Available online at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

Brain Research

Research report

Referential processing in the human brain: An Event-Related Potential (ERP) study

C. Barkley[a,*], R. Kluender[b], M. Kutas[c]

[a]University of Minnesota, Department of Experimental and Clinical Pharmacology, 308 Harvard Street SE, Minneapolis, MN 55455, United States

[b]University of California San Diego, Department of Linguistics, 9500 Gilman Drive, La Jolla, CA 92093, United States

[c]University of California San Diego, Department of Cognitive Science and Department of Neurosciences, 9500 Gilman Drive, La Jolla, CA 92093, United States

ARTICLE INFO

Article history:
Accepted 15 September 2015
Available online 9 October 2015

Keywords:
Long-distance dependencies
Referential processing
ERPs
LAN
Referential negativity (Nref)
P600

ABSTRACT

A substantial body of ERP research investigating the processing of syntactic long-distance dependencies has shown that, across languages and construction types, the second element in such configurations typically elicits phasic left anterior negativity (LAN). We hypothesized that these effects are not specific to syntactic dependencies, but rather index a more general cognitive operation in which the second (dependent) element in sentence-level linguistic long-distance relationships triggers a process of association with the first element. We tested this hypothesis with straightforward referential dependencies, comparing pronouns with proper name antecedents to those without, and proper names with and without preceding co-referring pronouns. We predicted phasic LAN effects in response to the second referential element in both comparisons, but observed them only in response to pronouns with antecedents; no differences were observed between responses to proper names with and without preceding co-referring pronouns. We argue that LAN effects observed at the pronoun index the cognitive operations necessary for the association of a pronoun with its antecedent, on which it depends for its reference. Similar but non-identical responses were elicited by the main clause verb following the gap position in object relative clause constructions compared to coordinate clause controls in an orthogonal manipulation. LAN effects were thus elicited by the second dependent element in both construction types, suggesting that long-distance syntactic and referential dependencies pose similar processing challenges. These findings help to clarify the cognitive processes indexed by anterior negative responses to associated dependent elements in a variety of language contexts.

Published by Elsevier B.V.

*Corresponding author.
E-mail addresses: barkl025@umn.edu (C. Barkley), rkluender@ucsd.edu (R. Kluender), mkutas@ucsd.edu (M. Kutas).

http://dx.doi.org/10.1016/j.brainres.2015.09.017
0006-8993/Published by Elsevier B.V.

1. Introduction

1.1. Long-distance dependencies

A pervasive property of human language that poses a great challenge to the comprehension system is that non-adjacent sentence elements often depend on each other for successful interpretation. This interpretation process may require the formation of interdependent relationships between such elements across arbitrarily long distances.[1] Modern linguistic theory has traditionally divided sentence-level long-distance dependencies into two main types, which we refer to here as "syntactic" and "referential." These have long been the subject of debates as to exactly which dependency types are handled exclusively by the syntax, and which are handled via a combination of syntactic constraints and discourse principles (Chomsky, 1980, 1981; Reinhart, 1983; Pesetsky, 1990; Cinque, 1990; Chung, 1994; Reinhart and Reuland, 1993).

English relative clauses are a common and oft-studied example of a syntactic long-distance dependency. In these constructions, a sentence constituent like *The proposal* in (1) appears at the left edge of its clause, rather than in its underlying or "base" position, where it would ordinarily appear in a simple declarative clause, indicated by *the proposal* in (2) and an underscore in (1).

(1) *The proposal* [that they're currently in the process of finalizing __] seems solid.

(2) They are currently in the process of finalizing *the proposal*.

In the psycholinguistic and neurolinguistic literatures, the displaced element (*the proposal* in (1)) is referred to as the "filler" and the underscore in (1) is referred to as a "gap". Forming a relationship between filler and gap, a so-called "filler-gap dependency" (Fodor, 1978), is necessary for comprehension of relative clauses like (1) because the filler must be interpreted as the direct object of the verb *finalize* and the undergoer of the *finalizing* action.

A common example of a referential dependency is the relationship between a pronoun and its antecedent, as in (3). Referential dependencies do not involve the dislocation of any sentence constituent, but do require the formation of a relationship between two non-local elements (the pronoun *it* in and its antecedent *the proposal* in (3)) for comprehension to succeed.

(3) *The proposal* is currently being finalized and *it* seems solid.

In this paper, we are concerned with the brain responses elicited by the second of the two elements in such long-distance dependencies (the gap position in (1) and the pronoun *it* in (3)). Specifically, we aim to determine the extent to which the brain treats syntactic and referential dependencies similarly or differently, as indexed by the brain responses elicited by the second elements in these two types of long-distance relationships.

We first review what is known about the brain's response to the second element in syntactic dependencies, focusing on studies using event-related brain potentials (ERPs). The precise temporal resolution of this technique makes it ideal for studying the real-time processes underlying language comprehension (see Kutas et al. (2006) for a review of language-related ERP research). Effects observed in these studies will provide useful points of comparison when examining the brain's response to the second element in referential dependencies.

1.2. ERP indices of syntactic dependency formation

The ERP component that would later become known as the phasic[2] left anterior negativity, or LAN, was first observed in response to morphosyntactic violations in Kutas and Hillyard (1983).[3] Phasic LAN effects have since been observed in a range of violation paradigms, but also in response to second elements in well-formed syntactic long-distance dependencies. These include wh-questions in English (Kluender and Kutas, 1993a, 1993b) and German (Felser et al., 2003), relative clauses in English (King and Kutas, 1995; Müller et al., 1997; Weckerly and Kutas, 1999), Japanese (Ueno and Garnsey, 2007), and Korean (Kwon et al., 2013), and in scrambling constructions that disrupt canonical word order in Japanese (Ueno and Kluender, 2003; Hagiwara et al., 2007), Korean (Kwon et al., 2013), and German ((Rösler et al., 1998; Matzke et al., 2002).[4] Although there are exceptions, whether processing relative clauses, wh-questions, or scrambled word order, the second dependent element typically elicits phasic LAN effects at the post-gap position (though see fn. 4 with regard to LAN responses in head-final relative clauses). Despite this consistency, the functional interpretations proposed for these effects differ, although not necessarily in mutually exclusive ways.

1.3. Functional interpretations of phasic LAN effects

Kluender and Kutas (1993a, 1993b) interpreted LAN effects following gap positions in English wh-questions as reflecting the working memory-based[5] operation of filler retrieval for purposes of

[1] Here we restrict ourselves to a discussion of sentence-level non-local relationships, while fully recognizing the existence of word-internal non-local relationships, such as phonological processes of vowel harmony (Rose and Walker, 2011) and long-distance consonant assimilation (Rose, 2011).

[2] The term "phasic" is used here to contrast this transient effect (between 300 and 500 or 600 ms post-stimulus) with longer-lasting sustained anterior negative potentials that often span multiple words.

[3] In order to restrict the scope of this review, we will exclude LAN effects elicited in such violation paradigms and other environments that engender "parsing difficulty" (for example Kutas and Hillyard, 1983; Neville et al., 1991; Osterhout and Holcomb 1992). We note that the functional significance of LAN responses elicited by morphosyntactic violations and those apparently related to verbal working memory processes in grammatical sentences, as discussed in the main text, has never been satisfactorily resolved (though see Martin-Loeches et al. (2005) for a discussion of this issue).

[4] In most cases, the second element in these dependencies is the gap site, with LAN elicited by the word in post-gap position. However, in head-final SOV languages (e.g. Japanese and Korean) with pre-nominal relative clauses, the head noun is the second element. Even in these cases, LAN-like responses are elicited (see for example Kwon et al., 2013).

[5] Here we remain largely agnostic on the specific architectural details of the working memory system and its internal operations, relying instead on a simple, more general model. Working memory is a short duration, limited capacity memory system capable of simultaneously storing, manipulating, and combining information in the service of accomplishing some task (Baddeley,

filler-gap association. King and Kutas (1995) likewise argued that the phasic LAN effect elicited by the matrix verb in English relative clauses indexes re-activation of the relative clause head noun for purposes of semantic role assignment. Other studies of wh-questions, topicalization (Matzke et al. 2002), and relative clauses (Weckerly and Kutas, 1999; Ueno and Garnsey, 2007; Kwon et al., 2013) similarly attributed phasic LAN effects to costs related to (some form of) retrieval.

Cross-linguistic studies of scrambling – i.e. leftward displacement of a case-marked noun phrase – have prompted additional interpretations of phasic LAN effects. This is because scrambling often involves nothing more than switching the order of two adjacent noun phrases, making the displacement relatively local (i.e. clause-internal) rather than long-distance (i.e. across clause boundaries). Critical for our purposes is what happens downstream from such scrambled constituents. Ueno and Kluender (2003) observed phasic LAN effects around the object gap position in scrambling constructions, which they interpreted as the parser's attempt to restore canonical word order (see Matzke et al. (2002) for similar observations with regard to clause-internal German topicalization).

In Hagiwara et al. (2007), another study of Japanese scrambling, both short- and long-distance scrambling conditions compared to a condition in which constituents were in canonical order elicited phasic LAN effects at the sentence-final main verb. The authors suggested that their sentence-final LAN effects reflected the costs of processing scrambled word order, with concomitant re-computing of relevant semantic roles (cf. King and Kutas, 1995) and grammatical relations.

In sum, the papers cited here agree that phasic LAN indexes a working memory-based operation, though the exact nature of this process appears to be an open question, and the inferences made about the underlying processes that LAN effects index rely on assumptions about construction-specific challenges posed to the comprehension system and the precise operations of the working memory system. Accounts based on memory processes such as retrieval, reactivation, (re-)computation and the processing burdens imposed by non-canonicity have been proposed, although many researchers are satisfied with an account that relies on an increase in load or cost, broadly construed. Below, we articulate a simplified functional interpretation of phasic LAN as indexing a simple operation of associating two distal elements participating in a long-distance dependency.

1.4. The "back association" hypothesis

Our account of the processes indexed by phasic LAN is not specific to syntactic relationships, does not rely on construction-specific details, nor on specific working memory operations. Rather, on our view, the phasic LAN indexes a process of "back association" or "association at a distance." This process is triggered at the second element in a long-distance dependency and consists of the presumed working memory-based operation of "looking back" through already processed material for the first dependent element for the purposes of "associating" it with the second.[6] This proposal aims to achieve broad empirical coverage by accounting for the totality of LAN effects in syntactic as well as non-syntactic long-distance dependencies.

Perhaps the most appealing aspect of this account of phasic LAN is that it generates testable predictions. Specifically, it predicts that the second of two dependent elements in a sentence-level long-distance dependency will elicit phasic LAN indexing the process by which the second element is associated with the (previously processed) first element. Moreover, this process is not limited to sentence-level syntactic relationships nor to relationships involving null elements (i.e. gaps). If we are correct, then a reassessment of the existing literature should reveal phasic LAN effects beyond those associated with the processing of syntactic filler-gap dependencies, and it is to these effects that we now turn.

1.5. LAN effects to 2nd elements in non-syntactic sentence-level dependencies

Shao and Neville (1998) examined two types of semantic violations that had previously received little attention in the ERP literature. The crucial comparisons were between sentences that violated hyponymy relations (4a) and their controls (4b), as well as between sentences that violated the grammatical requirement for licensing of negative polarity items (NPIs)[7] (5a) and their controls (5b).

(4a) #Jane does not eat any meat at all, and instead she eats lots of ***beef*** and vegetables.

