Interplay of meaning, syntax and working memory during pronoun resolution investigated by ERPs

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/brainres

BRAIN RESEARCH

Research Report

Interplay of meaning, syntax and working memory during pronoun resolution investigated by ERPs

Anke Hammer[a,*], *Bernadette M. Jansma*[b], *Monique Lamers*[c], *Thomas F. Münte*[a]

[a]*Department of Neuropsychology, Otto-von-Guericke-Universität Magdeburg, Germany*
[b]*Faculty of Psychology and Neuroscience, Department of Cognitive Neuroscience, Maastricht University, Maastricht, The Netherlands*
[c]*Department of Linguistics, University of Nijmegen, The Netherlands*

ARTICLE INFO

Article history:
Accepted 2 July 2008
Available online 9 July 2008

Keywords:
Language
Pronoun processing
Working memory
Syntax
Meaning
Event-related potential (ERP)
N400
P600

ABSTRACT

Event-related potentials were used to investigate the interaction of verbal working memory and gender information during pronoun resolution. Gender information is supposed to be disentangled using sentences about persons (semantic/syntactic) or things (syntactic) followed by gender congruent or incongruent pronouns. Memory was manipulated using differential distances (short distance (SD) and long distances with or without intermediate subject gaps (LD gap and LD no gap)) between the pronoun and the antecedent. Comparing incongruent to congruent conditions, person sentences with SD and LD no gap resulted in an N400-like effect indicating the involvement of semantic integration, whereas a P600 effect in LD gap (re-activated antecedents) sentences suggested the involvement of syntactic reanalysis. SD-thing sentences showed a P600 effect, whereas LD thing sentences revealed no effect at pronoun position. A delayed N400 effect for thing sentences was observed instead. Based on preceding and the current data, we present a working model on how the parser switches between the use of semantic and syntactic information to establish co-reference and how this switch depends on the type of antecedent, distance, or syntactic structure.

1. Introduction

Personal pronouns are linguistic devices which allow us to refer back to an earlier mentioned person or thing in discourse and thus, creating coherence in written or spoken language. Considering the sentence "The woman is smiling, because she is happy." it is easy to tell that *she* refers to the woman within the sentence. Several cognitive models outline the details of this process (Garrod and Sanford, 1994; Graesser et al., 1997; Gordon and Hendrick, 1998). Garrod and Sanford (1994) presented a model with two stages for pronoun interpretation (see also Garrod and Terras, 2000) with (1) a bonding or immediate recovery stage which is an automatic, more syntax-related process, and (2) a resolution or immediate integration stage which is evaluating the established link, i.e. by weighting syntactic and semantic discourse information. But, pronouns do not always occur in the same distance to the antecedent — in fact, they can be close or distant to their antecedent. Consequently, the question arises whether different demands on memory capacity might lead to different processing of close or distant pronouns — and more importantly whether differential distances alter syntactic and semantic integration processes.

* *Corresponding author*. Department of Neuropsychology, FNW/Institute of Psychology, Otto-von-Guericke-Universität Magdeburg, Post-Box 4120, 39 016 Magdeburg, Germany. Fax: +49 391 67 11947.
E-mail address: anke.hammer@ovgu.de (A. Hammer).

0006-8993/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.brainres.2008.07.004

There are two main classes of models on sentence processing to be differentiated. They agree that different sub-processes underlie the comprehension process. However, they differ in their characterization of the processes involved and the time course of the use of certain sorts of information (Frauenfelder and Tyler, 1987, Hagoort et al., 1999). One class of these models are the so-called serial or syntax-first models (e.g. Frazier and Fodor, 1978, Ferreira and Clifton, 1986) which claim that the parser initially builds up a syntactic structure independent of semantic information. Within these models the syntactic component is explicitly restricted to the domain of syntactic information. Non-syntactic information does not influence the analysis of the sentence at that early stage. Semantic and pragmatic information are represented in a different component and are used at a later stage. Thus, these models assume that syntactic and semantic processes are supported by different components within the comprehension system, and these may be implemented by distinct brain regions at the neuronal level. The second class of models, the so-called interactive models, assumes that syntactic and semantic information interact at any time during parsing (e.g. Marslen-Wilson and Tyler, 1980, Trueswell and Tanenhaus, 1994). In these models, all (available) sources immediately affect the parsing process without appealing a separate level of syntactic representation. Syntactic, lexical, and semantic/conceptual information are used interactively and continuously. For a debate concerning these controversial assumptions on a neural base, see Friederici (2002) and Hagoort (2003). For a review on neurobiology in sentence processing the reader is referred to Osterhout et al. (in press). Based on Vosse and Kempen (2000), Hagoort (2003, 2005) proposed a Unification model for language processing. According to this model each word in the linguistic memory (lexicon) is associated with a structural frame that specifies the possible structural environment of this particular word. All elements within the frame unify with each other to form a comprehensive sentence. Unification is a dynamic process over time based on syntactic but also on semantic/pragmatic and phonological rules. The second important assumption in the unification model is that unification links undergo a gradual decay over time. This decay can functionally be seen as decay in verbal working memory. As root nodes (single phrasal node, NP) have a syntactic function, this decay suggests that any type of syntactically based unification can only be carried out within the sentence and eventually can only span across a certain range of NPs. However, none of the models make direct predictions of the influence of differential working memory demands on semantic and syntactic processing during anaphoric resolution.

ERPs can be used to investigate language processes because different components have been found to be an index of semantic, syntactic and memory processes. In preceding ERP studies, syntactic integration difficulty became manifest in a positive shift that starts at about 400 to 500 ms after word onset, the so-called P600/SPS effect. The P600 related research has a long tradition starting with verb-agreement violations (Osterhout and Holcomb, 1992, Hagoort et al., 1993), and followed by a variety of syntactic tasks that were tested separately, i.e. verb inflection errors (Friederici et al., 1993, Gunter et al., 1997), case inflection errors (Münte et al., 1998a), and phrase structure violations (Neville et al., 1991, Friederici et al., 1999, Hahne and Friederici, 1999). These results have led researchers to assume that the P600 amplitude is related to "syntactic processing difficulty" and that this phenomenon might be related to syntactic reanalysis (Osterhout et al., 1994, Rösler et al., 1998), syntactic integration difficulty (Kaan et al., 2000, Kaan and Swaab, 2003), or the inability to assign the preferred structure to the words encountered (Hagoort et al., 1993, Friederici et al., 1999). In pronoun processing this component was found for a disagreement in syntactic gender information between pronoun and antecedent compared to congruent gender information indicating an involvement of syntactic gender information processing (Osterhout and Mobley, 1995, Osterhout et al., 1997, Schmitt et al., 2002, Hammer et al., 2005, Lamers et al., 2006). Semantic integration difficulties elicit a negative deflection with a peak around 400 ms and a right central-parietal negative maximum for incongruent compared to semantically congruent words, the so-called N400 component (Kutas and Hillyard, 1980, 1983). Disagreement in comparison to agreement in biological and syntactic gender reveals either an N400/P600 complex (Schmitt et al., 2002) or a large P600 effect (Hammer et al., 2005). Disagreement compared to agreement of pure syntactic gender violation resulted in a small P600. Moreover, pronoun integration based on pure syntactic gender violation is not finished right after pronoun position indexed by an N400 on the word following an incongruent pronoun (Hammer et al., 2005). The N400 was interpreted as an ongoing integration process characterized by a searching process for an alternative antecedent outside the sentence context to get the meaning of the pronoun. Taken together, these ERP studies refined Garrod and Sanford's two-stage model of pronoun resolution (Garrod and Sanford, 1994, Garrod and Terras, 2000) as they show bonding and resolution online on a word-by-word basis while the sentence structure and meaning unfold over time.

However, memory processes are not discussed within the approach of Garrod and Sanford (1994) and the study of Hammer et al. (2005). In order to be able to link an incoming pronoun to a potential antecedent, the information about possible antecedents is supposed to be re-activated (Nicol and Swinney, 1989, MacDonald and MacWhinney, 1990). ERP results showed that semantic and syntactic information are used to accomplish this goal. Next to semantic and syntactic aspects, the distance between pronoun and antecedent is a relevant issue. Clark and Sengul (1979) investigated pronoun resolution across one, two or three sentences back. They found that a larger distance between anaphor and antecedent leads to increased comprehension times (Clark and Sengul, 1979). This finding indicates that the integration process for long distant anaphors becomes more demanding with regard to working memory load during online sentence comprehension. An increased demand on working memory in sentence processing can give rise to the left anterior negativity (LAN) (Kluender and Kutas, 1993, Carpenter et al., 1995, Münte et al., 1998b, Streb et al., 1999, Fiebach et al., 2001). There is ERP evidence that working memory is involved in the process of pronoun resolution. For example Streb et al. (1999) investigated the effects of pronoun and proper name anaphors in both parallel (easy) and non-parallel (difficult) syntactic

discourse structures. In a parallel structure pronoun and noun have the same syntactic function, and the establishment of the co-reference is relatively easy (such as in: 'Peter visits Julia in the hospital. There he asked a question to the physician.'). In contrast, in a non-parallel structure pronoun and antecedent have a different syntactic function, which is more difficult to interpret (such as in: 'Peter visits Julia in the hospital. There she asked a question to the physician.'). Streb et al. assumed that the non-parallel conditions involve an increased working memory load due to the higher degree of syntactic processing difficulty. They reported two ERP results: (1) Pronoun anaphors evoked a more pronounced negativity than proper name anaphors between 270 and 420 ms over the frontal scalp electrodes. Another relative negativity occurred between 510 and 600 ms at parietal sites. (2) Anaphors in non-parallel positions compared to parallel positions were more negative at parietal sites. The authors suggested that these two effects are functionally related to discourse resolution including a possible search in working memory for the suitable antecedent. Later, Streb et al. (2004) compared far to near distance conditions and found differential anaphor resolution effects. Among others they used ellipses as in 'Werner gave Lisa a ring of shining gold and Joseph [] Anna a necklace.' (long distance) or 'Werner, a generous banker, gave Lisa a ring and Joseph [] Anna a necklace.' (short distance). Please be aware that these are literal translations from German to English. Ellipses, here omission of the verb 'gave' at [] position, resulted in a fronto-central negative shift with a short latency (approximately 120–200 ms) resembling a LAN which has been related to early syntactic processing. In contrast, pronouns and proper names resulted in a parietal/right occipital negative deflection (360–440 ms) resembling an N400 which is assumed to reflect semantic integration processes.

