

Early Anticipation Lies behind the Speed of Response in Conversation

Lilla Magyari¹, Marcel C. M. Bastiaansen^{1,2}, Jan P. de Ruiter³,
and Stephen C. Levinson^{1,4}

Abstract

■ RTs in conversation, with average gaps of 200 msec and often less, beat standard RTs, despite the complexity of response and the lag in speech production (600 msec or more). This can only be achieved by anticipation of timing and content of turns in conversation, about which little is known. Using EEG and an experimental task with conversational stimuli, we show that estimation of turn

durations are based on anticipating the way the turn would be completed. We found a neuronal correlate of turn-end anticipation localized in ACC and inferior parietal lobule, namely a beta-frequency desynchronization as early as 1250 msec, before the end of the turn. We suggest that anticipation of the other's utterance leads to accurately timed transitions in everyday conversations. ■

INTRODUCTION

The primary ecology for language use and for the acquisition of language by children is the give and take of conversation. This conversational setting is characterized by rapid turn-taking, mostly with minimal gaps (under 200 msec) between one speaker and the next (Stivers et al., 2009). Two additional properties make this coordination rather remarkable:

- (a) a conversational turn is of no fixed length, adapting to the open-ended or generative character of natural language syntax (Sacks, Schegloff, & Jefferson, 1974);
- (b) the language production system is quite slow, even a single word requiring 600 msec from conception to articulatory output (Indefrey & Levelt, 2004; Levelt, 1989), and multiword utterances considerably longer (see e.g., Schnurr, Costa, & Caramazza, 2006; Jescheniak, Schriefers, & Hantsch, 2003).

If we put these facts together, it is clear that a would-be speaker must begin the production of his or her turn half a second or more before the other speaker has stopped speaking and so must predict the end of the incoming turn though it is of no fixed length.

There have been various proposals about how this remarkable coordination might be achieved. Some authors have suggested that there are turn-ending signals (analogous to the “over and out” on a two-way half-duplex radio), either in prosody (Schegloff, 1996; Local, Kelly, &

Wells, 1986; Cutler & Pearson, 1985; Local, Wells, & Seba, 1985; Beattie, Cutler, & Pearson, 1982) or gaze (Kendon, 1967), but recent work does not support this for intonation (De Ruiter, Mitterer, & Enfield, 2006) or gaze (Rossano, Brown, & Levinson, 2009). Others have suggested that a composite bundle of turn-end features might be involved (Duncan, 1974). But all these suggestions run into problem (b) above, for the latency in the production system renders these signals too late to play a decisive role. Another suggestion is that turn-taking can be modeled by coupled oscillators (Wilson & Wilson, 2005) on the basis of the speaker's rate of syllable production, in a manner similar to emergent coordination in, for example, firefly synchronization (Camazine et al., 2001). This suggestion runs into problem (a) above, that turns are not fixed in size but have very varying durations. In addition, recent work shows that underlying even simple human synchronization there is a much more complex corepresentation of joint coordination (Sebanz, Bekkering, & Knoblich, 2006).

Thus, although we have a good grasp of the descriptive properties of the turn-taking system in conversation (Sacks et al., 1974) and evidence suggesting universal tendencies to minimize overlaps and gaps (Stivers et al., 2009), we do not understand the cognitive processes that make possible this virtuoso coordination, which we all practice on the order of 1200 times a day (extrapolated from Mehl, Vazire, Ramirez-Esparza, Slachter, & Pennebaker, 2007).

The aim of this study was to gain insight into the cognitive processes of the listener engaged in anticipating the ending of the incoming turn. We used the EEG signal of participants engaged in this task to explore the temporal dynamics of turn-end anticipation—how far from the end of the turn does the listener move from a passive

¹Max Planck Institute for Psycholinguistics, ²NHTV University of Applied Science, Breda, The Netherlands, ³Bielefeld University, ⁴Donders Institute

comprehension mode into a more active mode ready for the production of speech or action?

The current study builds especially on an earlier study (De Ruiter et al., 2006), which experimentally assessed the relative contribution of intonation and lexico-syntactic content to turn-end prediction using turns extracted from natural conversation. Participants listened to each of these out of context and tried to press a key exactly at the ending of the turn. In the different experimental conditions, participants listened to (a) the original recording of a turn, (b) a version with intonational contour removed, or (c) a version with no recognizable words but with intact intonation. When participants listened to the original recordings (a), they were able to press the key with an accuracy that paralleled turn-transitions in natural conversation, suggesting relatively little influence of pragmatic and context effects. Accuracy of the timing of key presses did not change significantly when the intonation was filtered out. In contrast, when the words were rendered incomprehensible but the intonation was intact, the accuracy was greatly reduced. The authors concluded that people rely mainly on lexical and syntactic information for anticipating turn-ends.

How might lexical and syntactic information play a decisive role in predicting turn-endings? Whereas prosodic cues are assumed to appear just before the turn-ends and to give only binary information to listeners whether a turn is ending soon or not yet, anticipated syntactic and lexical information is a good candidate for giving more fine-grained temporal information much earlier about when the turn is going to end. As a sentence unfolds the probabilities of continuations in different directions become ever narrower, a property exploited in nearly all modern machine processing of natural language (Manning & Schütze, 1999; Bates, 1995). Electrophysiological and eye-tracking studies have revealed that predictions are made during language comprehension at many different linguistic levels (DeLong, Urbach, Groppe, & Kutas, 2011; Altmann & Kamide, 2007; DeLong, Urbach, & Kutas, 2005; Van Berkum, Brown, Zwitterlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004; Kamide, Altmann, & Haywood, 2003). Listeners, as they process incoming turns, come to a point where they can actually predict the very next words (DeLong et al., 2005, 2011). Also, turns whose end points can be more accurately predicted allow the prediction of the final words (Magyari & De Ruiter, 2012).

It is clear that listeners can predict the end of a turn before it ends. But it is unclear how early they sense the imminence of ending and thus switch from a purely passive comprehending role into a more active role ready for speech or next action. These internal processes are not easy to get at through behavioral measures.