(4b) Jane does not eat any meat at all, and instead she eats lots of ***rice*** and vegetables.

(5a) *Fred believes that he has ***ever*** seen that woman before.

(5b) Fred believes that he has ***never*** seen that woman before.

Compared to controls, both violation types (4a and 5a) elicited anterior negativity (albeit with somewhat differing distributions and onset latencies). Shao and Neville relied on a working memory-based account to interpret these effects, noting that processing the hyponymy relation "may require the retrieval of the superordinate (*meat*) when the subordinate (*rice*) is processed," and that the processing of NPIs may require the parser to "retrieve the initial portion of the sentence to determine whether the negative polarity item fits within its scope" (i.e., within the scope of negation).

Anderson and Holcomb (2005) examined ERP effects associated with the processing of co-reference and synonymy. Participants read two-sentence mini-discourses in which a noun phrase in the second sentence was either intended to

(*footnote continued*)
1992). This formulation is sufficient for present purposes, especially in view of general disagreement in the literature about the specific architecture and mechanisms of verbal working memory (see for example Cowan, 2000; Just and Carpenter, 1992; Caplan and Waters, 1999; McElree et al., 2003; Lewis et al., 2006). We will however briefly return to the Lewis et al. (2006) model in the discussion section to aid in the interpretation of results obtained in the current experiment.

[6]This type of interpretation derives from ideas originally articulated in Kluender and Kutas (1993b) and Kutas et al. (2006).

[7]Negative polarity items (such as *ever* in (5a), or *a thing* in "I do not understand a thing") are linguistic elements that occur in negative contexts. Compare (5a), in which this requirement is not satisfied for the NPI *ever*, with "Fred does not believe that he has *ever* seen that woman before," in which it is. This contrast illustrates the restrictions on the occurrence of NPIs.

be co-referential[8] (occurring with the definite determiner *the*) or non-co-referential (occurring with the indefinite determiner *a/an*) with a noun phrase in the first sentence. A comparison between co-referential and non-co-referential noun phrases revealed LAN effects in response to nouns following definite determiners compared to those following indefinites. The authors argued that this effect reflected the increased working memory load triggered by the processing of the definite determiner, which signaled that the upcoming noun would require the retrieval of an antecedent in the preceding discourse. These effects, and their interpretations, are consistent with our association-based hypothesis of the functional significance of the LAN, and provide support for our claim that LAN effects may be elicited by the second element in any type of sentence-level relationship, as long as these relationships require the association of two non-adjacent elements.

The present study is explicitly designed to test this hypothesis by investigating ERPs elicited by the second element in simple unambiguous referential dependencies, such as the pronoun *he* in (6):

(6) **Bill** wondered whether **he** would be able to make it to work on time.

Next we summarize what is known about the real-time ERP processing of these referential relationships.

1.6. ERP studies of referential processing

In recent years there has been an increase in the number of ERP studies investigating referential processing (Van Berkum et al., 1999, 2003, 2007; Van Berkum, 2004; Anderson and Holcomb, 2005; Heine et al., 2006; Nieuwland and van Berkum, 2006; Nieuwland et al., 2007; Osterhout and Mobley, 1995). Van Berkum and colleagues typically have participants read sentences or small discourses, and record ERPs to target pronouns that either have a unique referent (7a), two possible referents (leading to "referential ambiguity," (7b)), or no referent (leading to "referential failure," (7c) – examples from Van Berkum (2004)):

(7a) *David* shot at *Linda* as ***he*** jumped over the fence.

(7b) *David* shot at *John* as ***he*** jumped over the fence.

(7c) *Anna* shot at *Linda* as ***he*** jumped over the fence.

Relative to its unambiguous counterpart in (7a), the ambiguous pronoun in (7b) elicited a sustained frontal negativity (approximately between 400 and 1100 ms), an effect subsequently dubbed the "referentially induced frontal negativity," or N*ref*. This effect has been elicited using both written and spoken ambiguous pronouns and noun phrases (Nieuwland and van Berkum, 2006; Van Berkum et al., 1999, 2003). The N*ref* has been argued to reflect situation model-level rather than superficial ambiguity (Nieuwland et al., 2007), correlates with reading span score (Nieuwland and van Berkum, 2006; Nieuwland, 2014), and has been tentatively localized to medial prefrontal cortex (Nieuwland et al., 2007). N*ref*-like effects have also been reported in response to pro-forms in partially elided noun phrases in ellipsis constructions (Martin et al., 2012, 2014).

Van Berkum and colleagues have shown that this response is distinct from the brain's response to referential failure, observed when comparing pronouns with unique antecedents to those with none, as shown in (7c). The "referentially failing" pronoun in (7c) did not elicit an N*ref* but rather a P600, a late positivity typically associated with ungrammaticality (Hagoort et al., 1993), parsing difficulty (Osterhout and Holcomb, 1992), and long-distance syntactic integration (Kaan et al., 2000). A similar response was reported in Osterhout and Mobley (1995) to anomalous reflexive pronouns, as in:

(8) The man prepared ***herself*** for the operation.

A few things about this literature are worth noting. None of the manipulations included the straightforward comparison of a pronoun with an overt antecedent to one lacking any kind of candidate antecedent in the discourse, focusing instead on referential ambiguity or highly salient gender mismatches between pronoun and antecedent. We believe that the latter type of manipulation, which is also a gender-based morpho-syntactic violation, may pose problems when interpreting the "referential failure" effects observed. In other words, it is unclear whether these late positivities index referential or morpho-syntactic processing difficulties. Accordingly, while it has been established that the brain responds differently to different types of problems posed by different referential configurations, the manner in which the brain processes straightforward referential relationships remains an open question.

To investigate this issue while providing a test of our back association hypothesis, we designed an ERP experiment in which participants read sentences containing pronoun or proper name main clause subjects that either did, or did not, refer back to co-referential elements in a preceding sentence-initial adjunct. In addition, our materials contained an orthogonal manipulation, comparing object relative clauses to their coordinate clause controls, in order to replicate previous findings of phasic LAN in response to the main clause verb in object relative clauses (i.e. King and Kutas, 1995). As such, our study was designed to investigate the ERP signatures of referential dependency formation, while also enabling us to determine the extent to which brain responses elicited by referential dependencies pattern with those elicited by syntactic dependencies.

On the basis of our hypothesis as to the functional identity of phasic LAN, we predicted that at the main clause subject position we would observe phasic LAN effects, i.e. enhanced negativity between 300 and 500 ms. over left anterior regions of scalp, in response to the two *Co-referent* conditions (conditions A and C, Table 1) compared to their *No Co-referent* (conditions B and D, Table 1) counterparts.

We also predicted that the onset and distribution of these negativities would be similar to those elicited by the main clause verb in the comparison between object relative clauses and their coordinate-clause controls – in other words, we expected the object relatives to elicit enhanced negativity with a similar scalp distribution.

The most commonly held interpretations (Van Berkum et al., 2007) of N*ref* and P600 effects typically elicited in referential paradigms (see Section 1.6) would lead one to predict a lack of N*ref* (or enhanced negativity) effects in response to the critical pronoun in the *Pronoun, Co-referent* condition, given that it contains a straightforward and unambiguous pronoun-antecedent relation. With regard to P600 effects, if any pronoun without an obvious intra-sentential antecedent leads to "referential failure," the antecedent-less

[8]Put simply, co-referring expressions refer to the same element in either the linguistic discourse or the world.

Table 1 – Experimental sentences: Pronouns and proper names with and without preceding co-referents.

A. Pronoun, Coreferent
After a covert mission that deployed *Will* for nine terrible months, **he** longed for home.
B. Pronoun, No coreferent
After a covert mission that required deployment for nine terrible months, **he** longed for home.
C. Name, Coreferent
After a covert mission that deployed *him* for nine terrible months, **Will** longed for home.
D. Name, No coreferent
After a covert mission that required deployment for nine terrible months, **Will** longed for home.

pronoun in the *Pronoun*, *No Co-referent* condition should elicit a P600 effect. If on the other hand the P600 response to "referentially failing" pronouns is due solely to a gender-based feature mismatch between a pronoun and a potential antecedent that triggers morpho-syntactic processing difficulty, no such effect should be observed.

2. Results

In this section, we first present the results of the orthogonal manipulation comparing object relatives to their coordinate clause controls in order to establish a point of reference for our primary experimental manipulations, in which we compare the two *Pronoun* conditions and the two *Name* conditions. We do this first for phasic responses in Section 2.1 before discussing more sustained responses in Section 2.2.

2.1. *Phasic responses*

2.1.1. *Object relatives versus coordinate clause controls*

At the critical verb (*smashed* in E and F of Table 2), object relatives elicited a negative-going ERP in comparison to coordinate clause controls (Fig. 1).

This effect was small in magnitude and had a narrow distribution, appearing maximal at left antero-lateral scalp sites. A one-tailed *t*-test on mean amplitude measurements at left anterior electrodes between 300 and 500 ms confirmed a significant difference between conditions [$t(19)=2.46$, $p=.02$]. The distributional ANOVA revealed no main effect of condition [$F<1.0$], but there were significant condition × hemisphere [$F(1.19)=6.16$, $p=.02$], condition × laterality [$F(1,19)=5.85$, $p=.0384$], and condition × hemisphere × laterality [$F(3,57)=8.57$, $p=.015$] interactions, as well as a marginal interaction of condition × anteriority × hemisphere × laterality [$F(3,57)=2.63$, $p=.08$]. These interactions appeared to be driven by a difference between conditions at left lateral scalp sites [$F(1,19)=9.82$, $p=.006$] that was not present at left medial, right medial, or right lateral sites (all $Fs<1.0$). Additionally, though the marginal four-way interaction suggested that the differences between conditions might vary at different levels of the anteriority factor, follow-up analyses showed this not to be the case, as comparing at individual levels of the anteriority factor over the left hemisphere yielded no significant results (all $Fs<2.0$).

Table 2 – Experimental sentences: object relative clauses and coordinate clause controls.

E. Object relative clauses
The soldier who the sailor roughly pushed **smashed** a bottle against the bar.
F. Coordinate-clause controls
The solider roughly pushed the sailor and **smashed** a bottle against the bar.

In sum, the comparison between object relatives and their coordinate clause controls (*E* vs. *F* in Table 2) yielded a LAN effect in response to the critical main verb in the object relative clause condition. This LAN effect, which exhibited a left lateral distribution (see Fig. 1 and Section 3.3), was significant in the traditional LAN time window of 300–500 ms, replicating previous findings in the literature. This effect at the second element of a syntactic dependency can now serve as a point of comparison for the brain response to the second element in our referential dependencies.

2.1.2. *Experimental sentences*

The results reported in this section focus on responses occurring at least 300 ms post-onset of the critical main clause subject. No comparisons undertaken in the N100 (50–150 ms) and P200 (150–250 ms) latency windows at this sentence position produced significant results.

2.1.2.1. *Pronouns with and without co-referents*

2.1.2.1.1.Responses between 300 and 500 ms. Visual inspection of the ERP response to the pronominal subject suggested that the response to the *Pronoun*, *Co-referent* condition was more negative than the response to its *No Co-referent* counterpart, as shown in Fig. 2.