Here we combine both research lines – semantic and syntactic gender processing and a manipulation of working memory load – to investigate the potential interaction during pronoun processing. Sentences with two types of gender manipulations (i.e. person sentences or thing sentences) were used for further distance manipulation. To manipulate verbal working memory (WM) in pronoun processing the distance between antecedent and pronoun was either short (short distance = SD condition) or long (long distance = LD condition; see material in Table 1). While subjects read eight different types of sentences their event-related potentials (ERP) were recorded. Following the logic of earlier ERP-research, we propose the following hypothesis: According to a strict modular view as proposed by serial models (e.g. Frazier and Fodor, 1978, Ferreira and Clifton, 1986, for a neurobiological account see Friederici, 2002), syntactic and semantic information as well as working memory should work independently of each other. In such a framework, we expect an N400/P600 complex for double violations in the person sentence types, a P600 for syntactic violation in the thing sentence types. Concerning working memory processes we would expect an increased LAN component for long distance compared to short distance conditions for congruent pronouns because of an increased demand on working memory resources due to increased distance. Alternatively, according to a more interactive view (e.g. Marslen-Wilson and Tyler, 1980, Trueswell and Tanenhaus, 1994, for a neurobiological account see Hagoort, 2003), working memory could influence semantic and syntactic

Table 1 – Example materials

Condition	Sentence
Experiment 1	
SP: Person as antecedent short distance	
C. congruent	Der Häuptling$_{MALE/mas}$ ist kriegerisch, weil er$_{MALE/mas}$ gewinnen will.
I. incongruent	Der Häuptling$_{MALE/mas}$ ist kriegerisch, weil sie$_{(FEMALE)/fem}$ gewinnen will.
English glosses	The chief is martial, because he/she win want.
ST: Thing as antecedent short distance	
C. congruent	Der Apfel$_{mas}$ ist süβ, weil er$_{(MALE)/mas}$ reif ist.
I. incongruent	Der Apfel$_{mas}$ ist süβ, weil sie$_{(FEMALE)/fem}$ reif ist.
English glosses	The apple is sweet, because he/she (it) ripe is.
LP: Person as antecedent long distance	
C. congruent	Der Häuptling$_{MALE/mas}$ greift bald an und ist kriegerisch, weil er$_{MALE/mas}$ gewinnen will.
I. incongruent	Der Häuptling$_{MALE/mas}$ greift bald an und ist kriegerisch, weil sie$_{(FEMALE)/fem}$ gewinnen will.
English glosses	The chief attacks soon and is martial, because he/she win want.
LT: Thing as antecedent long distance	
C. congruent	Der Apfel$_{mas}$ ist sehr saftig und ist süβ, weil er$_{(MALE)/mas}$ reif ist.
I. incongruent	Der Apfel$_{mas}$ ist sehr saftig und ist süβ, weil sie$_{(FEMALE)/fem}$ reif ist.
English glosses	The apple is very juicy and is sweet, because he/she ripe is.
Experiment 2: additional material	
LP no gap: Person as antecedent	
C. congruent	Der Häuptling$_{MALE/mas}$ ist von nun an besonders kriegerisch, weil er$_{MALE/mas}$ gewinnen will.
I. incongruent	Der Häuptling$_{MALE/mas}$ ist von nun an besonders kriegerisch, weil sie$_{(FEMALE)/fem}$ gewinnen will.
English glosses	The chief is henceforth especially martial, because he/she win want.

FEMALE/MALE = biological gender, mas = syntactic gender masculine, fem = syntactic gender feminine.

processing. In this case, we would expect a difference of the semantic and syntactic effect size of the N400 and/or P600 in LD compared to SD conditions.

In the second experiment, we focus on person antecedents only. The sentences of the first experiment where of the structure: "The chief attacks soon and [] is martial, because he/she wants to win." The reader might implicitly generate a zero-pronoun ('[he]') which is related to gapping (Hankamer and Sag, 1976). This reactivation would virtually shrink the long distance between antecedent and critical pronoun. We added LD sentences without such a gap. The second experiment aimed at investigating whether a gap vs. no gap construction alters the violation pattern. More specifically, the mismatch between this zero-pronoun (subject gap) and the experimentally manipulated pronoun ('[he]/she') in the subordinate clause for the incongruent LD sentence actually may be treated syntactically (virtual short distance) rather then semantically (long distance integration), resulting in P600 rather then N400 ERP effects.

2. Results

2.1. *Experiment 1*

The grand average ERPs ($N=16$) time locked to the onset of the pronouns are shown in Figs. 1 and 3. The Congruency effect (comparing congruent and incongruent pronouns) is shown in the Fig. 1. ERPs of pronouns referring to the person antecedent are shown in the upper panel and thing antecedents in the lower panel (short distance on the left and long distance on the right). Topographical distributions of the Congruency effects (incongruent minus congruent pronouns) are shown in Fig. 2.

All eight conditions elicited a N1–P2 complex which is typical for visually presented material. In the SD-person condition, incongruent compared to congruent pronouns were characterized by a rather widespread negativity between 200 and 400 ms after pronoun onset. The corresponding spline interpolated map (Fig. 2, upper left panel) shows a centro-parietal distribution. Past 400 ms the waveforms to incongruent and congruent pronouns were highly similar in the SD-person condition. In contrast, the SD-thing condition did not show a difference in the 200–400 ms time window. Rather, from 400 ms after pronoun onset and extending well into the ERP of the next word, the ERP to incongruent pronouns were more positive than that to congruent SD-thing-pronouns. Long distance conditions showed a different ERP pattern. LD person and thing conditions were similar between 200 and 400 ms. Pronouns in the LD person condition revealed a difference in the 400–700 ms range with incongruent pronouns being more positive than congruent pronouns whereas this effect was absent for LD thing sentences.

The Distance effect (comparing short and long distant pronouns) is shown in the Fig. 3. ERPs of pronouns referring to the person antecedent are shown in the upper panel and thing antecedents are shown in the lower panel (congruent pronouns on the left and incongruent pronouns on the right). Topographical distributions of the Distance effects (LD minus SD) are shown in Fig. 4. Congruent pronouns referring back to

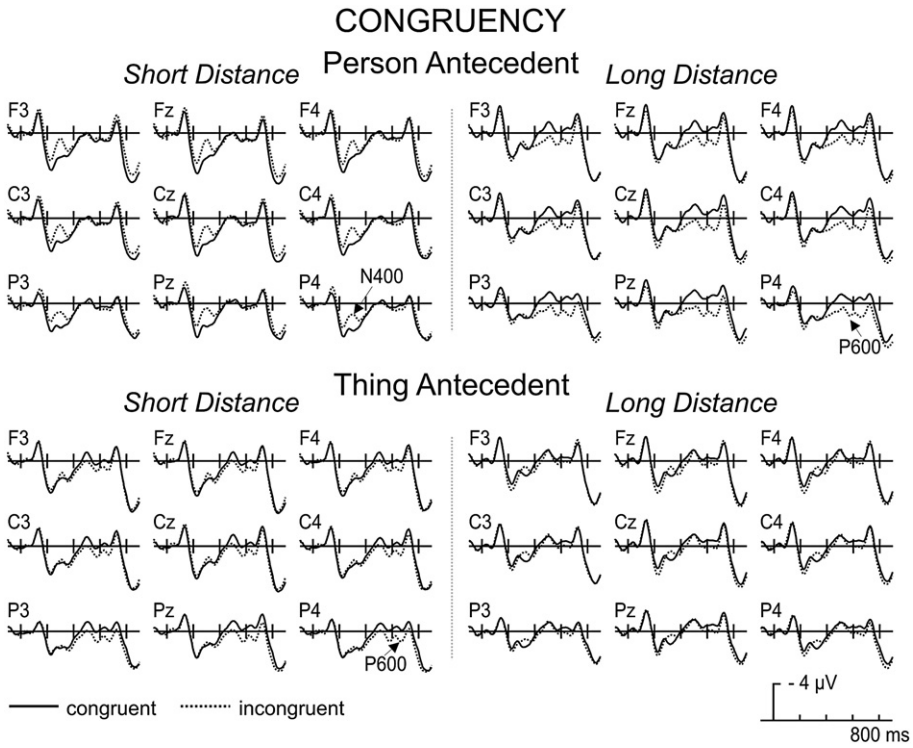

Fig. 1 – Grand average ERPs of Experiment 1 time locked to the onset of pronouns at a selected set of electrodes (negativity is plotted up and each hash mark represents 200 ms of activity in this and in the following figures). The upper panel shows the comparison of congruent and incongruent sentences for persons in SD and LD sentences and the lower panel thing-pronouns in SD and LD sentences.

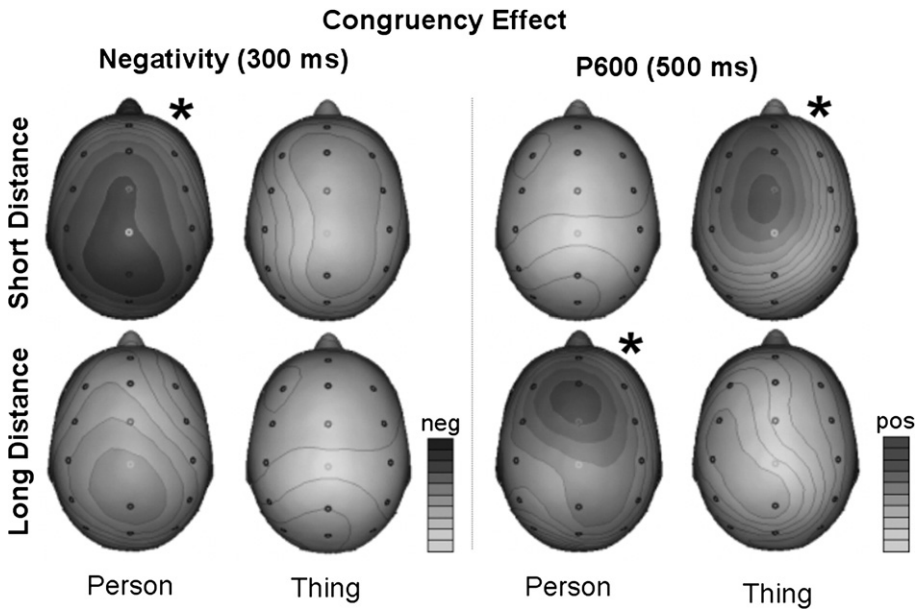

Fig. 2 – Spline interpolated maps based on the mean amplitude of the difference waves of person and thing antecedents (incongruent minus congruent) for the Negativity at 300 ms (left panel) and the P600 at 500 ms (right panel). Significant Congruency effects are marked with *. Contour lines are presented in steps of 0.10 μV.

persons showed an increased negativity for the LD condition compared to the SD condition between 200 and 700 ms. In contrast, incongruent person-pronouns showed a SD–LD difference in the P600 range (400–700 ms) with the LD signal being more positive than the SD signal. Sentences with congruent thing antecedents revealed a negativity for LD pronouns compared to SD pronouns between 200 and 400 ms. For sentences with incongruent thing-pronouns a similar negativity was observed between 200 and 700 ms. Topographical distributions of the Distance effects are shown in Fig. 4.

