 720–770 msec after pronoun onset. Isopotential line spacing = 0.15 μV.



effects (e.g., Hauk & Pulvermüller, 2004), the present study showed no modulation of P1 amplitude for *he/she* versus *they*. Word frequency, known to influence N400 amplitudes (Carreiras, Vergara, & Barber, 2005; Allen, Badecker, & Osterhout, 2003; Van Petten & Kutas, 1990, 1991), is an equally unlikely account of the greater negativity for *he/she* than *they* because all the pronouns have a high frequency of occurrence. Furthermore, *he* is more frequent than *she* and *they* [*he* = 8828 words per million (wpm), *she* = 4123 wpm, *they* = 4929 wpm; Baayen, Piepenbrock, & Gulikers, 1995], which, if anything, would lead to an opposite pattern of N400 results, as lower-frequency words have been shown to result in larger negativities. An analysis of repetition frequency of each pronoun also ruled out that the N400-like effect is simply due to *he* or *she* being repeated less often during the experiment than *they*.[3]

We propose that a more likely explanation of the different patterns of negativity is due to stable and inherent differences in how *he/she* and *they* are dealt with within the language processor. Our proposal is that both singular and plural pronouns initiate procedures that seek antecedents, and that they do this as soon as they are encountered, regardless of whether there is an explicit antecedent present in the prior discourse. However, there are distinctions between the singulars and plurals in terms of the computational resources allocated to the search process. In the case of singulars, incremental semantic interpretation is disrupted while an antecedent is sought, and disruption is great if an antecedent cannot be rapidly found. Effectively, singular pronouns require an explicit antecedent, and require one immediately. The search procedure associated with plural pronouns, on the other hand, is less demanding of computational resources, with less disruption to semantic interpretation processes while the search for a referent is underway, so there is no demand for immediate resolution. We believe that the N400-like effect is a direct reflection of differences in the initial computational resources or ''integration effort'' put into seeking an antecedent.

Previous research using eye tracking offers support for the proposal that there are differences in the antecedent search procedures associated with *they* and *he/she*, specifically, that under certain circumstances, plural pronouns do indeed cause less disruption to processing while a potential antecedent is being sought (e.g., Sanford & Filik, 2007; Moxey et al., 2004; Foertsch & Gernsbacher, 1997). For example, Moxey et al. (2004) found that antecedent-mismatch effects for plural pronouns were delayed compared to singulars. They offered the explanation that whereas singular pronouns are resolved immediately, the antecedents of plural pronouns may be initially underspecified, and only resolved later on. Such a delay may not be unreasonable as *they* typically can refer to a wider range of antecedent types than *he/she* can, and has a broader time period over which a potential antecedent might be found. Whereas Moxey et al. found evidence of delayed processing difficulty with *they* compared to *he/she*, we found no evidence of processing difficulty for *they* at all. However, in the Moxey et al. study, it was necessary for readers to ultimately resolve who *they* were (even if it was initially underspecified), whereas in the current study we argue that it is acceptable for the antecedent of *they* to remain underspecified. Earlier accounts of the mechanisms of pronoun resolution have emphasized the immediacy of a search for an antecedent, and the disruption to more general processing that results from difficulties in finding antecedents, but these accounts have been examined entirely in the context of singular pronouns such as *he* and *she* (e.g., Garrod, Freudenthal, & Boyle, 1994), and the different nature of *they* has thus far scarcely been studied in any detail.

These N400-like effects were independent of the presence of an antecedent, consistent with our argument that they reflect the launch of different basic search processes. However, the second major finding in the present work reflected different consequences of the presence or absence of an antecedent on singular and plural pronouns. There was a positive ERP deflection, maximal over fronto-central midline electrodes, peaking at about 750 msec after pronoun onset. This positivity was greater following *he/she* sentences with no explicit antecedent compared to all the other conditions, and mirrors the pattern of effects found by Sanford et al. (2008) in their eye-tracking study. This is consistent with the idea that although the procedure associated with *they* does not need to find an explicit antecedent early after the search is launched, *he/she* does, as argued earlier. Although it was possible that ERP and eye-tracking results might differ, with ERP evidence for some sort of processing difficulty being associated with *they* in the absence of an antecedent, this was not the case.

