Neuroscience Letters 665 (2018) 99–103

Contents lists available at ScienceDirect

Neuroscience Letters

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

Research article

Association of different neural processes during different emotional perceptions of white noise and pure tone auditory stimuli

Fumi Masuda[a], Yukiyoshi Sumi[a], Masahiro Takahashi[a], Hiroshi Kadotani[b], Naoto Yamada[a], Masahiro Matsuo[a,*]

[a] Department of Psychiatry, Shiga University of Medical Science, Japan
[b] Department of Sleep and Behavioral Science, Shiga University of Medical Science, Japan

ARTICLE INFO

Keywords:
Sounds of common loudness
Sound frequency characteristics
Event-related potentials
Subjectively experienced valence and arousal responses

ABSTRACT

Sound is a sensory stimulant ubiquitously found throughout our environment. Humans have evolved a system that effectively and automatically converts sound sensory inputs into emotions. Although different emotional responses to sounds with different frequency characteristics are empirically recognized, there is a paucity of studies addressing different emotional responses to these sounds and the underlying neural mechanisms. In this study, we examined effects of pure tone (PT) and white noise (WN) inputs at ordinary loudness levels on emotional responses. We found that WN stimuli produced more aversive responses than PT stimuli. This difference was endorsed by larger late posterior positivity (LPP). In a source localization study, we found increased neural activity in the parietal lobe prior to LPP. These findings show that WN stimuli produce aversive perceptions compared with PT stimuli, at typical loudness levels. In addition, different emotional responses were processed in a similar manner as visual stimulations, as reflected by increased LPP activation. Various emotional effects of WN and PT stimuli, at ordinary loudness levels, could expand our understanding of adverse effects of noise as well as favorable effects associated with music.

1. Introduction

Human life is surrounded by various kinds of sounds, from appetitive sounds like birds chirping to aversive sounds like a dog's growl. Human auditory systems can unconsciously and automatically convert these sounds to emotional responses, depending on sound characteristics.

Studies have shown that emotional responses to sound occur almost instantly, at speeds where only electroencephalography-based technologies are suitable [1]. For example, stimulation-linked neural activities can occur as fast as 200 ms after stimulation presentation, and these changes are recorded as an event-related potential (ERP). In addition, ERPs that follow rapidly changing electro-neural events are also linked to emotions; ERPs occurring 300 ms after stimulus presentation also correlate with valence responses.

Loudness is often linked to aversive emotions, and sounds with extraordinary loudness (up to 100 dB) are used for evoking aversive emotions [2,3]. However, loud sounds such as these elicit responses mediated by vestibular processes that directly link to the autonomic nervous system [4], in addition to cochlear functions that link to higher auditory cognitive processes.

In daily life, one is unlikely to routinely encounter 100 dB sounds; sounds at approximately 50 dB are most commonly heard, and most human conversations and mechanical noise occur at moderate loudness levels [5]. Even at moderate loudness levels, sound can induce different emotional responses. At typical loudness levels, vestibular processes do not play a major role, and emotional experiences coupled with these sounds cannot explain the range of observed emotional responses [6,7]. Thus, acoustic characteristics alone could play a role in this process. Indeed, pioneering work by Halpern reported the contribution of sound frequency spectra on perceived discomfort [8]. However, there is a paucity of studies addressing the neural response underling sound frequency dependent emotional responses.

Naturally occurring sounds comprise several resonant frequencies that follow a power law relationship [9,10]. These naturally occurring sounds can be affectively coupled with emotional (including violent) events, even if the sound itself does not convey emotion. This coupling effect occurs during emotional responses to certain sounds. To avoid

Abbreviations: WN, white noise; PT, pure tone; EPN, early posterior negativity; LPP, late posterior positivity; ERP, event-related potentials; LORETA, low resolution brain electromagnetic tomography

* Corresponding author at: Shiga University of Medical Science, Seta Tsukinowa-cho, Otsu, Shiga 520-2192, Japan.
E-mail address: masahiro.matsuo@gmail.com (M. Matsuo).

https://doi.org/10.1016/j.neulet.2017.11.046
Received 5 October 2017; Received in revised form 8 November 2017; Accepted 21 November 2017
Available online 24 November 2017

this affective coupling, PTs are often used [11,12]. WN is another example of an unnatural sound where similar sound intensities are delivered across a wide frequency range [10]. Because artificial sounds cannot naturally occur, they are less likely to relate to natural life events.

However, little is known about how the human brain codes different frequency properties or if sounds at ordinary loudness levels can induce emotionally distinct responses.

In this study, we examined differences in neural activity relative to sound frequency type, and the relation of this activity to subjective feelings at ordinary loudness. To address these problems, we used PT and WN as sound sources and observed neural activity by combining ERPs with low resolution brain electromagnetic tomography (LORETA).

2. Material and methods

2.1. Subjects

The subjects for this study were recruited by advertisements, and 17 healthy adult volunteers participated (10 men; age mean age ± standard error, 21.6 ± 0.50 years). For their participation, the subjects were each given a gift card equivalent to ¥2500. The subjects had no psychiatric disorders, hearing problems, or smoking history. None were on medication or took caffeine on the day of the study. All subjects were right-handed as confirmed by the Edinburgh handedness inventory. All subjects were informed about the purpose and design of the study and provided written informed consent prior to completing any study-related procedures. This study was approved by the ethical committee at Shiga University of Medical Science, Japan.