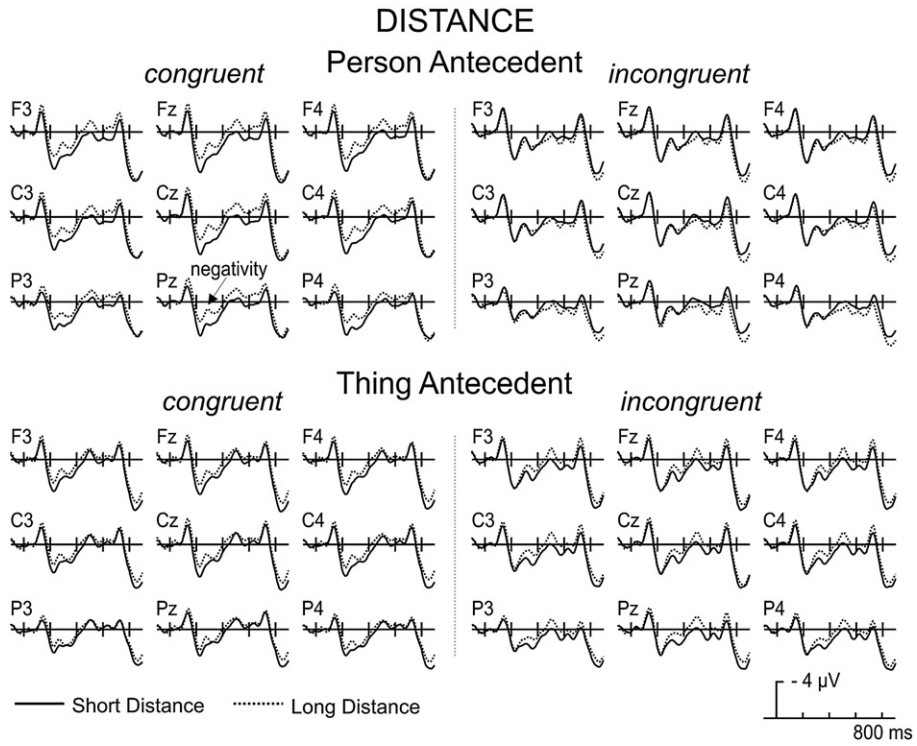

Fig. 3 – Grand average ERPs of Experiment 1 time locked to the onset of pronouns at a selected set of electrodes showing the comparison of both distances. The upper panel shows the person antecedent for congruent (left) and incongruent sentences (right) and the lower panel the thing antecedents for congruent (left) and incongruent sentences (right).

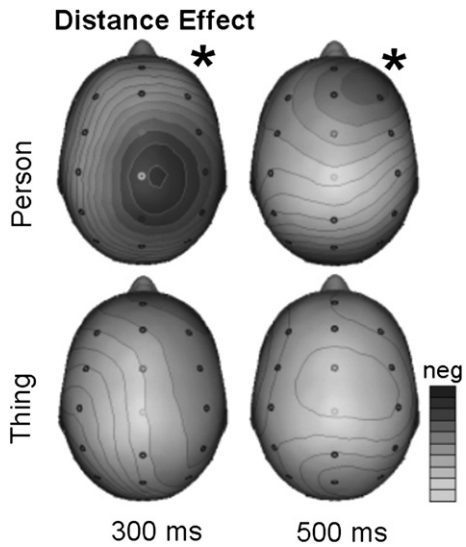

Fig. 4 – Spline interpolated maps based on the mean amplitude of the difference waves of the congruent conditions (long distance minus short distance) for the early (300 ms, left panel) and the late (500 ms, right panel) negativity. Significant Distance effects are marked with *. Contour lines are presented in steps of 0.10 μV.

Mean amplitude measurements were subjected to repeated measurement ANOVAs that crossed the factors Distance (2, LD vs. SD), Antecedent type (2, Person vs. Thing), Congruency (2, incongruent vs. congruent), and Electrode sites.

2.1.1. *Early negativity: 200–400 ms*

The statistics presented in Table 2 revealed that our pronoun manipulation showed a main effect for Distance plus interactions of Distance with Electrode, as well as with Congruency × Electrode and a three-way interaction of Distance × Congruency × Electrode in an omnibus ANOVA. The scalp distributions were subsequently explored in two separate ANOVAs, one with the factors Hemisphere (left vs. right) and lateral electrodes (11 left, 11 right), and one with factors front to back and electrodes (8 anterior electrodes, 8 posterior electrodes) design. These topographical analyses showed that the effects were neither different between hemispheres (Distance effect: $F(1,15) = 5.88$, $p \leq 0.03$; Distance × Hemisphere: $F(1,15) = 3.71$, $p \leq 0.07$), nor between anterior and posterior sites (Distance effect: $F(1,15) = 5.03$, $p \leq 0.04$; Distance × Ant Post: $F(1,15) = 3.68$, $p \leq 0.07$), indicating rather widespread effects (see also Fig. 2, left panel).

To further explore the interaction effects we performed planned pair-wise comparisons (Table 2, *Congruency*) revealing a significant effect for SD-person sentences with incongruent pronouns being more negative than congruent pronouns. This effect was absent in all other comparisons. Planned pair-wise comparisons of *Distance* (SD vs. LD sentences) are reported in Table 2 as well. Here, statistical analysis revealed an effect for sentences with congruent person-pronouns with the LD conditions being more negative than the SD conditions. This effect was absent in all other comparisons.

2.1.2. *P600: 400–700 ms*

The omnibus ANOVA with 29 electrodes (Table 2) revealed a significant main effect for Congruency. Furthermore, we found a three-way interaction Distance × Antecedent × Congruency, as well as an interaction between Distance and Electrodes. This interaction was due to the fact that in the SD condition things but not persons were sensitive to a congruency manipulation (see Fig. 1). This contrasted with the LD condition in which persons but not things were sensitive to congruency.

These effects were traced back by planned pair-wise comparisons to the following results (see Table 2): significant effects for *Congruency* in SD-thing-pronouns, and in LD person-pronouns (incongruent being more positive than congruent pronouns in both cases). With respect to *Distance effects* statistical analysis revealed a main effect for congruent person-pronouns with LD conditions being more negative than SD conditions (Fig. 3, upper left). Thus, this seems to be a continuation of the negative effect reported above. A post-hoc analysis comparing congruent LD with congruent SD person sentences within a single long time window between 200 and 700 ms ($F(1,15) = 8.57$, $p < 0.05$) confirmed this suggestion. As in the early time window, separate analyses for the other three conditions did not reveal significant differences between LD and SD type sentences. To exclude a possible confound of the distance comparison with sentence length, we compared a long period of congruent SD and LD sentences including the last six words of the sentences. We chose this period, because the words are identical for the compared conditions. As can be seen in Fig. 5, an increased negativity for LD sentences compared to SD sentences was only found at pronoun position confirming that the increased negativity found for distance comparisons is based on referential processing (see Table 3 for statistical analysis).

2.1.3. *Negativity at word following the pronoun*

Fig. 6 shows the grand average ERPs, time locked to the onset of the pronoun up to the end of the following word for SD person and thing conditions (panel A) and LD person and thing conditions (panel B). In addition to the ERP-patterns described before, Fig. 6 shows a more negative waveform for the word following the pronoun in incongruent sentences about things compared to congruent sentences in SD conditions. This negativity began approximately at 1000 ms and lasted up to 1300 ms (corresponding to 400–700 ms after target word onset) with an anterior maximum. LD thing conditions revealed a similar pattern with incongruent sentences being more negative than congruent sentences, but with a reduced effect size. However, these patterns were not visible for person sentences in SD and LD conditions.

Statistical analyses (Table 4) revealed that our pronoun manipulation elicited a significant main effect for Congruency and an interaction for Congruency × Electrode and Distance × Electrode in the 1000–1300 ms time window. This time window corresponds to the same time window for the late negative effect as in Hammer et al., 2005. Separate analysis splitting Person and Thing antecedents in SD and LD conditions revealed a main effect for Congruency for SD-thing condition (frontal and central analysis) and LD thing condition (frontal analysis), indicating an N400-like effect.

Table 2 – Experiment 1: Mean ERP amplitude ANOVAs for pronoun position

Source	*df*	Negativity (200–400 ms) *F*-value	P600 (400–600 ms) *F*-value
Omnibus ANOVA (29 electrodes)			
Congruency	1, 15	1.45	**8.07***
Antecedent	1, 15	0.28	2.54
Distance	1, 15	**5.58***	2.31
Distance × Electrode	28, 420	**2.91***	**4.34***
Congruency × Antecedent	1, 15	0.97	0.33
Congruency × Distance	1, 15	3.18	1.16
Congruency × Distance × Electrode	28, 420	**2.42***	**1.64***
Antecedent × Distance	1, 15	0.09	0.89
Congruency × Antecedent × Distance	1, 15	2.95	**6.02***
Planned pair-wise comparisons			
Congruency (29 electrodes)			
SD P: Congruency	1, 15	**6.40***	0.09
Congruency × El	28, 420	**3.96***	
SD T: Congruency	1, 15	0.09	**4.79***
LD P: Congruency	1, 15	0.20	**11.57***
Congruency × El			**3.38***
LD T: Congruency	1, 15	0.02	0.52
Distance (29 electrodes)			
CP: Distance	1, 15	**9.52***	**6.67***
Distance × El	28, 420	**3.25***	**3.71***
IP: Distance	1, 15	0.01	0.93
Distance × El	28, 420	1.46	**3.07***
CT: Distance	1, 15	2.06	0.61
Distance × El	28, 420	1.93°	1.46
IT: Distance	1, 15	2.28	3.03
Distance × El	28, 420	0.56	1.01

Note. El, Electrode; SD, short distance; LD, long distance; C, Congruent; I, Incongruent; P, Person; T, Thing; Congruency (congruent vs. incongruent); Antecedent (person vs. thing); Distance (short vs. long); °$p<0.08$; *$p<0.05$; Greenhouse–Geisser corrected *p*-value for *F* tests with more than one degree of freedom.

