Neuroscience Letters 609 (2015) 147–151

Contents lists available at ScienceDirect

Neuroscience Letters

journal homepage: www.elsevier.com/locate/neulet

Research paper

Turn-taking: From perception to speech preparation

Hendrik Wesselmeier*, Horst M. Müller

Experimental Neurolinguistics Group and Collaborative Research Center "Alignment in Communication" (SFB 673), Bielefeld University, Bielefeld, Germany

HIGHLIGHTS

- EEG language production study of a quasi-dialog.
- Readiness potential analysis of articulation start based on turn-anticipation.
- Syntactic stimuli structure modulates readiness potential onset intervals.
- Cognitive load modulates response time but not readiness potential.
- Readiness potential is not related to speech intention but to speech planning.

ARTICLE INFO

Article history:
Received 3 June 2015
Received in revised form 5 October 2015
Accepted 12 October 2015

Keywords:
Event-related potential
Readiness potential
Spoken language
Articulation preparation
Turn-taking
Human

ABSTRACT

We investigated the preparation of a spoken answer response to interrogative sentences by measuring response time (RT) and the response-related readiness potential (RP). By comparing the RT and RP results we aimed to identify whether the RP-onset is more related to the actual speech preparation process or the pure intention to speak after turn-anticipation. Additionally, we investigated if the RP-onset can be influenced by the syntactic structure (one or two completion points). Therefore, the EEG data were sorted based on two variables: the cognitive load required for the response and the syntactic structure of the stimulus questions. The results of the response utterance preparation associated event-related potential (ERP) and the RT suggest that the RP-onset is more related to the actual speech preparation process rather than the pure intention to speak after turn-anticipation. However, the RP-onset can be influenced by the syntactic structure of the question leading to an early response preparation.

1. Introduction

Turn-taking is the change of an interlocutor from listener to speaker in a natural spoken dialog. It has been studied for several decades, and a good understanding of its properties has been achieved. An universal property of a dialog is that it proceeds efficiently, and most turn-taking occurs with temporal precision without overlaps and almost no obvious gaps [1,2]. However, when an interlocutor produces a response utterance, it takes approximately one second between the appearance of the mental concept and the actual motoric articulation [3]. This process leads to the assumption that the following speaker anticipates the turn-end of the interlocutor to prepare their response. Nevertheless, it is unclear at what point a passive listener becomes an active interlocutor and starts to prepare their response, it is also uncertain if syntactic information influences turn-anticipation. Most recent EEG studies investigating language production focused on single word production [4] rather than utterance production in the form of a sentence or even a turn consisting of an illocutonary act. To our knowledge, only one study has investigated the conceptual planning of complex utterances in overt speech, elicited by presenting two pictures [5]. Each picture was associated with an action (e.g., book = reading) whereas a cue indicated if the associated action should be described in a chronological or reversed order. In that EEG experiment, they found a P300-related effect caused by the task complexity in linearization. In most EEG experiments investigating turn-taking, participants were prompted to indicate the anticipated turn-end with a button-press [6] or a simple verbal cue "ja" (yes) [7]. In this experiment, participants were prompted to utter a complete answer with one or more sentences.

The EEG readiness potential (RP, Bereitschaftspotential) is an event-related potential (ERP) in the form of a negative deflation which is related to selective motor preparation [8]. Previous EEG experiments comparing limb movements-RP to speech-RP [9] and experiments comparing the RP of vocal onset to RP of a button-press

* Corresponding author at: Experimental Neurolinguistics Group – SFB 673, Bielefeld University, P.O. Box 100131, 33501 Bielefeld, Germany.
E-mail address: hendrik.wesselmeier@uni-bielefeld.de (H. Wesselmeier).

http://dx.doi.org/10.1016/j.neulet.2015.10.033

during anticipation [7] have shown that the response-locked RP of vocal onset is similar to RP of a button-press. This evidence means that a voice-related RP-onset can be used for motor preparation in the same manner as it is used for button-press and is not affected by breathing in or other preparation factors for verbalization. In an earlier EEG experiment [10] on anticipation, the button-press-related RP in turn-end detection was influenced by syntactic or semantic violations in the stimulus sentence: participants had to press a button exactly when the sentence ended, and some sentences contained a syntactic or semantic violation. The results showed that both types of violations generated a shorter RP compared to intact sentences, whereas the RT was similar. Nonetheless, in responses after a syntactic violation, an argument for a shorter RP was that it could indirectly influence turn-anticipation by disturbing semantic integration. Additionally, recent behavioral studies [11] manipulated the presence and absence of syntactic information by low-pass filtering open- and closed-class words in turns; these studies confirm that syntactic information provides anticipation cues, even though semantic information is a more important cue for anticipation.