Relating back to our predictions, the fronto-central positivity effect in the ERP difference waveform (cf. Figure 4) is clearly distinct from language-related ERP components with negative polarity, such as the N400, and the sustained frontal negativity reported by Van Berkum et al. (1999, 2003, 2004). Also, given the fronto-central topography, this effect provides a marked contrast with the parietal P600/SPS effect found by Van Berkum et al. (2004, see also Van Berkum et al., 2007) in response to pronouns with no suitable antecedent in the text, and found by Osterhout et al. (1997) and Osterhout and Mobley (1995) for pronouns with antecedents that mismatched in gender. This difference in the topographic distribution informs recent discussions on proposed functional differences of anterior versus posterior P600 effects. It has been suggested that the P600 comprises a group of components which have differing topographies proposed to reflect different underlying processes. Within this framework, a posterior P600 may be an index of syntactic integration and repair processes, whereas a frontal P600 may reflect processes involved in reanalysis (e.g., Dwivedi, Phillips, Lague-Beauvais, & Baum, 2006; Kaan & Swaab, 2003; Friederici, Hahne, & Saddy, 2002; Hagoort et al., 1999).

Of particular relevance to the current finding is the study of Kaan and Swaab (2003), who report a so-called FP600, a frontal positivity emerging between 500 and 900 msec. They argued that this effect is distinct from the parietal P600, which emerges between 500 and 1100 msec, with the FP600 reflecting ambiguity resolution, and/or an increase in discourse-level complexity. Kaan and Swaab's FP600 emerged between 500 and 900 msec, with difference waves showing an effect in the 700–900 msec interval. This is a similar time interval to the antecedent effect observed in difference waves here (see Figure 4). In our study, the difference wave for *he/she* sentences differed reliably from zero in the 600–800 msec interval at electrodes Fz and FCz [$ts(20) > 2.1$, $ps < .05$]. As can be seen in ERP difference waveforms reflecting the antecedent effect for *they* sentences (cf. Figure 4, left panel), here a positivity appeared before 700 msec. However, this positivity was not significantly different from zero for the 600–800 msec interval (all $ts < 1.04$, $ps > .31$), nor the 550–650 msec time interval for which it was larger (all $ts < 1.54$, $ps > .13$). Nevertheless, when the *they* condition was included in the analysis with the *he/she* condition, statistically, the clearest effects were found between 720 and 770 msec.

Kaan and Swaab found larger FP600 effects for referentially more complex sentences. Consider the sentence, "I cut the cake beside the pizzas that were bought by Gill," which contains several entities that need to be added to the discourse model (i.e., cake and pizzas). Compare this with, "The man in the restaurant doesn't like the hamburgers that are on his plate." Here the sentence is simpler in terms of the number of discrete entities that have to be accommodated at the verb (i.e., only hamburgers). Kaan and Swaab suggested that setting up two discourse entities may be more effortful than just setting up one (Garrod & Sanford, 1994). They suggest that the FP600 may reflect the greater effort of setting up extra entities in the discourse representation.

However, it should be noted that Kaan and Swaab's materials were syntactically ambiguous as well as referentially ambiguous, hence, they suggested that further research is needed to identify the cognitive processes of which these different P600 effects are indices. Our results can inform this debate, contributing further evidence that the FP600 may reflect the difficulty involved with adding entities to the discourse model, in line with the results of Kaan and Swaab. Specifically, a greater fronto-central positivity 750 msec after pronoun onset for *he/she* sentences with no explicit antecedent may reflect the discourse complexity involved in inferring and adding an entity to the discourse model in this condition. This, we suggest, is because *he* and *she* set up procedures that require the specification of a specific antecedent token. In the case of *he/she* and *they* conditions, in which there is already an explicit antecedent in the context, no new entity needs to be introduced into the discourse model on encountering the pronoun. In the *they* case, where there is no explicit antecedent, we suggest, on the basis of previous studies, that no explicit new entity is introduced into the discourse model on encountering the pronoun (Sanford et al., 2008). Rather, a mapping between an implied antecedent and a representation of the situation depicted is made, but no attempt is made to identify a specific entity. For instance, given the sentence, "The meal on the flight was bad, and they served it cold," we suggest that the situation, being stereotyped, enables a mapping between a default agent that could be interpreted as *the airline*, or *the stewardesses*, and so forth, but where it is unimportant to actually determine which entity is specifically being referred to.