Fig. 7 shows the ERP-waveforms ($N=14$) for congruent and incongruent pronouns in SD sentences (upper), gap LD sentences (middle), and no gap LD sentences (lower). SD and no gap LD sentences elicited a negativity for incongruent compared to congruent pronouns between 200 and 400 ms.

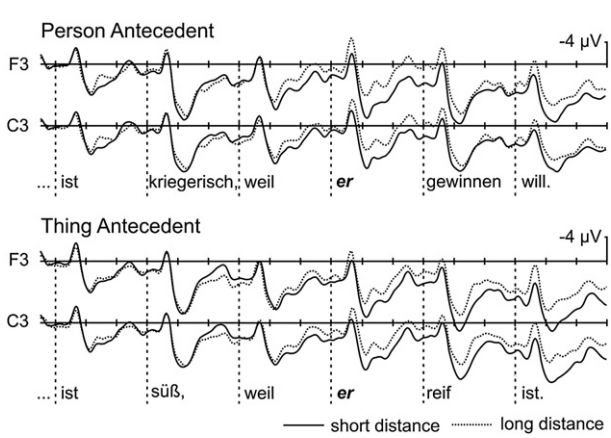

Fig. 5 – Grand average ERPs over the last six words of sentences comparing showing the distance comparison for congruent sentences. The upper panel shows person antecedents and the lower panel the thing antecedents. Vertical lines represent the beginning of a word. Each hash mark represents 200 ms.

This deflection was most prominent at frontal sites for the SD sentence and rather widespread for the LD no gap sentence. As in the first experiment, no prominent negative deflection was found for the gap LD sentence (zero-anaphor). From 400 ms onwards brain responses to all incongruent conditions showed a positive deflection compared to the congruent condition which was rather widespread. Fig. 8 shows the Distance comparison for congruent (upper panel) and incongruent sentences (lower panel). Compared to the SD condition, which has the lowest demands on verbal WM, both LD conditions showed an increased negativity between 200 and 400 ms for congruent and incongruent conditions. In congruent cases this negativity was largest for the gap condition. After 400 ms brain responses to the LD gap conditions accommodated with short distant sentences, whereas the LD no gap sentences appeared to show a positive deflection in frontal-central sites. For the incongruent sentences no pronounced difference was observed in that later time window. Mean amplitude measurements were subjected to repeated measurement ANOVAs that crossed the factors Distance (3, SD vs. LD gap vs. LD no gap), Congruency (2, congruent vs. incongruent), and Electrode sites (29 levels).

2.1.4. *Negativity: 200–400 ms*

To quantify the negativity, mean amplitudes in the 200–400 ms time range were obtained. The omnibus repeated measures ANOVA showed a main effect of Distance ($F(2,26)=4.28$;

Table 3 – Experiment 1: Mean ERP amplitude ANOVAS for last six words

	df	Pronoun −3	Pronoun −2	Conjunction	Pronoun	Pronoun+1	Final word
CP	1, 15	0.35	0.00	3.24	9.02*	0.09	0.27
CT	1, 15	0.10	1.02	2.07	11.60*	6.21*	1.55

Direct comparisons of congruent LD and SD sentences. Time window for each word was 200–600 ms. Tested electrode sites were Fz, Cz, Pz, F3/4, C3/4, P3/4. CP, congruent person; CT, congruent thing; *$p<0.05$.

$p(GG)\leq 00.03$), and a trend for congruency ($F(1,13)=4.21$; $p(GG)\leq 0.06$) and Congruency × Electrode ($F(28,364)=2.48$; $p(GG)\leq 0.06$). Separate analyses of *Congruency* (incongruent vs. congruent) revealed a significant difference for SD and LD no gap sentences on parietal sites but not for frontal and central electrodes (see Table 5). Separate analyses of *Distance* are reported in Table 5 as well. For the congruent conditions, there were significant differences for LD gap compared to SD in all tested regions. For LD gap and LD no gap a tendency was found at frontal and central sites. Incongruent conditions revealed differences between SD and LD gap for frontal, central, and parietal sites. Incongruent no gap differed from incongruent SD in central sites only. Both incongruent LD conditions did not differ significantly.

2.1.5. *P600: 400–700 ms*

For the P600 effect, mean amplitude values were analyzed for the time window between 400 and 700 ms. Although the positive deflection is apparent in the ERP-waveforms neither the omnibus ANOVA nor separate analyses revealed significant main effects (Distance: $F(2,26)=1.44$; $p(GG)>0.05$; Congruency: $F(1,13)=0.82$; $p(GG)>0.05$ or an interaction (Distance × Congruency: $F(2,26)=0.00$; $p(GG)>0.05$).[1]

3. Discussion

3.1. *Experiment 1*

Comparing the ERPs of incongruent and congruent pronouns for different distances allowed us to assess the influence of verbal working memory on the N400 and P600 patterns elicited by gender violations. As expected from an interactive point of view, we observed N400-like and P600 patterns that varied with distance. More specifically, we found an N400-like effect[2] with a centro-parietal maximum exclusively for short distance person conditions – but not for thing conditions – indicating hat verbal working memory interacts with semantic gender information processing. On the first glance, this finding is contrary to Hammer et al. (2005) where we reported a large P600 indexing semantic/biological and syntactic processing. One possible explanation would be that the comprehension system did react sensitive to linguistic and contextual changes (i.e. the change in experimental set up). This point will be addressed in more detail later. Concerning syntactic processing within thing sentences, we found a P600 effect for short thing sentences (incongruent > congruent) — but not for the long thing sentences. This differential P600 pattern for the thing conditions shows that distance manipulation does interact with syntactic parsing. Syntactic gender violation cannot be detected anymore for distant thing-pronouns indicating that both, congruent and incongruent pronouns are equally difficult to parse. In addition, we found a late N400-like effect on the word following the pronoun for incongruent compared to congruent thing sentences for both distances. Thus, discourse related post-pronoun processing in thing sentences is reliable and seems to be independent of working memory. This effect is interpreted as an ongoing integration process in order to resolve the meaning of the pronoun (see also Hammer et al., 2005) characterized by a searching process for an alternative antecedent out of sentence context. In contrast, in the person-case no such next word effect occurred, indicating that the integration process is finished right after pronoun processing.

The direct comparison between short and long distance conditions revealed an enhanced negativity (LD > SD) with an onset-latency of about 200 ms solely for the congruent person-case. This effect did not resemble a classical LAN component as it was rather widespread. At present, we are unable to determine a direct relationship of these negativities and the LAN reported in previous studies. But we would like to interpret the outcome based on design assumptions, i.e. the difference in distance between antecedent and pronoun and by that relate it to working memory load. We suggest, that the observed negativity reflects the more difficult establishment of co-reference for distant compared to close pronouns in sentences. However, this is only true for the person condition. Thing sentences did not reveal differential ERP distance effects indicating no direct influence of working memory on a pure syntactically based process.

3.2. *Experiment 2*

The aim of the second experiment was to test the alternative explanation, namely gap-processing between pronouns and person antecedents. We added a second distance condition without gapping to the person sentence material. As in Experiment 1, we found an N400-like effect for short person

[1] Analysis within a later time-window (500–700 ms) revealed no main effects as well: Distance: $F(2,26) = 1.82$; $p(GG) > 0.05$; Congruency: $F(1,13) = 1.75$; $p(GG) > 0.05$; (Distance Congruency: $F(2,26) = 0.05$; $p(GG) > 0.05$).

[2] Please note that the given effect does not resemble the classical N400 in terms of latency. However, it is known that the amplitude and the latency of the N400 component depend on word-class (Kutas and Hillyard, 1983, Osterhout et al., 1997, Münte et al., 2001), i.e. open class or content words (such as nouns, verbs, adjectives or most adverbs) and closed class or function words (such as pronouns, determiners, conjunctions or prepositions).

