Available online at www.sciencedirect.com
ScienceDirect
Neuroscience Research 61 (2008) 120–128
Neuroscience Research
www.elsevier.com/locate/neures

Feedback-based error monitoring processes during musical performance: An ERP study

Kentaro Katahira [a,b], Dilshat Abla [a,b], Sayaka Masuda [b], Kazuo Okanoya [a,b,*]

[a] Laboratory for Biolinguistics, RIKEN, Brain Science Institute (BSI), 2-1 Hirosawa, Wako, Saitama 351-0198, Japan
[b] Department of Cognitive and Information Sciences, Faculty of Letters, Chiba University, 1-33 Yayoicho, Inage-ku, Chiba City, Chiba 263-8522, Japan

Received 29 June 2007; accepted 5 February 2008
Available online 14 February 2008

Abstract

Auditory feedback is important in detecting and correcting errors during sound production when a current performance is compared to an intended performance. In the context of vocal production, a forward model, in which a prediction of action consequence (corollary discharge) is created, has been proposed to explain the dampened activity of the auditory cortex while producing self-generated vocal sounds. However, it is unclear how auditory feedback is processed and what neural mechanism underlies the process during other sound production behavior, such as musical performances. We investigated the neural correlates of human auditory feedback-based error detection using event-related potentials (ERPs) recorded during musical performances. Keyboard players of two different skill levels played simple melodies using a musical score. During the performance, the auditory feedback was occasionally altered. Subjects with early and extensive piano training produced a negative ERP component N210, which was absent in non-trained players. When subjects listened to music that deviated from a corresponding score without playing the piece, N210 did not emerge but the imaginary mismatch negativity (iMMN) did. Therefore, N210 may reflect a process of mismatch between the intended auditory image evoked by motor activity, and actual auditory feedback.

Keywords: Auditory feedback; ERP; Error detection; iMMN; Musical performance; N210

1. Introduction

When hearing self-produced sounds, our brains can detect differences between the intended sound and performed output, and use them to alter motor patterns so that proper sound production is eventually achieved (Houde and Jordan, 1998). In the context of speech production, it has been proposed that information relating to motor commands, called efference copy (Sperry, 1950; Jeannerod, 1998), is transformed into a prediction of auditory feedback (corollary discharge) and sent to the auditory cortex (Curio et al., 2000; Houde et al., 2002). This kind of model is termed a ‘‘forward model’’ (Jeannerod, 1998). Due to the cancellation of auditory response by corollary discharge, electric (N1) and magnetic (N1m) responses of the temporal lobe are suppressed (Ford et al., 2001). Furthermore, when auditory feedback is altered during speaking, the N1 reduction becomes weak, which supports the existence of a precise forward model modulating cortical response to self-generated speech (Heinks-Maldonado et al., 2005).

It is still unclear how auditory feedback is processed and what neural substrate is involved in such processes during other sound production behavior, such as musical performance. A number of studies have shown that language and music processes overlap in the brain (e.g., Patel, 2003; Koelsch et al., 2005). Clarifying the similarities and differences between the modules for language and music can help clarify the brain mechanisms involved in both cognitive functions. We investigated a neural correlate of human auditory feedback-based error detection using event-related potentials (ERPs) recorded during musical performances. How humans use auditory feedback during musical performances poses an interesting problem. Musical performance is one sound production behavior, but degrees of

* Corresponding author. Tel.: +81 48 467 7502; fax: +81 48 467 7503.
E-mail address: okanoya@brain.riken.jp (K. Okanoya).

0168-0102/$ – see front matter
doi:10.1016/j.neures.2008.02.001

proficiency differ widely across subjects, unlike speech. Investigating how such differences in proficiency affect strategies for error detection would give new insight into the error monitoring process, in which the brain detects differences between actual sound output and the intended sound image. We hypothesized that musically trained subjects would be able to make predicted auditory representations (i.e., corollary discharge) and detect their own errors by comparing them with actual auditory feedback, while non-trained subjects would lack this ability.

Our study compared the responses of trained and non-trained instrumentalists to congruent and deviant feedback during musical performances. In Experiment 1, we created an experimental situation in which the motor output was basically correct, but auditory feedback was occasionally wrong. For the purpose of controlling the action series of subjects without them having to memorize musical pieces, we inevitably had to use a musical score. Related to this problem, a magnetoencephalography (MEG) study (Yumoto et al., 2005) showed that a well-trained musician can create auditory imagery by merely seeing a musical score, so that they evoked a magnetic field, which Yumoto et al. termed ‘‘imagery mismatch negativity’’ (iMMN). To measure the ERP component that corresponds to the mismatch field (iMMN) observed by Yumoto et al., we conducted Experiment 2, in which subjects only saw the musical score without performing, similar to the experiment of Yumoto et al. (2005). The only difference between Experiment 1 and Experiment 2 was that the former involved actual musical performance. Therefore, the differences between the ERPs in two experiments should reveal how playing the keyboard affects the response to auditory feedback.