To explore the internal temporal dynamics, we used turns extracted from recordings of natural conversations as in the study (De Ruiter et al., 2006) earlier described. A prior offline gating task (see Methods), where partici-

pants had to complete actual turns cut short, was used to categorize turns as having either predictable (PRED) or unpredictable (UNPRED) final words during the last 600 msec before the turn-end (Figure 1). For the main task, participants were asked to listen to the full turns in both conditions and try to press the key exactly at the end of the turn. We expected key presses to be more accurate for predictable turn-ends. To reveal the temporal dynamics of turn-end anticipation, we measured the EEG of the participants while they were performing the experimental task. We expected to find anticipatory neural activity for predictable turn-ends, not for unpredictable turn-ends, appearing at least 600 msec before the turn-end. We focused on the dynamics of EEG oscillations, as oscillatory dynamics in the alpha and beta frequency ranges have been clearly associated with both motor and nonmotor anticipation in earlier research (Bastiaansen & Brunia, 2001; Pfurtscheller & Lopes da Silva, 1999; Pfurtscheller & Aranibar, 1977; Jasper & Penfield, 1949). Beta power and coherences changes have also been suggested to be related to syntactic and semantic processing (Wang, Zhu, & Bastiaansen, 2012; Bastiaansen, Magyari, & Hagoort, 2010; Weiss et al., 2005) and to reflect a close relationship between language comprehension and motor functions (Weiss & Mueller, 2012). We thus had two dependent measures, the timing of key

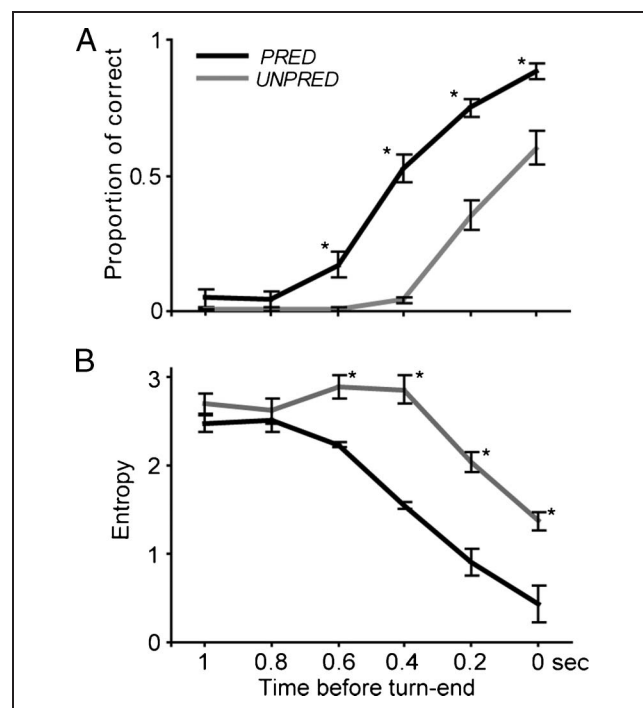


Figure 1. Averaged results of the gating study for turns selected into the PRED and UNPRED conditions. The x axis shows how many seconds before the end of the turn the recording was cut off. Error bars indicate the standard error; * indicates significant differences between conditions. (A) Proportion of correct answers averaged across turns of the two conditions at each gating points. (B) Entropy of the answers averaged across turns of the two conditions at each gating points.

presses and the time–frequency analysis of EEG power changes.

METHODS

Participants

Twenty-six participants (mean age = 25 years, range = 19–39 years; 7 men, 15 women) gave informed consent and were paid for their participation in the EEG experiment. All were right-handed, native speakers of Dutch with no history of neurological or language disorders. None of them took part in the pretest of the stimuli material. Data from four participants were discarded because of excessive blinking, left-handedness, or strikingly different key press results that suggested that the participant did not follow the instructions.

Pretest of Stimuli

The selection of the stimuli required a pretest using a gating paradigm. Forty-eight participants from the subject pool of the Max Planck Institute for Psycholinguistics participated in this study. None of them participated in the EEG study. Turns were used from Dutch, telephone-like conversations. The audio recordings of the conversations were made for another experiment (De Ruiter et al., 2006). The recordings were made in two soundproof cabins to separate the channels carrying the recordings of the two speakers. For the pretest, the audio recordings of 108 turns were selected. These turns were 2.25- to 10-sec long, were not followed by a laugh or breath, and were not interrupted by interjections from the other speaker. Each turn was cut 200, 400, 600, 800, and 1000 msec before the end. Each version of a turn (five shorter and a full version) was assigned to different experimental lists. Eight participants per list performed in the experiment. Each list started with 12 practice turns. The participants were asked to listen to each segment once. After hearing a segment, they had to type on a computer keyboard their guess about the continuation of the turn starting from the last word that they heard. For further information on the method, see a similar gating study in Magyari and De Ruiter (2012). The answers were evaluated with regards to two aspects. First, each answer was coded as correct or incorrect, where an answer was correct if it exactly matched the words used in the original uncut stimuli. Second, it was also coded as to whether the answers to the same segment given by different participants were the same or different. On the basis of this, we used entropy (Shannon, 1948) to measure the variety of the answers. Shannon entropy was calculated using this formula: $\text{entropy} = - \sum p_i \log_2(p_i)$ where p_i is the proportion of one kind of guess among the eight for each gating period (eight participants guessed the missing words from each gating). If guesses are similar to each other, the entropy is low (minimum: 0); if the answers are different, the entropy is high (maximum: 3).

Stimulus Material

On the basis of the results of the gating study, 30 turns with the highest proportion of correct answers (mean = 0.404, averaged across gating points) were selected into the PRED condition of the experiment. These turns had also a low entropy across all gating points (mean = 1.688, averaged across gating points). Later, another 30 turns with a low proportion of correct answers (mean = 0.169) and with high entropy (mean = 2.415) were added to the UNPRED condition (differences in proportion of correct answers: $t_{58} = 8.177, p < .001$; differences in averaged entropy: $t_{58} = -6.899, p < .001$). The entropy and proportion of correct answers was different between the two conditions from the 600 msec gating point before the turn-end ($t_{58} = 3.517, p = .001$, proportion of correct; $t_{58} = -5.028, p < .001$, entropy; Figure 1). Syllables were on average 178 msec, words 235 msec long. There was no significant difference in the duration of the turns in the two conditions (mean(PRED) = 4.25 sec, mean(UNPRED) = 3.84 sec, $t_{58} = 1.015, p = .314$).

