



The N400 Effect during Speaker-Switch—Towards a Conversational Approach of Measuring Neural Correlates of Language

Tatiana Goregliad Fjaellingsdal^{1*}, Esther Ruigendijk^{2,3}, Stefan Scherbaum⁴ and Martin G. Bleichner^{1,2}

¹ Department of Psychology, European Medical School, University of Oldenburg, Oldenburg, Germany, ² Cluster of Excellence Hearing4all, University of Oldenburg, Oldenburg, Germany, ³ Department of Dutch, University of Oldenburg, Oldenburg, Germany, ⁴ Department of Psychology, Technische Universität Dresden, Dresden, Germany

OPEN ACCESS

Edited by:

Tamara Swaab,
University of California, Davis, USA

Reviewed by:

Sara Bögels,
MPI for Psycholinguistics, Netherlands
Miguel Oliveira,
Universidade Federal de Alagoas,
Brazil

*Correspondence:

Tatiana Goregliad Fjaellingsdal
tatiana.goregliad@uni-oldenburg.de

Specialty section:

This article was submitted to
Language Sciences,
a section of the journal
Frontiers in Psychology

Received: 13 September 2016

Accepted: 09 November 2016

Published: 28 November 2016

Citation:

Goregliad Fjaellingsdal T, Ruigendijk E, Scherbaum S and Bleichner MG (2016) The N400 Effect during Speaker-Switch—Towards a Conversational Approach of Measuring Neural Correlates of Language. *Front. Psychol.* 7:1854. doi: 10.3389/fpsyg.2016.01854

Language occurs naturally in conversations. However, the study of the neural underpinnings of language has mainly taken place in single individuals using controlled language material. The interactive elements of a conversation (e.g., turn-taking) are often not part of neurolinguistic setups. The prime reason is the difficulty to combine open unrestricted conversations with the requirements of neuroimaging. It is necessary to find a trade-off between the naturalness of a conversation and the restrictions imposed by neuroscientific methods to allow for ecologically more valid studies. Here, we make an attempt to study the effects of a conversational element, namely turn-taking, on linguistic neural correlates, specifically the N400 effect. We focus on the physiological aspect of turn-taking, the speaker-switch, and its effect on the detectability of the N400 effect. The N400 event-related potential reflects expectation violations in a semantic context; the N400 effect describes the difference of the N400 amplitude between semantically expected and unexpected items. Sentences with semantically congruent and incongruent final words were presented in two turn-taking modes: (1) reading aloud first part of the sentence and listening to speaker-switch for the final word, and (2) listening to first part of the sentence and speaker-switch for the final word. A significant N400 effect was found for both turn-taking modes, which was not influenced by the mode itself. However, the mode significantly affected the P200, which was increased for the reading aloud mode compared to the listening mode. Our results show that an N400 effect can be detected during a speaker-switch. Speech articulation (reading aloud) before the analyzed sentence fragment did also not impede the N400 effect detection for the final word. The speaker-switch, however, seems to influence earlier components of the electroencephalogram, related to processing of salient stimuli. We conclude that the N400 can effectively be used to study neural correlates of language in conversational approaches including speaker-switches.

Keywords: mobile EEG, prediction, N400, social interaction, conversation, language, dialogue, turn-taking

INTRODUCTION

The exchange of ideas via language is a central part of human culture. Understanding how these verbal interactions relate to the underlying brain processes is a challenge. This challenge arises from the open nature of conversations on the one hand, and on the restrictions imposed by neuroimaging techniques on the experimental setup on the other hand. Consequentially, to date, the neural underpinnings of natural conversations are poorly understood. To gain a better understanding it is therefore necessary to develop new experimental paradigms and new brain imaging setups.

Neuroscientific studies have taught us a considerable amount about the neurobiological basis of language (Friederici, 2011). Central in this research endeavor is the study of language processing in individuals using controlled language material. Natural language, however, is typically used during social interactions with one or more participants and we could gain more knowledge on language processing itself, if we find ways to study their neural underpinnings during conversations (Levinson, 2016).

Language development and language usage are fundamentally interactive: language is acquired interactively and is used in conversations with other people. These interactions are characterized by a back and forth between the conversational partners; we constantly switch between listening and speaking, we plan our next utterances and we build expectations, and even concrete predictions, about what our conversational partner will say. We give the other person time to speak and we anticipate those moments when it is our turn to speak (Levinson and Torreira, 2015; Torreira et al., 2015).

The large number of processes involved in a dialogue makes it very challenging to study conversations using neuroscientific measures. One challenge is the poor controllability of natural conversations. There is an evident tradeoff between the naturalness of a dialogue and possible experimental manipulations; any experimental manipulation leads almost inevitably to an unnatural dialogue. Another challenge are the restrictions and requirements that brain imaging techniques pose on the experimental setup (Van Berkum, 2012). All imaging techniques are sensitive to movement artifacts (e.g., speech related mouth movements) and therefore require the participant to move as little as possible, which also leads to an unnatural dialogue situation. Further, the small magnitude of the signal of interest generally requires many repetitions to increase the signal-to-noise ratio by data averaging (De Vos and Debener, 2014). What is needed are, first, paradigms that capture the aspects of natural conversations but retain sufficient structure to allow for quantitative analysis (e.g., Magyari et al., 2014; Mandel et al., 2015; Bögels et al., 2015a) and second, recording setups that allow for a more natural interaction, with less restrictions on the participants (e.g., Bašnáková et al., 2015; Bögels et al., 2015c). Our research aim is to move toward neurophysiological studies of natural dialogues. As a first step this is targeted, on the one hand, by incorporating a conversational element in the experimental paradigm and, on the other hand, by using an EEG device that enables a less restrained recording setup.

We present an experimental approach that entails a number of aspects that play a role in a conversation: speaking and listening, the formation of expectations during a conversation (i.e., how a sentence will end), and an element of turn-taking between speakers (Coates, 1990; Purver et al., 2009), that is, a speaker-switch, comparable to a situation in which one person completes the utterance of another person. In this paradigm participants either read out or listen to the pre-recorded first part of a sentence, the last word of the sentence is in both conditions presented by another speaker and either completes the sentence in an expected or an unexpected way, i.e., it matches or violates the semantic expectation of the participant. Finally, we use a wireless EEG setup that allows EEG recordings in less restricted ways, i.e., the amplifier is directly attached to the cap and does therefore not restrain the participant. In the long run, the application of wireless EEG enables to move the experiment outside the strongly controlled lab environment (e.g., De Vos et al., 2014). In this study, further aspects such as sequential full turns and other conversational elements (cf. Levinson and Torreira, 2015) are left aside.