In this study, we investigated the influence of cognitive load, which has been shown to influence speech planning in turn-taking [12], for response preparation on the RT and RP-onset. Therefore, the stimulus questions were divided into two groups: those that could be answered without delay (low cognitive load) and those that could be answered with a short delay (high cognitive load). Furthermore, we investigated the influence of syntactic structure (one or two completion points) using two different types of stimulus questions with different types of syntactic complexity, whereas semantic processing is unaffected. To accomplish a more natural turn-taking situation, participants were asked to respond with an articulated answer. The aim of this study was to gain insight into the temporal aspects of the transition from speech perception to speech production and its preparation in turn-taking. Therefore, at first, we tested if the RP-onset was more related to the general intention to speak or more time related to the actual speech planning process after the decision of how to respond has already been made. Then, we tested if the syntactic structure of the stimulus questions influenced response preparation. The aim was to see whether syntax influences turn-anticipation rather than only indirectly influencing turn-anticipation by disturbing semantic integration. To test these hypotheses, we ran an EEG experiment in which participants were acoustically presented with questions to which the participants had to respond with a brief answer.

2. Material and methods

2.1. Participants

Thirty students (17 f, 20–35 years, mean 24.5) from Bielefeld University participated in our experiment, which lasted about one hour. All participants were native German speakers and right-handed with a lateralization quotient of 88.9 according to the Edinburgh Handedness Inventory [13]. According to their own accounts, they did not suffer from either auditory or motor restrictions or diseases that could have influenced the results. Written consent was obtained from all participants, and the study was approved by the Ethics Committee of Münster University.

2.2. Stimuli

The acoustically presented stimuli consisted of 25 interrogative clauses that varied from 1300 ms to 6643 ms in length with a mean of 3989 ms and a SD of 1421 ms. Questions with two syntactic completion points were significantly longer compared to questions with only one syntactic completion point, $F(1,23) = 11.65$, $p < .005$. For the 12 sentences with a possible completion point within the sentence, the time interval from the first possible syntactic completion point to the utterance end ranged from 1194 ms to 2637 ms (mean = 2048 ms, SD = 508 ms). The duration of the stimuli between the two groups selected based on the RT were not significantly different at the .05 level. All sentences in the experiment were spoken at a speed of 400–450 syllables per minute by a professional female speaker with natural intonation and were recorded in a sound studio.

Stimulus examples:
The "#" indicates the first possible syntactic completion point.
Question, to be answered without delay:
Finden Sie die Mietpreise in Bielefeld angemessen?
(Do you consider the rental prices in Bielefeld as appropriate?)
Question, to be answered with a delay:
Sehen Sie im neuen Bachelor/Master-System eine Verbesserung?
(Do you consider the new Bachelor/Master-system an improvement?)
Question, with the only possible syntactic completion point at the end of the sentence:
Müssen Sie neben dem Studium arbeiten gehen?#
(Do you need to hold a job in addition to your study?#)
Question, with a first possible syntactic completion point within the sentence:
Wohnen Sie alleine# oder in einer WG?
(Do you live alone# or in a shared flat?)

2.3. Rating of the stimuli

A four-point scale for the 25 stimulus sentences was used to verify that the measured RT difference depends on the cognitive load necessary to answer the question rather than the complexity of the sentence and the difficulty of understanding. To distinguish a question that could be answered quickly or with a short delay to think, 58 students (55 f, mean age 23.7, SD = 4.3) who were all native German speakers participated in this rating.

2.4. Procedure

Our experiment was conducted in an electromagnetically shielded and sound-proof booth. Each trial started with a fixation cross presented in the middle of the screen. After the fixation cross appeared, the spoken stimulus sentence started after a random inter-trial interval (range 1000–2500 ms) so that the participants could not anticipate the sentence onset time after several trials. The fixation cross was continuously shown until 6000 ms after the spoken sentence ended. Participants were instructed to give a quick and short answer to the interrogative clause. The participants were informed that the responses would not be recorded or judged. After a short practice block with five different sentences matched in complexity and length, all participants became comfortable with the task. The mean stimulus intensity ranged between 55 and 60 dB, which matches a normal face-to-face conversation.

2.5. EEG recording

EEG was continuously recorded from 32 active scalp electrodes (ActiCap, Brain Products) placed at locations based on the International 10/20 system [14] with the reference at FCz. Signals were sampled at 1000 Hz, amplified with a bandpass of 0.16–80 Hz and a 50 Hz notch filter by amplifiers (QuickAmp, Brain Products) and recorded with BrainVision Recorder software (Version 1.20). The impedance remained below 5 kΩ for all channels prior to recording.