An example from the PRED condition:

“Eh ik woon in een huis met vier vrouwen en nog een andere man” (Dutch)
 (“Eh I live in the same house with four women and with another man.” (Translation))

An example from the UNPRED condition:

“Oe en toen was ze weer eh s solo in eh in het noorden” (Dutch)
 (“Uh and then, she was again eh alone in eh in the north.” (Translation))

Experiment and Procedure

On the basis described above, 30 turns were selected into the PRED and 30 turns were selected into the UNPRED conditions. There were 22 other items that were selected originally for a third condition and 18 turns for practice. Data from these trials were not used for further analysis. Four experimental lists were created with different orders of the experimental trials. The practice trials were always at the beginning of each list, in the same order. Instructions and experimental task were similar to the instructions and task in De Ruiter et al.’s key press experiment (De Ruiter et al., 2006). Instructions appeared on the computer screen and contained the following (in Dutch): “The aim is that you should press the button PRECISELY at the moment the speaker finishes his turn. This means that you must try to predict the end of the fragment. You should not wait until the fragment has finished and then press the button.” Participants were also instructed to avoid blinks and movements other than the key press during a trial. When participants pressed a green button, the next trial started and a red button measured the responses. When an experimental trial started, a fixation cross appeared on the screen, 1500 msec after which

the audio fragment was played. A fixation cross was present until 2000 msec after the fragment finished or until the red button was pressed with the right hand. A blank screen was presented for a minimum of 1500 msec after the fixation cross indicating that the participant was allowed to blink. When the participants pressed the red button, the audio stimuli stopped. When the black screen changed, a screen appeared with the instruction: "Press the green button!" Then the participants were free to start with the next trial. After the first half of the trials, there was a break. Then, the experimenter went into the room and checked the participant and the electrodes. The experiment continued after the experimenter pressed a button outside the room.

Participants were tested in a sound-proof, electromagnetically shielded room. They were seated at a distance of approximately 60 cm from a computer screen mounted on a table, next to a key box with green and red response keys. The visual and auditory stimuli were played by Presentation software (version 12.1.03.24.08; Neurobehavioral Systems, Inc., Albany, CA). Key presses and the EEG were both recorded.

EEG Recordings

EEG was recorded from 61 active Ag/AgCl electrodes using an actiCap (Radnor, PA). Fifty-nine of the 61 electrodes were mounted in the cap with equidistant electrode montage referenced to the left mastoid. Two separate electrodes were placed at the left and the right mastoid outside the cap. Blinks were monitored through an electrode on the intraorbital ridge below the left eye. Horizontal eye movements were monitored through two electrodes in the cap placed approximately at each outer canthus. The ground electrode was placed on the forehead. Electrode impedance was kept below 10 k Ω . EEG and EOG recordings were amplified through BrainAmp DC amplifiers. DC recording was applied with a low-pass filter of 100 Hz. The recording was digitized online with a sampling frequency of 500 Hz and stored for offline analysis.

Data Preprocessing

Segmentation and artifact rejection of the EEG data were performed with Brain Vision Analyzer (version 1.05.0005; Gilching, Germany) software. The data were segmented in epochs of 5000 msec, -3000 msec before, and 2000 msec after key press. A baseline between -2000 msec and -1500 msec before the key press was used for artifact rejection. Approximately 23% of the trials were rejected. The average number of trials was 22.5 in PRED and 23.8 in the UNPRED conditions.

Behavioral Data

The temporal offset between the end of a turn and the key presses was measured. The averaged RT indicates

how accurately participants could anticipate the turn-ends. The averaged time is positive when participants press the key too late, and it is negative when participants press the key before the turn-end.

Time-Frequency Analysis of Power

Time-frequency representations (TFRs) of single trial data were computed by using the multitaper approach (Mitra & Pesaran, 1999) with FieldTrip software package (Oostenveld, Fries, Maris, & Schoffelen, 2011). TFRs show the power of the different frequency ranges at multiple time points. Multitaper was applied first in a wider frequency range, and then the multitaper parameters were optimized for the beta frequency range. The final time-frequency analysis was done between 6 and 31 Hz in 1.25-Hz step size and time steps of 10 msec with 5-Hz frequency smoothing and 800-msec time smoothing. A relative baseline was applied on the TFRs between -2000 msec and -1700 msec before key press. As a result of this, the power values were expressed as the relative increase or decrease compared with baseline.

Source Reconstruction

To identify the sources in the beta band, we used a beamforming approach, Dynamic Imaging of Coherent Sources (Gross et al., 2001). We were interested in localizing power differences between the conditions at the beginning and in the middle of the trials. Therefore, we created trials in both conditions that contained data from 2 to 1.5 sec before key press (preperiod) and from 1.2 to 0.7 sec before the key press (postperiod). On the basis of the results of the time-frequency analysis, frequency analysis was applied using the multitaper method based on discrete prolate spheroidal sequences (Slepian sequences) on the trials at 15 Hz with a frequency smoothing of ± 3 Hz. Electrodes were aligned to a volume conduction model that was made based on a template brain using the boundary element method (Oostenveld, Praamstra, Stegeman, & van Oosterom, 2001). A common spatial filter was then computed at 15 Hz for the different conditions and the pre- and postperiod together. The spatial filter was projected to all trials. Power values were calculated on an equidistant template 3-D grid with a 5-mm resolution. Trials were averaged in the pre- and postperiods of the different conditions, and the relative differences between conditions were calculated using the following formula: $(\text{power}_{\text{postperiod}} - \text{power}_{\text{preperiod}}) / \text{power}_{\text{postperiod}}$. Finally, the grand averages were computed and interpolated on the template brain.

Statistical Analysis of Behavioral Results

Statistical significance of the differences between conditions in RTs was evaluated by PASW Statistics 18 (Quarry Bay, Hong Kong). Repeated-measures ANOVAs were computed

on the averaged RTs of each participant. Participants' averages were calculated for the two conditions and for the first and second half of the experiment. The ANOVA had two factors: Condition (PRED vs. UNPRED) and Order (first vs. second half of the experiment).

Statistical Analysis of EEG Results

For evaluating the differences between conditions in the EEG, we used a cluster-based random permutation procedure (Maris & Oostenveld, 2007) that is implemented in FieldTrip. We used this statistical approach because it elegantly handles multiple comparison problems. First, for every data point (sensor–time–frequency point) a simple dependent samples *t* test was performed that gave uncorrected *p* values. All data points that did not exceed a preset *p* value (here .05) were zeroed. Clusters of adjacent nonzero data points were computed, and for each cluster, cluster-level test statistics were calculated by taking the sum of all *t* statistics within that cluster. A null distribution was created by randomly assigning the participant averages to one of the two conditions 1000 times, and for each of these randomizations a cluster-level statistic was computed. Then the largest cluster-level statistics of each randomization were entered into the null distribution. The observed cluster-level statistic was compared against the null distribution and clusters falling under the 2.5% of the two sides of the difference distribution were considered to be significant. The statistical test was carried out between 2000 msec before and until the key press.