The variable of interest in this paradigm is the N400 event-related potential (ERP) in response to the last word of the sentence. The N400 is a useful tool to study expectation violation during language processing and is characterized by a negative deflection in the averaged electroencephalogram (EEG) peaking around 400 ms after a semantic violation (Kaan, 2007; Kutas and Federmeier, 2011). The amplitude is increased (i.e., more negative) for words that do not match the previously given semantic context and are hence unexpected (e.g., “*The pizza is too hot to cry.*”, Hillyard and Kutas, 1983). A small number of studies have investigated the N400 mimicking differential aspects of conversations, such as, for example, the effect of long or short pauses after questions in conversations (Bögels et al., 2015b). However, whether the N400 effect, the difference between expected and unexpected items, is a useful means for studying prediction building during conversation is not well studied.

Our design allows to study the effect of a semantic manipulation *during* a turn take within a sentence context on the N400, while still providing sufficient control to allow for a quantitative analysis (e.g., predefined sentences and precise timing information of the final word). Speech utterances in one of our conditions (reading aloud) are usually related to large muscle artifacts in the EEG, masking the signal of interest. We are able to circumvent this problem in this paradigm by analyzing only the sentence final word, i.e., the word that is generated by another speaker, providing us with an artifact free data segment.

Central to this paradigm is the semantic violation for the sentence's final word during the switch from one speaker to another (i.e., an element of turn-taking). For this we expected to observe the N400 effect for incongruent sentence endings compared to congruent endings. We expected an N400 effect despite speaker-switch and independent of the participant's previous role (speaking vs. listening) as in both cases one has to process how well the last word fits the formed expectation.

MATERIALS AND METHODS

Sixteen healthy German-native speakers (eight female, mean age 24.1 years) took part in this study. All were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). Participants gave written informed consent prior testing and were paid for participation. The study was approved by the Ethics Committee of the University of Oldenburg and conducted in accordance with the declaration of Helsinki.

EEG was recorded from the participants in two turn-taking modes: (1) “Listening,” the mode of listening to the pre-recorded sentence fragment, which was also visible on the screen, and (2) “Reading aloud,” the mode of reading aloud the sentence fragment presented on the screen (Figure 1). To keep the two conditions as similar as possible, the sentence fragments were presented visually in both modes. The last word of the sentence (from now on critical word or CW) was, in both conditions, presented auditorily by another speaker and either completed the sentence in an expected semantically congruent or an unexpected semantically incongruent way. This resulted in a 2 (Listening vs. Reading aloud) × 2 (Congruent vs. Incongruent) design with

the four conditions Listening congruent, Listening incongruent, Reading aloud congruent, and Reading aloud incongruent.

To control for the participants task involvement, a control word was presented after each sentence. Participants had to indicate whether this word was present or absent in the sentence by either pressing the left or the right screen button, respectively.

Linguistic Material

The linguistic material was adapted from the study of Ruigendijk et al. (2015). The sentences were balanced for word frequency and word length and a plausibility judgment for congruent and incongruent sentences was performed by native speakers. In brief, the material consists of 40 German eight-word sentences with *subject-predicate-direct object-goal/direction* pattern (e.g., “The teacher takes the kids into this school,” from German: “Der Lehrer bringt die Kinder in diese Schule”). For each of these sentences a set of four sentences was generated, where the verb and the goal/direction (i.e., last word) was constant, while the subjects and direct objects were changed but still semantically closely related (e.g., “The father takes the sons into this school,” “The mother takes the daughters into this school,” and “The

