Hearing Research 319 (2015) 25–31

Contents lists available at ScienceDirect

Hearing Research

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

Research paper

Neural correlates of perceptual grouping effects in the processing of sound omission by musicians and nonmusicians

Kentaro Ono*, Christian F. Altmann, Masao Matsuhashi, Tatsuya Mima, Hidenao Fukuyama

Human Brain Research Center, Graduate School of Medicine, Kyoto University, Kyoto, 606-8507, Japan

ARTICLE INFO

Article history:
Received 12 May 2014
Received in revised form
21 October 2014
Accepted 28 October 2014
Available online 7 November 2014

ABSTRACT

Perceptual grouping is the process of organizing sounds into perceptually meaningful elements. Psychological studies have found that tones presented as a regular frequency or temporal pattern are grouped according to gestalt principles, such as similarity, proximity, and good continuity. Predictive coding theory suggests that this process helps create an internal model for the prediction of sounds in a tone sequence and that an omission-related brain response reflects the violation of this prediction. However, it remains unclear which brain areas are related to this process, especially in paying attention to the stimuli. To clarify this uncertainty, the present study investigated the neural correlates of perceptual grouping effects. Using magnetoencephalography (MEG), we recorded the evoked response fields (ERFs) of amateur musicians and nonmusicians to sound omissions in tone sequences with a regular or random pattern of three different frequencies during an omission detection task. Omissions in the regular sequences were detected faster and evoked greater activity in the left Heschl's gyrus (HG), right postcentral gyrus, and bilateral superior temporal gyrus (STG) than did omissions in the irregular sequences. Additionally, an interaction between musical experience and regularity was found in the left HG/STG. Tone-evoked responses did not show this difference, indicating that the expertise effect did not reflect the superior tone processing acquired by amateur musicians due to musical training. These results suggest that perceptual grouping based on repetition of a pattern of frequencies affects the processing of omissions in tone sequences and induces more activation of the bilateral auditory cortex by violating internal models. The interaction in the left HG/STG may suggest different styles of processing for musicians and nonmusicians, although this difference was not reflected at the behavioral level.

1. Introduction

In an orchestral performance, a musical piece is produced by multiple sequences of tones played in parallel. The auditory system can extract the structural components of the piece, such as its melody and rhythm, from this mixture of tones using processes that integrate acoustic information over time. Together, these processes are called perceptual grouping, and psychological studies have identified the rules for grouping sound features, such as similarity, proximity or good continuity (Bregman, 1990; Deutsch, 2012; Koffka, 1935). Bregman (1990) suggested that two types of perceptual grouping exist: one is stimulus-driven and works preattentively in a short time window, while the other requires higher cognitive functions such as attention and/or experience-based knowledge, and has a longer time window.

A wording to predictive coding theory, cortical circuits create internal models to generate predictions about incoming stimuli (Friston and Kiebel, 2009a, 2009b; Friston, 2005). An evoked response may occur reflecting the transient expression of a prediction error, which results from comparison between the bottom-up inputs from lower cortical/subcortical areas and top-down predictions from higher cortical areas. Several studies have applied this theory to explain the early stages of auditory processing (Bendixen et al., 2012; Winkler and Czigler, 2012; Winkler, 2007; Winkler et al., 2009). Key to this explanation is that an internal model is created by spectral or temporal regularity, which is extracted from a tone sequence. Mismatch negativity (MMN) and the omission-related response (OR), both of which are elicited by deviation (a deviant tone or the omission of a tone) from a sequence of repetitive tone stimuli, can be interpreted as resulting from the violation of the

Abbreviations: OR, omission-related response; MEG, magnetoencephalography; HG, Heschl's gyrus; STG, superior temporal gyrus; MMN, mismatch negativity; ISI, inter-stimulus interval; AC, auditory cortex; MRI, magnetic resonance imaging; RMS, root mean square; ANOVA, analysis of variance; RT, reaction time; SD, standard deviation

* Corresponding author. Tel.: +81 75 751 3695; fax: +81 75 751 3202.
E-mail address: onoken@kuhp.kyoto-u.ac.jp (K. Ono).

http://dx.doi.org/10.1016/j.heares.2014.10.013
0378-5955/

prediction. In particular, the OR is suitable for investigating prediction-related brain activity because it does not overlap with the response elicited by the stimulus. Previous studies have shown that an OR can be elicited by a tone omission in an unattended tone sequence at an inter-stimulus interval (ISI) of less than 200 ms (Alain et al., 1989; Hughes et al., 2001; Raij et al., 1997; Snyder and Large, 2005; Tarkka and Stokic, 1998; Todorovic et al., 2011; Wacongne et al., 2011; Yabe et al., 2001, 1997). Together with Bregman's idea and the predictive coding theory, these results can be interpreted as resulting from a violation of a prediction based on pre-attentive perceptual grouping based on temporal regularity. The OR in the absence of attention is localized in the auditory cortex (AC) (Raij et al., 1997; Todorovic et al., 2011), which may be involved in prediction and pre-attentive perceptual grouping.

