Biological Psychology 92 (2013) 410–416

Contents lists available at SciVerse ScienceDirect

Biological Psychology

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

Prediction errors in self- and externally-generated deviants

Franziska Knolle[a,*], Erich Schröger[b], Sonja A. Kotz[a]

[a] Research Group "Subcortical Contributions to Comprehension", Dept of Neuropsychology, Max Planck Institute of Human Cognition and Brain Science, Leipzig, Germany
[b] Cognitive incl. Biological Psychology, Institute of Psychology, University of Leipzig, Leipzig, Germany

ARTICLE INFO

Article history:
Received 30 March 2012
Accepted 22 November 2012
Available online 14 December 2012

Keywords:
Prediction
Forward model
Prediction error
Self-generated and externally-generated sounds

ABSTRACT

Sounds generated by one's own action elicit attenuated brain responses compared to brain responses to identical sounds that are externally-generated. The present study tested whether the suppression effect indexed by the N1- and P2-components of the event-related potential (ERP) is larger when self-generated sounds are correctly predicted than when they are not. Furthermore, sounds violating a prediction lead to a particular prediction error signal (i.e., N2b, P3a). Thus, we tested whether these error signals increase for self-generated sounds (i.e., enhanced N2b, P3a). We compared ERPs elicited by self- and externally-generated sounds that were of frequent standard and of infrequent deviant pitch. The results confirmed an N1- and P2-suppression effect elicited by self-generated standard sounds. The N1-suppression was smaller in response to self-initiated deviant sounds, indicating the specificity of predictions for self-generated sounds. In addition, an enhancement of N2b and P3a for self-generated deviants revealed the saliency of prediction error signals.

1. Introduction

The notion of forward predictions addresses a relevant aspect of everyday life. As sensations we perceive are often self-generated, it is necessary to distinguish between the input produced by one's own actions and the input that is generated by external sources. The internal forward model (Blakemore et al., 1998a,b; Wolpert, 1997) that provides a theoretical framework for forward predictions suggests that if an action is self-produced, an efference copy of the motor command is generated to predict the sensory consequences of an action and to prepare the respective cortical areas to receive sensory input (Wolpert et al., 1995). The processing activity directed to the expected sensory input (e.g., self-generated sensations) is consequently reduced, providing more resources for processing externally-generated sensory input (Blakemore et al., 1998b; Chen et al., 2011; Creutzfeldt et al., 1989; Hesse et al., 2010; Martikainen et al., 2005).

Recently, we were able to show that the cerebellum is involved in generating motor-to-auditory predictions when processing self-generated sounds (Knolle et al., 2012). We used an N1-suppression paradigm (Schäfer and Marcus, 1973) that allows comparing self- and externally-generated sounds. The sound types are identical but differ in the way of production: self-generated sounds are generated via a finger tap, whereas externally-generated sounds are played back externally. In this patient study (Knolle et al., 2012), participants with focal cerebellar lesions showed a strongly diminished N1-suppression effect in response to self-generated sounds. This indicates that the cerebellum is involved in generating forward predictions. Testing the same phenomenon in healthy participants, several electroencephalographic (EEG) and magnetoencephalographic (MEG) studies used the N1/M1-suppression paradigm to investigate motor-to-sensory predictions (Baess et al., 2008, 2011; Lange, 2011; Martikainen et al., 2005; Schäfer and Marcus, 1973): These studies consistently report an N1/M1 suppression effect in response to self-generated sounds and confirm that the paradigm reliably tests motor-to-auditory predictions.

However, as we encounter acoustic information in a dynamic environment, acoustic features of an auditory stimulus may be unpredictable even when we produce these auditory sensations ourselves. Thus, it is relevant to ask whether motor-to-auditory predictions are specific. If this is the case, the slightest perturbation of a prediction (i.e., a prediction error) should be salient. For example, at a party where the noise level is so high that you can barely hear yourself speak, you may be unable to control whether you speak loud enough for your conversation partner to understand you well. Although a prediction concerning the sensory consequences of your speech output is generated, this prediction may be imprecise as the monitoring of your voice is altered by the noise around you. If the expectation of the loudness of your voice is violated, a prediction error is created. In such a situation the benefit of a prediction is smaller compared to a conversation in a quiet room where predictions can be specific.