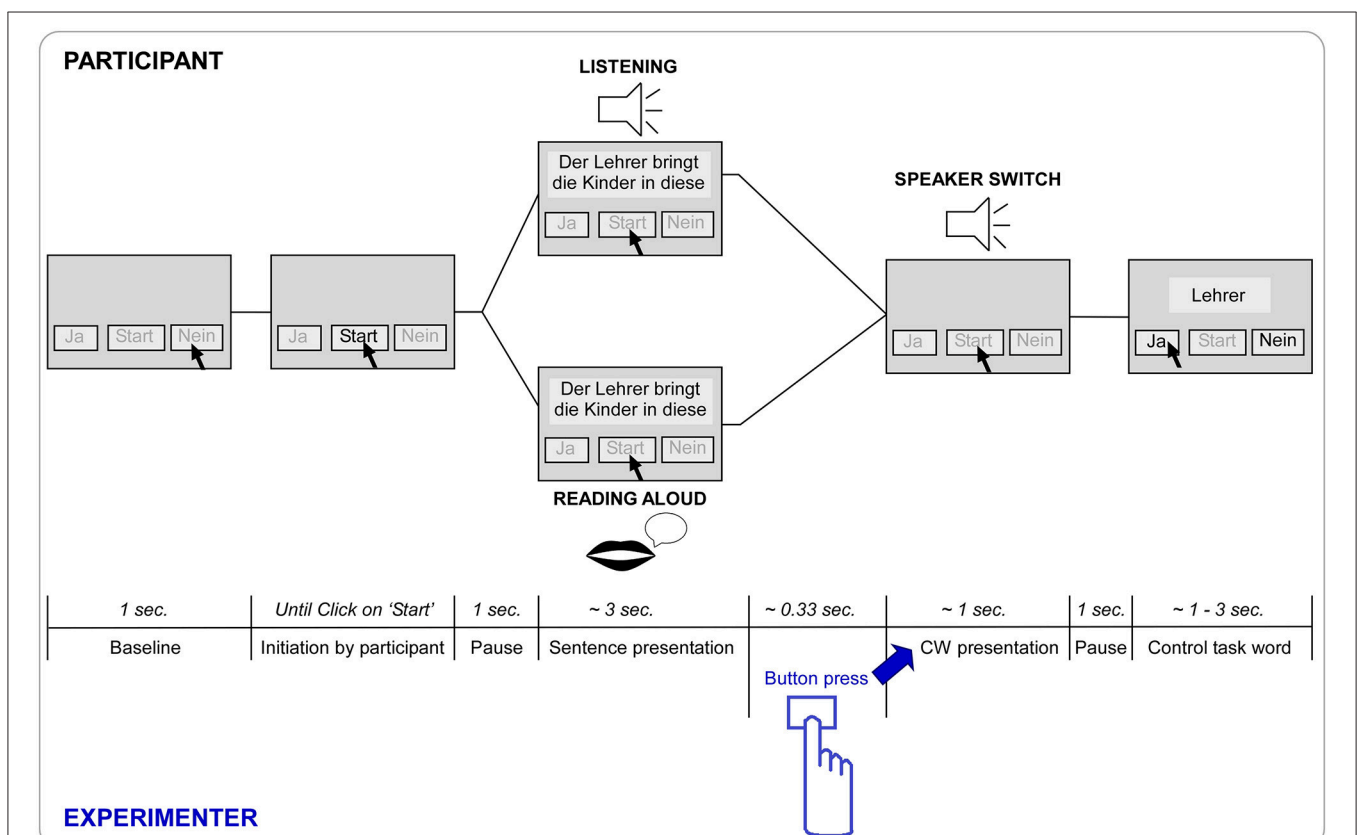


FIGURE 1 | Paradigm trial timeline in seconds and respective screen state. After a baseline of 1 s the “Start” button is enabled and the trial is initiated by a click of the participant. After a pause of 1 s the respective first seven words of the sentence are presented simultaneously on the screen. Depending on block the participant listens to a recording of the first speaker saying the first part of the sentence (Listening) or reads the first part of the sentence out aloud on its own (Reading aloud). The sentence is then completed by the second recorded speaker (speaker switch) saying the final (eighth) word of the sentence (= critical word CW), triggered by the button press of the experimenter (see highlights in blue). After a pause of 1 s the control task word appears on the screen and the “yes” (“Ja”) and “no” (“Nein”) buttons are enabled. As soon as the participant has clicked or maximum after 3 s, the word disappears. After a baseline of 1 s the next trial can be initiated again.

bus driver takes the pupils into this school”). This resulted in 160 semantically congruent sentences. Semantically incongruent sentences were created by replacing the final word with another final word (out of the 40 final words used, e.g., “*The teacher takes the kids into this wound*”). In this study we used in total a set of 80 incongruent sentences, providing us 40 incongruent sentences for the listening condition and 40 incongruent sentences for the reading aloud condition. We created two lists, each containing 160 sentences in total: 120 semantically congruent (with three presentations of the same CW) and 40 semantically incongruent sentences (i.e., the incongruent sentences were non-repetitive for each participant, whereas there was some repetition of congruent sentences). The order of the sentences was pseudo-randomized in five different versions for each list, to avoid order effects but to control for sequential presentation of semantically incongruent endings (≤ 5). Eight further unique sentences and unique final words were used for a practice block only.

The auditory material used for the sentences in the listening condition and the critical words for both conditions were recorded by two female speakers. The complete sentences were all recorded by one speaker (instructed to read in natural sentence intonation) and later cut, resulting in an auditory file with the first seven words of the sentence (mean sentence fragment length: 2.69 s, range 1.88–3.98 s). The final (eighth) word closing the sentence was recorded by the second speaker (mean word length: 0.71 s, range 0.46–0.98 s). To improve natural sentence intonation for the final words, an additional person was present during the recordings reading out the first part of the sentence. All auditory stimuli (44.1 kHz, 16 bit, WAV format) were edited with Praat (Boersma, 2001). A cosine ramp with a duration of 5 ms was implemented at the end of the seven-word sentences to remove any clicks after cutting and to smoothen the transition to the final word (see Auditory Material in the Supplementary Material for examples). All stimuli—sentences and final words—were adjusted to the same volume. Voice onset times for the final words were defined within millisecond range and set as stimulus onset time for analysis.

EEG Recording

Brain electrical activity was recorded with a wireless EEG system (<http://www.mbraintrain.com>). The EEG is recorded from 24 sintered Ag/AgCl electrodes (international 10/20: Fp1, Fp2, F7, Fz, F8, FC1, FC2, C3, Cz, C4, T7, T8, TP9, TP10, CP5, CP1, CPz, CP2, CP6, P3, Pz, P4, O1, and O2, reference: FCz, ground: AFz) with a small wireless amplifier (Smarting mBrainTrain, Belgrade, Serbia) attached to the back of the cap (EasyCap, Herrsching, Germany). Recordings were digitized (Smarting Software 2.0.0, Smarting mBrainTrain, Belgrade, Serbia) with a sampling rate of 500 Hz and sent to a computer via Bluetooth. Electrode impedances were kept below 10 k Ω .

Procedure

The EEG cap was set up and the individual participant was seated in the sound-attenuated testing cabin in front of the computer screen (≈ 60 cm). A short practice block was used to familiarize the participants with the task.

Each participant underwent two testing blocks—Listening and Reading aloud. The order of the testing blocks was randomized over participants. Two different stimuli lists were used for the first and second block (see Section “Linguistic Material”).

Each trial was initiated by the participant by clicking the “Start” button (Figure 1). The first seven words of the sentence were shown on the screen as text, and either presented also auditorily via loudspeakers (Listening) or read aloud by the participant (Reading aloud), which was followed by the auditory presentation of the critical word (CW, and speaker-switch) presented via loudspeakers. The final word was triggered by the experimenter with a button press in both modes (to distribute variance of gaps similarly). After the end of the critical word there was a pause of 1 s and the control word was presented on the screen (Figure 1). The participant was instructed to click “yes”/“no” indicating presence/absence of the control word in the sentence. In 50% of the trials the control word was present in the first part of the sentence and in the other 50% of the trials it was absent from the sentence. The control word was never the critical word. The task was implemented to make sure the participant was paying attention to the stimuli, as indicated by correctly categorized words (Duncan et al., 2009). In case of no answer, the word disappeared after 3 s. Participants were instructed not to swallow or move and to blink as little as possible from the onset of each trial (click on “Start”) until completion of the control word task.

An analysis of the gaps between sentence fragment and critical word was performed to ensure that no systematic differences in gap length were present between the modes. The average gap for Listening was 0.339 s long ($SD = 0.019$, range 0.163–1.048 s) and for Reading aloud 0.322 s long ($SD = 0.020$, range 0.130–1.531 s). No significant differences in gap length between the two turn-taking modes were present [paired t -test, $t_{(15)} = -1.83$, $p = 0.088$].

The experiment was programmed in Matlab R2012a (Mathworks Inc., Natick, MA), using the psychophysics toolbox (Brainard, 1997; Pelli, 1997). A microphone in the testing cabin recorded the trials and enabled the experimenter to listen to the utterances of the participant in the testing cabin. Lab Streaming Layer (<https://github.com/sccn/labstreaminglayer>, Lab Recorder, version 1.05) was used to record and synchronize all data streams (EEG signal, EEG markers, and microphone signal).

EEG Analysis

All trials were incorporated in the analysis as the participants rated on average 99.26% words correct in the control task (≈ 158.81 out of 160 trials; average Listening congruent 99.11%, Listening incongruent 99.38%, Reading aloud congruent 99.53%, and Reading aloud incongruent 98.75%).

EEG analysis was performed with EEGLAB (version 13.4.4b, Delorme and Makeig, 2004).

