Brief communication

Conflicts in language processing: A new perspective on the N400–P600 distinction

Sabine Frenzel, Matthias Schlesewsky, Ina Bornkessel-Schlesewsky

A R T I C L E   I N F O

Article history:
Received 20 November 2009
Received in revised form
13 November 2010
Accepted 2 December 2010
Available online 9 December 2010

Keywords:
Language comprehension
Event-related potentials
Conflict
Animacy
N400
P600
Actor
Well-formedness
Categorization

A B S T R A C T

Conflicts in language processing often correlate with late positive event-related brain potentials (ERPs), particularly when they are induced by inconsistencies between different information types (e.g. syntactic and thematic/plausibility information). However, under certain circumstances, similar sentence-level interpretation conflicts (inanimate subjects) engender negativity effects (N400s) instead. The present ERP study was designed to shed light on this inconsistency. In previous studies showing monophasic positivities (P600s), the conflict was irresolvable and induced via a verb, whereas N400s were elicited by resolvable, argument-induced conflicts. Here, we therefore examined irresolvable argument-induced conflicts (pronoun case violations) in simple English sentences. Conflict strength was manipulated via the animacy of the first argument and the agreement status of the verb. Processing conflicts engendered a biphasic N400-late positivity pattern, with only the N400 sensitive to conflict strength (animacy). These results suggest that argument-induced conflicts engender N400 effects, (which we interpret in terms of increased competition for the Actor role) whereas irresolvable conflicts elicit late positivities (which we interpret as reflecting well-formedness categorisation).

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Electrophysiological measures are well suited to shedding light on the neural mechanisms of language processing, since they allow us to dissociate qualitatively different processing mechanisms while tracking language comprehension/production in real time (e.g. Kutas, Van Petten, & Kluender, 2006). A common strategy for the isolation of these mechanisms is to compare stimuli which induce a conflict in a certain domain to minimally differing, "conflict free" control conditions. A conflict, in this sense, refers to some dimension of the incoming linguistic stimulus that is unexpected based on the previous input; it could thus either result from an outright violation or from a possible, but less expected continuation.

Traditionally, lexical-semantic conflicts were associated with modulations of a centro-parietal negativity occurring approximately 400 ms post stimulus onset (the “N400”, cf. Kutas & Federmeier, 2000), while syntactic conflicts were thought to engender later parietal positivity effects (the “P600”; Osterhout & Holcomb, 1992; Hagoort, Brown, & Groothuisen, 1993). Recently, however, this dichotomy has been challenged by the observation of N400 effects in response to conflicts that would typically be considered syntactic (e.g. Bornkessel, McElree, Schlesewsky, & Friederici, 2004; Choudhary, Schlesewsky, Roehm, & Bornkessel-Schlesewsky, 2005; Haupt, Schlesewsky, Roehm, Friederici, & Bornkessel-Schlesewsky, 2008) and P600 effects in response to conflicts that would typically be considered semantic (e.g. Hoeks, Stowe, & Doedens, 2004; Kim & Osterhout, 2005; Kolk, Chwilla, van Herten, & Oor, 2003; Kuperberg, Sitnikova, Caplan, & Holcomb, 2003). Consequently, the functional interpretation of the P600 in particular has been subject to a substantial amount of debate (for a review, see Bornkessel-Schlesewsky & Schlesewsky, 2008). The fact that the P600 does not appear to reflect a particular linguistic domain has led to proposals that it correlates with domain-general mechanisms. For example, Kolk and colleagues have suggested that P600s reflect “conflict monitoring” during language comprehension, i.e. a check for a possible processing error which is initiated when the processing system encounters conflicting information (e.g. Kolk et al., 2003; van Herten, Chwilla, & Kolk, 2006; Vissers, Chwilla, & Kolk, 2006).1 However, if the

1 In a somewhat similar vein, it has been proposed that the P600 is a domain-general component belonging to the P300-family, since it is sensitive to event-probability manipulations (e.g. Coulson, King, & Kutas, 1998; Gunter, Stowe, & Mulder, 1997).
P600 is indeed a general marker of conflict processing during language comprehension, a late positivity should be expected to occur for every linguistic conflict. Yet, this does not appear to be the case. Rather, as noted above, non-lexical, sentence-level processing conflicts sometimes engender N400 effects (e.g., for word order reanalysis in German, Bornkessel et al., 2004; Haupt et al., 2008; or case/aspect violations in Hindi, Choudhary et al., 2009). Findings such as these thus suggest that sentence-level processing conflicts can correlate with both N400 and P600 effects. Clearly, if we are to gain a fuller understanding of the functional significance of these components, we will need to clarify under which circumstances one or the other is observable.