This effect appeared larger over the left than the right hemisphere, and also appeared to be slightly larger at anterior electrodes. A one-tailed *t*-test on mean amplitude measurements at left anterior electrodes between 300 and 500 ms confirmed a significant difference between conditions [$t(19)=4.42$, $p<.0005$]. The distributional analysis showed no main effect of condition [$F<1.0$], but did reveal a significant condition × hemisphere interaction [$F(1,19)=5.78$, $p=.0049$], driven by a main effect of condition over the left [$F(1,19)=5.61$, $p=.02$] but not the right hemisphere [$F<1.0$]. There were also marginal condition × laterality [$F(1,19)=4.14$, $p=.0577$] and condition × hemisphere × anteriority [$F(3,57)=2.22$, $p=.09$] interactions in the distributional analysis. Investigating these interactions at individual levels of these factors, there were main effects of condition at left frontal [$F(1,19)=4.13$, $p=.05$] and left temporal [$F(1,19)=8.51$, $p=.0096$] sites, but not at left prefrontal [$F(1,19)=1.07$, n.s.], left occipital [$F(1,19)=2.74$, n.s.], or right hemisphere sites (all $Fs<1.0$). Likewise, effects were significant at left medial [$F(1,19)=6.94$, $p=.024$, but not left lateral, right medial or right lateral sites (all $p>.27$). The isovoltage map in Fig. 2C shows the left fronto-central distribution indicated by the results of these comparisons.

2.1.2.1.2.Responses between 500 and 800 ms. Visual inspection of the data also suggested increased positivity in response to the *Pronoun*, *No Co-referent* condition, as predicted by the work of Van Berkum and colleagues. Results of planned *t*-tests confirmed significant differences between conditions. In the

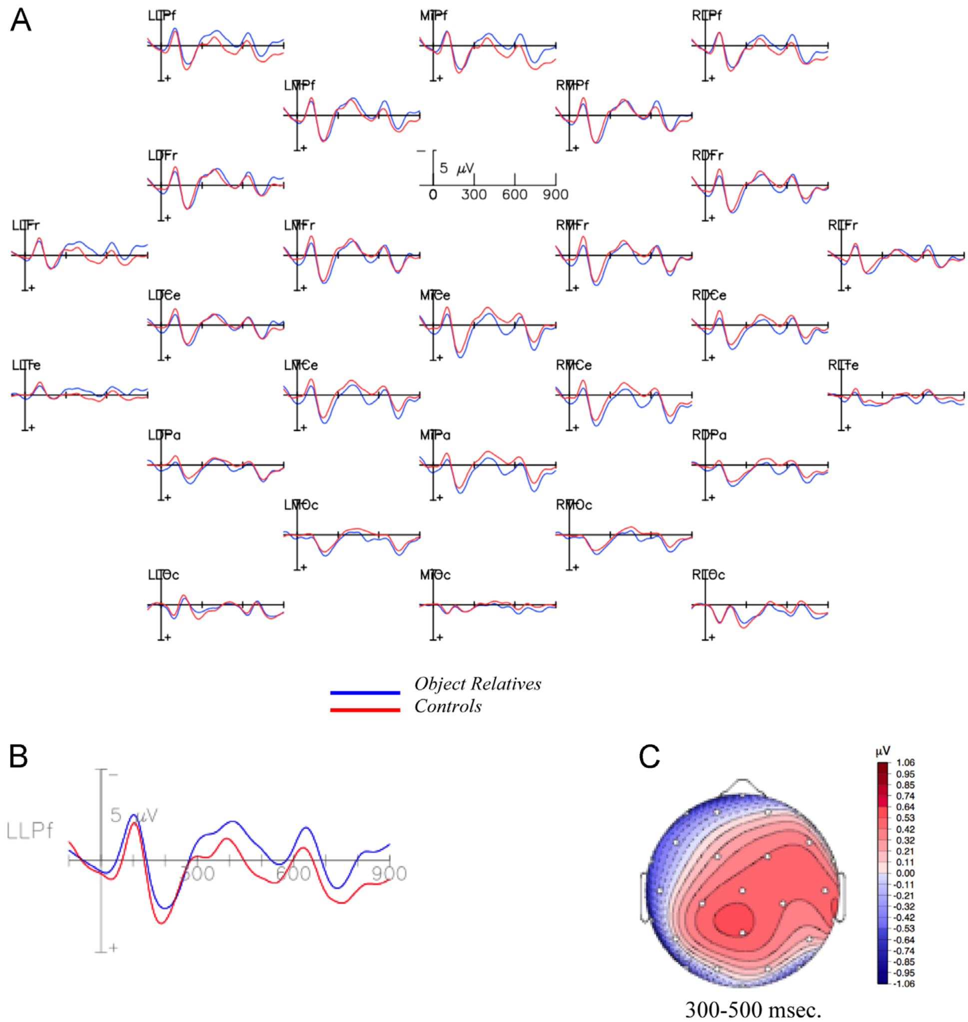

Fig. 1 – (A) Grand average ERP waveforms for *object relative* and *coordinate clause controls* at all 26 electrode sites (1000 ms epoch including 100 ms baseline). (B) Grand average ERP waveform for object relative and coordinate clause control conditions at electrode LLPf (1000 ms epoch including 100 ms baseline). (C) Topographic scalp isovoltage map of the mean difference between conditions between 300 and 500 ms.

500–800 ms time window, the distributional analysis correspondingly revealed a marginal main effect of condition [$F(1,19)=3.99$, $p=.067$], as well as a significant condition × hemisphere interaction [$F(1,19)=4.85$, $p=.0417$] and a marginal condition × hemisphere × anteriority interaction [$F(3,57)=2.10$, $p=.10$] – in other words, the same interactions with distributional factors as were reported for the 300–500 ms time window, minus the marginal interaction of condition × laterality. Perhaps unsurprisingly, the isovoltage map in Fig. 2D indicates that the distribution of the response in the later time window of 500–800 ms was virtually identical to the distribution of the LAN effect observed in the earlier 300–500 ms window. We therefore provisionally conclude that what might initially appear to be a late positivity in response to the *Pronoun, No Co-referent* condition is in fact another phasic LAN response to the word subsequent to the critical main clause subject position. This conclusion was supported by a lack of significant differences ($p>.35$) between the two *Pronoun* conditions in an unplanned analysis between 400 and 600 ms post-stimulus onset of the critical main clause subject position (i.e. the last 100 ms of the critical word and the first 100 ms of the following word). We return to this issue in Section 3.1.1.

2.1.2.1.3.*Summary*. In sum, while the onset latency, polarity, and lateralization of the phasic LAN effects elicited by both object relatives and pronouns with antecedents were similar as predicted, the effects did differ subtly in terms of their distribution across the scalp. The phasic LAN effect elicited by object relative clauses was significant at left lateral scalp sites (with no difference between conditions at left medial scalp sites), but did not differ reliably along the anterior–posterior dimension. In contrast, the LAN effect elicited by pronouns with antecedents in the *Pronoun* comparison was significant at left medial but not at left lateral scalp sites, and *did* differ along the anterior–posterior dimension. Last but not least, based on the virtually identical statistical interactions and scalp distribution of the potential observed in the 500–800 ms time window, we interpret this response as a phasic LAN response to the word one position downstream from the critical position rather than as a late positivity.

2.1.2.2. Proper name subjects with and without co-referents. In comparing the *Name, Co-referent* and *Name, No Co-referent*

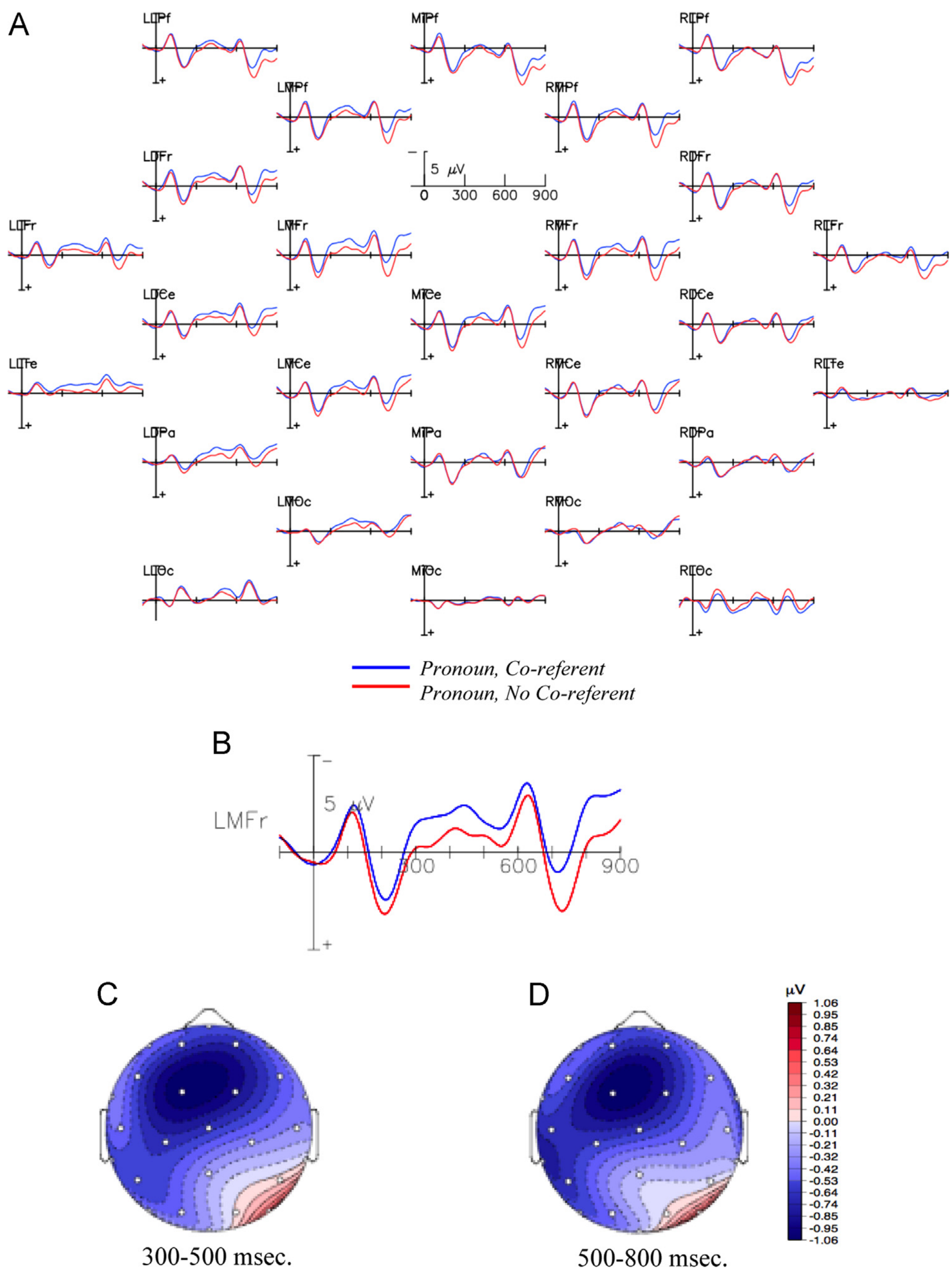

Fig. 2 – (A) Grand average ERP waveforms for the *Pronoun, co-referent* and *Pronoun, No co-referent* conditions at all 26 electrode sites (1000 ms epoch including 100 ms baseline). (B) Grand average ERP waveforms for both *Pronoun* conditions at electrode LMFr (1000 ms epoch including 100 ms baseline). (C) Topographic scalp isovoltage map of the mean difference between conditions in the 300–500 ms time window. (D) Topographic scalp isovoltage map of the mean difference between conditions in the 500–800 ms time window.

conditions at the matrix clause subject position, not only were there no effects in the N100 and P200 time windows, but additionally, neither the results of planned *t*-tests nor the full or distributional analyses in any of the later time windows (i.e 300–500 or 500–800 ms) tested showed main effects of condition or any interactions with electrode or distributional factors (see Fig. 3).