For artifact attenuation with independent component analysis (ICA) data were high-pass filtered at 1 Hz and low-pass filtered at 60 Hz [Finite impulse response filter (FIR), window type “Hann,” cutoff frequency -6 dB]. Dummy epochs (1000 ms) unrelated to the task structure were generated and all epochs displaying two or more standard deviations from the mean signal were rejected.

Extended infomax ICA was applied and unmixing ICA weights were saved with the raw data.

For ERP analysis the raw data with ICA weights was high-pass filtered at 0.1 Hz and low-pass filtered at 30 Hz (FIR filter, window type “Hann,” cutoff frequency -6 dB). Artifactual ICA components were identified by visual inspection and removed. Data was then re-referenced to the arithmetic mean of the electrodes that were located on the left and right mastoid (TP9, TP10), epoched around CW onset (-50 ms to $+1000$ ms), and baseline corrected (-50 to $+50$ ms). This peristimulus baseline (compare Blackford et al., 2012) was used to ensure that no prior influence of either turn-taking mode (i.e., speech output in the reading aloud mode or auditory input in the listening mode) was present in the baseline signal. Analysis of the gaps between sentence fragment and critical word (compare “Procedure”) confirmed prior offset of speech (Reading aloud) and recording (Listening) to the selected baseline time window (-50 ms to CW onset).

An automatic artifact rejection procedure was used to exclude trials. Epochs displaying three or more standard deviations from the mean signal were rejected. The mean number of trials entering statistical analysis for Listening congruent were 111, Listening incongruent 37, Reading aloud congruent 111, and Reading aloud incongruent 33.

Statistical Analysis

For the definition of an unbiased time window for analysis of the N400 effect, we created a grand average across all participants including all conditions (Listening congruent and incongruent, Reading aloud congruent and incongruent) and electrodes (compare Keil et al., 2014, p. 7). Three components were present in the ERP: a negative peak at 140 ms, a positive peak at 254 ms, and a negative peak at 446 ms. Based on this grand average (see Supplementary Figure 1 in Data Sheet 1) we defined a window of analysis relying on the zero crossings of the ERP: a window for analysis of the N400 from 370 to 530 ms and for the P200 from 166 to 336 ms, relative to CW onset. Both effects, N400 and P200, were analyzed using repeated measures analysis of variance (ANOVA).

For analysis of the N400 effect, we conducted separate repeated measures ANOVAs along the midline and quadrants to evaluate its topographical distribution (compare Kos et al., 2010). The midline analysis included the within-subjects factors electrode (Fz, Cz, CPz, Pz), turn-taking mode (Listening, Reading aloud), and congruency (congruent, incongruent). In case of a significant interaction between electrode and congruency, follow-up single ANOVAs for each electrode were computed. The quadrant analysis included four quadrants (averaging over four electrodes), with the within-subjects factors quadrant (left anterior: Fp1, F7, FC1, C3, left posterior: CP5, CP1, P3, O1, right anterior: Fp2, F8, FC2, C4, right posterior: CP6, CP2, P4, O2), and the main factors turn-taking mode (Listening, Reading aloud) and congruency (congruent, incongruent). In case of a significant interaction between quadrant and congruency, follow-up single ANOVAs for each quadrant were computed.

To address concerns of differing signal-to-noise ratios between congruent and incongruent conditions (~ 120 – 40

trials, respectively) the midline and quadrant analyses of the N400 effect were repeated with matching numbers of trials, i.e., 40 trials for the congruent condition and 40 trials for the incongruent condition (see Supplementary Material, Data Sheet 1).

The P200 component has a fronto-central distribution with no clear lateralization (cf. Lee et al., 2012; Schierholz et al., 2015). Therefore, the analysis of this component was limited to frontal electrodes. The selection of the specific frontal electrodes for analysis of the P200 was based on the grand average topography including all conditions (Listening congruent and incongruent, Reading aloud congruent and incongruent) which led to the electrodes Fz, FC1, FC2, and Cz. The P200 statistical analysis included the within-subjects factors electrode (Fz, FC1, FC2, Cz), turn-taking mode (Listening, Reading aloud), and congruency (congruent, incongruent). In case of a significant interaction between electrode and congruency or electrode and turn-taking mode, follow-up single ANOVAs for each electrode were computed.

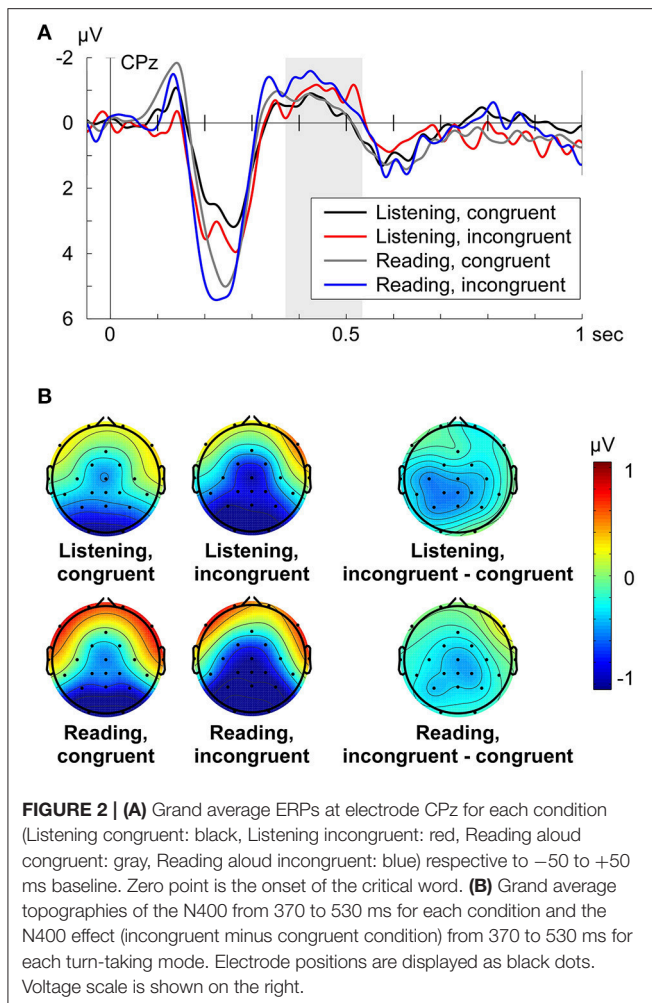
For all analyses, whenever Mauchly’s test indicated a violation of sphericity, the Greenhouse-Geisser corrected p -values with the original degrees of freedom will be reported. In all cases, we focused on the results of the main effects of congruency, turn-taking mode, and their interaction; as well as the possible interaction of these main effects with the factors electrode or quadrant. Estimates of the effect sizes are reported in terms of partial eta squared (η_p^2) values.