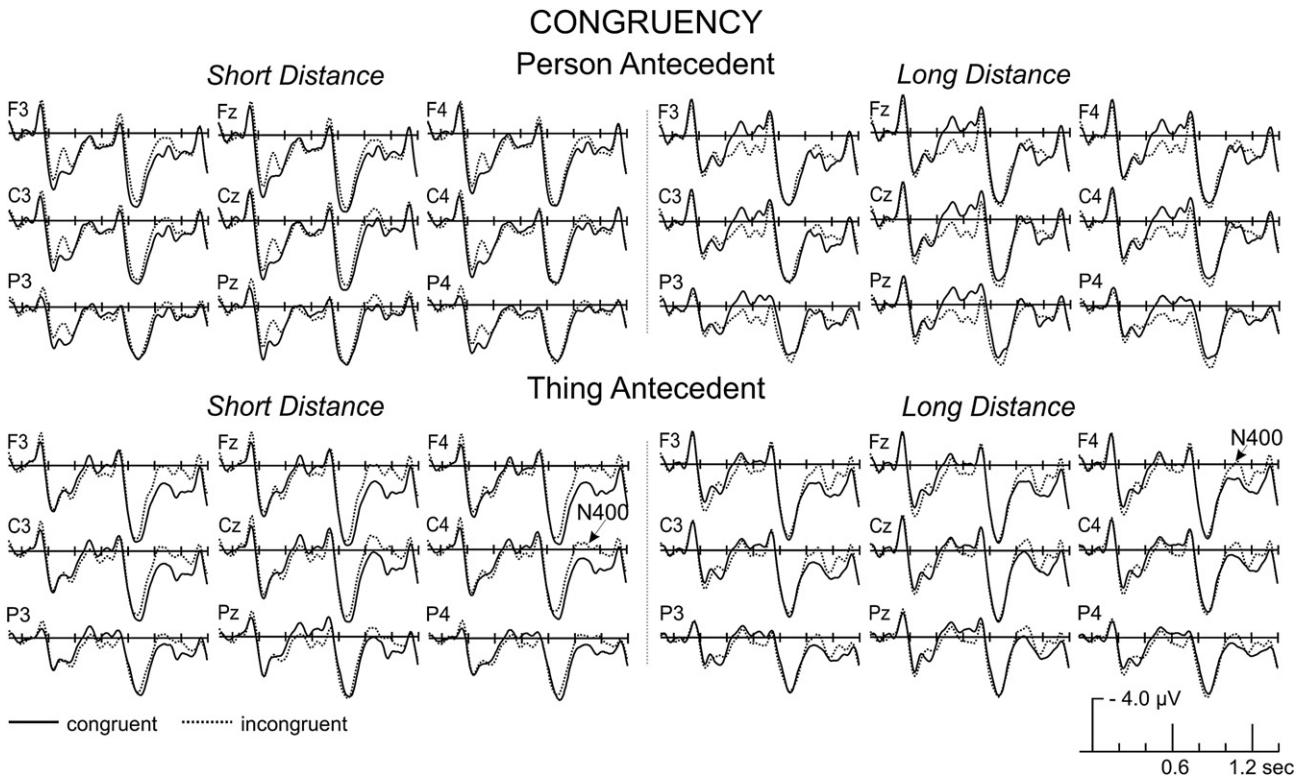

Fig. 6 – Grand average ERPs over two words of Experiment 1 time locked to the onset of pronouns. The upper panel shows the comparison of congruent and incongruent sentences for persons in SD and LD sentences and the lower panel thing-pronouns in SD and LD sentences.

sentences with an increased negativity for incongruent compared to congruent pronouns. And again, as in Experiment 1, this effect was absent in long distant sentences with a gap between antecedent and pronoun. Interestingly and as predicted, the sentence-type without a zero-pronoun revealed an N400-like effect as well. This finding together with the observed lack of any P600 pattern indicates that the direct mismatch between person antecedent and pronoun is rather associated with semantic integration problems.[3] This finding is comparable to that reported by Streb et al. (1999, 2004). They found an N400 but no P600 modulation for differential distances between pronouns referring back to persons suggesting a reflection of rather semantic integration processes. In contrast to these N400 effects, others found a P600-like response for gender anomalies involving a pronoun and a person (e.g. Osterhout and Holcomb, 1992, Osterhout et al., 1997). However, a direct comparison is difficult due to the difference in design, studied language and material. These studies were performed in English which has no syntactic gender marking on the determiners of the antecedents. Directly comparing congruent conditions, a rather widespread negativity for long distant sentences with an intermediate gap relative to short distant conditions with an onset-latency of 200 ms was found. However, interpreting the negativity as belonging to the N400 family, this result may be a reflection of an increased effort of semantic co-establishment with increasing distance. This negativity was not as long lasting as in the first experiment. Distance effects in incongruent conditions could interact with syntactic processing difficulties due to violations in both cases, thus not showing pure working memory effects in violated sentences. Overall, the data (lack of P600 effects, but N400 like effects) suggest that establishing co-reference between person antecedents and pronouns is rather semantically than syntactically driven, at least in the German language.

3.2.1. *Interactions of syntax, semantics and working memory*

A rather simple and plausible explanation for the observed results would be that our brain overall uses semantic integration whenever it has to link a pronoun to an animate antecedent (person), and syntactic integration whenever it has to link a pronoun to an inanimate antecedent (thing). Moreover, the distance between a person antecedent and the pronoun matters such that with increasing distance the semantic integration gets more demanding (as shown in an increase of the N400-like effect between short and long distance conditions). This effect is even more striking if one considers the overall N400 effect size across sentences in general, which usually decreases with more incoming information towards sentence ending, reflecting an increase in

[3] Contrary to the first experiment the comparison of incongruent to congruent LD gap sentences did not reveal a clear P600. Here as well, the comprehension system might react sensitive to contextual changes, i.e. changes of the experimental set up. The second experiment only incorporated sentences with person-antecedents and no thing-antecedents which might alter the general processing. Importantly and as expected, no N400 – and thus no pure semantic reflex – was found for LD gap sentences.

Table 4 – Experiment 1: Mean ERP amplitude ANOVAs for the following word (late N400)

Source: 1000–1300 ms (late N400)	*df*	*F*
Omnibus ANOVA (29 electrodes)		
Congruency	1, 15	**4.89***
Congruency × Electrode	28, 420	**3.58***
Antecedent	1, 15	0.06
Distance	1, 15	0.20
Distance × Electrode	28, 420	**4.49***
Congruency × Antecedent	1, 15	4.08°
Congruency × Distance	1, 15	0.21
Antecedent × Distance	1, 15	0.15
Congruency × Antecedent × Distance	1, 15	0.08
Planned pair-wise comparisons		
Congruency (29 electrodes):		
SD P: Congruency	1, 15	0.28
Congruency × El	28, 420	1.80
SD T: Congruency	1, 15	3.61°
Congruency × El	28, 420	**2.71***
LD P: Congruency	1, 15	0.17
Congruency × El	28, 420	0.88
LD T: Congruency	1, 15	2.68
Congruency × El	28, 420	1.43
Congruency: Frontal (F3, Fz, F4)		
SD T	1, 15	**4.65***
LD T	1, 15	**4.35***
Congruency: Central (C3, Cz, C4)		
SD T	1, 15	**4.77***

El, Electrode; SD, short distance; LD, long distance; C, Congruent; I, Incongruent; P, Person; T, Thing; n.s., not significant; Congruency (congruent vs. incongruent); Antecedent (person vs. thing); Distance (short vs. long); °$p<0.08$; *$p<0.05$; Greenhouse–Geisser corrected *p*-value for *F* tests with more than one degree of freedom. For the topographical pair-wise comparisons only significant results are presented.

efficient semantic integration (Van Petten and Kutas, 1990). Our data also suggest that a purely syntactic integration gets less likely with increasing distance as shown by a decrease of P600 from short to long distance thing sentences. The P600 is related to syntactic integration difficulty (Kaan et al., 2000, Kaan and Swaab, 2003) or the inability to assign the preferred structure to the words encountered (Hagoort et al., 1993, Friederici et al., 1999). In German pronoun processing syntactic gender integration depends on syntactic structure (P600 in simple SD>complex LD; P600 in gap>no gap) and the type of antecedent (person or thing). The P600 pattern in the SD-thing sentences suggests that syntactic processes try to solve syntactic gender violation between two adjacent noun phrases during unification (see below). This syntactic attempt to unify also takes place when an antecedent is salient (person) and when it becomes re-activated via sentence structure (P600 in LD gap person condition). However, changes in the adjacency status of the NPs modulate this process. When the adjacency status for the NPs is not given as in long distance conditions without a reactivating gap the syntactic parser is not trying to solve the conflict but rather semantic integration kicks in (N400-like effect in LD no gap person conditions). When the adjacency status for the NPs is less clear as in long distance condition in which thing antecedents become less re-activated due to a lower saliency status (things compared to persons; Gernsbacher, 1991), the integration process is delayed (N400-like effect in LD gap thing conditions at the following word).

In a series of studies we investigated several types of pronouns in comparable settings and we would like to review the main results here shortly. Pronouns that refer back to things show a P600 effect whenever the syntactic gender assignment is violated (Hammer et al., 2005, current work). Taken the P600 as a syntactic marker proper for a moment (but see Münte et al., 1998a), this result can be interpreted as a syntactically driven establishment of co-reference, or as a syntactically driven check in case the link cannot be established immediately. Interestingly, the word following the gender violated pronoun generates an N400, which can be

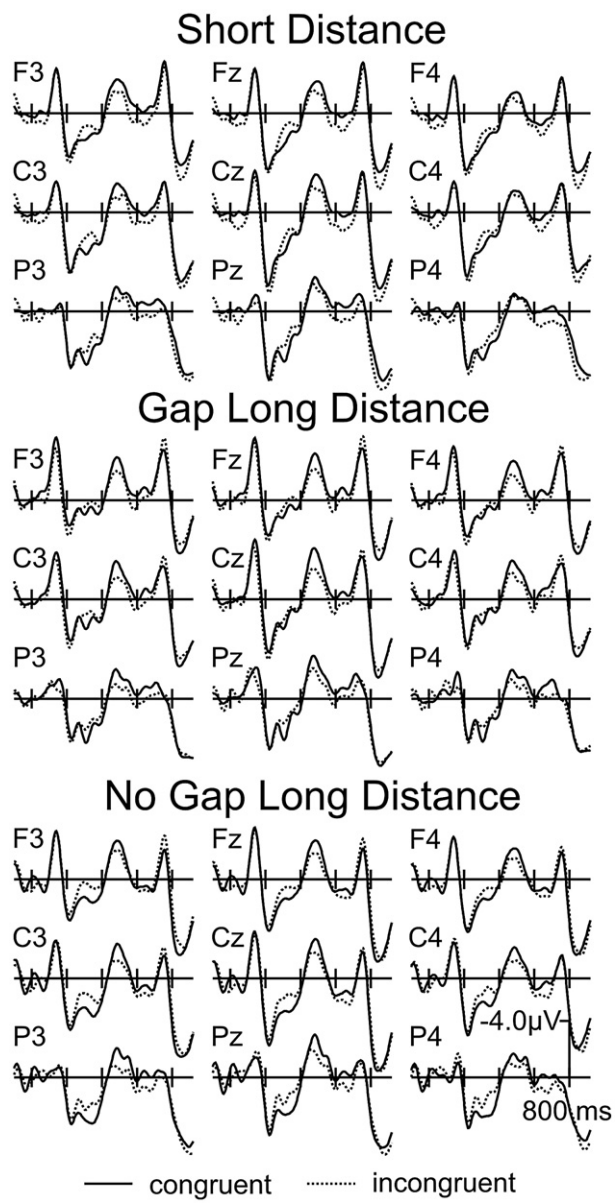

Fig. 7 – Grand average ERPs of Experiment 2 time locked to the onset of the critical pronouns comparing incongruent and congruent conditions for SD conditions (upper panel), LD gap conditions (i.e. intermediate zero-pronoun; middle panel), and LD no gap conditions (i.e. no intermediate zero-pronoun, lower panel).