For the statistical analysis of the source reconstruction, one-sided dependent sample *t* statistics were used comparing the power values of the trial-averaged participant data of PRED and UNPRED conditions at each source point, which fall in the 3-D grid within the template brain. There were 15,711 grid points inside the brain, and for each grid point, there were six neighbors (except at points at the edges of the brain where neighboring locations fall outside the brain). Then, as a way of clustering, for source points that reached significance (uncorrected, $p < .05$, $df = 21$), we examined whether all of their neighboring points were also significant. Voxels that had only significant neighbors were accepted as showing an effect. For localizing the spatial coordinates of the significant areas, the *t* values of the significant, clustered source points and zeros at all other points were interpolated to a template brain (Oostenveld et al., 2001). We identified brain areas using a template atlas (Lancaster et al., 1997).

RESULTS

Behavioral Data

Participants pressed the key on average 70 msec before the end of the turn in the PRED condition, but for the

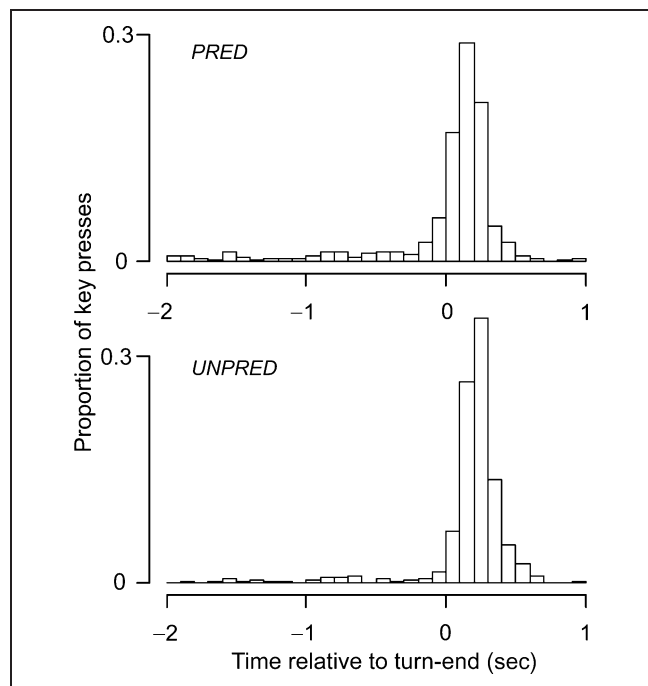


Figure 2. Histogram of RTs in the PRED and UNPRED condition. RTs were measured as the temporal offset between the key presses and the end of turns. When the key was pressed before the turn ended, the RT is negative, when it was pressed after the turn-end, the RT is positive. The percentage of trials is shown on the y axis, and time in seconds before and after the key press (key press is at 0) is shown at the x axis. The bars show the percentage of trials that falls into a 100-msec time bin. Most of the key presses fall into the 100–200 msec bias bin in the PRED and into the 200–300 msec bin in the UNPRED condition. (Outlier responses smaller than –2 sec and larger than 1 sec are not shown.)

UNPRED condition they pressed the key on average 139 msec after the turn-end (see Figure 2). Figure 2 shows that there is a long negative tail in the distribution of the key presses relative to the turn-end. Note, however, that the very early responses (1000 msec) before the turn-end, which might be considered premature, occurred only in a small percentage of the cases (5.3%). Moreover, all responses occurred after turn onset, and so, even in the case of very early responses, participants probably tried to predict the turn-end. The Experimental Condition showed a significant effect ($F = 35.388$, $p = 0$), but not the Order of the presentation of the stimuli ($F = 1.867$, $p = .186$), and there was no significant interaction between Condition and the Order of stimulus presentation ($F = 0.255$, $p = .619$). Thus, as expected, those turns whose actual final words could be predicted in a prior gating study proved more predictable in an online RT task.

EEG Data

Time–Frequency Analysis of Power Changes

The EEG signal showed a significant ($p = .033$) difference between the two conditions in the lower beta

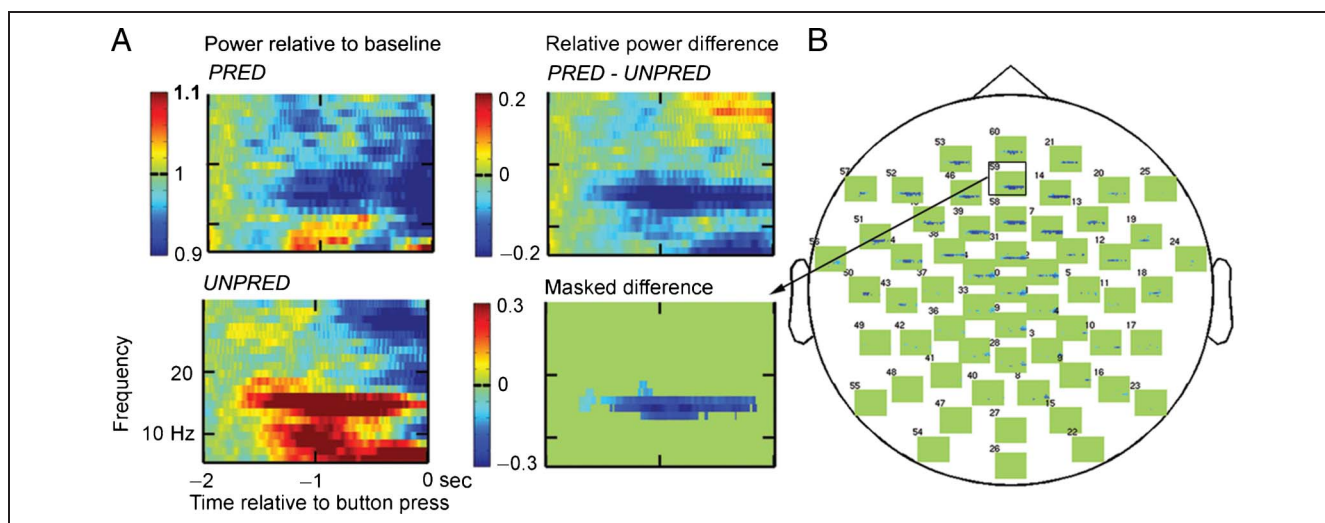


Figure 3. TFRs of EEG power changes. (A) TFRs at electrode 59. The color bars show the power values relative to baseline (from -2 sec until -1.7 sec). The first column shows the TFRs for each condition (PRED, UNPRED). The upper figure in the second column shows the relative power difference between conditions (PRED $-$ UNPRED). The lower figure shows the significant power differences (MASKED). (B) Schematic head with statistically masked TFRs at the corresponding electrode positions. The rectangle shows electrode 59.

frequency range (11–18.5 Hz), starting around 1800 msec and lasting all the way up until the key press (Figure 3). A larger power decrease can be observed in the PRED condition. This difference was most prominent over midfrontal areas (Figure 4).