RESULTS

The grand average ERP (e.g., CPz shown in **Figure 2A**) for Listening and Reading aloud was characterized by three dominant peaks: A negative deflection around 140 ms, a large positive deflection at around 200 ms (P200), and the negative deflection at around 400 ms, the time window in which the N400 effect can be expected to occur. For both turn-taking modes (Listening, Reading aloud) the amplitude of the N400 was larger for the incongruent condition compared to the congruent condition (**Figures 2A,B**).

N400 Effect

For analysis of the N400 effect, a separate midline (factor electrode: Fz, Cz, CPz, Pz) and lateral (factor quadrant) repeated measures ANOVA for the time window 370–530 ms was computed with the main common factors congruency (congruent vs. incongruent) and turn-taking mode (Listening vs. Reading aloud). The midline analysis of the N400 revealed that congruency (congruent vs. incongruent) had a statistically significant effect on the N400 amplitude [midline: $F_{(1, 15)} = 6.07$, $p = 0.026$, $\eta_p^2 = 0.288$], where incongruent conditions led to a larger N400 (**Figure 2A**). The turn-taking mode (Listening vs. Reading aloud) did not significantly affect the N400 amplitude [midline: $F_{(1, 15)} = 0.11$, $p = 0.741$] and there was no significant interaction between congruency and mode on the N400 amplitude [midline: $F_{(1, 15)} = 0.26$, $p = 0.62$]. We found a significant interaction between electrode and congruency [midline: $F_{(3,45)} = 4.54$, $p = 0.015$, $\eta_p^2 = 0.232$] pointing to



a distributional difference of the N400 effect along the midline (**Figure 2B**). Follow-up analyses revealed a significant N400 effect (i.e., a main effect of congruency) at electrodes Cz [$F_{(1, 15)} = 5.13, p = 0.039, \eta_p^2 = 0.255$], Cpz [$F_{(1, 15)} = 7.63, p = 0.015, \eta_p^2 = 0.337$], and Pz [$F_{(1, 15)} = 6.82, p = 0.02, \eta_p^2 = 0.313$]. To check whether these N400 effects were similar across turn-taking modes (Listening, Reading aloud) a follow-up Pearson's linear correlation analysis of the mean N400 effect (i.e., incongruent minus congruent mean ERP amplitude) across electrodes Cz, Cpz, and Pz was computed. It showed a significant correlation of the N400 effect ($r = 0.43, p = 0.049$) across turn-taking modes.

The analysis of the N400 among quadrants (left anterior: Fp1, F7, FC1, C3, left posterior: CP5, CP1, P3, O1, right anterior: Fp2, F8, FC2, C4, right posterior: CP6, CP2, P4, O2) also displayed a significant main effect of congruency on the N400 amplitude [quadrant: $F_{(1, 15)} = 5.45, p = 0.034, \eta_p^2 = 0.266$]. The turn-taking mode did not significantly affect the N400 [quadrant: $F_{(1, 15)} = 0.04, p = 0.844$] and there was no significant interaction between congruency and turn-taking mode [quadrant: $F_{(1, 15)} = 0.93, p = 0.35$]. We did not find a significant interaction between quadrant and congruency [quadrant: $F_{(3, 45)} = 1.329, p = 0.28$] suggesting a widespread N400 effect distribution. The results of

the same trial number analysis were in line with the present results (see Supplementary Material, Data Sheet 1).

Visual inspection of the N400 effect topographies (Listening and Reading aloud: incongruent minus congruent, **Figure 2B**) suggested a slightly left-lateralized and parietal distribution. The grand average topography of the incongruent condition showed a more central and posterior distribution of the N400 within the analyzed time window (**Figure 2B**: Listening, incongruent and Reading, incongruent). This slightly posterior gradient of the N400 effect was confirmed by the midline analysis (compare above). These findings are in line with the reported distributions of the N400 component (cf. Hagoort and Brown, 2000; Kos et al., 2010; Wang et al., 2012; for a review see Kutas and Federmeier, 2011).

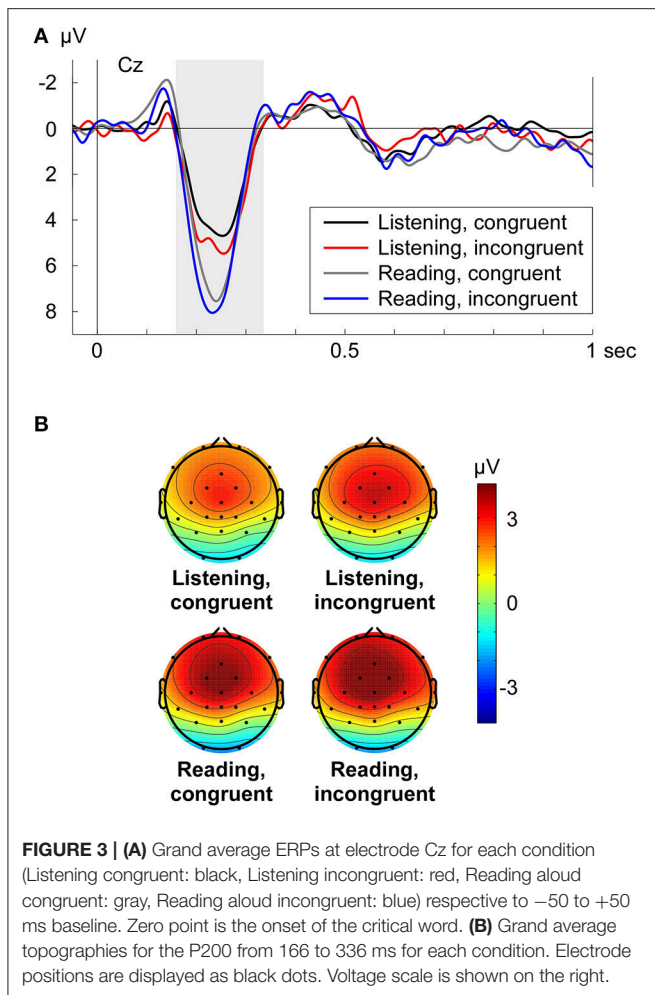
P200

The P200 (Cz shown in **Figure 3A**) was analyzed with a repeated measures ANOVA in the time window from 166 to 336 ms along frontal electrodes (factor electrode: Fz, FC1, FC2, Cz) with the main factors congruency (congruent vs. incongruent) and turn-taking mode (Listening vs. Reading aloud). Congruency was marginally significant on the P200 amplitude [$F_{(1, 15)} = 4.56, p = 0.05, \eta_p^2 = 0.233$], where incongruent conditions showed an enhanced positivity (**Figure 3A**). In contrast to the N400 effect, the turn-taking mode (Listening/Reading aloud) had a significant effect on the P200 amplitude [$F_{(1, 15)} = 7.71, p = 0.014, \eta_p^2 = 0.340$], visible in higher amplitudes for the Reading aloud mode compared to the Listening mode (**Figure 3A**). There was no significant interaction between congruency and mode on the P200 [$F_{(1, 15)} = 0.38, p = 0.548$]. Further, no significant interactions between electrode and congruency [$F_{(3, 45)} = 1.33, p = 0.276$] or between electrode and turn-taking mode [$F_{(3, 45)} = 1.64, p = 0.214$] were present.