2.2. *Sustained responses (300–2000 ms)*

2.2.1. *Object relatives versus coordinate clause controls*

As has previously been reported for LAN effects elicited by second elements in English object relative clauses (King and Kutas, 1995) and wh-questions (Kluender and Kutas 1993a), the response to the main clause verb in our orthogonal manipulation appeared to continue past the critical word (see Fig. 4A). The full multi-word[9] analysis conducted between 300 and 2000 ms (including the critical word plus three additional words: *smashed a bottle against* in E and F of Table 2) revealed a significant condition × electrode interaction [$F(25,475) = 2.22$, $p = .0007$], and the distributional analysis showed a marginal condition × anteriority interaction [F

[9]See Section 4.5 for a discussion of the distinction between single- and multi-word analyses of sustained effects.

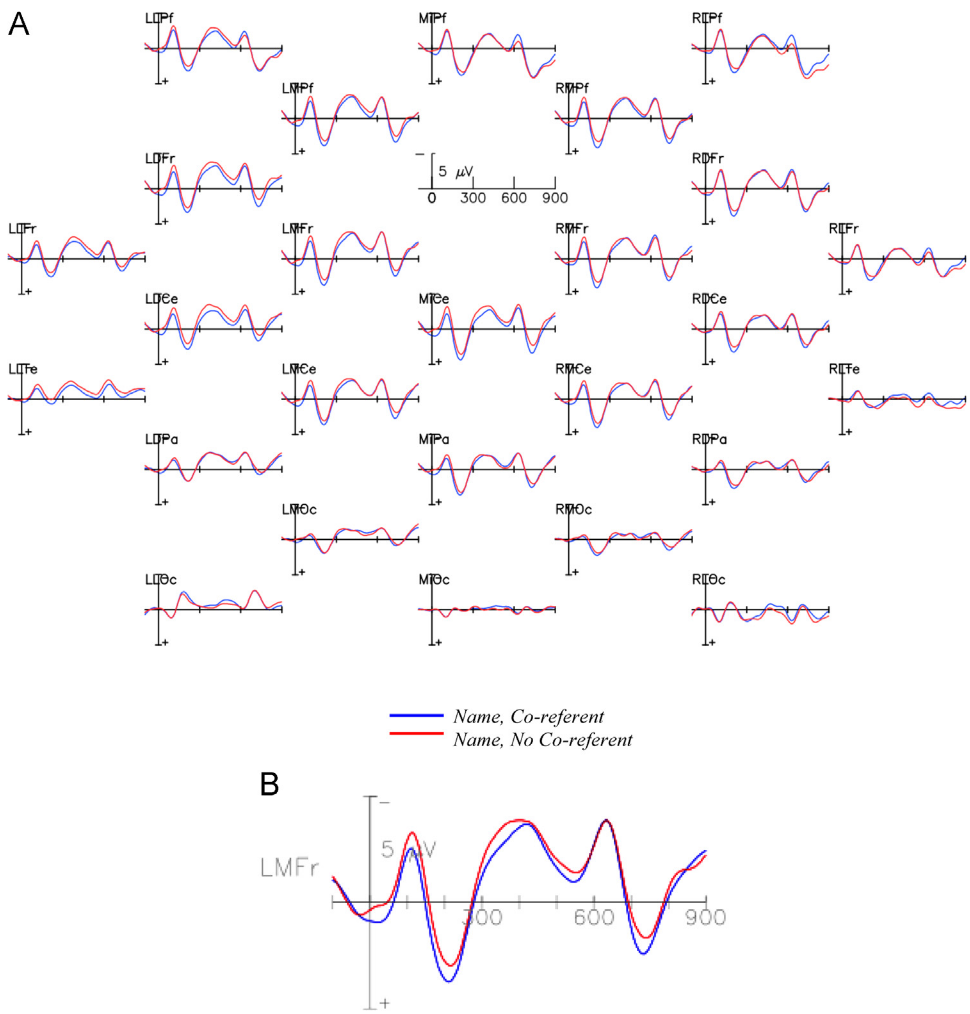

Fig. 3 – (A) Grand average ERP waveforms for the *Name, co-referent* and *Name, No co-referent* conditions at all 26 electrode sites (1000 ms epoch including 100 ms baseline). (B) Grand average ERP waveforms for both *Name* conditions at channel LMFr (1000 ms epoch including 100 ms baseline). No isovoltage map is provided because of a lack of observable differences between conditions.

(3,57)=2.48, p=.0718] as well as a significant condition × anteriority × hemisphere interaction [F(3,57)=4.41, p=.0078]. Further investigation of these interactions indicated that the difference between object relatives and controls was larger over the left hemisphere, and slightly larger over frontal sites. However, the individual word analyses of the three words subsequent to the critical position revealed no significant effects after re-baselining (all Fs<1.0; see isovoltage maps in Fig. 4B).

2.2.2. *Pronouns with and without co-referents (300–2000 ms)*
Similar to the LAN effect elicited by object relatives vs. coordinate clause controls, the LAN effect elicited by the main clause pronominal subject with a preceding co-referent also seemed to extend over multiple word positions.

The full multi-word analysis between 300 and 2000 ms (including the entire main clause of conditions A and B in Table 1: *he longed for home*) again showed a significant condition × electrode interaction [F(25,475)=1.98, p=.037], and the distributional multi-word analysis again showed a significant condition × anteriority × hemisphere interaction [F(3,57)=3.69, p=.0177] attributable to large differences between conditions over the front of the head, predominantly over the left hemisphere. In this respect, the sustained effects elicited by pronouns with antecedents were very similar to those elicited by the main clause verbs in object relative clauses.

However, in contrast, individual word analyses of the sustained effect elicited by pronouns with antecedents (Fig. 5B) showed a significant main effect of condition at word positions 13 (*longed* in conditions A and B of Table 1) [F(1,19)=3.89, p=.04] and 14 (*for* in conditions A and B of Table 1) [F(1,19)=4.57, p=.0473], as well as condition × laterality [F(1,19)=6.14, p=.02] and a marginal condition × laterality × anteriority F(3,57)=3.94, p=.11] interactions at word position 14. These interactions were caused by differences between conditions mainly over fronto-central scalp sites.

3. Discussion

Generally, the results of this study provide evidence in support of our back association hypothesis. Consistent with our predictions, we observed greater left anterior negativity in response to pronouns with preceding antecedents compared to those

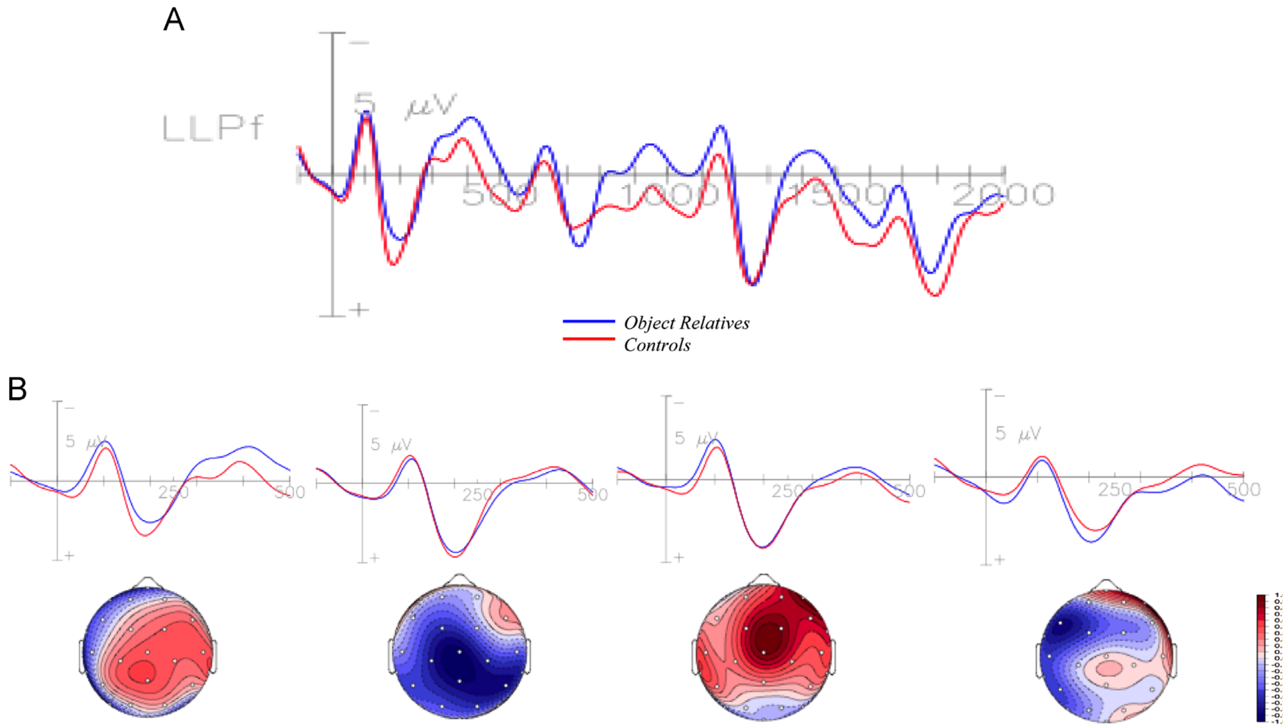

Fig. 4 – Grand average ERP waveforms for *object relative* and *coordinate clause controls* at electrode LLPf. (A) The *multi-word analysis* consists of a 2000 ms epoch (shown with the final 100 ms of the 500 ms baseline) time-locked to the critical matrix clause verb. (B) The *single-word analysis* consists of four individual 500 ms epochs and associated isovoltage maps for the mean difference between conditions for the 300–500 ms time window. These plots begin at the critical position and are then re-baselined at the three subsequent word positions.

without, arguably indexing the process of associating a pronoun with its antecedent, a process necessary for the establishment of a straightforward referential dependency, much like the syntactic dependency between a filler and its gap. Somewhat to our surprise, we observed no such effect in our *Name* comparison, an issue to which we return below in Section 4.2.

The similar (albeit non-identical) within-participant effects observed at the second element in both syntactic and referential dependencies – at least as far as simple antecedent-pronoun relations are concerned – suggest that phasic LAN-like effects reflect a cognitive process that is largely independent of the linguistic level of analysis at which it occurs. However, subtle differences in the duration of effects and scalp topography suggest that back association processes and their attendant brain responses may be modulated to some extent by the type of long-distance relationship under construction. We return to this issue in Section 4.3.

3.1. *Pronominal main clause subjects (300–500 ms)*

Based on our hypothesis regarding the functional identity of LAN effects, we predicted enhanced LAN in response to a pronoun preceded by a co-referring nominal antecedent in the same sentence when compared to an antecedent-less pronoun, and this is the pattern of results that we observed. Accordingly, we maintain that the LAN effect observed in the comparison of the *Pronoun* conditions can plausibly be interpreted as an index of the association processes necessary for dependency formation (Fig. 2A and C).