Interestingly, the time course of beta power showed a different pattern over motor versus midfrontal areas for the two conditions (Figure 5). Although beta power decreases were small (PRED) or nonexistent (UNPRED) over the motor cortex, over midfrontal areas a strong

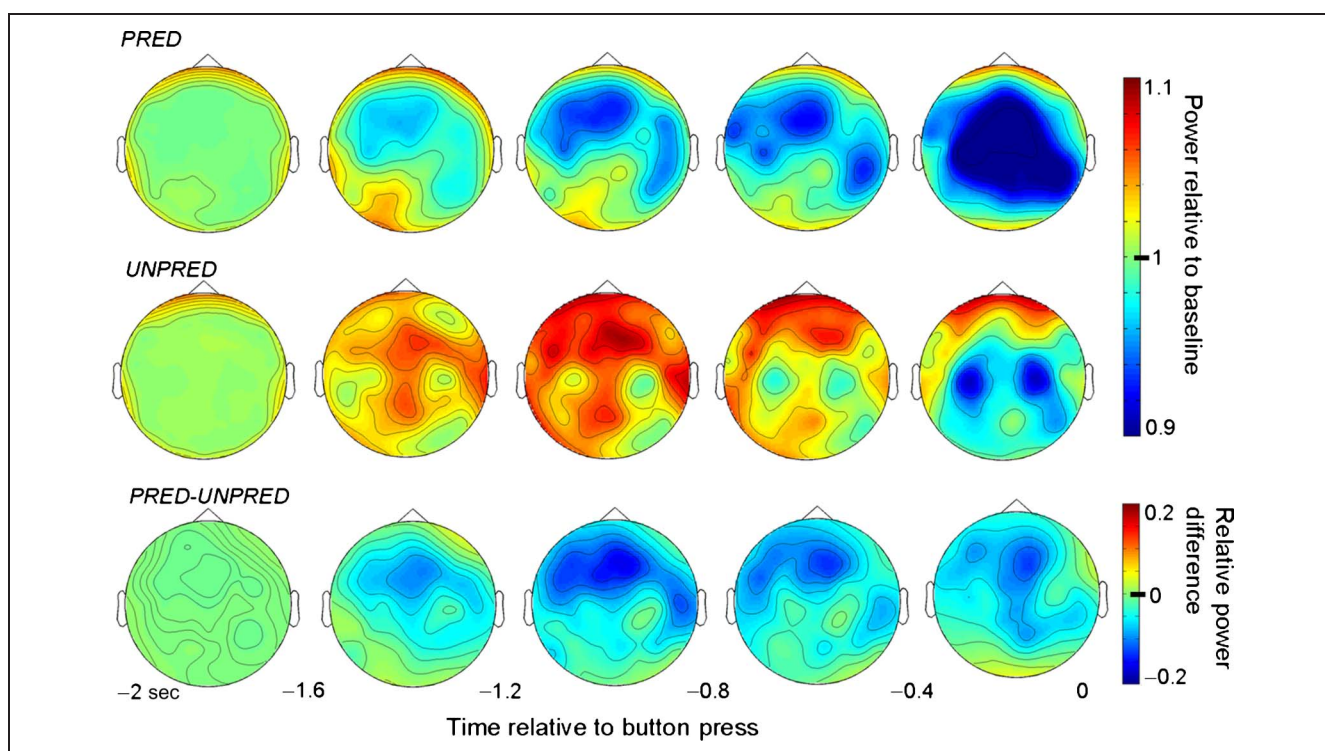


Figure 4. Topographical distribution of beta band power (11–18.5 Hz) in subsequent bins of 400 msec. The upper and middle rows show beta power relative to baseline in the PRED and UNPRED conditions, respectively. The lower row shows the differences in power between the two conditions.

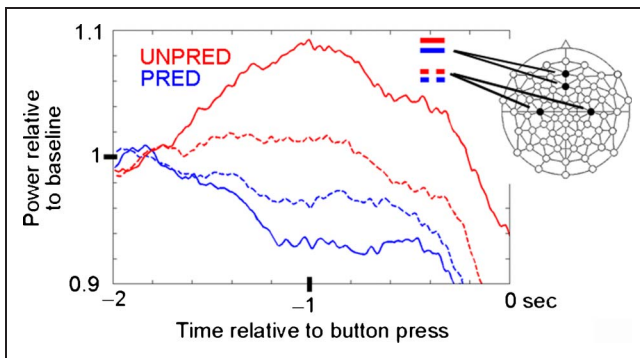


Figure 5. Power values in the beta frequency range (11–18.5 Hz). Beta power is averaged across pairs of midfrontal (electrodes 58, 59, straight lines) and lateral central (electrodes 37, 5, dotted lines) electrodes. Time is on the x axis, in seconds, before the key press (at 0); relative power values on the y axis. Power is shown in red in the UNPRED condition and in blue in the PRED condition.

decrease was associated with the PRED condition and a strong increase with the UNPRED condition.

Source Reconstruction of the Power Changes

The source locations of the relative power changes were estimated with a beamformer technique and compared in both conditions for two time windows: 1.2–0.7 sec (the interval in which the beta power difference between the conditions was largest) versus 2–1.5 sec before the key press (the baseline interval). The areas that show a difference in source strength between the two conditions are shown in Figure 6B. The relative power decrease in the PRED condition, compared with the UNPRED condition, was estimated to originate from frontal and left parietal areas (Figure 6A). Frontally, a source is located in

the anterior part of the left and right superior frontal gyrus that extends into the left middle and inferior frontal gyrus (BA 11 and BA 47) and to the left and right ACC. The parietal source is located in the left inferior parietal lobule (IPL, BA 39, and BA 40) and in the posterior part of the left middle and inferior temporal gyrus (BA 37; Figure 6B).

DISCUSSION

Given the latency of the speech production process, if speakers are going to come in on time, they must begin the production process well before the end of the other's turn—and to time that, would-be speakers must predict the end point of the incoming turn. As described, we used a prior gating task to sort turns into two kinds, relatively predictable or unpredictable, on the basis of whether their last words could be exactly predicted (Figure 1). In the main experiment, as expected, participants more accurately predicted the turns that were more easily completed in the gating study. The corresponding EEG signal showed that predictable turns, compared with less predictable turns, were accompanied by a power decrease in the beta band, which is estimated to originate from left medial frontal, left superior frontal, left inferior parietal, and left posterior temporal brain areas.