The aim of the present study was to contribute to answering this question by clarifying the role of two factors which have been implicated in the generation of conflict-induced P600 and N400 effects: conflict resolvability and the nature of the conflict-inducing element (argument vs. verb). In the following, we first motivate the choice of these two factors, before describing the present study.

The findings that led to the conflict monitoring interpretation of P600 effects are based on the processing of so-called “semantic reversal anomalies,” i.e., implausible sentences (e.g., the hearty meals were devouring… Kim & Osterhout, 2005) in which the interpretation called for by the syntax (hearty meals as subject) is contradicted by the most plausible interpretation (hearty meals as object). These effects have hitherto been observed in irresolvable, implausibilities induced at the position of a verb in English and Dutch (e.g., Hoeks et al., 2004; Kolk et al., 2003; Kuperberg et al., 2003; Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007; van Herten et al., 2006; van Herten, Kolk, & Chwilla, 2005) and there is a relatively general consensus that they reflect a conflict between different information types (e.g., between syntactic and thematic or plausibility-based representations).

However, cross-linguistic results indicate that conflicts between syntax and interpretation may also engender N400s rather than P600s. For example, Frisch and Schlesewsky (2001; see also Roehm, Schlesewsky, Bornkessel, Frisch, & Haider, 2004) observed an N400 for inanimate vs. animate subjects in verb-final object-before-subject constructions in German (1), i.e., when the subject analysis for NP2 was contradicted by animacy.2 The cross-linguistic generalizability of this result is supported by similar findings in Mandarin Chinese (Philipp, Bornkessel-Schlesewsky, Bisang, & Schlesewsky, 2008) and English (Weckerly & Kutas, 1999).

Table 1

<table>
<thead>
<tr>
<th>Condition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>–CON, +ANIM, –AGR</td>
<td>The pilot took her to paradise.</td>
</tr>
<tr>
<td>–CON, +ANIM, +AGR</td>
<td>The pilot has taken her to paradise.</td>
</tr>
<tr>
<td>+CON, –ANIM, –AGR</td>
<td>*The pilot took she to paradise.</td>
</tr>
<tr>
<td>+CON, –ANIM, +AGR</td>
<td>*The pilot has taken she to paradise.</td>
</tr>
<tr>
<td>–CON, –ANIM, –AGR</td>
<td>The plane took her to paradise.</td>
</tr>
<tr>
<td>–CON, –ANIM, +AGR</td>
<td>The plane has taken her to paradise.</td>
</tr>
<tr>
<td>+CON, –ANIM, –AGR</td>
<td>*The plane took she to paradise.</td>
</tr>
<tr>
<td>+CON, –ANIM, +AGR</td>
<td>*The plane has taken she to paradise.</td>
</tr>
</tbody>
</table>

CON=conflict (+CON: conflict; –CON: no conflict); ANIM=animacy (+ANIM: NP1 animate; –ANIM: NP1 inanimate); AGR=agreement (+AGR: overt subject–verb agreement with NP1; –AGR: no overt subject–verb agreement with NP1).

CON=conflict (+CON: conflict; –CON: no conflict); ANIM=animacy (+ANIM: NP1 animate; –ANIM: NP1 inanimate); AGR=agreement (+AGR: overt subject–verb agreement with NP1; –AGR: no overt subject–verb agreement with NP1).

2 Note that the N400 for the twig in (1) does not appear to be lexically-driven, since a comparison of the identical animate and inanimate subjects in the sentence-initial position revealed no comparable effect (Ott, 2004). Rather, we have argued previously that it reflects an expectation mismatch with regard to prototypical properties of event participants (Schlesewsky & Bornkessel, 2004): an unambiguously case marked initial object is interpreted as an affected participant (undergoer) and leads to the prediction of a prototypical, and thereby animate event initiator (actor). When a subsequent inanimate argument is encountered (e.g., the twig), this expectation is not borne out and an N400 is evoked.