An alternative, albeit less likely, interpretation of this LAN-like negativity in the current study is that it is an N*ref* indexing referential ambiguity, an interpretation buttressed by the results of an analysis in the traditional N*ref* tine window of 400–1100 ms that showed significant differences between conditions (condition × anteriority × hemisphere interaction [$F(3,57)=4.17$, $p=.0113$]). If this were the case, however, one would be forced to adopt the position that the parser diagnoses the main clause subject in the Pronoun, Co-referent condition as referentially ambiguous, with the ambiguity (perhaps) arising because of indecision as to whether to associate the pronoun with (a) the co-referring nominal expression in the preceding sentence-initial adjunct or (b) an extra-linguistic referent. However, post-experiment debriefing questionnaires provided no evidence that participants experienced these items as ambiguous, or even difficult to process.[10] Moreover, during debriefing, no participants indicated that they had been associating pronouns with extra-linguistic co-referents. Nor do we believe that this would have been a reasonable default parsing strategy, given the presence of an explicit co-referent with matching number and gender features within the same sentence.

If we pursue the line of reasoning that the enhanced negativity in response to the pronoun is an index of referential ambiguity (i.e. an N*ref*), we are inevitably led to the

[10] In fact, during debriefing, the majority of subjects reported having the most difficulty processing object relative clauses, assuming that it was these constructions that were the focus of the experiment.

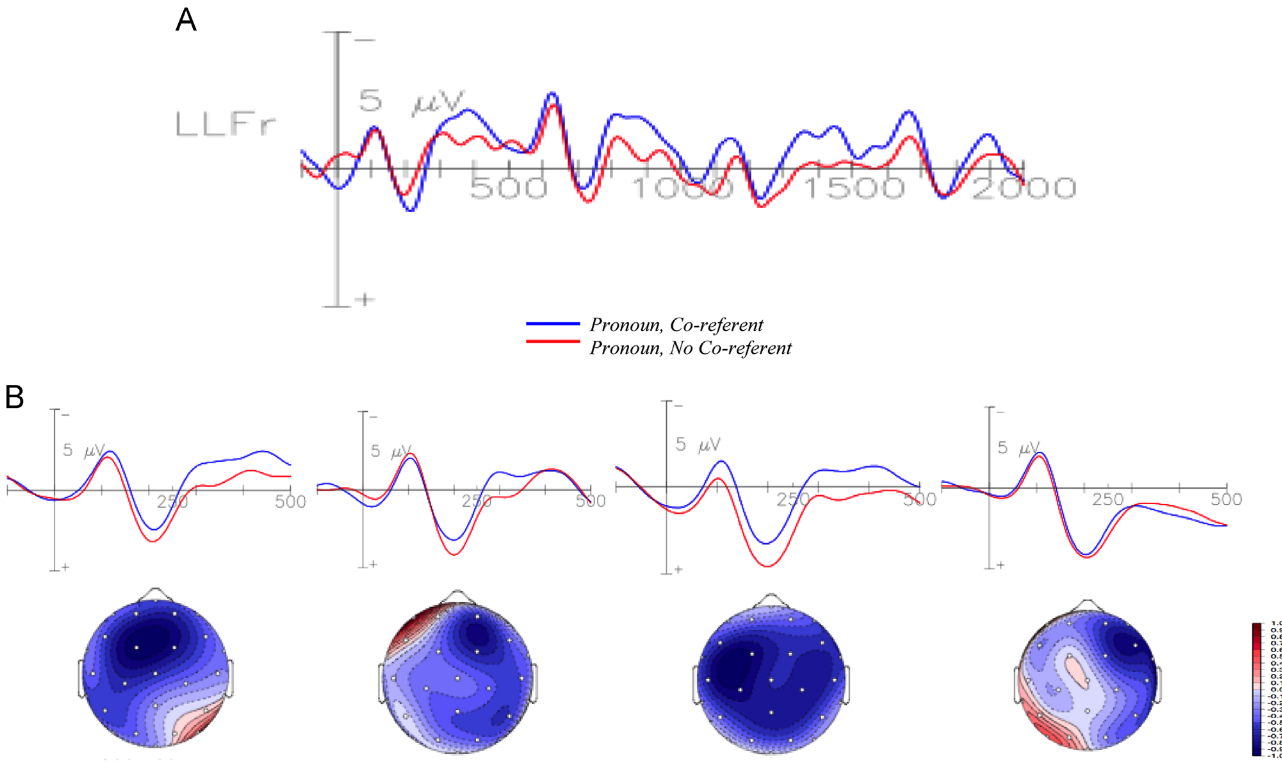

Fig. 5 – Grand average ERP waveforms for the ***Pronoun, Co-referent*** **and** ***Pronoun, No Co-referent*** **at electrode LLFr. (A) The** ***multi-word analysis*** **consists of a 2000 ms epoch (shown with the final 100 ms of the 500 ms baseline) time-locked to the critical matrix clause subject. (B) The** ***single-word analysis*** **consists of four individual 500 ms epochs and associated isovoltage maps showing the mean difference between conditions between 300 and 500 ms These plots begin at the critical position and are re-baselined at three subsequent word positions. A. Cumulative analysis. B. Non-cumulative analysis.**

conclusion that the parser initially experiences *every* non-bound-variable pronoun (Reinhart, 1983) as ambiguous – a position that does not seem parsimonious or consistent with evidence that the processor often initially pursues the simplest parse available (Frazier, 1979). We therefore maintain that the negativity observed in response to the pronominal subject in the *Pronoun, Co-referent* condition is a LAN most plausibly interpreted as indexing the process of associating two non-adjacent elements.

3.1.1. *Pronominal main clause subjects (500–800 ms)*

While visual inspection of the ERPs to *Pronouns* with and without co-referring antecedents (Fig. 2) might suggest a relative positivity in response to the *No Co-referent condition* , careful examination of the data renders this interpretation unlikely. We instead interpret this effect as a LAN in response to the word subsequent to the critical position. Support for this view comes from the virtually identical statistical interactions and scalp topographies of this comparison in these two adjacent time windows (see the isovoltage maps in Fig. 2C and D), the fact that our analyses showed, after re-baselining, a significant LAN response to the subsequent word, and the fact that effects observed in the 500–800 ms time window were preceded by an epoch (400–600 ms) during which the comparison between the two pronoun conditions did not reach significance ($p > .35$). Moreover, our late effect does not resemble the late positivities observed in response to antecedent-less pronouns in the work of Van Berkum and colleagues, which all present with a clear posterior maximum (Nieuwland and van Berkum, 2006; Nieuwland, 2014).

The apparent absence of a late positivity to the *Pronoun, No Co-referent* condition calls into question the assertion that pronouns lacking antecedents universally elicit late positivity. As mentioned in the introduction, in the most widely used manipulation intended to induce referential failure, co-referents are typically rendered unavailable with a gender or number manipulation that results in a highly salient feature-based mismatch between pronouns and their practically adjacent (potential) antecedents. This mismatch could plausibly trigger either referential or morphosyntactic processing difficulty. We suggest that it is the latter that leads to the late positive response – arguably a P3b (Donchin, 1981; Kok, 2001) or a P600 effect, elicited not by a referentially failing pronoun, but rather by a salient gender-based morphosyntactic anomaly (Osterhout and Mobley, 1995; Coulson et al., 1998).

As also noted in the introduction, the ERP literature on referential processing contains many manipulations that, though they may not always create outright violations *per se*, at a minimum engender processing difficulty as the parser attempts to construct referential relationships. The most frequently cited response to such manipulations is the *Nref*, a frontal negativity linked to processing costs associated with referential ambiguity (or at a minimum referential processing difficulty, broadly construed). In our study, we elicited negativity that was very similar in timing (see Section

3.3.1) and scalp distribution (see Section 3.3.2) to the N*ref* in response to the second element in an unambiguous, straightforward referential dependency. This suggests that N*ref*, and N*ref*-like responses) may not be a specific response to referential ambiguity or referential processing difficulty. Thus while we agree that the N*ref* indexes referential processing, we question whether the presence of referential elements that generate ambiguity or trigger processing difficulty are a necessary condition for its elicitation. Instead, we would like to suggest that the N*ref* is just one instance of a family of anterior negativities elicited by association at a distance in language contexts more generally (this conjecture is not very different from one found in Van Berkum et al. (2007)).

We believe that the results of our study can be reconciled with the totality of findings reported across studies by Van Berkum and colleagues with a slight modification of the cognitive processes that the N*ref* is assumed to index. Specifically, the LAN response elicited by pronouns with antecedents in our materials was found to be larger than that elicited by pronouns without antecedents. Similarly, Van Berkum and colleagues have shown that the Nref, with very similar timing parameters and scalp topography to our data, is larger in response to pronouns (and other anaphoric expressions) with two possible antecedents than to pronouns with only one possible antecedent. Under this view, LAN effects in response to pronouns with antecedents are interpreted as indexing straightforward dependency formation, while N*ref* effects elicited in response to pronouns with multiple candidate antecedents index *attempts* to form such a dependency, with increases in amplitude the result of the difficulty associated with selecting among multiple candidate antecedents.

In other words, the amplitude of the anterior negative response in each case indexes association at a distance of two dependent sentence elements, with the amplitude of the response determined by the ease of association. In line with our study showing similarities between syntactic and referential dependencies, this view is consistent with data reported in King and Kutas (1995:389, Figure 6) showing monotonic increases in the amplitude of LAN responses from coordinate (no dependency) to subject relative (one dependency between main clause subject and verb) to object relative clauses (two dependencies: between main clause subject and verb, and between filler and gap). If this interpretation is accurate, it would predict that a pronoun with three possible antecedents might likewise elicit greater anterior negativity than a pronoun with two, which is an empirical question for future research.

3.2. Proper name main clause subjects

Counter to our expectations, we observed no differences between the two *Name* conditions at the main clause subject position. This leads us to conclude that – unlike in the *Pronoun* conditions – back association operations were not triggered by the proper names in main clause subject position of our stimulus sentences.

This result would seem to be in direct conflict with data showing an increase in reading time at the second element (i.e. a proper name) in constructions in which the pronoun precedes its co-referent (Kazanina et al., 2007). This increase is argued to reflect processes underlying the formation of the referential dependency between a pronoun and a subsequent co-referring proper name. However, Kazanina et al. (2007:406) argue that if the parser encounters evidence (in our case the initial adverbial adjunct) predicting an upcoming clause that will contain a subject (as was always the case in our materials), this information could influence the parse before the subject is actually encountered. If this is the case, then there is no reason to assume that the proper name subjects would necessarily trigger back association operations.[11]

Put simply, under our hypothesis, if the process of back association is not triggered, then no LAN effect should be elicited. This account gains support from both theoretical and experimental work on reference form, as well as from recent influential models of the working memory processes subserving sentence comprehension (McElree et al., 2003; Lewis et al., 2006).