In a previous study comparing self- and externally-generated sounds, Baess et al. (2008) modified the predictability of sounds.

---

* Corresponding author.
E-mail address: fknolle@cbs.mpg.de (F. Knolle).

0301-0511/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.biopsycho.2012.11.017

The authors reported that even when the predictability of self-generated sounds is reduced, the N1 is suppressed when compared to externally-generated sounds. However, the authors did not discuss the impact of prediction specificity in terms of the N1 nor the possibility that such modifications may create a prediction error due to a less specific (or incorrect) prediction. Such a prediction error may not only modulate the N1 suppression effect but also leads to other ERP responses when sounds are self-generated deviants, such as the N2b and P3a that signal deviance detection. Relatedly, the concept of prediction specificity has been tested in speech perception comparing natural and altered auditory feedback (i.e., Behroozmand et al., 2011; Heinks-Maldonado et al., 2005). In these studies, participants listened to their own, unchanged voice as well as to acoustically altered feedback. Thus, they were able to form a specific prediction with respect to their own, unchanged vocalizations. However, these predictions were altered when the auditory feedback was changed (change in voice onset or frequency). In the latter case, when a prediction is less specific, the N1-suppression effect was smaller compared to unchanged feedback. The less specific prediction creates a prediction error. Therefore, we hypothesize that the degree of suppression varies as a function of prediction specificity. In other words, the more specific a prediction is (i.e., a specific representation of the to-be-predicted sensation), the larger the suppression effect becomes. Additionally, the more specific a prediction is, the more efficient the processing of prediction errors is (i.e., even subtle violations are detected). Thus, we speculated that other components, such as the N2b or the P3a, may reflect the detection of attentional orienting towards infrequent, unexpected sounds.

To address the possibility that a prediction error affects the processing of self-generated stimuli, we adapted the standard N1-suppression paradigm by comparing self-generated with externally-generated sounds, of which 30% were altered in pitch. The pitch-altered, self- and externally-generated deviants induce predictions errors. On the one hand, this set-up allowed testing whether specific auditory predictions are generated when a sound is self-produced as seen in the modulation of the N1-suppression effect. Moreover, it allowed investigating whether deviant sounds that induce a prediction error, as a specific prediction is violated, elicit additional ERP components known to respond to deviance detection such as the MMN/N2b and P3a.

Considering previous studies (e.g., Baess et al., 2008; Behroozmand et al., 2011; Knolle et al., 2012), we expected to find an N1-suppression effect in response to self-generated standard sounds when predictions are fulfilled. Furthermore, we expected a reduced N1-suppression effect in response to self-generated deviants because predictions regarding a specific acoustic feature of a sound (i.e., frequency) are not fulfilled, and are therefore less specific. In other words, we expected a modulation of the N1 in the case of a prediction error, which is seen in a minimized N1-suppression effect, due to an increase in the N1 in response to self-generated deviants. Relatedly, we expected to find a P2-reduction in response to self-generated standard and deviant sounds as previous evidence has shown that while the N1-suppression may reflect specific predictions, the P2-reduction could be a conscious reflection of such a prediction, or rather, the conscious detection of a sound that is self-generated (Knolle et al., 2012; Sowman et al., 2012).