The behavioral measure, the timing of key presses, is in line with the hypothesis that turn-end estimation matches the ability of participants to predict the actual last words of many turns starting from c. 600 msec before turn-ending as shown in our prior gating study. It suggests that turn-end anticipation is built on predicting the actual forthcoming words. It would also allow just enough time for the production system to produce the first word of the response, given a 600-msec production

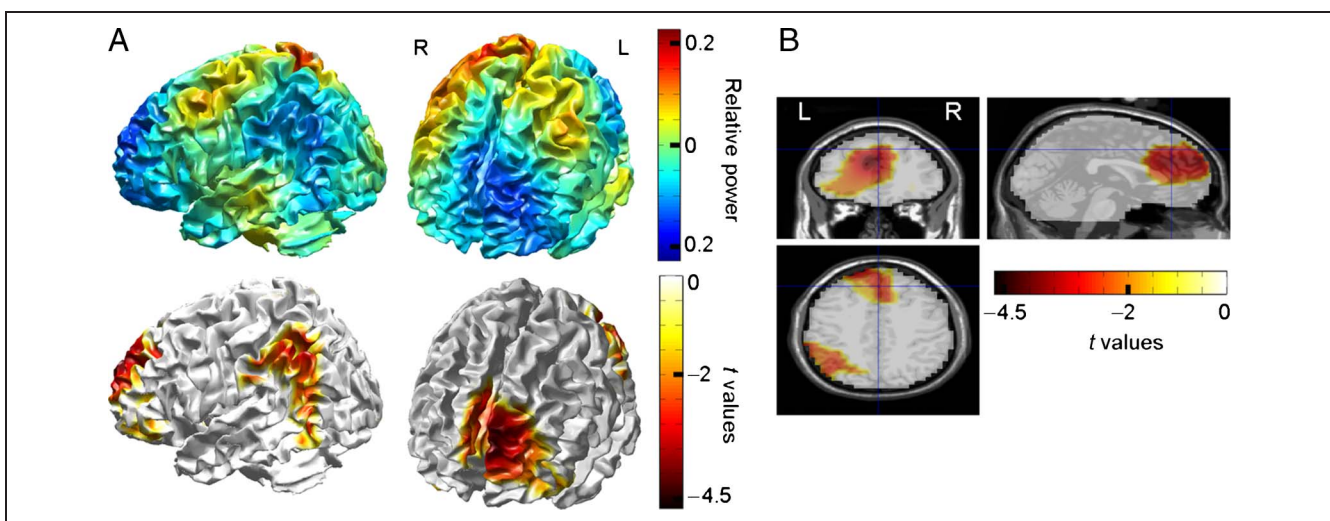


Figure 6. Source reconstruction of the lower beta effect. (A) Relative power changes (first row) and t values of the source points (second row) interpolated onto a 3-D template brain surface. (B) t values of the source points interpolated onto a template MRI. Slices are shown at $x = 0$, $y = 39$, $z = 42$ MNI coordinates.

latency and an average turn gap of 200 msec. It would already rule out any role for late cues of turn-ending, such as turn-final prosodic cues.

However, the EEG signal shows a much earlier anticipation of turn-ending. We found beta power differences during the anticipation of predictable (vs. unpredictable) turn-ends already 1.8 sec before the button press. Allowing for the time smoothing inherent to the time–frequency analysis (± 400 msec) and the latencies of key pressing (around +140 msec in the UNPRED condition), the observed differences in the EEG signal between conditions occurred on average at least 1250 msec before turn-ending. This means that people were anticipating the turn-ends in the predictable condition at least more than five words before the turn-end on average (see average syllable and word duration in Methods, Stimulus Material).

Turning to the interpretation of the EEG signals, it is well established that power decreases in the beta band can be observed during preparation for a movement above the sensorimotor areas (Alegre et al., 2006; Rektor, Sochůrková, & Bočková, 2006; Pfurtscheller & Lopes da Silva, 1999; Pfurtscheller & Aranibar, 1977; Jasper & Penfield, 1949). Furthermore, beta power decreases have been associated with the temporal predictability of stimulus occurrence (Alegre et al., 2003, 2006). Beta power and coherence changes have also been suggested to be related to syntactic and semantic processing (Wang et al., 2012; Weiss & Mueller, 2012; Bastiaansen et al., 2010; Weiss et al., 2005).

The key press results show that the difference in entropy (confidence in predictions) correlated with turn-end predictions. More confident responses in the PRED condition could have resulted in differences in motor preparation. However, we found beta power decreases above the motor areas in both conditions as expected, but there were no differences across conditions above the motor areas. This indicates that a relative decrease in beta power in both conditions reflects motor preparation associated with key pressing and that motor preparation processes are not different across the two conditions. Above frontal areas, however, there was a large beta power decrease during the predictable turns and a large increase during the unpredictable turns. These results show that neuronal correlates related to the anticipation of turn-endings are distinct from those related to the anticipation of action.

The observed beta-band effects in the condition comparison might be thought to be a result of the differences in the predictability of the turn's content itself. However, empirical evidence shows that lexical predictability induces changes in gamma-band power, not beta-band power (Rommers, Dijkstra, & Bastiaansen, 2013; Wang et al., 2012). Another possibility is that the turns in the unpredictable condition are less coherent, which could lead to differences in the oscillatory activity. Bastiaansen et al.'s (2010) study shows that beta power increases throughout a correct sentence (correct and also coherent

condition) compared with words presented in a random order (a less coherent condition). Therefore, if coherence plays a role in the observed EEG effect across conditions, we would expect to find higher beta power during predictable turns compared with unpredictable ones. However, instead of an increase we found beta decrease in the predictable condition. Therefore, the observed differences in the beta power across conditions most probably relate to the experimental manipulation, namely to turn-end predictions and not to differences in coherence.