One further possibility is that the fronto-central positivity found here reflects a P2 on the word following the pronoun. However, there is no overall agreement over the functional significance of the P2 component for information processing, which makes it difficult to relate present fronto-central positivity findings to known modulations of the P2. For example, an anterior P2 has been reported to be larger to open-class than closed-class words (e.g., Brown, Hagoort, & ter Keurs, 1999) and action words (e.g., Preissl, Pulvermüller, Lutzenberger, & Birbaumer, 1995), whereas a left-lateralized anterior P2 was found to have a larger amplitude for unambiguous verbs as compared to nouns (e.g., Federmeier, Segal, Lombrozo, & Kutas, 2000). Moreover, Brown et al. (1999) argued that their P2 effects might be due to overall word length differences (see Hauk & Pulvermüller, 2004) and/or possible visual–spatial attention effects rather than language-related effects. In our study, the word following the pronoun was identical across experimental conditions, hence, ruling out an explanation in terms of a word-specific effect. Alternatively, one might argue that the word following *he/she* without antecedent received more attention. If, indeed, more attention had been allocated to the word following an antecedentless *he/she*, we would have expected sensory potentials P1/N1 to reveal such an effect as well. However, P1/N1 were uninfluenced by antecedent conditions. We therefore believe it to be more likely that our fronto-central positivity was elicited by the pronoun itself, rather than the following word.

Before concluding, it is necessary to discuss some methodological issues. It is possible that the way in which materials were presented (i.e., one word at a time) might have affected the results obtained. Unlike normal reading, presentation rate was relatively slow (one word per 500 msec, which is typical for an ERP study, compared to normal reading rate of ∼5 words per second), and not under the control of the reader. However, there are numerous examples of studies using this procedure revealing comparable effects to those found during auditory sentence presentation at a normal input speed (see, e.g., Van Berkum et al., 2003, compared to Van Berkum et al., 1999; Federmeier & Kutas, 1999, compared to Federmeier, McLennan, De Ochoa, & Kutas, 2002), thus it may be reasonable to expect similar results with auditory sentence presentation. In the current study, participants were instructed to read normally, and for comprehension. To ensure that they remained alert throughout, they were given regular breaks. Comprehension questions were not used, in order to keep the task as close to normal reading as possible, and to avoid the possibility that participants may adopt reading strategies based on what they think the questions might be probing. Crucially, the presence of reliable effects would suggest that participants had been engaged in the task.

In conclusion, for singular pronouns *he* and *she*, a missing antecedent is a problem, producing an increased fronto-central positivity peaking at about 750 msec after pronoun onset. We suggest that this positivity reflects a process of adding a new, inferred, discourse referent to the ongoing discourse model, which does not occur for *they*. This pattern of effects matches that found in eye tracking (e.g., Sanford et al., 2008), thus providing converging evidence strongly supporting the view that the pronoun *they* does not create processing disruption in the case where it has no explicit antecedent, provided it is possible to link the pronoun to a situation. Thus, the data support a fundamental distinction between the behavior of the pronouns *he/she* and *they*. The additional finding of a larger negativity in the 300–500 msec interval for *he/she* than *they*, regardless of whether there is an explicit antecedent for the pronoun, clearly supports this distinction, and offers a way of elucidating it. We suggest that the larger negativity found for *he/she* reflects a difference in the immediate resources made available to aid the integration effort on encountering *he/she* and *they*. *He/she* led to resource-consuming procedures being launched to determine a specific antecedent: If one is not recovered, then processing disruption occurs until one is inferred. This process of inferring a specific antecedent is then reflected in the late fronto-central positivity. The procedure associated with *they* is not so resource consuming, and in the case of a missing explicit antecedent, with *they*, no explicit entity is introduced into the discourse model so there is no resultant late fronto-central positivity. Taken together with the purely behavioral findings of Sanford et al. (2008), we believe that there is strong evidence for a distinction in the processes brought to bear when *he/she* and *they* are encountered in text.