2. Materials and methods

2.1. Experiment 1

2.1.1. Subjects

Subjects were 15 healthy, right-handed adults aged 21–30 years. They were divided into two groups based on their piano experience. In the piano-trained group (six males and four females), subjects had received formal piano lessons for at least 3 years before the age of 12, while the subjects in the non-trained group (five males) could read music notation and play the melodies on the keyboard, but had not been formally trained. Subjects were informed of the experimental procedures and gave their written informed consent for participation.

2.1.2. Procedure

Subjects played a simple melody on the keyboard with the right hand while seeing the score and listening to the MIDI piano tones as auditory feedback. They were instructed to play at a slow tempo. The mean(±standard deviation, S.D.) interval of key press was 474.3(±88.63) ms for the piano-trained group and 605.05(±138.59) ms for the non-trained group. When subjects pressed the key on the MIDI keyboard, the MIDI signal was sent to a computer program developed using Visual C++ 6.0 (Microsoft Inc.), which generated the MIDI piano tones presented to subjects through binaural earphones (Fig. 1a). Tone intensity was adjusted for each subject in the 60–90 dB SPL range. The pieces were selected from a workbook of music dictation (Ryuginsha, Tokyo, Japan, 1982) and Chorübungen (Zenon, Tokyo, Japan, 1956). All pieces used in experiments consisted of whole notes, half notes, dotted half notes, quarter notes, dotted quarter notes, and eighth notes. The meters were 3/4 or 4/4 time. Each phrase consisted of 8 bars, and 32 phrases including 582 notes in total were used. Fig. 1b shows the typical shape of the presented melodies. None of the subjects had heard these melodies previously, all of which consisted of tones in C major or A minor to avoid using the black keys. With a probability of 5%,

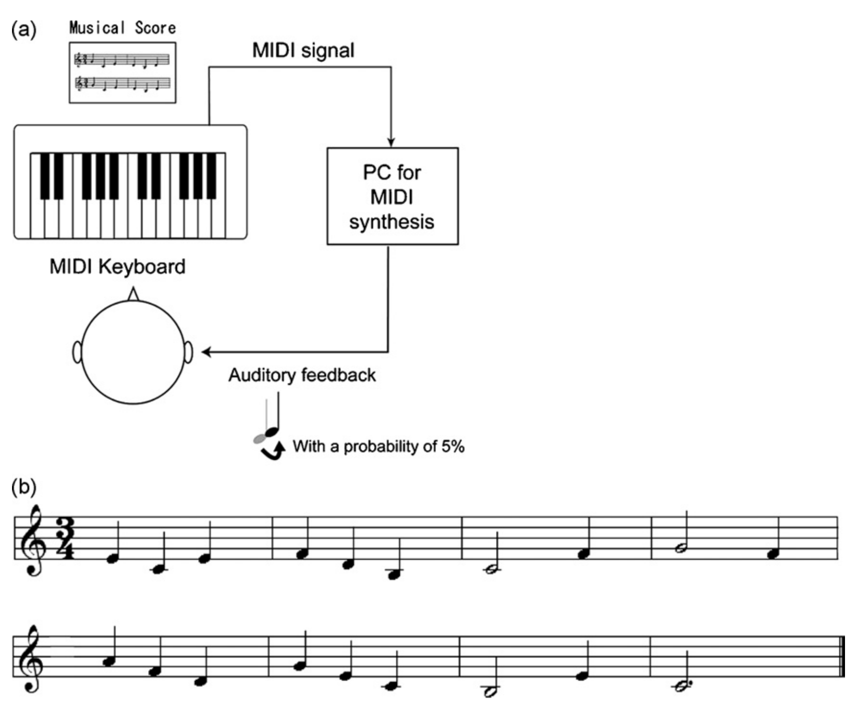

Fig. 1. The experimental setup for Experiment 1. (a) When subjects pressed the key on the MIDI keyboard, the MIDI signal was sent to the computer program, which generated the piano tone given to subjects as auditory feedback through the binaural earphones. The EEG was recorded from a 32-channel cap through an amplifier while the keyboard was being played. (b) A score depicting typical phrases used in the experiment. See text for details.