In our study, we localized most of the beta power decrease to the left superior and middle frontal areas and the ACC. This activation extended until the left middle frontal gyrus and the left inferior frontal gyrus (BA 47). Another large locus of activation was found in the left IPL and in the left (posterior) middle and inferior temporal gyrus. During turn-end anticipation, the temporal estimation is based on the incoming linguistic information, which offers a different basis for prediction than other studies that have used time estimation tasks (see, e.g., Bastiaansen & Brunia, 2001). It is interesting therefore to try and delineate the functional brain network that subserves turn-end anticipation. The pFC and ACC are well known for being involved in anticipation and in time processing (Bubic, von Cramon, & Schubotz, 2010; Aarts, Roelofs, & van Turenout, 2008; Lewis & Miall, 2003; Macar et al., 2002; Fuster, 2001), constituting a network of attentional control (MacLeod & MacDonald, 2000), verbal action planning (Hagoort, 2005), and speech act comprehension (Egorova, Pulvermüller, & Shtyrov, 2014). A left frontoparietal network involving the left intraparietal sulcus and left inferior premotor cortex has been suggested to be recruited particularly for directing attention toward a particular moment in time (Coull & Nobre, 1998). The IPL has been associated with the integration of incoming information into current syntactic and contextual frames (Lau, Phillips, & Poeppel, 2008). BA 47 has been involved in semantic unification, for example, in the integration of word meaning into the unfolding discourse context (Hagoort, 2005). The left posterior middle temporal gyrus and inferior temporal gyrus have been related to the activation and storage of lexical representations (Lau et al., 2008; Hagoort, 2005; Pulvermüller, 2005). Taking all these findings together, our present observation that the frontal, left parietal, and temporal areas desynchronize in the same frequency range as the motor cortical areas suggest a close coordination between brain areas subserving language comprehension processes, more general anticipatory behavior, and the motor network, during the execution of the experimental task.

The EEG data therefore show a clear, interpretable signal of early anticipation of turn-ending, based on the involvement of areas associated with syntactic, semantic and temporal processing. Although our experiment does not directly address the issue whether anticipation of turn-ends are based on prosodic or lexical/syntactic information, we selected our stimuli such that there was a

difference between the predictability of the turn's lexical content between the different conditions from at least 600 msec before the turn-end. Prosodic cues are assumed to give information to listeners just before turn-ends on whether (1) the turn is ending soon or (2) it is not ending yet. In contrast, syntactic and lexical information are good candidates to give more fine-grained temporal information about when the turn is going to end. On the basis of our results, it seems likely that this information is available much earlier than turn-yielding prosodic cues. Syntax provides an architectural framework into which lexical material must slot, and as mentioned earlier, it provides ever narrowing completion probabilities as the incoming sentence is parsed (a process that seems to be reflected in our EEG measure toward the end of the turn), until a point where the precise final words can be anticipated (a point that seems to be reflected in our behavioral measure). Therefore, it is likely that turn-ends can be anticipated early based on lexical-syntactic information. These findings fit well into a Bayesian model of language processing, where the incoming linguistic material provides constant updating of expectations and narrowing likelihoods for alternative continuations (Friston, 2010; Chater & Manning, 2006; Christiansen & Chater, 2001). However, follow-up studies are needed to further narrow down the possible range of interpretations of the effects observed in this study.

This study has probed a little understood domain, namely how language is actually processed in its prime natural habitat, conversation. It suggests that, underlying the rapid turn-exchange system, anticipatory processing is required relatively early in the comprehension of a turn to achieve the apparently effortless coordination that is so commonly observed.

Acknowledgments

This research was supported by the Max Planck Institute for Psycholinguistics. The original data are archived at the Max Planck Institute for Psycholinguistics (hdl:1839/00-0000-0000-0017-3713-6). The last author was supported by ERC Advanced Grant 269484 "INTERACT."

Reprint requests should be sent to Lilla Magyari, Language and Cognition Department, Max Planck Institute for Psycholinguistics, P.O. Box 310, 6500 AH Nijmegen, The Netherlands, or via e-mail: lilla.magyari@mpi.nl.

REFERENCES

- Aarts, E., Roelofs, A., & van Turennout, M. (2008). Anticipatory activity in anterior cingulate cortex can be independent of conflict and error likelihood. *Journal of Neuroscience*, *28*, 4671–4678.
- Alegre, M., Gurtubay, I. G., Labarga, A., Iriarte, J., Malanda, A., & Artieda, J. (2003). Alpha and beta oscillatory changes during stimulus-induced movement paradigms: Effect of stimulus predictability. *NeuroReport*, *14*, 381–385.
- Alegre, M., Imirizaldu, L., Valencia, M., Iriarte, J., Arcocha, J., & Artieda, J. (2006). Alpha and beta changes in cortical oscillatory activity in a go/no go randomly-delayed-response choice reaction time paradigm. *Clinical Neurophysiology*, *117*, 16–25.
- Altmann, G. T. M., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, *57*, 502–518.
- Bastiaansen, M. C. M., & Brunia, C. H. M. (2001). Anticipatory attention: An event-related desynchronization approach. *International Journal of Psychophysiology*, *43*, 91–107.
- Bastiaansen, M. C. M., Magyari, L., & Hagoort, P. (2010). Syntactic unification operations are reflected in oscillatory dynamics during on-line sentence comprehension. *Journal of Cognitive Neuroscience*, *22*, 1333–1347.
- Bates, M. (1995). Models of natural language processing. *Proceedings of the National Academy of Sciences, U.S.A.*, *92*, 9977–9982.
- Beattie, G., Cutler, A., & Pearson, M. (1982). Why is Mrs. Thatcher interrupted so often? *Nature*, *300*, 744–747.
- Bubic, A., von Cramon, D. Y., & Schubotz, R. I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, *4*, 1–15.
- Camazine, S., Deneubourg, J.-L., Franks, N. R., Sneyd, J., Theraulaz, G., & Bonabeau, E. (2001). *Self-organization in biological systems*. Princeton: Princeton University Press.
- Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, *10*, 335–344.
- Christiansen, M. H., & Chater, N. (2001). Connectionist psycholinguistics: Capturing the empirical data. *Trends in Cognitive Sciences*, *5*, 82–88.
- Coull, J. T., & Nobre, A. C. (1998). Where and when to pay attention: The neural systems for directing attention to spatial locations and to time intervals as revealed by both PET and fMRI. *The Journal of Neuroscience*, *18*, 7426–7435.
- Cutler, A., & Pearson, M. (1985). On the analysis of prosodic turn-taking cues. In C. Johns-Lewis (Ed.), *Intonation in discourse* (pp. 139–155). London: Croom Helm.
- De Ruiter, J. P., Mitterer, H., & Enfield, N. J. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, *82*, 515–535.
- DeLong, K. A., Urbach, T. P., Groppe, D. M., & Kutas, M. (2011). Overlapping dual ERP responses to low cloze probability sentence continuations. *Psychophysiology*, *48*, 1203–1207.
- DeLong, K. A., Urbach, T. P., & Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, *8*, 1117–1121.
- Duncan, S. (1974). On the structure of speaker-auditor interaction during speaking turns. *Language in Society*, *3*, 161–180.
- Egorova, N., Pulvermüller, F., & Shtyrov, Y. (2014). Neural dynamics of speech act comprehension: An MEG study of naming and requesting. *Brain Topography*, *27*, 375–392.
- Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138.
- Fuster, J. M. (2001). The prefrontal cortex—An update: Time is of the essence. *Neuron*, *30*, 319–333.
- Gross, J., Kujala, J., Hamalainen, M., Timmermann, L., Schnitzler, A., & Salmelin, R. (2001). Dynamic imaging of coherent sources: Studying neural interactions in the human brain. *Proceedings of the National Academy of Sciences, U.S.A.*, *98*, 694–699.
- Hagoort, P. (2005). On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences*, *9*, 416–423.
- Indefrey, P., & Levelt, W. J. M. (2004). The spatial and temporal signatures of word production components. *Cognition*, *92*, 101–144.