The grand average topographies for the analyzed P200 time window displayed the same fronto-central distribution over all four conditions (**Figure 3B**). In line with the statistical analysis, it was also visible that the P200 was increased in the Reading aloud mode compared to the Listening mode. Incongruent conditions too, had a slightly increased P200 in contrast with the congruent conditions (**Figure 3B**). Grand average difference topographies of the P200 can be found in Supplementary Figure 2 (Supplementary Material, Data Sheet 1).

DISCUSSION

Our research aim is to move toward neurophysiological studies of natural dialogues. The present study is a first step with regard to loosening some of the classical constraints that are posed on experimental paradigms by the recording setup. We investigated the N400 effect in a controlled turn-taking scenario using wireless EEG. Our results show that the expectancy violation in response to incongruent words is measurable as an amplitude increase in the N400 compared to congruent words. This was the case for both measured scenarios: when reading aloud the first sentence fragment and listening to the final word, as well as when listening to the first sentence fragment and the final word



from different speakers. In other words, this introduced speaker-switch for the final (analyzed) word did not prevent detecting the N400 effect. The effect on the N400 was further present independent of whether the sentence fragment is uttered by oneself or by another person (i.e., reading aloud or listening). The topographies we found for the N400 effect are in line with earlier studies (Hagoort and Brown, 2000; Kos et al., 2010; for a review see Kutas and Federmeier, 2011; Van den Brink et al., 2012; Wang et al., 2012; Erlbeck et al., 2014). Unexpectedly, we found a large P200 component, which was also enhanced for the reading aloud condition compared to the listening condition. We ascribe this difference to the saliency of the new stimulus (auditory speaker-switch) and the differences between the modes in sound source and sound generation. In summary, this new paradigm has proven valuable in investigating the effects of a speaker-switch on the N400. The findings of this study provide a framework for future studies of neural correlates of language processing in more natural conversational setups, that inevitably incorporate turn-taking (and hence speaker-switches). A direct comparison of a speaker-switch condition and a non-speaker-switch condition might also be valuable for further research.

The amplitude difference between expected and unexpected sentence endings of the observed N400 effect within this turn-taking paradigm was moderate in comparison to other literature findings (Hagoort and Brown, 2000; Kos et al., 2010, but see also Gunter et al., 2000; Rueschemeyer et al., 2014). It is therefore possible that the participants did not process the sentences in depth. While the control task (word present or absent in sentence) ensured attention to the sequence of words, it did not ensure that the sentence was processed semantically (for a study of effects of task instructions on the N400 see Erlbeck et al., 2014). We expect that questions that require a more in-depth semantic analysis of the sentences (e.g., “Who took the children to the school?”) would lead to an increased level of engagement, in turn, this would drive expectations for the continuation of the sentences (cf. Van Berkum, 2012) which should also affect the N400 amplitude. Due to the design of the language material (see Section Linguistic Material), the N400 effect could also have been influenced by variations in the expectedness of specific sentence endings (i.e., the cloze probability for the specific critical words). A further reason for the rather small N400 effect could also be that the unexpected sentence endings of our study are not only unexpected but anomalous. As a result, participants might have stopped semantic analysis in some cases, leading to no enhanced integration costs and again to a less pronounced overall N400 effect (cf. Kutas and Federmeier, 2011).

Apparent in the data is the strong P200 for both turn-taking conditions. In general, the P200 (and N100, auditory evoked potentials) have been linked to higher order processes of attention and perception (Beauducel et al., 2000; Federmeier et al., 2005; Heim and Alter, 2006) and respond strongly to salient sound stimuli (i.e., novel or differing in intensity). We expect that this attention response is reflected in our P200 findings, that is, a strong orientation toward the new auditory stimulus: the speaker-switch. Further, we find an enhanced P200 for the reading aloud condition in contrast to the listening condition. Several factors might play a role here. First, in the reading aloud condition the participant has to switch from an active (i.e., reading aloud the first part of the sentence) to a passive task (i.e., listening to the critical word), while in the listening condition this switch is absent. Second, in the reading aloud condition there is a change in the sound source. On the one hand, the origin of the sound changes from the mouth of the participant to the loudspeaker, and on the other hand, the sound changes from self-generated to externally generated. Taken together, the sound properties in the reading aloud condition between the first part of the sentence and the last word differ more dramatically than in the listening condition. As the difference between the sentence and the critical word is more pronounced in the reading aloud condition, the higher amplitudes in this condition can be ascribed to that.

Our focus in this study was on the effect of a speaker switch and a semantic manipulation on the N400. Note, however, that we found congruency to be marginally significant on the P200 ($p = 0.05$), such that we cannot exclude an influence of congruency on the differences of the P200 amplitude. Previous studies presenting written material have reported similar results of an

enhanced early positivity for semantically unexpected sentence endings compared to expected ones, which seem to interact with several factors, such as word frequency and repetition effects (Lee et al., 2012; Wang et al., 2012). In the present study, the critical word is presented auditorily but preceded by written (Listening and reading aloud condition) and spoken (Reading aloud condition) material. The interaction of these factors and their impact on the P200 remains to be studied. The absent interaction of congruency and turn-taking mode on the P200 points to two possibly different processes of congruency and physical stimulus analysis that are reflected in the P200.