However, if specific predictions are generated but violated, a prediction error should be reflected in an ERP response. We hypothesize that prediction errors elicited in response to self-generated deviants sounds should not only elicit a reduced N1-suppression effect (Baess et al., 2008; Behroozmand et al., 2011), but also components indicating that an error was detected such as an N2b and a P3a (for reviews see, e.g., Escera et al., 2000; Folstein and Van Petten, 2008; Näätänen, 1990). Even though the MMN also reflects deviance detection, we will not further consider this component. Firstly, the MMN is assumed to be automatic and does not depend on agency (Rinne et al., 2001), and secondly, it may be difficult to separate it from a temporally overlapping N2b that is expected in response to self-generated deviant sound. In general, the N2b, however, reflects the cognitive detection, or controlled registration, of an infrequent variation of stimulus properties (Horváth et al., 2008; Näätänen et al., 1982; Ritter et al., 1992). Furthermore, we speculated that a self-generated deviant should also produce a P3a response as this component reflects the detection of infrequent and unexpected sounds indicating attentional orienting (Linden, 2005; Polich, 2007; Snyder and Hillyard, 1976). In contrast, we expected that strongly reduced N2b and P3a components will be elicited in response to externally-generated deviants, indicating that self-generated deviants are more salient than externally-generated deviants (Ford et al., 2010). Taken together, the current study aimed to (1) investigate the specificity of predictions when a sound is self-generated and (2) explore how prediction errors are reflected in the ERP.

2. Methods

2.1. Participants

Sixteen volunteers (8 females) participated in the current study. All were right-handed according to the Edinburgh Handedness Inventory (Oldfield, 1971). The mean age was 24.9 years (SD: 1.8 years) and the age range was 23–27 years of age. Participants were recruited from the database of the Max-Planck Institute for Human Cognitive and Brain Sciences in Leipzig. None of the participants reported any neurological dysfunctions. All subjects reported normal or corrected-to-normal visual acuity, and normal hearing. All participants gave their written informed consent and were paid for their participation. The experiment was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the University of Leipzig.

2.2. Experimental conditions and procedure

The study contained two experimental conditions and one control condition (Fig. 1). In the auditory-motor condition (AMC) participants induced finger taps approximately every 2.4 s (see Knolle et al., 2012, for a detailed description). Each tap elicited an immediate presentation of a sinusoidal tone via headphones (maximal delay between the finger tap and the presentation of the sound was 2 ms due to the loading of the sound). 30% of the taps elicited a deviant sound that was either higher or lower in frequency than the standard sound (the allocation of standard and deviant sound was counterbalanced: 680 Hz vs. 1000 Hz). Counterbalancing the allocation of frequencies allowed comparing two different sound types (i.e., standard and deviant) that had identical sound qualities. Otherwise, possible differences in ERPs between the sound types could also result from such physical differences. The acoustic stimulation of AMC was recorded on-line. This recorded sound stream was used as the "external sound sequence" in the auditory-only condition (AOC). Thus, the participants received exactly the same set of stimuli in both experimental conditions, containing standard and deviant sounds. As AOC is a passive listening condition, participants did not produce finger taps, but were simply asked to listen attentively to the auditory stimuli. Lastly, participants took part in a control condition: the motor-only condition (MOC), in which they also performed self-paced finger taps every 2.4 s. However, in contrast to the AMC no auditory feedback was given. This final condition served as a control condition for motor activity in AMC.

The experimental run was preceded by a learning block and a training block. In the learning block, the interval of 2.4 s was externally presented via a metronome and the participants were asked to tap along. After having acquired a basic understanding of the tapping interval, participants performed a training block. The training block included visual feedback to indicate whether a trial was too slow (tapping interval longer than 3 s) or too fast (tapping interval shorter than 1.8 s). The feedback allowed participants to adjust their tapping interval. Thus, the training block was included to ensure that participants had learned to estimate 2.4 s between two successive finger taps without counting. For that reason neither the learning nor the training block contained deviant stimuli. During the experimental run, no feedback was given.

Participants were comfortably seated in an electrically shielded and sound-attenuated experimental chamber. A fixation cross was displayed in the middle of a computer screen. To ensure that the motor activation pattern was similar across participants, they were instructed to change the index finger they were using whenever indicated on the screen. Hence, all participants tapped in equal parts with their left and right hand. The order of tapping hands was randomized across participants; additionally, there was a sign on the screen which indicated the hand they should be using. The participants performed the combination of AMC and AOC twice: during