Acknowledgments

This research was supported by EPSRC grant GR/S76434/01 and ESRC grant RES-062-23-0002. We thank four anonymous reviewers for their helpful comments and suggestions.

Reprint requests should be sent to Ruth Filik, Centre for Cognitive Neuroimaging, Department of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, Scotland, UK, or via e-mail: r.filik@psy.gla.ac.uk.

Notes

1. The high-pass filter was set to 0.3 Hz in order to reduce large drifts in the ERP signal in some of the participants. An analogous analysis with the low-pass filter set to 0.01 Hz did not show any additional slow-wave effects.

2. One reviewer's concern was that the pronoun effect during the N1 time interval might be due to residual EEG alpha activity overlapping with the ERP waveform. Whereas we cannot completely rule out this possibility, we consider this possibility unlikely for the following reasons. Firstly, EEG alpha power (8–12 Hz), as determined from ERPs that were transformed to the frequency domain with fast Fourier transform, was symmetrically distributed over posterior scalp electrodes, with a power maximum at Pz. Secondly, an analysis of low-pass filtered (8 Hz) ERP data still indicated a reliably larger N1 for *he/she* than *they* sentences over posterior electrodes (−2.4 vs. −1.6 μV) [$F(1, 20) = 5.9, p < .05$].

3. Across the four stimulus files, *they* appeared an average of 96 times, *he* an average of 89 times, and *she* an average of 55 times in each file, including the filler/nonexperimental trials. It is apparent from this that *they* and *he* appeared a roughly equal number of times, with *she* appearing less frequently. In contrast, the ERP results show that during the 300–500 msec interval *he* and *she* elicit a similar sized negativity, whereas it is significantly smaller for *they*. Therefore, we do not believe that it is the number of times that each pronoun is repeated that underlies this N400-like effect. Furthermore, if we examine the number of experimental trials which are immediately preceded by a trial containing the same pronoun (i.e., an immediate repetition effect), we find that *they* is immediately repeated an average of 6.5 times, *he* an average of 7.5 times, and *she* 3.75 times in each stimulus file. Again, this would not appear to match the pattern of the N400-like effect.

REFERENCES

Allen, M., Badecker, W., & Osterhout, L. (2003). Morphological analysis in sentence processing: An ERP study. *Language and Cognitive Processes, 18,* 405–430.

Baayen, R. H., Piepenbrock, R., & Gulikers, L. (1995). *The CELEX lexical database* [CD-ROM]. Philadelphia: University of Pennsylvania, Linguistic Data Consortium.

Barber, H., Salillas, E., & Carreiras, M. (2004). Gender or genders agreement? In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension; eye-tracking, ERPs and beyond* (pp. 309–328). Brighton, UK: Psychology Press.

Brown, C. M., Hagoort, P., & ter Keurs, M. (1999). Electrophysiological signatures of visual lexical processing: Open- and closed-class words. *Journal of Cognitive Neuroscience, 11,* 261–281.

Carreiras, M., Vergara, M., & Barber, H. (2005). Early event-related potential effects of syllabic processing during visual word recognition. *Journal of Cognitive Neuroscience, 17,* 1803–1817.

Dien, J., & Santuzzi, A. M. (2005). Application of repeated-measures ANOVA to high-density ERP datasets: A review and tutorial. In T. Handy (Ed.), *Event related potentials: A methods handbook*. Cambridge: MIT Press.

Ditman, T., Holcomb, P., & Kuperberg, G. (2005). Examining anaphor resolution using event related potentials. *Journal of Cognitive Neuroscience, 16(Suppl.),* 228.