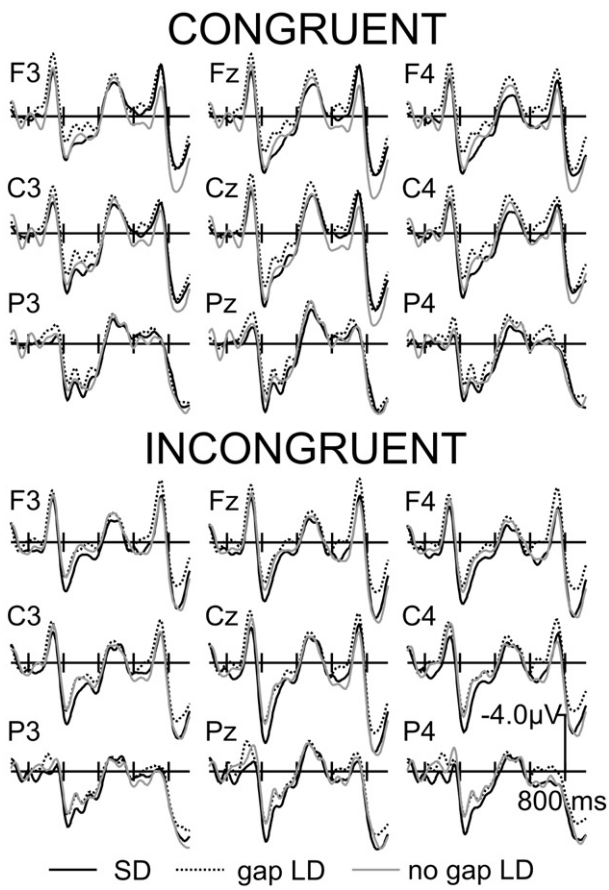

Fig. 8 – Grand average ERPs of Experiment 2 time locked to the onset of the critical pronouns comparing the different distances for congruent conditions (upper panel) and incongruent conditions (lower panel).

read as an ongoing semantic search for an alternative outside the sentence as the violation was not resolved at pronoun position. Pronouns that refer back to persons can be separated in studies that have a rather deterministic gender assignment or a probabilistic one. The latter type are pronouns that refer back to diminutives (das Bübchen$_{\text{NEUTER}}$ [little boy]). Usually a speaker or writer uses a biological gender marked pronoun in 63% of the cases (er $_{\text{MALE}}$), and a neuter gender marked pronoun in 37 % (es $_{\text{NEUTER}}$) to refer back to a diminutive. We observed P600 effects in gender violations, suggesting syntactically driven resolution processes for diminutives (Schmitt et al., 2002). The deterministic category of pronouns includes those that refer back to a non-diminutive person as antecedent, at least in German. These always refer back in terms of biological gender and syntactic gender. A gender violation in this case either elicited a N400/P600 ERP pattern suggesting a semantic–syntactic driven co-referencing (Schmitt et al., 2002) or a large P600 compared to a rather small P600 in single violations (Hammer et al., 2005), again showing a semantically driven manipulation of syntactic integration. Both types of ERP results were interpreted as showing syntactically driven co-establishment that also involves semantic integration. But interestingly, syntactic and semantic processing is signaled by two different ERP signatures across those studies. One explanation may be related to attentional (proportion of violations in the experiment) and linguistic context (proportion of antecedent types) based on the configuration of the design. *Proportions of violations:* It is well known that proportions in experimental trials matter. Different proportions may effect the amount of attention that is driven by the stimuli, especially in mismatch or violation paradigms (for a discussion in the P3/P600 context see Coulson et al., 1998). The proportion of violations differed between the studies of 2002 and 2005. In 2002, the diminutive study, 67 % were violations whereas in 2005 50 % of the sentences were violated. As a result the salience of the violations may be different, and in turn the P600 may have varied, and may have affected any visibility of the N400 (visible in Schmitt et al., 2002, not visible in Hammer et al., 2005). However, this does not change the overall conclusion: semantics (either at pronoun position or at the following word) and syntax play a role in linking a pronoun to an antecedent. *Proportion of antecedent types:* From a linguistic point of view, it has been shown that the linguistic context matters for syntactic and semantic integration. For example, the linguistic system carries out syntactic reanalysis, but only if it is receiving meaningful words, not when it is receiving non-words (Münte et al., 1998a). The brain obviously adjusts expectations and syntactic assembly according to available sentence context. Next to the P600, the N400 is modulated with context too (Bonte et al., 2006 for an auditory demonstration, Mestres-Misse et al., 2006 for conceptual

Table 5 – Experiment 2: Mean ERP amplitude ANOVAs in the 200–400 ms latency range

Source	*df*	*F*
Congruency (congruent vs. incongruent)		
Parietal (P3, Pz, P4)		
SD	1, 13	**4.74***
LD gap	1, 13	1.26
LD no gap	1, 13	**5.68***
Distance		
Frontal (F3, Fz, F4)		
Congruent SD vs. LD gap	1, 13	**8.92***
Congruent SD vs. LD no gap	1, 13	1.49
Congruent LD gap vs. LD no gap	1, 13	**3.92***
Incongruent SD vs. LD gap	1, 13	**10.63***
Incongruent SD vs. LD no gap	1, 13	**7.72***
Incongruent LD gap vs. LD no gap	1, 13	3.10
Central (C3, Cz, C4)		
Congruent SD vs. LD gap	1, 13	**7.42***
Congruent SD vs. LD no gap	1, 13	0.39
Congruent LD gap vs. LD no gap	1, 13	4.56°
Incongruent SD vs. LD gap	1, 13	**12.96***
Incongruent SD vs. LD no gap	1, 13	**6.79***
Incongruent LD gap vs. LD no gap	1, 13	2.52
Parietal (P3, Pz, P4)		
Congruent SD vs. LD gap	1, 13	**6.10***
Congruent SD vs. LD no gap	1, 13	0.50
Congruent LD gap vs. LD no gap	1, 13	3.06
Incongruent SD vs. LD gap	1, 13	**11.74***
Incongruent SD vs. LD no gap	1, 13	2.89
Incongruent LD gap vs. LD no gap	1, 13	2.41

Note. SD, short distance; LD, long distance; n.s. not significant; GG, Greenhouse–Geisser corrected *p*-value for *F* tests with more than one degree of freedom.

integration in reading). Our studies of 2002 and 2005 had different contexts as well. Schmitt et al. (2002) used 100% trials with Persons as referents (as its purpose was to compare pronoun resolution for diminutives vs. non-diminutives). Hammer et al. (2005) used only 50% trials with persons as referent (as it aimed to compare pronoun processing for person vs. thing antecedents). From discourse processing we learned that animacy of the antecedent correlates with saliency (Sridhar, 1988, Gernsbacher, 1991, Kuperberg, 2007). On the first view, this means that in the 2005 study the referent was less salient compared to 2002. On the second view, however, this imbalance probably is counteracted by the fact that the referent types changes in 2005 at random manner (between persons and things), and did not change in 2002. From the attention literature, we know that switching attracts attention (e.g. Wylie and Allport, 2000, Barcelo et al., 2006, Gruber et al., 2006, Philipp et al., 2007). Due to switching between persons and thing sentences, the referent may have been attended more in Hammer et al. (2005). Both views may hold and explain differences in the ERP signal between studies. Why exactly those differences in proportions of violation and type of antecedent are reflected in the ERP signature were not purpose of these studies, but should be looked at in the future.

3.2.2. *Unification in sentences with pronouns*

Neuronal organization of language processing has been addressed by several proposals (Bornkessel and Schlesewsky, 2006, Grodzinsky and Friederici, 2006, Bornkessel and Schlesewsky, in press, for an integrative review see Osterhout et al., in press). The extended Argument dependency Model (eADM) is a hierarchically organized incremental processing model (Bornkessel and Schlesewsky, 2006, in press) including concepts of autonomous thematic processing ("prominence"/ "linking"). The given results are not in contrast with this approach. However, as the results indicate interaction between semantics, syntax and working memory, we would like to embed our findings into an interactive sentence processing model including a direct link to time; for example the elaborated unification model by Vosse and Kempen (2000, Hagoort, 2005). According to this model each word in the linguistic memory (lexicon) is associated with a structural frame that specifies the possible structural environment of this particular word. All elements within the frame unify with each other to form a comprehensive sentence. With regard to ERPs, the P600 is modulated by competitions between alternative unification options (ambiguity), syntactic complexity and semantic influences (Hagoort, 2003). The various results indicating a P600 modulation for semantic alternations (Münte et al., 1998a, Kolk et al., 2003, Kuperberg et al., 2003) support this idea. Our findings fit into the framework as well. The two antecedent types differed in saliency due to animacy (Sridhar, 1988; for a related review on ERP components see Kuperberg, 2007). Less semantically salient antecedents (things) obviously bias pronoun resolution on syntactic grounds (P600). Salient animate antecedent rather elicit semantic unification (N400 or modulation of the P600). The unification model assumes a decay of root node availability over time. We manipulated distance and hence we varied the time between the processing of the antecedent and the pronoun. The gradual decay within a certain range over NPs can functionally be seen as decay of syntactic information in working memory. This has been observed by others, showing that violations in discourse plausibility trigger an N400 effect rather than a P600 effect (van Berkum et al., 1999, 2003), or that violations over longer distances elicited N400 rather than P600 effect (Streb et al., 2004). The present study, too, supports this view with an increase of the N400 effect with increasing distance, and a decrease or lack of P600 in LD sentences without a gap. The unification model also explains the different ERP results between gap (P600) and no gap (N400) conditions. The gap at an intermediate sentence position reactivates the underlying information about the antecedent in terms of root nodes and semantic background (van Berkum et al., 1999, van Berkum et al., 2003). By doing so, this information is now again relatively close to the target pronoun. This reactivation leads to a decrease in syntactic complexity and distance and hence increases the probability to unify on syntactic grounds, reflected in the observed P600 effect in LD gap person conditions. In turn, long distances without a gap suffer from decay of root node structures and rather rely on semantic unification (N400). This idea is also supported when we look into the results of the thing sentence conditions. We found a P600 which was higher for more difficult violation cases compared to congruent cases for thing SD sentences (Fig. 1, lower left panel) as one would expect from unification. A similar effect is visible but non-significant for the thing LD condition. This syntactic effect gets weaker with longer distance, again supporting the unification idea. This pattern suggests a syntactically driven process when a thing is the antecedent and it supports the idea that syntactic unification strength decreases with complexity and distance (time).