3 Note that the study by Coulson et al. (1998) was not suited to examining our hypotheses because the critical violation was averaged together with another pronoun case violation (e.g., *Ray fell down and skinned he knee,) and a verb-agreement violation (e.g., *Every Monday he mow the lawn/they suns themselves on the beach.).
violations in German (sentences with two nominative-marked arguments) engender N400 effects when both arguments are animate but not when one argument is animate and the other is inanimate. They interpreted this result as showing that, when two arguments bear identical case marking, the processing system cannot decide which argument is the Actor. By contrast, when the two arguments differ in animacy, competition for the Actor role is reduced because only animate arguments are prototypical Actors (e.g. Silverstein, 1976; Comrie, 1989; Croft, 2003). The interaction between case and animacy in the N400 effect thus provides strong converging support for the interpretive nature of case information. In the present study, the +CON conditions thus induced a conflict between the Actor choice implied by the case marking and the Actor choice implied by the syntax (word order).

In order to examine the relationship between the effects observed and conflict processing in more detail, we included an additional manipulation of conflict strength. To this end, the animacy of the first argument and the agreement status of the predicate were manipulated to either weaken or strengthen the Actor interpretation of the first argument. Animate Actors are more prototypical than inanimate Actors (see above); they also occur more frequently than their inanimate counterparts (for corpus counts on English, see Jäger, 2007). Furthermore, Actors are realized as grammatical subjects in transitive active sentences in English and thus agree with the verb in number and person (though, due to the language's relatively impoverished morphological system, not all number-person combinations are distinguished). Thus, we assumed that the conflict for Actorhood that is induced by a post-verbal nominative pronoun should be more pronounced for sentences with an animate initial argument (pilot) or overt subject–verb agreement (has taken; here, no agreement with the post-verbal argument is possible) in comparison to sentences with an inanimate initial argument (plane) or non-overt agreement (took; here, the verb could agree with either the first or the second argument). Should the conflict conditions engender multiple ERP effects (e.g. a biphasic N400 – late positivity pattern), we should expect the component that directly reflects the processing of the conflicting information to be modulated by the degree of conflict strength.

3. Methods

3.1. Participants

Thirty-one native English speakers participated in the experiment after giving informed consent (13 women; mean age: 25 years, range: 21–33). All were right-handed with normal or corrected-to-normal vision. Seven participants were excluded from the final data analysis due to EEG-artifacts or a multilingual background.

3.2. Materials

320 stimulus sentences were constructed using 80 noun phrase (NP; e.g. the pilot) and verb combinations (e.g. has taken/took). The verb was followed by a subject pronoun he/she/ve/they in the conflict condition (+CON) or by an object pronoun him/her/us/them in the control condition (~CON). All sentences ended with a prepositional phrase (e.g. to paradise) to avoid wrap-up effects. The sentence materials were subdivided into two lists of 160 sentences each, which were presented in a pseudo-randomized order. List presentation was counterbalanced across participants.

3.3. Procedure

Sentences were presented in a word-by-word manner in the center of a computer screen. Each trial began with the presentation of an asterisk (500 ms) followed by an inter-stimulus interval (ISI) of 300 ms. Single words were presented for 200 ms (ISI: 300 ms). Each sentence was followed by 500 ms of blank screen, after which participants performed two tasks. First, a question mark cued them to judge the acceptability of the preceding sentence (maximal reaction time: 2000 ms). Following a further 500 ms of blank screen, a probe word appeared, for which participants decided whether it had occurred in the preceding sentence (maximal reaction time: 2000 ms). Both tasks were performed by pressing one of two hand-held push-buttons, with the assignment of “acceptable”/“yes” and “unacceptable”/“no” answers to the left and right buttons counterbalanced across participants.

3.4. EEG recording and preprocessing

The EEG was recorded from 24 Ag/AgCl-electrodes fixed at the scalp via an elastic cap (Electro-Cap International; Eaton, OH). Electrodes were arranged according to the International 10–10 system (average impedances < 3.5 kΩ). The electrooculogram (EOG) was monitored by means of electrodes at the outer canthi of participants’ eyes and above and below the right eye. EEG and EOG signals were recorded using a Twente Medical Systems DC amplifier (sampling rate: 250 Hz; 50 mV/V; DC – 67.5 Hz) and referenced to the left mastoid (re-referenced to linked mastoids offline). The ground electrode was positioned above the sternum.

