Investigation of optimal auditory signal for visually-challenged people using auditory evoked magnetic responses

Y. Soeta, Y. H. Kim, S. Nakagawa
Health Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Midorigaoka 1-8-31, Ikeda, Osaka 563-8577, Japan
y.soeta@aist.go.jp

Abstract: Birdsongs are used as an auditory signal for visually-challenged people in public spaces. More than 40% of the visually-challenged people reported that the auditory signal is hard to identify. We tried to find the auditory signal that is easy to identify based on auditory evoked magnetic field (AEF) responses. We choose eight birdsongs as experimental stimuli. The birdsong was presented in silent and noisy conditions in Experiment 1 and in different interaural time delay (ITD) and reverberation time (RT) conditions in Experiment 2. The most prominent AEF response, N1m, and correlation between the envelope of the birdsong and the AEF was analyzed. The results indicated that the N1m amplitudes and the correlation showed maximum when the participants listened to the song of Cuckoo and Oriental scops owl, suggesting that the song of Cuckoo and Oriental scops owl is a candidate for optimal auditory signal that is easy to identify.

Keywords: Auditory signal, Auditory evoked magnetic field, birdsong

1. Introduction

Auditory signals are classified by the functions. Main auditory signals are alarm, which informs the events involving human lives, attention, which inform basic point in the plant, such as stairs, toilets, and ticket gate, and announcement, which covey information by speech. We focused on the attention in this study.

Some signals are used as an auditory signal for visually-challenged people in public spaces in Japan. Security guidelines for public transportation system in Japan recommend that sounds with fundamental frequency between 100 and 2500 Hz and broader frequency bands are used as the auditory signals. However, more than 40% of the visually-challenged people reported that the auditory signal is hard to identify although the auditory signals were set according to the guideline that was established based on lots of previous psychophysical findings. Neurophysiological responses, such as brain magnetic response, can inform us what types of sounds are processed in human brain. The amplitude of the response is correlated with loudness, pitch salience, and localization performance. The degree of correlation between the waveform of the response and envelope of the speech is correlated with speech intelligibility. Thus we tried to find the auditory signal that is easy to identify based on the amplitude and the correlation of auditory evoked magnetic fields (AEFs). We focused on songbirds as an auditory signal that informs the location of stairs in train stations. The larger amplitude of the response or higher correlation between the response and envelope of the sound suggests that we perceive birdsongs as louder and stronger pitch or clearer temporal waveform of the birdsongs.

2. Method

We analyzed 32 kinds of birdsongs to find appropriate birdsongs as an auditory signal, and selected eight birdsongs that meet the specification shown by the guideline as candidate auditory signals. They are Black Paradise Flycatcher (BP), Cuckoo (CC), Japanese Grosbeak (JG), Jay (JY), Jungle Nightjar (JN), Little Cuckoo (LC), Oriental Scops Owl (OS), and Ruddy Kingfisher (RK). Figure 1 shows the temporal waveform and spectrogram of the eight birdsongs.

Ten normal-hearing participants (23-40 years old) took part in the experiment. The birdsongs were presented binaurally through insert earphones at 60 dB SPL with the duration of 500-1000 ms. In Experiment 1, the eight birdsongs were presented in two background noise conditions. One was silent and another one was simulated noise environments in averaged subway station (58 dB SPL). In Experiment 2, four birdsongs (BP, CC, OS, and RK) was presented with interaural time delay (ITD) of 0.0 and 0.7 ms or reverberation time (RT) of 0.5 and 2.0 s.
was controlled as envelope ITDs. RT was simulated by same volume rooms with different absorptive materials. The impulse responses obtained in the simulated room were convolved with birdsongs.
AEFs were recorded using a 122-channel whole-head magnetometer in a magnetically shielded room. Brain activities evoked by the sound stimuli were recorded using a 122 channel whole-head magnetometer (Neuromag-122\(^{TM}\)) in a magnetically shielded room.\(^1\) In order to maintain a constant vigilance level, the participants were instructed to concentrate on a self-selected silent movie that was being projected on a screen in front of them and to ignore the stimuli. Magnetic data were sampled at 400 Hz after being bandpass-filtered between 0.03 and 100 Hz, and then averaged approximately 100 times. The averaged responses were digitally filtered between 1.0 and 30.0 Hz. The analysis time was 0.7 s from 0.2 s prior to onset of the stimulus. The pre-stimulus period (average of the 0.2 s) was used as the baseline level. The root mean square (RMS) amplitudes were obtained with a subset of 18 channels over left and right temporal areas. The peak amplitude and latency having the maximum value of the RMS amplitudes in the amplitude and latency range of 70-140 ms over each left and right hemisphere was defined as the latency of the N1m in each subject. The cross-correlation function between the RMS and envelope of the birdsongs also obtained. The envelopes of the birdsongs were calculated by Hilbert transform.

### 3. Results

Figure 2 (a) shows the normalized N1m amplitude averaged for all participants in Experiment 1. The N1m amplitude was affected by the types of songbirds and decreased in the simulated noise conditions. The N1m amplitude for CC was largest in both silent and simulated noise conditions.

Figure 2 (b) shows the normalized N1m amplitude averaged for all participants in Experiment 2. The N1m amplitude for CC was largest in both ITD and RT conditions. The N1m amplitude for CC and OS with ITD of...
0.7 ms was larger than that for CC and OS with ITD of 0.0 ms, suggesting that CC and OS from lateral direction are well localized. The N1m amplitude for CC with longer RT became larger, suggesting that reflections can strengthen pitch salience for CC.

Figure 3 shows the maximum correlation coefficients between the waveform of the response and envelope of the birdsong’s waveform. There was little difference between silent and simulated noise conditions. The correlation coefficient showed maximum for CC and OS in silent, simulated noise, and ITD conditions. The correlation coefficient showed maximum for OS in RT conditions.

4. Discussion

The N1m amplitude for CC was largest in the all conditions. The larger amplitude of N1m suggests that we perceive louder, stronger pitch, and localize clearly.4-7 In reverberant conditions, localization performance worsens. Thus, it was assumed that the N1m amplitude decreases with increasing RT. However, the N1m amplitude increased with increasing for RT when the participant heard CC. Birdsongs are often used in underground train stations that has long RT. The results suggests that song of cuckoo is a candidate for an auditory signal that is easy to identify.

The AEF amplitude for CC and OS is highly correlated with the envelope of the sound waveform. The higher correlation indicates that envelope information of sound that is important for perception of speech and music10 is transmitted to auditory cortex clearly.8 This suggests that auditory information of cuckoo and Oriental Scops Owl is well conveyed to human auditory cortex.

As shown in the temporal waveforms of birdsongs (left hand side of Fig .1), the early part of CC and OS has longer early component and longer silent intervals. The longer early components more than 50 ms and longer silent interval more than 200 ms may cause larger N1m response and higher corelation between envelope of the birdsongs and the AEF amplitude. As shown in the spectrogram of birdsongs (right hand side of Fig .1), the cuckoo and oriental scops owl has salient frequency component that is not stable. These frequency characteristics may be important for our perception in noise or reverberant conditions and localization performance.

Acknowledgments

This work was supported by a Grant-in-Aid for Young Scientists (A) from the Japan Society for the Promotion of Science (No. 23686086), Kawai Foundation for Sound Technology & Music, and the Maeda Engineering Foundation.

References