Several neurophysiological studies have elicited ORs at an ISI longer than 200 ms when the participants paid attention to the stimuli (Alain et al., 1989; Joutsiniemi and Hari, 1989; Penney, 2004). These results suggest that an OR to tone sequences with long ISI can occur as a result of a violation of a prediction based on attentive perceptual grouping. However, the neural correlates of this phenomenon remain unclear. Thus, we aimed to find the neural correlates of prediction based on attentive perceptual grouping in a tone sequence with a regular frequency pattern. We hypothesized that, when participants paid attention to the stimuli, a repetitive frequency pattern would cause perceptual grouping and help create stronger predictions about incoming stimuli, compared to a tone sequence with a random pitch pattern. Thus, a violation of this prediction by an omission in a tone sequence with a pitch pattern would evoke a stronger OR than would an omission in a random tone sequence. To clarify this issue, we compared the brain magnetic responses evoked by omissions in regular and random tone sequences using magnetoencephalography (MEG).

In addition, we evaluated the impact of musical experience on the grouping effect. Musical training normally includes the structural analysis of musical pieces, which should improve the ability to extract regular patterns from a tone sequence because the structural components of a piece (e.g., melody, chord progression, meter, etc.) are established by pitch and/or rhythm patterns. Although psychological studies have demonstrated that perceptual grouping depends on experience (Bhatara et al., 2013; Dewar et al., 1977; Gobet and Simon, 1996; Idson and Massaro, 1976; Iversen et al., 2008; Saariluoma, 1989; Simon and Chase, 1973), no study has investigated the neural correlates of this phenomenon. Thus, we tested the hypothesis that musical training influences the brain mechanisms involved in the perceptual grouping of frequency patterns, leading to more pronounced patterns of cortical activation in musicians than in nonmusicians.

2. Methods

2.1. Participants

The participants consisted of 13 amateur musicians (7 males and 6 females) who regularly played musical instruments, such as piano, guitar, violin, and cello, with an average experience of 13 $\pm$ 5 years (mean $\pm$ standard deviation [SD]), and 14 nonmusicians (11 males and 3 females) who had no instrumental experience, except for lessons in school. All participants were right-handed with an average age of 22 $\pm$ 2 years and provided written informed consent to participate in the experiment. Although we did not measure the participants' hearing thresholds, none of them reported difficulty in discriminating the stimuli. The participants also did not report any neurological or hearing problems. The experiment was performed in accordance with the ethical standards of the Declaration of Helsinki and the guidelines approved by the local ethics committee of the Graduate School of Medicine and Faculty of Medicine, Kyoto University.

2.2. Stimuli

Pure tones (50-ms duration, 5-ms onset/offset ramps, 65 dB SPL) with three different frequencies (C5: 523 Hz, E5: 659 Hz, and G5: 784 Hz) were created as wave files using the Audacity software program (ver. 2.0.3; http://audacity.sourceforge.net/). A silent period with a length of 500 ms was created as the omission stimulus. Each tone sequence was constructed of these tones, presented either in a regular pattern of "CEG" (regular sequence) or pseudorandomly (irregular sequence), with an ISI of 450 ms (Fig. 1A). In the irregular sequence, randomization was controlled so as not to present the same frequency more than three times consecutively, and at least three tones were presented between omissions.

2.3. Procedure

Participants were seated in a chair in a magnetically shielded room. The tone sequences were presented through earphones, which was coupled to the ear by a silicon tube and the ear insert (E-A-R-tone 3A, Aearo Corporation, Indianapolis, USA). The earphone was connected to an amplifier (Roland SRQ-2031, Roland Corporation, Hamamatsu, Japan) outside of the shielded room. Participants were instructed to press a button with their right index finger as quickly as possible upon noticing any omission in the sequence. Because perceptual grouping facilitates the processing of deviant stimuli in a tone sequence (Idson and Massaro, 1976; Jones et al., 1982; Mondor and Terrio, 1998; Royer and Garner, 1970), the response time was used to characterize the effect of perceptual grouping.

Each sequence was presented in three separate blocks. Six blocks were conducted in total, and the order of the blocks was randomized between participants. In each block, approximately 7% of the tones were replaced with a silent period. In total, 2520 tones and 180 omissions (60 omissions for each tone) were presented in regular and irregular sequences. An additional restriction for the regular sequence was that, after each omission, the sequence started again from the C tone (e.g., CEGCEGCE_CEG …) to maintain the repetition of the CEG pattern.

At the end of the experiment, we asked the participants whether they had recognized the regular sequence as a CEG pattern, and all participants reported that they had.

2.4. MEG acquisition

Event-related fields (ERFs) were recorded with a 306-channel whole-head magnetoencephalography (MEG) system (Vectorview, Elekta Neuromag Oy, Finland). The head position was determined using four indicator coils attached to the scalp. In addition, three head landmarks (the nasion and bilateral preauricular points) and head shape were recorded for each participant using a spatial digitizer (Polhemus Inc., Colchester, VT, USA) before the experiment. These data were used for co-registration with the T1 anatomical image of each participant obtained using a 0.2 T magnetic resonance imaging (MRI) machine (Signa Profile, GE Health Care, Waukesha, WS, USA). The ERFs were recorded with a band-pass filter (0.1–200.0 Hz) and a sampling rate of 600 Hz. To reduce external noise, we used spatiotemporal signal space separation (tSSS) methods (MaxFilter, Elekta Neuromag Oy, Helsinki, Finland) with a correlation window of 900 s, which covered the entire length of each block, and a correlation limit of 0.980. The acquired data were low-pass filtered using a fifth-order Butterworth zero-phase filter with a cut-off frequency of 40 Hz. The time window of each epoch lasted between 50 ms prestimulus and