3.5. Data analysis

For the ERP data, averages at the pronoun position (baseline: −100 to 0 ms) were calculated per condition, participant and electrode, before grand averages were computed over all participants. The data were analyzed statistically using repeated measures ANOVAs with the condition factors conflict (~CON, i.e. case violation vs. grammatical control), agreement (~ARGT, i.e. overt vs. non-overt agreement with NP1) and animacy (~ANIM; i.e. animate vs. inanimate NP1) and the topographical factor Region of Interest (ROI). Lateral regions of interest were defined as follows: left-anterior (F7 F3 F1), right-anterior (F8 F4 F2), left-posterior (FC5 CP5 C1), right-posterior (FC6 CP6 C2), left-posterior (P7 P3 O1) and right-posterior (P8 P4 O2). Midline analyses are not reported here for the sake of brevity, but confirmed the results for the lateral ROIs. Trials containing eye movements or other artifacts were excluded from further analysis, as were trials for which the probe detection task had not been performed correctly. Huynh and Feldt (1970) corrected p-values are reported for all effects involving more than one degree of freedom in the numerator.

4. Results

4.1. Behavioral results

The accuracy for both behavioral tasks was very high (means: 96% for the acceptability judgment; 98% for the probe-identification task), thus attesting to the fact that participants processed the sentences accurately. We refrained from analyzing reaction times as the tasks were not directly time-locked to the critical sentence region.

4.2. ERP results

Visual inspection of Fig. 1 suggests that both violation conditions (~ANIM/~ANIM + CON) elicited a widely distributed late positivity (500–1000 ms). For the animate violation (~ANIM + CON) condition, the positivity was preceded by an N400 effect (350–450 ms, cf. the enlarged insets in Fig. 1). Statistical analyses were conducted for both time windows.

4.2.1. Early time window (350–450 ms)

A repeated measures ANOVA for the early time window revealed an interaction of ROI and CON (~F(5,115) = 0.01, p < 0.04); however, no individual ROI showed a main effect of CON. In accordance with our hypothesis that interpretation conflicts may yield N400s that are modulated by conflict strength, we conducted an additional hypothesis-driven analysis of the posterior ROIs (i.e. the regions in which N400 amplitude is typically largest) in order to examine possible influences of animacy and agreement. This analysis revealed an interaction of ~ANIM and CON for both posterior regions (left: F(1,23) = 4.56, p < 0.05; right: F(1,23) = 5.94, p = 0.03); A further resolution by animacy yielded an effect of CON only in the ~ANIM condition in the left-posterior ROI (~CON: F(1,23) = 5.60, p < 0.03).

4.2.2. Late time window (500–1000 ms)

For the late time window, a repeated-measures ANOVA revealed a main effect of CON (~F(1,23) = 51.89, p < 0.001) and an interaction of ROI and CON (~F(5,115) = 18.51, p < 0.001). Resolving this interaction by ROI yielded a main effect of CON in all ROIs (all ps < 0.001; minimal F(1,23) = 17.43 in the left-anterior ROI; maximal F(1,23) = 75.17
5. Discussion

The present ERP study, which examined irresolvable argument-induced interpretation conflicts in English, revealed a biphasic N400-late positivity pattern for the conflict-inducing conditions. The N400 was only observable when NP1 was animate, while the positivity did not differ between the conflict conditions.

As illustrated in Table 2, these results serve to fill a gap in the ERP literature on interpretation conflicts in English. The overall data pattern shown in the table suggests that ERP responses reflect both the type of conflict-inducing constituent (argument vs. verb) and conflict resolvability: argument-induced, but not verb-induced, conflicts consistently engender N400 effects, while irresolvable, but not resolvable, conflicts elicit late positivities (P600 effects). The biphasic pattern observed in the present study thus results from the fact that the conflicts examined here were both argument-induced and irresolvable. Strikingly, it was the N400 rather than the P600 that was modulated by conflict strength, since an N400 effect was only apparent when the initial noun phrase was animate. This finding, which resembles Frisch and Schlesewsky’s (2001) results on case marking and animacy in German, strongly suggests that the interpretation conflict itself is reflected in the N400. The late positivity, by contrast, appears to constitute a categorical response to sentence ill-formedness. In the following, we provide a more detailed discussion of both of these effects.
5.1. The N400 and competition for the Actor role