- Jasper, H., & Penfield, W. (1949). Electrocorticograms in man: Effect of voluntary movement upon the electrical activity of the precentral gyrus. *Archiv Für Psychiatrie Und Zeitschrift Neurologie*, *183*, 163–174.
- Jescheniak, J. D., Schriefers, H., & Hantsch, A. (2003). Utterance format affects phonological priming in the picture-word task: Implications for models of phonological encoding in speech production. *Journal of Experimental Psychology: Human Perception and Performance*, *29*, 441–454.
- Kamide, Y., Altmann, G. T. M., & Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, *49*, 133–156.
- Kendon, A. (1967). Some functions of gaze-direction in social interaction. *Acta Psychologica*, *26*, 22–63.
- Lancaster, J. L., Rainey, L. H., Summerlin, J. L., Freitas, C. S., Fox, P. T., Evans, A. C., et al. (1997). Automated labeling of the human brain: A preliminary report on the development and evaluation of a forward-transform method. *Human Brain Mapping*, *5*, 238–242.
- Lau, E. F., Phillips, C., & Poeppel, D. (2008). A cortical network for semantics: (De)constructing the N400. *Nature Reviews Neuroscience*, *9*, 920–933.
- Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lewis, P. A., & Miall, R. C. (2003). Brain activation patterns during measurement of sub- and supra-second intervals. *Neuropsychologia*, *41*, 1583–1592.
- Local, J., Kelly, J., & Wells, B. (1986). Towards a phonology of conversation: Turn-taking in Tyneside English. *Journal of Linguistics*, *22*, 411–437.
- Local, J., Wells, B., & Seba, M. (1985). Phonetic aspects of turn delimitation in London Jamaican. *Journal of Pragmatics*, *9*, 309–330.
- Macar, F., Lejeune, H., Bonnet, M., Ferrara, A., Pouthas, V., Vidal, F., et al. (2002). Activation of the supplementary motor area and of attentional networks during temporal processing. *Experimental Brain Research*, *142*, 475–485.
- MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. *Trends in Cognitive Sciences*, *4*, 383–391.
- Magyari, L., & De Ruiter, J. P. (2012). Prediction of turn-ends based on anticipation of upcoming words. *Frontiers in Psychology*, *3*, 376.
- Manning, C. D., & Schütze, C. T. (1999). *Foundations of statistical natural language processing*. Cambridge, MA: MIT Press.
- Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*, 177–190.
- Mehl, M. R., Vazire, S., Ramirez-Esparza, N., Slachter, R. B., & Pennebaker, J. W. (2007). Are women really more talkative than men? *Science*, *317*, 82.
- Mitra, P. P., & Pesaran, B. (1999). Analysis of dynamic brain imaging data. *Biophysical Journal*, *76*, 691–708.
- Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J.-M. (2011). FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Computational Intelligence and Neuroscience*, *2011*, 1–9.
- Oostenveld, R., Praamstra, P., Stegeman, D., & van Oosterom, A. (2001). Overlap of attention and movement-related activity in lateralized event-related brain potentials. *Clinical Neurophysiology*, *112*, 477–484.
- Pfurtscheller, G., & Aranibar, A. (1977). Event-related cortical desynchronization detected by power measurements of scalp EEG. *Electroencephalography and Clinical Neurophysiology*, *42*, 817–826.
- Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology*, *110*, 1842–1857.
- Pulvermüller, F. (2005). Brain mechanisms linking language and action. *Nature Reviews Neuroscience*, *6*, 576–582.
- Rektor, I., Sochůrková, D., & Bočková, M. (2006). Intracerebral ERD/ERS in voluntary movement and in cognitive visuomotor task. In C. Neuper & W. Klimesch (Eds.), *Event-related dynamics of brain oscillations* (Vol. 159, pp. 311–330). Amsterdam: Elsevier.
- Rommers, J., Dijkstra, T., & Bastiaansen, M. C. M. (2013). Context-dependent semantic processing in the human brain: Evidence from idiom comprehension. *Journal of Cognitive Neuroscience*, *25*, 762–776.
- Rossano, F., Brown, P., & Levinson, S. C. (2009). Gaze, questioning and culture. In J. Sidnell (Ed.), *Conversation analysis: Comparative perspectives* (pp. 187–249). Cambridge: Cambridge University Press.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, *50*, 696–735.
- Schegloff, E. A. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 52–133). Cambridge: Cambridge University Press.
- Schnurr, T. T., Costa, A., & Caramazza, A. (2006). Planning at the phonological level during sentence production. *Journal of Psycholinguistics Research*, *35*, 189–213.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and mind moving together. *Trends in Cognitive Sciences*, *10*, 70–76.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *76*, 379–423.
- Stivers, T., Enfield, N. J., Brown, P., Englert, C., Hayashi, M., Heinemann, T., et al. (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences, U.S.A.*, *106*, 10587–10592.
- Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., & Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 443–467.
- Wang, L., Zhu, Z., & Bastiaansen, M. C. M. (2012). Integration or predictability? A further specification of the functional role of gamma oscillations in language comprehension. *Frontiers in Psychology*, *3*, 187.
- Weiss, S., & Mueller, H. M. (2012). “Too many betas do not spoil the broth”: The role of beta brain oscillations in language processing. *Frontiers in Psychology*, *3*, 201.
- Weiss, S., Mueller, H. M., Schack, B., King, J. W., Kutas, M., & Rappelsberger, P. (2005). Increased neuronal communication accompanying sentence comprehension. *International Journal of Psychophysiology*, *57*, 129–141.
- Wicha, N. Y. Y., Moreno, E. M., & Kutas, M. (2004). Anticipating words and their gender: An event-related brain potentials study of semantic integration, gender expectancy and gender agreement in Spanish sentence reading. *Journal of Cognitive Neuroscience*, *16*, 1272–1288.
- Wilson, M., & Wilson, T. P. (2005). An oscillator model of the timing of turn-taking. *Psychonomic Bulletin and Review*, *12*, 957–968.