It has been noted (Ariel, 1990; Gundel et al., 1993) that all known languages contain a range of referential forms that vary in terms of lexical specificity/elaboration. For example, in English, these range from null forms to proper names, with several additional intermediate forms. It has been argued that these referential forms constitute a hierarchy, with fully specified forms at one end of the spectrum and attenuated (or even elided or implicit) forms at the other, with position on the hierarchy argued to predict, most importantly for our purposes, *referential role*. Crucial for present purposes is the contrast between pronouns and proper names, elements that occupy opposite poles of the referential hierarchy and that are argued to fulfill the roles of *referential maintenance* and *referential establishment*, respectively (Silverstein, 1976). If we adopt this logic, then we can straightforwardly argue that while pronouns typically "look backward" to maintain reference, proper names function to establish reference and intrinsically "look forward." This notion gains support from experimental work by Marslen-Wilson et al. (1982) and Vonk et al. (1992) that demonstrates the different roles these forms play: (1) pronouns are associated with the continuation of existing topics and the maintenance of focus on antecedents, while (2) proper names tend to introduce new topics and referents. We therefore believe that the referential hierarchy provides a simple explanation of the null effect in the *Name* condition comparison: we observed no LAN effect in the *Name Co-referent* vs. *Name, No Co-referent* comparison because no back association operations were initiated.

In addition to this work on reference form, the model in Lewis et al. (2006), based on well-established memory constructs such as interference, encoding, and retrieval, is intended to explain numerous effects in the sentence processing literature (though as stated above, here we do not commit to a specifically retrieval-based functional interpretation of observed LAN effects). Most pertinent to our purposes

[11]In fact, as Kazanina et al. argue, it is probably the processing of the first element in the dependency, namely the pronoun, which imposes the processing burden on the parser. Unfortunately, as discussed below (see fn. 15), there was despite our best intentions no suitable comparison for this pronoun preceding its proper name co-referent in our materials, and it was thus impossible for us to determine the electrophysiological indices of the processing burden imposed by these constructions. It seems reasonable to us to suppose that such an electrophysiological response may have been present in our data, but this issue will need to be resolved in future research.

is the notion of cue-based content-addressable retrieval (a notion first elaborated in McElree et al. (2003)). Simplifying for ease of exposition, features of certain words in the language input trigger the retrieval of previously processed items in order to associate them. The retrieval operation is “content addressable,” (i.e., access is direct rather than serial), and cue-based, such that the retrieval is cued by features, and retrieval efficiency is dependent on the number of items in memory that share features with the cue.

Under this set of assumptions, it becomes easy to explain the contrast between our *Pronoun* and *Name* comparisons. The second element in the *Pronoun*, *Co-referent* condition is a pronoun, which could plausibly be assumed to have a feature that cues the retrieval of a suitable antecedent (presumably in addition to, for example, number and gender cues), particularly when a candidate antecedent is available in the same sentence. In contrast, no such [+antecedent] feature is associated with the second element (a proper name) in the *Name Co-referent* condition, and therefore retrieval can plausibly be assumed not to be cued.

Additionally, we believe that the model in Lewis et al. (2006) provides further support for our assertion that the LAN effect in response to our *Pronoun*, *Co-referent* does not index the mechanisms underlying the processing of referential ambiguity, but rather a negativity that is part of a continuum, the amplitude of which is determined by the difficulty of forming a referential dependency. Specifically, there is no need for the parser to search for an antecedent outside the current discourse in our experimental paradigm (but see Nieuwland (2014) for an experimental paradigm very different from ours in which participants appear to be doing just that), given that there is a readily available, featurally appropriate and syntactically licensed antecedent within the current sentence. In sum, this architecture helps to explain our pattern of results, and in combination with the work on reference form, supports our contention that no back association operation is triggered in the *Name*, *Co-referent* condition.

3.3. *Relation between syntactic and referential dependencies*

An important part of our argument for a domain-general back association interpretation of the LAN rests on its elicitation by the process of association in the comparison between object relatives and their controls, as well as in the comparison between the two *Pronoun* conditions. In support of this view, the second element in both syntactic and referential dependencies elicited LAN in our data. At the same time, these negativities were not identical in either their extended time course or in their amplitude distributions across the scalp, even within the same group of participants. Although this may seem non-optimal for our position, it is also not sufficient to undermine our argument, as discussed in the following sections.

3.3.1. *The time course of anterior negativity*

Brain responses to the second elements in syntactic dependencies (i.e. the main clause verb following the gap position at the end of the relative clause; Table 2, condition E) and referential dependencies (i.e. the main clause pronoun subject following the sentence-initial adjunct; Table 1, condition A) appeared to persist beyond the critical word, as confirmed by interactions of these LAN effects with factors of anteriority and hemisphere in the multi-word analyses (see Section 2.2) between 300 and 2000 ms post-onset of the critical word (Section 3.2 and Figs. 4 and 5A). The responses to syntactic and referential dependencies were thus very similar in eliciting apparently sustained effects.

However, the responses to syntactic and referential dependencies differed in the individual word analyses undertaken between 300 and 500 ms post-onset of the words following the critical word. In the individual word analyses of the object relative vs. conjoined clause comparison, there were no further significant differences after the critical word when the onsets of subsequent words were rebaselined. In other words, only the critical word, the main clause verb in object relative clauses, elicited a LAN effect in this comparison; subsequent words apparently sustained the effect but did not contribute to it independently. On the other hand, the individual word analyses of pronouns with and without antecedents showed a significant difference between 300 and 500 ms not only in response to the critical word, but also in response to the two words following it, even after rebaselining. This suggests that the two words following the pronoun subject in this comparison made independent contributions to the LAN effect.

However, we believe that the response to the critical word in the pronoun comparison was entirely independent of the responses to the two following words, for the following reason: when we measured 400–600 ms post-onset of the critical word (i.e. the last 100 ms of the response to the critical word and the first 100 ms of the response to the following word), we found no significant differences. This suggests that the LAN effect in response to the critical word did not spill over into the following word (*longed* in Table 1), but rather that this word appears to have elicited an independent LAN effect of its own. We similarly found no significant differences in an analysis of the 400–600 ms time window post-onset of word 13, suggesting that the LAN effect at this position was also independent of the LAN effect elicited by word 14 (*for* in Table 1). This difference in the individual word analyses suggests that incoming words following the second element in a referential dependency may induce additional processing difficulty, while those following the second element in a syntactic dependency do not.[12] In any case, the differences in time course elicited during the processing of referential and syntactic dependencies point to similar but non-identical mechanisms underlying the processing of these long-distance relationships.

3.3.2. *The scalp distribution of anterior negativity*

Differences in the scalp topography of anterior negativities, both phasic and sustained, in response to syntactic dependencies are not without precedent in the literature: though often

[12] Though we do not wish to make strong commitments to the nature of such possible additional processing costs, there is evidence in the literature that the formation of relationships between antecedents and pronouns under various discourse conditions (Sanford et al., 1983; Garrod and Sanford, 1990) and between referents and open discourse roles (Garrod and Terras, 2000) actually occurs in two processing stages. If this is in fact the case, it could account for the differences in the time course of the anterior negativities observed across our dependency types, as to our knowledge no such explicit claims have been made regarding the formation of syntactic dependencies.

reported with canonical left anterior maxima (see Section 1), bilateral anterior (King and Kutas, 1995; Ueno and Kluender, 2003) and right anterior (Müller et al., 1997; Ueno and Kluender, 2009; Dillon et al., 2012) distributions are also common. Moreover, although one cannot infer the location of neural generators for ERP effects from their amplitude distribution at the scalp, the fMRI literature suggests reasonable, overlapping but non-identical neural sources for the negativities elicited in response to the second element in a long-distance dependency. Recall (Section 1.6) that Nieuwland et al. (2007) tentatively localized the generators of N*ref* effects in response to referential ambiguity to medial prefrontal cortex. In contrast, an array of fMRI and PET studies have typically reported increased activation in response to long-distance syntactic dependencies in traditional language areas of left lateral cortex. For example, Just et al. (1996) compared object relative clauses, subject relatives, and coordinate clause controls (such as were used in our experiment) and reported the greatest activation in and around Broca's and Wernicke's areas of the left hemisphere. Other neuro-imaging studies (Stromswold et al., 1996; Caplan et al., 1998, 1999; Caplan et al., 2000, 2001; Cooke et al., 2001; Ben-Schachar et al., 2003, 2004) have likewise identified anterior and posterior (though not always both) language-related areas of activation on the lateral surface of the left hemisphere in response to the more complex of two constructions, in most cases when comparing object to subject dependencies. Crucially for our purposes, none of the reported loci of activation were in medial frontal areas. Though these inferences remain somewhat speculative, these neuro-imaging studies offer a plausible account of the differences in the scalp distributions of the negativities observed in response to the second element in referential and syntactic dependencies, which thus far appear to be have medial and lateralized generators, respectively.

It is possible that that differences in both the time course and the scalp distribution of the effects elicited by the second element in our syntactic and referential dependencies could turn out to be specific to the group of participants we tested and/or the set of materials we used,[13] and therefore that the underlying mechanisms involved in processing these dependencies are more similar than our data suggest. After all, contra to our findings, previous studies investigating syntactic object dependencies have reported LAN effects extending beyond critical words following gapped positions, even after re-baselining (Kluender and Kutas, 1993a:205–6; King and Kutas, 1995:386–7). Likewise, it is possibile that the somewhat atypical left-lateral scalp topography of the LAN response to object relative sentences in our results[14] is specific to the particular experimental circumstances of our study. However, even if this should turn out to be the case, either in terms of time course or of scalp distribution, it would only lend further credence to our back association hypothesis, which predicted LAN effects in response to the second element in both dependency types. If future work using different materials shows the time course and distribution of these effects to be even more similar than those elicited in our study, this would still be entirely consistent with our claims.

As stated in the introduction, there has been a longstanding debate within the theoretical linguistics literature (Chomsky, 1980, 1981; Reinhart, 1983; Reuland and Reinhardt, 1995) as to exactly which types of referential relationships are handled within the syntax and which are handled via pragmatics and principles of discourse. While it is too early to make any concrete claims based on electrophysiological evidence, we believe that experiments like the one reported here represent a potentially fruitful way to help tease apart these issues and begin to understand the cognitive processes involved in the formation of various types of long-distance relationships.

4. Experimental procedures

4.1. *Participants*

20 monolingual native speakers of English participated in the experiment (9 females, 11 males; mean age=20.7), and received either course credit or $7 an hour for their participation. All students were enrolled as students at the University of California, San Diego. All participants were right-handed with no neurological or psychiatric disorders and normal or corrected-to-normal vision.

4.2. *Materials*

A total of 160 sets of the four conditions shown in Table 1 were created for the purposes of this experiment. Each experimental item was fifteen words long, and consisted of an eleven-word sentence-initial adjunct followed by a four-word main clause. Co-referents of the main clause subject occupied the same position in the sentence-initial adjunct across the two *Co-referent* conditions (condition A: co-referential antecedent, condition C: preceding co-referential pronoun), and the four sentence positions intervening between the first and second co-referents were identical in the main comparisons, establishing a clean baseline for our critical comparisons, which were made at the subject position introducing the main clause, and in all cases consisted of comparisons between identical lexical items.

The sentence-initial adjuncts in the primary experimental conditions all had either *after, because, since, until,* or *during* in initial position, with the occurrence of these adverbials balanced across items (32 items per initial adverbial). The verbs internal to these adjuncts, crucial to the experimental manipulation, were selected in one of two ways. First, as in the example in Table 1, we selected transitive verbs that, when nominalized, could serve as the internal argument of verbs like *require* or *provide*. In this manipulation, co-referents served as direct objects of the verb

[13]Somewhat atypically, our materials did not include subject relative clauses as a point of comparison for object relative clauses, but rather coordinate clause controls matched to the relative clause sentences in lexical content (which was not the case for example in the materials used by King and Kutas, 1995). Additionally, our object relative clauses were by design and intent embedded in a set of sentences containing various types of referential relationships, which obviously provides a very different context from those used in previous studies examining the processing of syntactic dependencies.