The results summarized in Table 2 suggest that N400 effects reliably correlate with argument-induced interpretation conflicts. Recall from the introduction that these findings from English are supported by a range of cross-linguistic results from typologically diverse languages (e.g., German, Chinese, Tamil; for reviews see Bornkessel-Schlesewsky & Schlesewsky, 2008, 2009). These findings provide strong converging support for the assumption that N400 effects in sentence-level interpretation conflicts reflect competition for the Actor role (Bornkessel & Schlesewsky, 2006; Bornkessel-Schlesewsky & Schlesewsky, 2009; Frisch & Schlesewsky, 2001; Schlesewsky & Bornkessel, 2004). In English, the sentence-initial position is a strong cue for Actorhood (MacWhinney, Bates, & Kliegl, 1984). In the conflict conditions employed here, however, the assumption that the initial argument is the Actor is contradicted by the case marking of the second argument. The competition for the Actor role thus resulting is reflected in an N400 effect. By contrast, when there is an additional animacy difference between the arguments, the degree of competition is reduced because inanimate arguments are less prototypical Actors than their animate counterparts. Hence, N400 amplitude is reduced. Note that this does not mean that the overall conflict induced by the sentence is resolvable: the sentence is still ill-formed, since a nominative pronoun can only occur in the (pre-verbal) subject position in English.

Overall, these results suggest that the correlation between argument-induced interpretation conflicts and the N400 (see Table 2) may be a byproduct of the underlying competition for the Actor role rather than reflecting a true categorial distinction.

5.2. The late positivity and well-formedness categorization

In contrast to the N400, the late positivity (P600) effects in the present study were not modulated by conflict strength. Rather, the issue of conflict resolvability appears to play a key role, i.e., late positivities are only observable for irresolvable sentence-level interpretation conflicts (cf. Kolk et al., 2003; Kuperberg, 2007; Bornkessel-Schlesewsky & Schlesewsky, 2008). We thus assume that, rather than reflecting the conflict itself, the late positivity indexes a categorization process by means of which sentences are classified as ill-formed (Bornkessel & Schlesewsky, 2006; Kretzschmar, 2010; Bornkessel-Schlesewsky et al., 2010). A possible alternative interpretation is provided by the conflict monitoring account of late positivity effects (e.g., Kolk et al., 2003; van Herten et al., 2006; Vissers et al., 2006) with the proviso that the conflict monitoring mechanism would need to be restricted to irresolvable conflicts rather than applying to any type of conflict in language processing. To some degree, this issue has already been raised in recent studies conducted from the perspective of the conflict monitoring account: van de Meerendonk et al. (2010), for example, report that P600 effects are only engendered by “strong” conflicts. However, in contrast to the well-formedness categorization account, the conflict monitoring account does not provide a principled explanation for the N400 effects in the present study and in previous experiments.

Finally, on the basis of previous results, it appears likely that the late positivity effects observed here could be modulated by a change of task. For example, Vissers, Chwilla, and Kolk (2007) showed that the amplitude of semantic P600 effects in Dutch (i.e., a P600 induced by a semantic reversal anomaly) can be modulated by different task instructions. This is in line with the notion of well-formedness categorization, since categorization processes are clearly task dependent. Likewise, a reduction of the effect should be expected in the absence of a judgement task, since stimulus categorization would not be explicitly required.

5.3. Summary and conclusions

To conclude, the present findings support the perspective that there are multiple neural correlates of sentence-level interpretation conflicts. These reflect several dimensions including the nature of the conflict and whether it is resolvable or not. We have proposed that competition for the Actor role correlates with N400 modulations, thus accounting for the fact that argument-induced conflicts tend to engender N400 effects. By contrast, late positivities mark irresolvable sentence-level interpretation conflicts independently of conflict strength. We therefore assume that they reflect categorization processes by means of which a sentence is classified as ill-formed.

Acknowledgements

The research reported here was performed while S.F. and I.B.S. were at the Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, and was conducted in collaboration with the Clinic for Audiology and Phoniatry (Prof. Manfred Gross) of the Charité Berlin. We are grateful to Katja Bruening for invaluable assistance in data acquisition and to Dietmar Roehm and Markus Philipp for help in experiment preparation and data analysis.

References