Dwivedi, V. D., Phillips, N. A., Lague-Beauvais, M., & Baum, S. R. (2006). An electrophysiological study of mood, modal context, and anaphora. *Brain Research, 1117,* 135–153.
Erku, F., & Gundel, J. K. (1987). The pragmatics of indirect anaphors. In J. Verschueren & M. Bertuccelli-Papi (Eds.), *The pragmatic perspective: Selected papers from the 1985 international pragmatics conference* (pp. 533–545). Amsterdam: John Benjamins.
Federmeier, K. D., & Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language, 41,* 469–495.
Federmeier, K. D., McLennan, D. B., De Ochoa, E., & Kutas, M. K. (2002). The impact of semantic memory organisation and sentence context information on spoken language processing by younger and older adults: An ERP study. *Psychophysiology, 39,* 133–146.
Federmeier, K. D., Segal, J. B., Lombrozo, T., & Kutas, M. (2000). Brain responses to nouns, verbs and class-ambiguous words in context. *Brain, 12,* 2552–2566.
Foertsch, J., & Gernsbacher, M. A. (1997). In search of gender neutrality: Is singular *They* a cognitively efficient substitute for generic *he*? *Psychological Science, 8,* 106–111.
Friederici, A., Hahne, A., & Saddy, D. (2002). Distinct neurophysiological patterns reflecting aspects of syntactic complexity and syntactic repair. *Journal of Psycholinguistic Research, 31,* 45–63.
Garnham, A. (2001). *Mental models and the interpretation of anaphora*. Sussex: Psychology Press.
Garrod, S., & Sanford, A. J. (1994). Resolving sentences in a discourse context: How discourse representation affects language understanding. In M. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 675–698). New York: Academic Press.
Garrod, S. C., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language, 33,* 39–68.
Gerrig, R. J., Horton, W. S., & Stent, A. (2006). *Unheralded pronouns and theories of pronoun resolution*. Paper presented at 16th annual meeting of the Society for Text and Discourse, Minneapolis, MN, July 13–15.
Greene, S. B., Gerrig, R. J., McKoon, G., & Ratcliff, R. (1994). Unheralded pronouns and management by common ground. *Journal of Memory and Language, 33,* 511–526.
Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69,* 274–307.
Hagoort, P., Brown, C. M., & Groothusen, J. (1993). The syntactic positive shift (SPS) as an ERP measure of syntactic processing. *Language and Cognitive Processes, 8,* 439–483.
Hagoort, P., Brown, C. M., & Osterhout, L. (1999). The neurocognition of syntactic processing. In C. M. Brown & P. Hagoort (Eds.), *The neurocognition of language* (pp. 273–316). New York: Oxford University Press.
Hammer, A., Jansma, B. M., Lamers, M., & Münte, T. F. (2005). Pronominal reference in sentences about persons or things: An electrophysiological approach. *Journal of Cognitive Neuroscience, 17,* 227–239.
Hauk, O., & Pulvermüller, F. (2004). Effects of word length and frequency on the human event-related potential. *Clinical Neurophysiology, 115,* 1090–1103.
Kaan, E., & Swaab, T. Y. (2003). Repair, revision, and complexity in syntactic analysis: An electrophysiological differentiation. *Journal of Cognitive Neuroscience, 15,* 98–110.
Kennison, S. M. (2003). Comprehending the pronouns her, him, and his: Implications for theories of referential processing. *Journal of Memory and Language, 49,* 335–352.
King, J. W., & Kutas, M. (1997). *Is she an engineer? Brain potentials and anaphora*. Poster presented at the Fourth Annual Meeting of the Cognitive Neuroscience Society, Boston.
Klin, C. M., Guzman, A. E., Weingarter, K. M., & Ralano, A. S. (2006). When anaphor resolution fails: Partial encoding of anaphoric inferences. *Journal of Memory and Language, 54,* 131–143.
Kutas, M., & Federmeier, K. D. (2000). Electrophysiology reveals semantic memory use in language comprehension. *Trends in Cognitive Sciences, 4,* 463–470.