2.2. Experimental design and settings

The affective stimulus was a 500-ms burst of 50 dB[A] WN with instantaneous (10 ms) rise/fall times. WN includes all frequency bands within the audible range. As a control, we used 1000-Hz PT stimuli because sounds with 1000 Hz peaks are observed most often [9] and 1000 Hz peaks are less affected by age-related losses in hearing sensitivity [13]. We presented stimulations in a passive task context where subjects were instructed to simply view the fixation point that was presented.

Auditory stimulations were provided through headphones (AKG closed-back headphones, K404). The stimulus sound was programmed to randomly produce each frequency 75 times, with randomized stimulus intervals (2000 ± 200 ms) using E-Prime v 2.0 software (Psychology Software Tools, 2013).

During the experiment, subjects remained seated on a chair that was placed 70 cm in front of a cathode ray tube (CRT) display in a sealed room. The illumination in the room was maintained at 80 lx. Subjects were directed to look at a white cross fixation point that appeared against the black CRT background. The first 5 min were designated as the silent condition in which no sound was administered. After this, the stimulus sound was administered for approximately 5 min.

We scored subjective emotional responses using the Self-Assessment Manikin (SAM), which is a two-dimensional subjective scoring system used for assessing affective stimuli using the International Affective Picture System [14]. This is a nine-point rating scale consisting of two sets of figures for measuring valence responses (1 = unpleasant; 9 = pleasant) and arousal responses (1 = arousing; 9 = calming). Study participants scored SAM scores for both WN and PT stimuli, immediately after each experiment.

2.3. Electroencephalography data acquisition

Electroencephalography (EEG) signals were recorded using NetStation software (Electrical Geodesics Inc., Eugene OR, USA), and 64-channel recordings were made through the HCGSN v.1.0. gel cap. Data were sampled at 500 Hz and referenced to Cz. Electrode impedances were kept at < 60 kΩ throughout the experiments, according to methods described elsewhere [15]. Subjects were asked to remain awake, and we confirmed vigilance by online observation of the EEG signal.

2.4. ERP data processing

EEG data were processed using EEGLAB (version 13.4.4b) [16], an open source toolbox that runs on MATLAB (The Mathworks, Inc. version 2015a). EEG data were re-referenced to the average of the left and right mastoids. First, data were bandpass filtered offline by 0.1–50 Hz, and gross artifacts were visually rejected following independent component analysis, excluded components produced by eye movements, and EMG. We epoched all data segments from 500 ms prior to and 1500 ms post stimulations, and baseline corrections were done by subtracting the average 100–0 ms prior to stimulation. The number of epochs used for each individual ERP calculation ranged from 38 to 73 epochs (average: 55.76 epochs) for the PT condition and from 37 to 71 epochs (average: 53.94 epochs) for the WN condition. To compare component amplitudes, we compared the averages of the Pz, P3, and P4 electrode potentials between the WN and PT conditions to capture potential fluctuations that are maximal at the centro-parietal sites [17]. To investigate the regions involved in these differential processes, we used time-series LORETA analysis for every 2 ms, to estimate the current source density distribution for each ERP component [18].

2.5. Statistical analyses

Data are shown as mean: M and standard error of mean: SE, unless otherwise stated. Student's *t* tests were used for between-group comparisons, and we calculated Cohen's *d* to estimate effect size [19]. For comparison of ERP component amplitudes between PT and WN conditions, we used the two-way repeated measure ANOVA analysis. Statistical analyses were performed using IBM SPSS Statistics for Macintosh, Version 22.0 (IBM Corp. Armonk, NY). The effect size was estimated using Cohen's *d*, following definition and criteria described elsewhere [19]. LORETA images were statistically compared between sound conditions by using voxel-by-voxel *t*-test, which were corrected by the calculation of exact randomization probabilities (5000 randomizations). The threshold for *p* value to determine statistical significance was set at $p < 0.05$.

2.6. Correlational analysis

We examined correlations between subjective feelings and ERPs. To assess the role of time-specific components, we subdivided ERPs into five components: N1P2, early posterior negativities (EPN); P3; early LPP (eLPP); and late LPP (lLPP) which span 150–200, 200–300, 300–450, 450–650, and 650–900 ms post stimulations. These divisions are described elsewhere [20]. In addition, we defined ERP 900–1100 ms after stimulations as POST. The intra-individual differences in ERP amplitudes between PT and WN were compared against intra-individual self-reported differences in emotional ratings as previously reported [21].

3. Results

3.1. Subjective emotional ratings of WN and PT

First, we examined emotional responses evoked by WN and PT. Both sound stimuli evoked comparable arousal responses [WN: M = 4.00, SE = 1.58; PT: M = 4.71, SE 1.26, t(16) = 1.73, $p = 0.104$, paired Student's *t*-test, two-tailed, Cohen's $d = 0.25$]. However, subjects perceived significantly aversive responses for WN, and a moderate effect size is indicated by a Cohen's *d* of about 0.5 [WN: M = 2.76, SE = 1.15;