3.2.3. *Conclusion*

Our experiments on pronoun resolution reflect dynamic pattern of P600 related to syntactic and N400 effect related to semantic integration processes. To understand the pattern of ERP data, we propose to integrate pronoun processing into the unification model in which online unification of incoming sentence elements are bound together in a flexible way. Here, unification is based on syntactic rules when the antecedent is relatively non-salient (inanimate, no biological gender) and the distance between the antecedent and the pronoun is small. In turn, unification is rather based on semantic grounds when the antecedent is salient (animate), the distance between antecedent and pronoun is rather large (decay of syntactic root nodes increased with increasing load in working memory), and unification has to go outside the sentence (discourse search process at words following the pronoun).

4. Experimental procedures

4.1. *Experiment 1*

4.1.1. *Participants*

Twenty German native speakers gave written consent to take part in the experiment and were paid for their participation. All had normal or corrected-to-normal vision, were right-

handed and neurologically healthy. Four datasets were not included in later analysis because of too many artifacts (the cut off for rejection was more than 30% of trials in one condition) or technical problems, leaving data of sixteen participants (12 women, mean age 22 years).

4.1.2. Material and design

For half of the material we used sentences from an earlier study (Hammer et al., 2005). Two-hundred-forty sentences, 120 about persons and 120 about things, had two clauses each. The first main clause described a state of a person or a thing. The person or the thing was the subject of the main clause, and formed the antecedent of the pronoun following later. Care was taken to guarantee equal word frequencies for the antecedents (persons: 272.98 and things 277.28 using the CELEX (Baayen et al., 1995) frequency count per million, $t(237{,}862) = -0.48$; $p = 0.961$, two-sample t-test). The second clause was a subordinate clause introduced by the conjunction *weil (because)*. This conjunction was followed by the critical word, a pronoun referring to the person or the thing. All sentences were then copied and the congruent pronoun was replaced by an incongruent pronoun. These sentences comprised the SD conditions of the current experiment. To manipulate verbal working memory (WM) during pronoun processing the distance between antecedent and pronoun was increased by adding four words after the antecedent in the same sentences resulting in a total of 480 different sentences (see Table 1). Altogether this resulted in a $2 \times 2 \times 2$ repeated measures design with working memory (short vs. long distance SD/LD), antecedent type (person/thing P/T), and congruency (congruent/incongruent C/I) as factors.[4] To minimize repetition of antecedents, the sentences were distributed across two different lists counterbalancing antecedent type, congruency, and distance leading to 60 sentences per condition for each participant. The sentences of each list were then pseudo-randomized over four blocks taking care that the two repeated antecedents in one list were in different blocks. We measured eight subjects per list which were pooled again for analysis.

To ensure reading and sentence processing, content related questions were presented after 10 ± 5 sentences. Subjects were asked to answer the questions with 'yes' or 'no' by pushing one of two mouse buttons. This task was independent of the pronoun manipulation in order to avoid interference of the explicit task performance with the implicit electrophysiological measures of interest. An example question that required a yes/no response following the sentence 'The apple is sweet because it is ripe' could have been: Yes towards the question 'Was the apple ripe?' or No towards 'Was the apple sour?' Only subjects with at least 80% correct answers in the comprehension task were accepted.

The words were presented in white on a black background in a 16-point font size (Arial) at a viewing distance of about 100 cm (subtending approximately 0.5 by 1.2° of visual angle for critical words).

[4] Repetitive violation designs are different from natural reading. However, violations enable us to reliably evoke the processes of interest in an experimental setting that is a good starting point to study the neural underpinning of pronoun resolution.

4.1.3. Procedure

Subjects were tested in a dimly lit sound-attenuating room while sitting in a comfortable chair. Subjects were asked to read the sentences concentrated and carefully, because they were supposed to answer question concerning the content of the sentences. They should fixate on the screen and avoid possible movements during the word-by-word presentation of the sentences. Between sentences and during the question subjects were free to blink and move their eyes.

A trial started with a fixation cross in the middle of the screen (duration 2050 ms) which was followed by the word-by-word presentation of the sentence. Each word was presented for 350 ms with a 250 ms inter-stimulus interval. Sentence ending words were presented with a dot indicating the end of a sentence, followed by a blank screen for 850 ms.

Questions were presented on the screen for 4 s. One block lasted approximately 20 min. The entire experiment, including instructions, electrode application, and removal took about 2.5 h.

4.1.4. EEG recording and analysis

Electroencephalography (EEG) signals were registered with a digitization rate of 250 Hz and filtered with a band pass of 0.01–50 Hz. Twenty-nine tin electrodes mounted in an elastic cap were positioned according to the 10/20 system (Fz, Fcz, Afz, Cz, Cpz, Pz, Oz, Fp1/2, F3/4, F7/8, Fc3/4, C3/4, T3/4, Cp3/4, Tp7/8, P3/4, T5/6, O1/2). Bio-signals were recorded with a left mastoid reference, and were offline re-referenced to the mean of the activity at the two mastoids. Blinks and vertical eye-movements were recorded with electrodes placed at the sub- and supraorbital ridge of the left eye. Lateral eye-movements were monitored by a bipolar montage using two electrodes placed on the right and left external canthus. Eye movements were recorded in order to allow for later offline rejection. All electrode impedances (EEG and EOG) were kept below 5 kΩ.

From the continuous signal epochs were created of 1024 ms starting 100 ms prior to pronoun onset. These epochs were monitored for artifacts – such as eye-movements – by an automated procedure. Trials with a base-to-peak EOG-amplitude of more than 75 μV, or baseline shifts exceeding 200 μV/s were rejected from further analysis (10 % rejection on average, no difference between conditions, cut off for rejection was more than 30% of trials in one condition). By averaging the remaining artifact-free epochs per condition the event-related potentials (ERPs) were derived. Waveforms were quantified by mean amplitude measures (relative to a 100 ms pre-stimulus baseline) to assess LAN, N400 (time window 200 to 400 ms) and P600 (400 to 700 ms) components. These windows were derived from previous pronoun experiments of our group and corroborated by visual inspection of the grand average ERPs. These measures were subjected to repeated measures analysis of variance that crossed the factors Distance (2, LD vs. SD), Antecedent type (2, Person vs. Thing), Congruency (2, incongruent vs. congruent), and Electrode sites. Additionally, in case of main effects and interaction, we carried out planned pair-wise comparison of the mean amplitudes for Congruency as well as for Distance. Additionally, ERPs were averaged for 2048 ms epochs with a 100-ms pre-stimulus interval in order to analyze the processing of the word following the pronoun.

These measures were subjected to repeated analysis of variance following the same logic as above and an additional analysis for the frontal (F3, Fz, F4), central (C3, Cz, C4) and parietal line (P3, Pz, P4) in a time window from 1000 to 1300 ms (corresponding to a 400–700 ms time window for the word after the pronoun). The Greenhouse–Geisser correction for inhomogeneity of covariance was applied whenever an evaluated effect had more than one degree of freedom in the numerator.

4.2. *Experiment 2*

The second experiment followed the same procedures as the former experiment unless otherwise specified.

4.2.1. *Participants*

Twenty-two German native speakers gave written consent to take part in the experiment. Eight datasets were not included in later analysis remaining in a final population of subjects consisted of 14 native German speakers, age range 18 to 28 years, mean age 22.6 years.

4.2.2. *Material and design*

For the second experiment we used the person sentences only (congruent and incongruent SD and LD; 120 sentences each condition). Additionally, we created a "LD no gap" condition without an '*and*' in the first clause of the sentence. The distance between antecedent and pronoun is the same as in the former LD condition (henceforth LD gap) and in the new LD no gap. Thus, this resulted in a 3×2 repeated measured design with Distance (SD/LD gap/LD no gap), and Congruency (congruent/incongruent) as factors. The sentences were distributed across two different lists leading to 60 sentences per condition per participant, counterbalancing Distance and Congruency. The sentences were pseudo-randomized over five blocks. Seven subjects were measured per list.

4.2.3. *EEG recording and analysis*

Twenty-nine tin electrodes (Fz, Cz, Pz, Fp1/2, F3/4, F7/8, Fc1/2, Fc5/6, C3/4, T3/4, Cp1/2, Cp5/6, P3/4, P7/8, O1/2, Po3/4). Rejection rate of trials including artifacts was 6% on average, no differences between conditions. The mean amplitude measurements were subjected to repeated measures analysis of variance that crossed the factors Distance (3, SD vs. LD gap vs. LD no gap), Congruency (2, congruent vs. incongruent), and Electrode sites (29 levels). Additionally, we compared the mean amplitudes of conditions with separate analysis for Congruency and Distance.

Acknowledgments

We thank Irene Nagel and Marian Scholz for their help in collecting the data. The research reported in this paper was supported by a grant of the Dutch Science Foundation (NWO) to BMJ (she also published under her maiden name Schmitt), and by a grant MU1311/9-1 of the German Science Foundation (DFG) to TFM, as a bilateral co-operation project DFG/NWO. Reprint Requests should be sent to Bernadette Jansma, Maastricht University, Faculty of Psychology, Department of Neurocognition, P. O. Box 616, 6200 MD Maastricht, The Netherlands. Email: B.Jansma@psychology.unimaas.nl.

REFERENCES

Baayen, R.H., Piepenbrock, R., Van Rijen, H., 1995. The Celex lexical database (CD-ROM). University of Pennsylvania, Linguistic Data Consortium, Philadelphia.

Barcelo, F., Escera, C., Corral, M.J., Perianez, J.A., 2006. Task switching and novelty processing activate a common neural network for cognitive control. J. Cogn. Neurosci. 18, 1734–1748.