[14]In fact, subsequent unpublished work conducted in our lab comparing these exact same object relative and coordinate clause sentences elicited left anterior responses with a more standard scalp distribution when not mixed with conditions containing referential dependencies.

(e.g. *deployed Will/him)* in the two *Co-referent* conditions, while the nominalization eliminated the co-referent (e.g. *required deployment*) in the two *No Co-referent* conditions. Second, verbs were selected from the list of "*pro-arb*" verbs in Levin (1993). These verbs, which select for optional direct object arguments, were employed in their transitive forms in the two *Co-referent* conditions (e.g. *startled Katherine*) and were changed into an adjectival predicate form in the two *No Co-referent* conditions (e.g. *was startling)*. These verb selection strategies ensured that there was minimal repetition of adjunct-internal verbs across items, with a single verb repeated no more than four times across experimental items. The object of the sentence-initial preposition, an adjective-noun combination like *covert mission* in Table 1 (which also served as the head noun of the subject relative clause internal to the sentence-initial adjunct) was unique in each item set and held constant across conditions.

80 additional sentences constituted an orthogonal manipulation of crucial interest, namely 40 object relative clauses (condition E, Table 2) and their coordinate clause controls (condition F, Table 2); cf. King and Kutas (1995). By including these syntactic dependencies in the materials, we were able to look for similarities and differences in the brain responses elicited by these dependencies compared to those elicited by referential dependencies in the same group of participants.

The set of materials also contained 80 filler sentences, shown in Table 3. In an attempt to provide a point of comparison for the adjunct-internal pronoun in the *Name, co-referent* condition C, subjects read 40 fillers in which an adjunct-internal pronoun referred back to a preceding sentence-initial antecedent (condition G, Table 3). In retrospect, this comparison turned out to be confounded and will not be discussed further.[15] 40 of the other fillers were similar to the experimental sentences, and were designed to add variety to the materials. 30 of these 40 contained sentence-medial (rather than sentence-initial) adverbials (*after, because, since, until, during,* and, in addition, *before*) and varied in terms of both number and type (common versus proper) of nouns present (condition G, Table 3). The remaining ten filler items were identical in structure to the experimental items, except that the sentence-initial adjunct began with *before*[16] (condition H, Table 3).

The resulting set of 320 sentences was rotated through a Latin square design into ten lists, with two orders per list, for a total of 20 lists. The sentences in each list were pseudo-randomized such that three sentences of the same condition never appeared consecutively, and there were never more than five items from the same condition within a sequence of ten items.

4.3. *Procedures*

Subjects were run in a single session that lasted approximately 2.5 h, including preparation. The experiment consisted of five blocks, each containing 64 sentences and lasting approximately 15–17 min. During the session, subjects were seated comfortably in a chair in a sound-attenuated booth. 1500 ms before the first word of a sentence, a red square appeared in the middle of the screen and remained throughout sentence presentation for fixation purposes. Sentences were visually presented above fixation, with each word presented for 300 ms (500 ms stimulus onset asynchrony (SOA)). True/false comprehension questions were presented at the end of 25% of the materials (probing both experimental manipulations and the filler sentences); sentences without a comprehension question ended with a "Press for next" message that remained on the screen until the subject pressed either response button. Comprehension questions focused on the content of the preceding sentence, but crucially not any co-referential relationships contained within the sentence (e.g. "*Was Will deployed for a full year?*"). Comprehension questions appeared 2000 ms after the end of the sentence-final word and remained on the screen until the participant made a response with one of two hand-held buttons. Response hands were balanced across participants to control for handedness, and the correct response was balanced across comprehension questions. The next stimulus sentence began 3000 ms after the subject's response. In order to familiarize participants with the task, they completed a ten-sentence practice session before beginning the experiment.

Table 3 – Filler sentences.

G. Sentence-medial adverbials
Lori nailed the interview after Lawrence coached her on how to impress the recruiter.
H. Sentence-initial adverbials (before)
Before Marcus had even started his dissertation, Vanessa graduated with honors.
I. PP-internal anaphor with sentence-initial antecedent
Will joined a mission that deployed **him** for nine terrible months, and longed for home.

4.4. *Electrophysiological recording*

The electroencephalogram (EEG) was recorded from 26 tin electrodes mounted geodesically in a commercially available Electro-Cap. These sites included midline prefrontal (MiPf), left and right lateral prefrontal (LLPf and RLPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral frontal (LLFr and RLFr), left and right medial frontal (LMFr and RMFr), left and right medial lateral frontal (LDFr and RDFr), left and right medial central (LMCe and RMCe), midline central (MiCe), left and right medial lateral central (LDCe and RDCe), left and right lateral temporal (LLTe and RLTe), left and right medial lateral parietal (LDPa and RDPa), midline parietal (MiPa), left and right lateral occipital (LLOc and RLOc), left and right medial occipital (LMOc and RMOc), and midline occipital (MiOc). Each electrode was referenced online to the reference electrode at the left mastoid and later re-referenced offline to the average of the two mastoids. To monitor blinks and eye movements, electrodes were placed on the outer canthi and under each eye. Impedances were kept below 5 KΩ during recording. The EEG was amplified using James Long amplifiers with an online band-pass filter (.01–100 Hz), and digitized at a sampling rate of 250 Hz.

[15]In constructing Condition G as a control for Condition C, we failed to take into account the fact that these two conditions might elicit responses for different reasons – one has a preceding antecedent, much like the main clause subject comparison, while condition C contains a pronoun preceding its antecedent, also known as a cataphor.

[16]The temporal re-sequencing of events required when processing *before* is itself known to elicit LAN (Münte et al, 1998), and therefore this type of sentence was not included in the experimental materials.

4.5. Data analysis

The main analyses, as described below, focused on the critical main clause subject position (i.e. *he* in conditions A and B, and *Will* in conditions C and D). Mean amplitude measurements were taken over 1000 ms epochs (including a 100 ms pre-stimulus baseline) with primary focus on the LAN time window between 300 and 500 ms post-stimulus onset. Analyses between 500 and 800 ms were also conducted in all comparisons, but were of special interest in the comparison of the two *Pronoun* conditions, due to the possibility that the antecedent-less pronoun might elicit late positivity, as discussed in Section 1.6.

When LAN effects in response to the critical word position appeared to extend beyond 800 ms, across several word positions, mean amplitude measurements of 2000 ms epochs (including a 500 ms pre-stimulus baseline) were taken. These multi-word averages enabled us to measure effects of sustained negativity across the course of the sentence. However, multi-word averages cannot distinguish between initial processing costs elicited by the first (critical) word in the average and additional processing costs contributed by subsequent words. For this reason, we followed King and Kutas (1995) and Phillips et al. (2005) in computing mean amplitude measurements relative to a 100 ms baseline re-established *at each word position in the region spanned by the sustained effect*. If a sustained effect is due solely to an initial processing cost triggered by the first word in a multi-word average that remains constant over time, there should be no observable differences in the responses to the individual words spanned by the sustained effect following the re-baselining procedure. On the other hand, if any subsequent word in the multi-word average contributes to the sustained effect, a significant difference should be evident in that particular individual word average.

Trials contaminated by excessive muscle activity, amplifier blocking, or eye movements were discarded before averaging. This led to the rejection of 8.3% of trials in the single-word averages and 33.8% of trials in the multi-word averages.[17] The averaged data were algebraically re-referenced off line to the average of the activity recorded at the two mastoid sites. ERP waves were smoothed offline using a low-pass filter with a 7 Hz cutoff for visualization purposes only.

Because of the design of the materials, a full factorial ANOVA was deemed inappropriate.[18] Rather, we first conducted one-tailed *t*-tests for predicted effects.[19] For unpredicted effects, we used one-way ANOVAs to make the critical comparisons between (1) the two *Pronoun* conditions and between (2) the two *Name* conditions. Each analysis had the repeated measures of experimental condition (two levels) and electrode (26 levels). This will be referred to as the full analysis. In addition (independent of whether we first ran *t*-tests or omnibus ANOVAs), a subsequent distributional analysis was conducted for each of the three main comparisons, including condition (two levels), hemisphere (left vs, right), anteriority (prefrontal vs. frontal vs. temporal vs. occipital), and laterality (lateral vs. medial). Electrodes included in the distributional analysis were left and right lateral prefrontal (LLPf and RLPf), left and right medial prefrontal (LMPf and RMPf), left and right lateral frontal (LLFr and RLFr), left and right lateral temporal (LLTe and RLTe), left and right medial lateral parietal (LDPa and RDPa), left and right lateral occipital (LLOc and RLOc), and left and right medial occipital (LMOc and RMOc). This distributional analysis was conducted in order to investigate further those effects that appeared to be spatially localized. We felt that this type of analysis was warranted because one of the stated goals of the experiment was to examine the ways in which the brain's responses to referential and syntactic dependencies patterned together or differently, and one dimension along which these effects may differ is in their topographies over the scalp. Furthermore, when it was necessary to corroborate local effects suggested by the distributional analysis, ANOVAs were performed at individual levels of these factors. The Huynh-Feldt (1976) correction for lack of sphericity was applied, and corrected *p*-values are reported with the original degrees of freedom.

[17]This high rejection rate was a consequence of averaging over the last four words (a period of two seconds) of sentences that were 15 words in length.

[18]This type of analysis would force us to compare pronouns to proper names. This comparison would elicit differences driven solely by lexical factors of word class (Müller and Kutas, 1996) that were not the focus of this experiment.

[19]This type of analysis was conducted on effects that were predicted in terms of the sentence position at which they were elicited, the polarity of the effect and its distribution at the scalp (see Section 1.6 for explicit predictions).

REFERENCES

Anderson, J., Holcomb, P., 2005. An electrophysiological investigation of the effects of coreference on word repetition and synonymy. Brain Lang. 43, 1–17.

Ariel, M., 1990. Accessing Noun Phrase Antecedents. Routledge, London.

Baddeley, A., 1992. Working memory. Science 255 (5044), 556–569.

Ben-Schachar, M., Hendler, T., Kahn, I., Ben-Bashat, D., Grodzinsky, Y., 2003. The neural reality of syntactic transformations: evidence from functional magnetic resonance imaging. Psychol. Sci. 14 (5), 433–440.

Ben-Schachar, M., Palti, D., Grozinsky, Y., 2004. Neural correlates of syntactic movement: converging evidence from two fmri experiments. NeuroImage 21, 1320–1336.

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

Caplan, D., Alpert, N., Waters, G., 1998. Effect of syntactic structure and propositional number on patterns of regional cerebral blood flow. J. Cogn. Neurosci. 10, 541–552.

Caplan, D., Alpert, N., Waters, G., 1999. PET studies of syntactic processing with auditory presentation. NeuroImage, 43–51.

Caplan, D., Alpert, N., Waters, G., Olivieri, A., 2000. Activation of Broca's area by syntactic processing under conditions of concurrent articulation. Hum. Brain Mapp. 9, 65–71.

Caplan, D., Vijayan, S., Kuperberg, G., West, C., Waters, G., Greve, D., Dale, A., 2001. Vascular responses to syntactic processing: event-related fmri study of relative clauses. Hum. Brain Mapp. 15, 26–38.

Chomsky, N., 1980. On binding. Linguist. Inq. 11 (1), 1–46.