The overall goal of our research is to approach neural correlates of natural conversational situations. As a first step, we integrated a speaker-switch for each congruent or incongruent ending word of a sentence within this turn-taking paradigm. The participant read aloud the first part of the sentence or listened to the first part of the sentence, but the last word was always presented by another speaker. An advantage of this speaker-switch setup was the separation of overt speech and the word of interest. This way, artifacts arising from mouth movements during speech could be circumvented for the signal of interest. To ensure the transition between the sentence fragment and the final word in this reading aloud condition, a human involvement for triggering the final word was necessary. As a result, the presentation of the last word of each sentence was triggered by the experimenter; only after she pressed the button the final word was presented. Engaging the experimenter had the advantage that the participant was able to read in the own preferred pace. Further, artifactual motor components arising from button presses could be evaded in the participant's EEG signal of interest. A drawback of human involvement is certainly the variability in reaction times. However, this should not have any effect on the studied EEG components. The analysis of the time between the sentence fragment and the final word shows no significant difference in the gap length for the two conditions. In addition, variable gap lengths between sentence fragment and final word are advantageous in terms of ecological validity. The gaps between utterances also vary during natural verbal interaction (Levinson and Torreira, 2015). Similarly, natural conversations include scenarios where the own utterance is completed by another person or one listens to someone else being completed (Coates, 1990; Purver et al., 2009). Both of these scenarios were included in the present paradigm (listening mode and reading aloud mode). In conversations, finishing each other's sentences is usually attested in situations where speakers seem to have lost their track of argumentation (e.g., helping out with a specific word) or to emphasize specific aspects of a discussion (Purver et al., 2009; Levinson and Torreira, 2015). Here, we mimicked the first case of helping out with a word in an active (reading aloud) and passive (listening) mode. In both modes, we predicted that expectation violations (in our case semantic violations) are visible in an N400 effect. Our results support this hypothesis.

Evidently, the paradigm used here still lacks many aspects of natural conversations. Conversations are primarily characterized by two (or more) speakers taking consecutive full turns

(i.e., a full sentence or number of sentences), one at a time, with no or almost no overlap (Levinson and Torreira, 2015). This aspect could not be addressed in the present setup.

Our study provides a promising starting point for future research. Future studies will have to increase the naturalness of conversational settings while keeping the constraints that are necessary to ensure quantitative analysis. These would include measuring two persons at the same time that interact with one another, moving from controlled language material to self-generated speech and natural (i.e., not pre-defined) turn-taking behavior. Some of these aspects already have been addressed separately (e.g., Jiang et al., 2012; Kawasaki et al., 2013; Himberg et al., 2015; Mandel et al., 2015; Torreira et al., 2015; Bögels et al., 2015a,b,c; Bašnáková et al., 2015) but it remains a challenge to combine these with neuroimaging. An advantage of a wireless (and hence mobile) EEG—as used here—is the possibility to measure electrophysiology outside the strongly controlled and restricted lab environment (e.g., Debener et al., 2012) and, moreover, in interactive situations of two or more participants in immediate proximity (meeting the described criteria for more naturalness described previously).

CONCLUSION

The study of neural correlates of language in conversational settings is challenging. Here we have made an attempt to study the neurophysiological underpinnings of language in a controlled turn-taking scenario using wireless EEG. We have shown that we can measure the N400 effect reliably despite a speaker-switch. Future studies can build upon the approach presented here and move toward more conversational settings to study neural mechanisms of language processing.

ETHICS STATEMENT

The study was approved by the Ethics Committee of the University of Oldenburg. All participants signed written informed consent prior measurement.

AUTHOR CONTRIBUTIONS

Study design: TG, MB, ER, SS. Conduction of study: TG, MB. Data analysis: TG, MB. Data interpretation: TG, MB, ER, SS. Manuscript writing: TG, MB, ER, SS.

FUNDING

This research was funded by the Volkswagen Foundation (European platform) and funds from the Task Group 7 “BCI for Hearing Aids,” DFG Cluster of Excellence 1077 Hearing4all, Oldenburg, Germany.

ACKNOWLEDGMENTS

We thank Jana Müller for her technical help, Rebecca Carroll for her assistance in recordings, and Jan Patrick Zeller for his help with the language material. We are grateful to Stefan Debener for his support.

REFERENCES

- Bašnáková, J., van Berkum, J., Weber, K., and Hagoort, P. (2015). A job interview in the MRI scanner: how does indirectness affect addressees and overhearers? *Neuropsychologia* 76, 79–91. doi: 10.1016/j.neuropsychologia.2015.03.030
- Beauducel, A., Debener, S., Brocke, B., and Kayser, J. (2000). On the reliability of augmenting/reducing: peak amplitudes and principal component analysis of auditory evoked potentials. *J. Psychophysiol.* 14, 226–240. doi: 10.1027//0269-8803.14.4.226
- Blackford, T., Holcomb, P. J., Grainger, J., and Kuperberg, G. R. (2012). A funny thing happened on the way to articulation: N400 attenuation despite behavioral interference in picture naming. *Cognition* 123, 84–99. doi: 10.1016/j.cognition.2011.12.007
- Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glott Int.* 5, 341–345.
- Bögels, S., Barr, D. J., Garrod, S., and Kessler, K. (2015a). Conversational interaction in the scanner: mentalizing during language processing as revealed by MEG. *Cereb. Cortex* 25, 3219–3234. doi: 10.1093/cercor/bhu116
- Bögels, S., Kendrick, K. H., and Levinson, S. C. (2015b). Never say no ... How the brain interprets the pregnant pause in conversation. *PLoS ONE* 10:e0145474. doi: 10.1371/journal.pone.0145474
- Bögels, S., Magyar, L., and Levinson, S. C. (2015c). Neural signatures of response planning occur midway through an incoming question in conversation. *Sci. Rep.* 5, 1–11. doi: 10.1038/srep12881
- Brainard, D. H. (1997). The psychophysics toolbox. *Spat. Vis.* 10, 433–436. doi: 10.1163/156856897X00357
- Coates, J. (1990). Modal meaning: the semantic-pragmatic interface. *J. Semant.* 7, 53–63. doi: 10.1093/jos/7.1.53
- Debener, S., Minow, F., Emkes, R., Gandras, K., and de Vos, M. (2012). How about taking a low-cost, small, and wireless EEG for a walk? *Psychophysiology* 49, 1617–1621. doi: 10.1111/j.1469-8986.2012.01471.x
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- De Vos, M., and Debener, S. (2014). Mobile EEG: towards brain activity monitoring during natural action and cognition. *Int. J. Psychophysiol.* 91, 1–2. doi: 10.1016/j.ijpsycho.2013.10.008
- De Vos, M., Gandras, K., and Debener, S. (2014). Towards a truly mobile auditory brain-computer interface: exploring the P300 to take away. *Int. J. Psychophysiol.* 91, 46–53. doi: 10.1016/j.ijpsycho.2013.08.010
- Duncan, C. C., Barry, R. J., Connolly, J. F., Fischer, C., Michie, P. T., Näätänen, R., et al. (2009). Event-related potentials in clinical research: guidelines for eliciting, recording, and quantifying mismatch negativity, P300, and N400. *Clin. Neurophysiol.* 120, 1883–1908. doi: 10.1016/j.clinph.2009.07.045
- Erlbeck, H., Kübler, A., Kotchoubey, B., and Vesper, S. (2014). Task instructions modulate the attentional mode affecting the auditory MMN and the semantic N400. *Front. Hum. Neurosci.* 8:654. doi: 10.3389/fnhum.2014.00654
- Federmeier, K. D., Mai, H., and Kutas, M. (2005). Both sides get the point: hemispheric sensitivities to sentential constraint. *Mem. Cognit.* 33, 871–886. doi: 10.3758/BF03193082
- Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiol. Rev.* 91, 1357–1392. doi: 10.1152/physrev.00006.2011
- Gunter, T. C., Friederici, A. D., and Schriefers, H. (2000). Syntactic gender and semantic expectancy: ERPs reveal early autonomy and late interaction. *J. Cogn. Neurosci.* 12, 556–568. doi: 10.1162/0899892900562336
- Hagoort, P., and Brown, C. M. (2000). ERP effects of listening to speech: semantic ERP effects. *Neuropsychologia* 38, 1518–1530. doi: 10.1016/S0028-3932(00)00052-X