Bonte, M., Parviainen, T., Hytonen, K., Salmelin, R., 2006. Time course of top-down and bottom-up influences on syllable processing in the auditory cortex. Cereb. Cortex 16, 115–123.

Bornkessel, I., Schlesewsky, M., 2006. The extended argument dependency model: a neurocognitive approach to sentence comprehension across languages. Psychol. Rev. 113, 787–821.

Bornkessel, I., Schlesewsky, M., in press. An alternative perspective on "semantic P600" effects in language comprehension. Brain Res. (Electronic publication ahead of print). doi:10.1016/j.brainresrev.2008.05.003.

Carpenter, P.A., Miyake, A., Just, M.A., 1995. Language comprehension: sentence and discourse processing. Annu. Rev. Psychol. 46, 91–120.

Clark, H.H., Sengul, C.J., 1979. In search of referents for nouns and pronouns. Mem. & Cog. 3, 35–41.

Coulson, S., King, J.W., Kutas, M., 1998. ERPs and domain specificity: beating a straw horse. Lang. Cogn. Processes 13, 653–672.

Ferreira, F., Clifton, C., 1986. The independence of syntactic processing. J. Mem. Lang. 25, 348–368.

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

Frauenfelder, U.H., Tyler, L.K., 1987. The process of spoken word recognition: an introduction. Cognition 25, 1–20.

Frazier, L., Fodor, J.D., 1978. The sausage machine: a new two-stage parsing model. Cognition 6, 291–325.

Friederici, A.D., 2002. Towards a neural basis of auditory sentence processing. Trends Cogn. Sci. 6, 78–84.

Friederici, A.D., Pfeifer, E., Hahne, A., 1993. Event-related brain potentials during natural speech processing: effects of semantic, morphological and syntactic violations. Brain Res. Cogn. Brain. Res. 1, 183–192.

Friederici, A.D., Steinhauer, K., Frisch, S., 1999. Lexical integration: sequential effects of syntactic and semantic information. Mem. Cognit. 27, 438–453.

Garrod, S.C., Sanford, A.J., 1994. Resolvong sentences in a discourse context. In: Gernsbacher, M.A. (Ed.), Handbook of Psycholinguistics. Academic Press, New York, pp. 675–698.

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

Gernsbacher, M.A., 1991. Comprehending conceptual anaphors. Lang. Cogn. Processes 6, 81–105.

Gordon, P.C., Hendrick, R., 1998. The representation and processing of coreference in discourse. Cogn. Sci. 22, 389–424.

Graesser, A.C., Millis, K.K., Zwaan, R.A., 1997. Discourse comprehension. Annu. Rev. Psychol. 48, 163–189.

Grodzinsky, Y., Friederici, A.D., 2006. Neuroimaging of syntax and syntactic processing. Curr. Opin. Neurobiol. 16, 240–246.

Gruber, O., Karch, S., Schlueter, E.K., Falkai, P., Goschke, T., 2006. Neural mechanisms of advance preparation in task switching. NeuroImage 31, 887–895.

Gunter, T.C., Stowe, L.A., Mulder, G., 1997. When syntax meets semantics. Psychophysiology 34, 660–676.

Hagoort, P., 2003. How the brain solves the binding problem for language: a neurocomputational model of syntactic processing. NeuroImage 20 (Suppl 1), S18–29.
Hagoort, P., 2005. On Broca, brain, and binding: a new framework. Trends Cogn. Sci. 9, 416–423.
Hagoort, P., Brown, C.M., Groothusen, J., 1993. The syntactic positive shift (SPS) as an ERP measure of syntactic processing. Lang. Cogn. Proc. 8, 439–483.
Hagoort, P., Brown, C.M., Osterhout, L., 1999. The neurocognition of syntactic processing. In: Brown, C.M., Hagoort, P. (Eds.), The Neurocognition of Language. University Press, Oxford, pp. 273–316.
Hahne, A., Friederici, A.D., 1999. Electrophysiological evidence for two steps in syntactic analysis. Early automatic and late controlled processes. J. Cogn. Neurosci. 11, 194–205.
Hammer, A., Jansma, B.M., Lamers, M., Münte, T.F., 2005. Pronominal reference in sentences about persons or things: an electrophysiological approach. J. Cogn. Neurosci. 17, 227–239.
Hankamer, J., Sag, I., 1976. Deep and surface anaphora. Linguist. Inq. 7, 391–426.
Kaan, E., Harris, A., Gibson, E., Holcomb, P.J., 2000. The P600 as an index of syntactic integration difficulty. Lang. Cogn. Proc. 15, 159–201.
Kaan, E., Swaab, T.Y., 2003. Repair, revision, and complexity in syntactic analysis: an electrophysiological differentiation. J. Cogn Neurosci. 15, 98–110.
Kluender, R., Kutas, M., 1993. Bridging the gap: Evidence from ERPs on the processing of unbounded dependencies. J. Cogn. Neurosci. 5, 196–214.
Kolk, H.H., Chwilla, D.J., van Herten, M., Oor, P.J., 2003. Structure and limited capacity in verbal working memory: a study with event-related potentials. Brain Lang. 85, 1–36.
Kuperberg, G.R., 2007. Neural mechanisms of language comprehension: challenges to syntax. Brain Res. 1146, 23–49.
Kuperberg, G.R., Holcomb, P.J., Sitnikova, T., Greve, D., Dale, A.M., Caplan, D., 2003. Distinct patterns of neural modulation during the processing of conceptual and syntactic anomalies. J. Cogn. Neurosci. 15, 272–293.
Kutas, M., Hillyard, S.A., 1980. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 207, 203–205.
Kutas, M., Hillyard, S.A., 1983. Event-related brain potentials to grammatical errors and semantic anomalies. Mem. Cognit. 11, 539–550.
Lamers, M.J., Jansma, B.M., Hammer, A., Munte, T.F., 2006. Neural correlates of semantic and syntactic processes in the comprehension of case marked pronouns: evidence from German and Dutch. BMC Neurosci. 7, 23.
MacDonald, M.C., MacWhinney, B., 1990. Measuring inhibition and facilitation from pronouns. J. Mem. Lang. 29, 469–492.
Marslen-Wilson, W., Tyler, L.K., 1980. The temporal structure of spoken language understanding. Cognition 8, 1–71.
Mestres-Misse, A., Rodriguez-Fornells, A., Munte, T.F., 2006. Watching the brain during meaning acquisition. Cereb. Cortex.
Münte, T.F., Heinze, H.J., Matzke, M., Wieringa, B.M., Johannes, S., 1998a. Brain potentials and syntactic violations revisited: no evidence for specificity of the syntactic positive shift. Neuropsychologia 36, 217–226.
Münte, T.F., Schiltz, K., Kutas, M., 1998b. When temporal terms belie conceptual order. Nature 395, 71–73.
Münte, T.F., Wieringa, B.M., Weyerts, H., Szentkuti, A., Matzke, M., Johannes, S., 2001. Differences in brain potentials to open and closed class words: class and frequency effects. Neuropsychologia 39, 91–102.
Neville, H., Nicol, J., Barss, A., Forster, K., Garrett, M., 1991. Syntactically based sentences processing classes: Evidence from event-related potentials. J. Cogn. Neurosci. 3.
Nicol, J., Swinney, D., 1989. The role of structure in coreference assignment during sentence comprehension. J. Psycholinguist Res. 18, 5–19.
Osterhout, L., Bersick, M., McKinnon, R., 1997. Brain potentials elicited by words: word length and frequency predict the latency of an early negativity. Biol. Psychol. 46, 143–168.
Osterhout, L., Holcomb, P.J., 1992. Event-related brain potentials elicited by syntactic anomaly. J. Mem. Lang. 34, 785–806.
Osterhout, L., Holcomb, P.J., Swinney, D.A., 1994. Brain potentials elicited by garden-path sentences: evidence of the application of verb information during parsing. J. Exp. Psychol. Learn Mem. Cogn. 20, 786–803.
Osterhout, L., Kim, A., Kuperberg, G., in press The neurobiology of sentence comprehension. In: M. Spivey, M. J., K. McRae (Ed.), The Cambridge Handbook of Psycholinguistics. Cambridge University Press, Cambridge.
Osterhout, L., Mobley, P.J., 1995. Event-related brain potentials elicited by failure to agree. J. Mem. Lang. 34, 785–806.
Philipp, A.M., Jolicoeur, P., Falkenstein, M., Koch, I., 2007. Response selection and response execution in task switching: evidence from a go-signal paradigm. J. Exp. Psychol. Learn Mem. Cogn. 33, 1062–1075.
Rösler, F., Pechmann, T., Streb, J., Hennighausen, E., 1998. Parsing of sentences in a language with varying word order: word-by-word variations of processing demands are revealed by event-related brain potentials. J. Mem. Lang. 38, 15–176.
Schmitt, B.M., Lamers, M., Munte, T.F., 2002. Electrophysiological estimates of biological and syntactic gender violation during pronoun processing. Brain Res. Cogn. Brain Res. 14, 333–346.
Sridhar, 1988. Cognition and Sentence Production: A Cross-linguistic Study. Springer Verlag, New York Berlin.
Streb, J., Hennighausen, E., Rosler, F., 2004. Different anaphoric expressions are investigated by event-related brain potentials. J. Psycholinguist. Res. 33, 175–201.
Streb, J., Rosler, F., Hennighausen, E., 1999. Event-related responses to pronoun and proper name anaphors in parallel and nonparallel discourse structures. Brain Lang. 70, 273–286.
Trueswell, J.C., Tanenhaus, M.K., 1994. Toward a lexicalist framework for constraint-based syntactic ambiguity resolution In: Clifton, C., et al. (Ed.), Perspectives on Sentence Processing. Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 155–179.
van Berkum, J.J., 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., Hagoort, P., Brown, C.M., 1999. Semantic integration in sentences and discourse: evidence from the N400. J. Cogn. Neurosci. 11, 657–671.
Van Petten, C., Kutas, M., 1990. Interactions between sentence context and word frequency in event-related brain potentials. Mem. Cognit. 18, 380–393.
Vosse, T., Kempen, G., 2000. Syntactic structure assembly in human parsing: a computational model based on competitive inhibition and a lexicalist grammar. Cognition 75, 105–143.
Wylie, G., Allport, A., 2000. Task switching and the measurement of “switch costs”. Psychol. Res. 63, 212–233.