Kutas, M., Federmeier, K. D., Coulson, S., King, J. W., & Münte, T. F. (2000). Language. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), *Handbook of psychophysiology* (pp. 576–601). Cambridge: Cambridge University Press.
Kutas, M., & Hillyard, S. A. (1984). Brain potentials during reading reflect word expectancy and semantic association. *Nature, 307,* 161–163.
Moxey, L. M., Sanford, A. J., Sturt, P., & Morrow, L. I. (2004). Constraints on the formation of plural reference objects: The influence of role, conjunction, and type of description. *Journal of Memory and Language, 51,* 346–364.
Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory & Cognition, 25,* 273–285.
Osterhout, L., & Holcomb, P. J. (1995). Event-related potentials and language comprehension. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind* (pp. 171–215). Oxford: Oxford University Press.
Osterhout, L., & Mobley, L. A. (1995). Event-related brain potentials elicited by failure to agree. *Journal of Memory and Language, 34,* 739–773.
Preissl, H., Pulvermüller, F., Lutzenberger, W., & Birbaumer, N. (1995). Evoked potentials distinguish between nouns and verbs. *Neuroscience Letters, 197,* 81–83.
Sanford, A. J., & Filik, R. (2007). "They" as a gender-unspecified singular pronoun: Eye-tracking reveals a processing cost. *Quarterly Journal of Experimental Psychology, 60,* 171–178.
Sanford, A. J., Filik, R., Emmott, C., & Morrow, I. (2008). They're digging up the road again: The processing cost of Institutional They. *Quarterly Journal of Experimental Psychology, 61,* 372–380.
Schmitt, B. M., Lamers, M. J. A., & Münte, T. F. (2002). Electrophysiological estimates of biological and syntactic gender violation during pronoun. *Cognitive Brain Research, 14,* 333–346.
Sereno, S. C., Brewer, C. C., & O'Donnell, P. J. (2003). Context effects in word recognition: Evidence for early interactive processing. *Psychological Science, 14,* 328–333.
Sereno, S. C., Posner, M. I., & Rayner, K. (1998). Establishing a time-line of word recognition: Evidence from eye movements and event-related potentials. *NeuroReport, 9,* 2195–2200.
Sereno, S. C., & Rayner, K. (2003). Measuring word recognition in reading: Eye movements and event related potentials. *Trends in Cognitive Sciences, 7,* 328–333.
Streb, J., Rösler, F., & Hennighausen, E. (1999). Event-related responses to pronoun and proper name anaphors in parallel and nonparallel discourse structures. *Brain and Language, 70,* 273–286.
Swaab, T. Y., Camblin, C. C., & Gordon, P. C. (2004). Electrophysiological evidence for reversed lexical repetition effects in language processing. *Journal of Cognitive Neuroscience, 16,* 715–726.

Van Berkum, J. J. A., Brown, C. M., & Hagoort, P. (1999). Early referential context effects in sentence processing: Evidence from event-related brain potentials. *Journal of Memory and Language, 41,* 147–182.
Van Berkum, J. J. A., 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., Brown, C. M., Hagoort, P., Zwitserlood, P., & Kooijman, V. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 443–467.
Van Berkum, J. J. A., Koornneef, A. W., Otten, M., & Nieuwland, M. S. (2007). Establishing reference in language comprehension: An electrophysiological perspective. *Brain Research, 1146,* 158–171.
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. *Journal of Cognitive Neuroscience, 16,* 70.
Van Gompel, R. P. G., & Majid, A. (2004). Antecedent frequency effects during the processing of pronouns. *Cognition, 90,* 255–264.
Van Petten, C., & Kutas, M. (1990). Interactions between sentence context and word frequency in event-related brain potentials. *Memory & Cognition, 18,* 380–393.
Van Petten, C., & Kutas, M. (1991). Influences of semantic and syntactic context on open and closed class words. *Memory & Cognition, 19,* 95–112.
Yule, G. (1982). Interpreting anaphora without identifying reference. *Journal of Semantics, 1,* 315–322.