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2016.01854/full#supplementary-material>

- Heim, S., and Alter, K. (2006). Prosodic pitch accents in language comprehension and production: ERP data and acoustic analyses. *Acta Neurobiol. Exp. (Wars)* 66, 55–68.
- Hillyard, S. A., and Kutas, M. (1983). Electrophysiology of cognitive processing. *Annu. Rev. Psychol.* 34, 33–61. doi: 10.1146/annurev.ps.34.020183.000341
- Himberg, T., Hirvenkari, L., Mandel, A., and Hari, R. (2015). Word-by-word entrainment of speech rhythm during joint story building. *Front. Psychol.* 6:797. doi: 10.3389/fpsyg.2015.00797
- Jiang, J., Dai, B., Peng, D., Zhu, C., Liu, L., and Lu, C. (2012). Neural synchronization during face-to-face communication. *J. Neurosci.* 32, 16064–16069. doi: 10.1523/JNEUROSCI.2926-12.2012
- Kaan, E. (2007). Event-related potentials and language processing: a brief overview. *Lang. Linguist. Compass* 1, 571–591. doi: 10.1111/j.1749-818X.2007.00037.x
- Kawasaki, M., Yamada, Y., Ushiku, Y., Miyauchi, E., and Yamaguchi, Y. (2013). Inter-brain synchronization during coordination of speech rhythm in human-to-human social interaction. *Sci. Rep.* 3, 1–8. doi: 10.1038/srep01692
- Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., et al. (2014). Committee report: publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology* 51, 1–21. doi: 10.1111/psyp.12147
- Kos, M., Vosse, T., van den Brink, D., and Hagoort, P. (2010). About edible restaurants: conflicts between syntax and semantics as revealed by ERPs. *Front. Psychol.* 1:222. doi: 10.3389/fpsyg.2010.00222
- Kutas, M., and Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annu. Rev. Psychol.* 62, 621–647. doi: 10.1146/annurev.psych.093008.131123
- Lee, C.-Y., Liu, Y.-N., and Tsai, J.-L. (2012). The time course of contextual effects on visual word recognition. *Front. Psychol.* 3:285. doi: 10.3389/fpsyg.2012.00285
- Levinson, S. C. (2016). Turn-taking in human communication – origins and implications for language processing. *Trends Cogn. Sci.* 20, 6–14. doi: 10.1016/j.tics.2015.10.010
- Levinson, S. C., and Torreira, F. (2015). Timing in turn-taking and its implications for processing models of language. *Front. Psychol.* 6:731. doi: 10.3389/fpsyg.2015.00731
- Magyar, L., Bastiaansen, M. C., de Ruiter, J. P., and Levinson, S. C. (2014). Early anticipation lies behind the speed of response in conversation. *J. Cogn. Neurosci.* 26, 2530–2539. doi: 10.1162/jocn_a_00673
- Mandel, A., Bourguignon, M., Parkkonen, L., and Hari, R. (2015). Sensorimotor activation related to speaker vs. listener role during natural conversation. *Neurosci. Lett.* 614, 99–104. doi: 10.1016/j.neulet.2015.12.054
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the edinburgh inventory. *Neuropsychologia* 9, 97–113. doi: 10.1016/0028-3932(71)90067-4
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat. Vis.* 10, 437–442. doi: 10.1163/156856897X00366
- Purver, M., Howes, C., Healey, P. G. T., and Gregoromichelaki, E. (2009). “Split utterances in dialogue: a corpus study,” in *Proceeding SIGDIAL 2009 Conference the 10th Annual Meeting of the Special Interest Group in Discourse and Dialogue* (London), 262–271.
- Rueschmeyer, S.-A., Gardner, T., and Stoner, C. (2014). The Social N400 effect: how the presence of other listeners affects language comprehension. *Psychon. Bull. Rev.* 22, 128–134. doi: 10.3758/s13423-014-0654-x
- Ruigendijk, E., Hentschel, G., and Zeller, J. P. (2015). How L2-learners’ brains react to code-switches: an ERP study with Russian learners of German. *Second Lang. Res.* 32, 197–223. doi: 10.1177/0267658315614614

- Schierholz, I., Finke, M., Schulte, S., Hauthal, N., Kantzke, C., Rach, S., et al. (2015). Enhanced audio-visual interactions in the auditory cortex of elderly cochlear-implant users. *Hear. Res.* 328, 133–147. doi: 10.1016/j.heares.2015.08.009
- Torreira, F., Bögels, S., and Levinson, S. C. (2015). Breathing for answering: The time course of response planning in conversation. *Front. Psychol.* 6:284. doi: 10.3389/fpsyg.2015.00284
- Van Berkum, J. J. A. (2012). “The electrophysiology of discourse and conversation,” in *The Cambridge Handbook of Psycholinguistics*, eds M. J. Spivey, K. McRae, and M. F. Joannisse (New York, NY: Cambridge University Press), 589–614.
- Van den Brink, D., Van Berkum, J. J., Bastiaansen, M. C., Tesink, C. M., Kos, M., Buitelaar, J. K., et al. (2012). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Soc. Cogn. Affect. Neurosci.* 7, 173–183. doi: 10.1093/scan/nsq094
- Wang, L., Zhu, Z., and Bastiaansen, M. (2012). Integration or predictability? A further specification of the functional role of gamma oscillations in language comprehension. *Front. Psychol.* 3:187. doi: 10.3389/fpsyg.2012.00187

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Goregliad Fjaellingsdal, Ruigendijk, Scherbaum and Bleichner. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.