Chomsky, N., 1981. Lectures on Binding and Government. Foris Publications, Dordrecht.

Chung, S., 1994. Wh-agreement and referentiality in Chamorro. Linguist. Inq. 25, 1–44.

Cinque, G., 1990. Types of A'-dependencies. MIT Press, Cambridge, MA.

Cooke, A., Zurif, E., DeVita, C., Alsop, D., Koenig, P., Detre, J., Gee, J., Pinango, N., Balogh, J., Grossman, M., 2001. Neural basis for sentence comprehension: grammatical and short-term memory components. Hum. Brain Mapp. 15, 80–94.
Coulson, S., King, J., Kutas, M., 1998. Expect the unexpected: event-related brain response to morphosyntactic violations. Lang. Cogn. Process. 13 (1), 21–58.
Cowan, N, 2000. The magical number 4 in short-term memory: a reconsideration of mental storage capacity. Behav. Brain Sci. 24, 87–185.
Dillon, B., Nevins, A., Austin, A., Phillips, C., 2012. Semantic and syntactic predictors of tense in Hindi: An ERP investigation. Lang. Cogn. Process. 27 (3), 313–344.
Felser, C., Clahsen, H., Münte, T., 2003. Storage and integration in the processing of filler-gap dependencies: An ERP Study of topicalization and wh-movement in German. Brain Lang. 87, 345–354.
Fodor, J., 1978. Parsing strategies and constraints on transformations. Linguist. Inquiry 9 (3), 427–473.
Donchin, E., 1981. Surprise!.....surprise?. Psychophysiology 18 (5), 493–513.
Frazier, L., 1979. On Comprehending Sentences: Syntactic Parsing Strategies. University of Connecticut Dissertation.
Garrod, S., Sanford, A., 1990. Referential processing in reading: focusing on roles and individuals. In: Balota, D.A., Flores d'Arcias, G.B., Rayner, K. (Eds.), Comprehension Processes in Reading. Erlbaum, Hillsdale, NJ, pp. 465–486.
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.
Gundel, H., Hedberg, N., Zacharski, R., 1993. Cognitive status and the form of referring expressions. Language 69 (2), 274–307.
Hagoort, P., Brown, C., Groothusen, J., 1993. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. Lang. Cogn. Process. 8 (4), 439–483.
Hagiwara, H., Soshi, T., Isihara, M., Imanaka, K., 2007. A topographical study on the event-related potential correlates of scrambled word order in japanese complex sentences. J. Cogn. Neurosci. 19 (2), 175–193.
Heine, A., Tamm, S., Hoffman, M., Hutzler, F., Jacobs, A., 2006. Does the frequency of the antecedent noun affect the resolution of pronominal anaphors? An ERP study. Neurosci. Lett. 400, 7–12.
Huynh, H., Feldt, L., 1976. Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. J. Educ. Behav. Stat. 1 (1), 69–82.
Just, M.A., Carpenter, P.A., 1992. A capacity theory of comprehension: individual differences in working memory. Psychol. Rev. 98, 122–149.
Just, M., Carpenter, P., Keller, T., Eddy, W., Thulborn, K., 1996. Brain activation modulated by sentence comprehension. Science 274, 114–116.
Kaan, E., Harris, T., Gibson, E., Holcomb, P., 2000. The P600 as an Index of syntactic integration difficulty. Lang. Cogn. Process. 15 (2), 159–201.
Kazanina, N., Lau, E., Lieberman, M., Yoshida, M., Phillips, C., 2007. The effect of syntactic constraints on the processing of backwards anaphora. J. Mem. Lang. 56, 384–409.
King, J., Kutas, M., 1995. Who did what and when? Using word- and clause-level ERPS to monitor working memory usage in reading. J. Cogn. Neurosci. 7 (3), 376–395.
Kluender, R., Kutas, M., 1993a. Bridging the gap: evidence from ERPs on the processing of unbounded dependencies. J. Cogn. Neurosci. 5 (2), 196–214.
Kluender, R., Kutas, M., 1993b. Subjacency as a processing phenomenon. Lang. Cogn. Process. 8 (4), 573–633.
Kok, A., 2001. On the utility of P300 amplitude as a measure of processing capacity. Psychophsyiology 38, 557–577.
Kutas, M., Hillyard, S., 1983. Event-related brain potentials to grammatical errors and semantic anomalies. Mem. Cogn. 11 (5), 539–550.
Kutas, M., Van Petten, C., Kluender, R., 2006. Psycholinguistics electrified II (1994–2005). In: Gernsbacher, M.A., Traxler, M. (Eds.), Handbook of Psycholinguistics, 2nd ed.. Elsevier Press, New York.
Kwon, N., Kluender, R., Kutas, M., Polinsky, M., 2013. Subject/object processing asymmetries in Korean relative clauses: evidence from ERP data. Language 89 (3), 537–585.
Levin, B., 1993. English Verb Classes and Alternations: A Preliminary Investigation. University of Chicago Press, London.
Lewis, R., Vasishth, S., Van Dyke, J., 2006. Computational principles of working memory in sentence comprehension. Trends Cogn. Sci. 10 (10), 447–454.
Marslen-Wilson, W., Levy, E., Tyler, L.K., 1982. Producing interpretable discourse: the establishment and maintenance of reference. In: Jarvella, R.J., Klein, W. (Eds.), Speech, Place, and Action. Wiley, Chicester.
Martin, A.E., Nieuwland, M.S., Carreiras, M., 2012. Event-related brain potentials index cue-based retrieval interference during sentence comprehension. NeuroImage 59, 1859–1869.
Martin, A.E., Nieuwland, M.S., Carreiras, M., 2014. Agreement attraction during comprehension of grammatical sentences: ERP evidence from ellipsis. Brain Lang. 135, 42–51.
Martin-Loeches, M., Munoz, F., Casado, P., 2005. Are the anterior negativities to grammatical violations indexing working memory?. Psychophysiology 42 (5), 508–519.
Matzke, M., Mai, H., Nager, W., Russeler, J., Münte, T., 2002. The costs of freedom: an ERP study of non-canonical sentences. Clin. Neurophysiol. 113, 844–852.
McElree, B., Foraker, S., Dyer, L., 2003. Memory structures that subserve sentence comprehension. J. Mem. Lang. 48 (1), 67–91.
Müller, H., Kutas, M., 1996. What's in a name? Electrophysiological differences between spoken nouns, proper names and one's own name. Neuroreport 8 (1), 221–225.
Müller, H., King, J., Kutas, M., 1997. Event-related potentials elicited by spoken relative clauses. Cogn. Brain Res. 5, 193–203.
Münte, T., Schlitz, S., Kutas, M., 1998. When temporal terms belie conceptual order. Nature 395, 71–74.
Neville, H., Nicol, J., Barss, A., Forster, K., Garrett, M., 1991. Syntactically based sentence processing classes: evidence from event related brain potentials. J. Cogn. Neurosci. 3 (2), 151–165.
Nieuwland, M., van Berkum, J., 2006. Individual differences and contextual bias in pronoun resolution. Brain Res. 1118, 155–167.
Nieuwland, M., Petersson, K.M., Van Berkum, J., 2007. On sense and reference: examining the functional neuroanatomy of referential processing. NeuroImage 37, 993–1004.
Nieuwland, M., Otten, M., Van Berkum, J., 2007. Who are you talking about? tracking discourse level referential processing with event-related brain potentials. J. Cogn. Neurosci. 19, 228–236.
Nieuwland, 2014. "Who's he?" Event-related brain potentials and unbound pronouns. J. Mem. Lang. 76, 1–28.
Osterhout, L., Mobley, L., 1995. Event-related brain potentials elicited by failure to agree. J. Mem. Lang. 34 (6), 739–773.
Osterhout, L., Holcomb, P., 1992. Event-related potentials elicited by syntactic anomaly. J. Mem. Lang. 31, 785–806.
Pesetsky, D., 1990. Wh-in-situ: movement and selective binding. In: Reuland, Martin (Eds.), The Representation of (in-)definiteness. MIT Press, Cambridge, MA.
Phillips, C., Kazanina, N., Abada, S., 2005. ERP Effects of the processing of syntactic long-distance dependencies. Cogn. Brain Res. 22, 407–428.
Reinhart, T., 1983. Coreference and bound anaphora: a restatement of the anaphora questions. Linguist. Philos. 6, 47–88.

Reinhart, T., Reuland, E., 1993. Reflexivity. Linguist. Inquiry 24, 657–720.
Rose, S., 2011. Long Distance Assimilation of Consonants. The Blackwell Companion to Phonology. In: Marc van Oostendorp, Colin, Ewen, Elizabeth, Hume, Rice, Keren (Eds.), Wiley-Blackwell.
Rose, S., Walker, R., 2011. Harmony systems. In: Goldsmith, J., Riggle, J., Yu, A. (Eds.), Handbook of Phonological Theory 2nd ed. Blackwell.
Rösler, F., Pechmann, T., Streb, J., Roder, B., Henninghausen, 1998. Parsing of sentences in a language with varying word order: word-by-word variations of processing demands are revealed by event-related potentials. J. Mem. Lang. 38 (2), 150–176.
Sanford, A., Garrod, S., Lucas, A., Henderson, R., 1983. Pronouns without explicit antecedents. J. Semant. 2, 303–318.
Shao, J., Neville, H., 1998. Analyzing Semantic Processing Using Event-Related Brain Potentials. The Newsletter of the Center for Research in Language.
Silverstein, M., 1976. Hierarchy of features and ergativity. In: Dixon, R.M.W. (Ed.), Grammatical Categories in Australian Languages. Humanities Press, New Jersey.
Stromswold, K., Caplan, D., Alpert, N., Rausch, S., 1996. Localization of syntactic comprehension by positron emission tomography. Brain Lang. 52, 452–473.
Ueno, M., Garnsey, S., 2007. An ERP study of the processing of subject and object relative clauses in Japanese. Lang. Cogn. Process. 23 (5), 646–688.
Ueno, M., Kluender, R., 2003. Event-related brain indices of Japanese scrambling. Brain Lang. 86, 243–271.
Ueno, M., Kluender, R., 2009. On the processing of japanese Wh-questions. Brain Res., 1–29.
Van Berkum, J., Brown, C., Hagoort, P., 1999. Early referential context effects in sentence processing: evidence from event-related brain potentials. J. Mem. Lang. 41, 147–182.
Van Berkum, J., Brown, C., Hagoort, P., Zwisterlood, P., 2003. Event-related brain potentials reflect discourse-referential ambiguity in spoken language comprehension. Psychophysiology 40, 235–248.
Van Berkum, J.J.A. van, 2004. Sentence comprehension in a wider discourse: can we use erps to keep track of things?. In: Carreiras Jr., M., Clifton, C. (Eds.), The On-line Study of Sentence Comprehension: Eyetracking, ERPs and Beyond. Psychology Press, New York.
Van Berkum, J., Koorneef, A., Otten, M., Nieuwland, M., 2007. Establishing reference in language comprehension: an electrophysiological perspective. Brain Res. 1146, 158–171.
Vonk, W., Hustinx, L., Simons, W., 1992. The use of referential expressions in structuring discourse. J. Verbal Learn. Verbal Behav. 18, 545–567.
Weckerly, J., Kutas, M., 1999. An electrophysiological analysis of animacy effects in the processing of object Relatives. Psychophysiology 36, 559–570.