Estimation of optimal auditory signal for visually-challenged people using auditory evoked magnetic response

Y. Soeta, 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, s-nakagawa@aist.go.jp

ABSTRACT

Birdsongs are used as an auditory signal for visually-challenged people in public spaces in Japan. The auditory signal informs them where stairs, toilets, and ticket gates are located. 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. Security guidelines for public transportation facility in Japan recommend some acoustic characteristics of auditory signals. We analyzed 32 different birdsongs and choose eight birdsongs as experimental stimuli according to the guideline. The birdsong was presented in silent and simulated noisy underground train stations in Experiment 1. The birdsongs was presented with interaural time delay (ITD) of 0 and 0.7 ms and with reverberation time (RT) of 0.5 and 2.0 s in Experiment 2. The most prominent AEF response, N1m, which was found around 100 ms after the onset of the stimuli, was analyzed. The results indicated that the N1m amplitudes showed maximum when the participants listened to the song of Cuckoo in silent, simulated noise, ITD, and RT conditions. The findings suggest that the song of Cuckoo is a candidate for an auditory signal that is easy to identify.

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 (Ministry of Land, Infrastructure, Transport and Tourism 2007) 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 (Foundation for Promoting Personal Mobility and Ecological Transportation 2009) 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 (Hämäläinen et al. 1993). The amplitude of the response is correlated with loudness (Soeta and Nakagawa 2008; Uppenkamp and Röhl 2013), pitch salience (Soeta et al. 2005), and localization performance (Soeta and Nakagawa 2006). The degree of correlation between the waveform of the response and envelope of the speech is correlated with speech intelligibility (Aiken and Picton 2008). 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. 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.

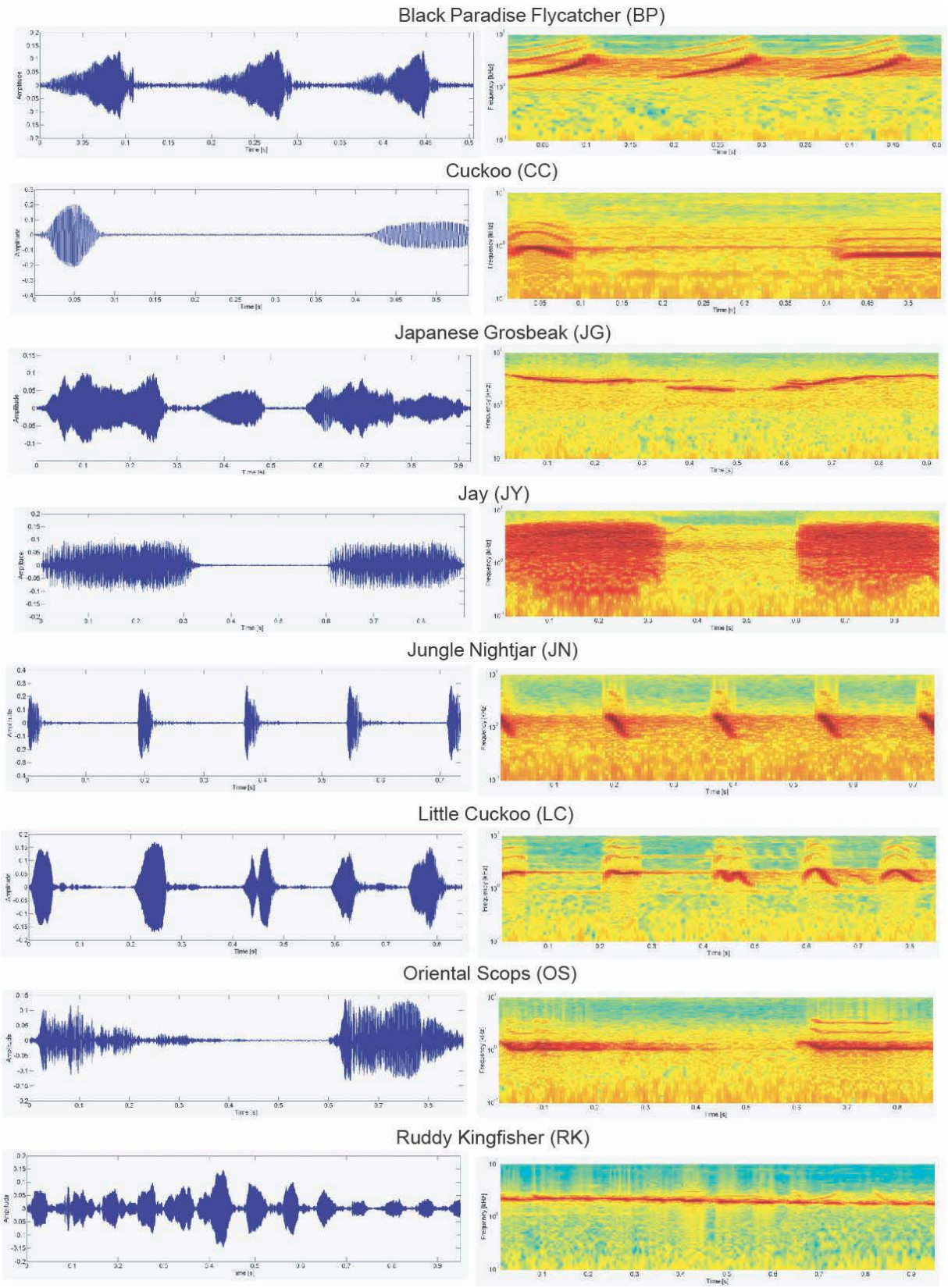

Figure 1: Temporal waveform and spectrogram of the birdsongs used.

METHODS

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 (Ministry of Land, Infrastructure, Transport and Tourism 2007). 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) (Shimokura and Soeta 2011). 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. ITD was controlled as envelope ITDs (Soeta and Nakagawa 2006). 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. 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 latency having the maximum value of the RMS amplitudes in the 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.

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. The N1m amplitude for CC with longer RT became larger.

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.

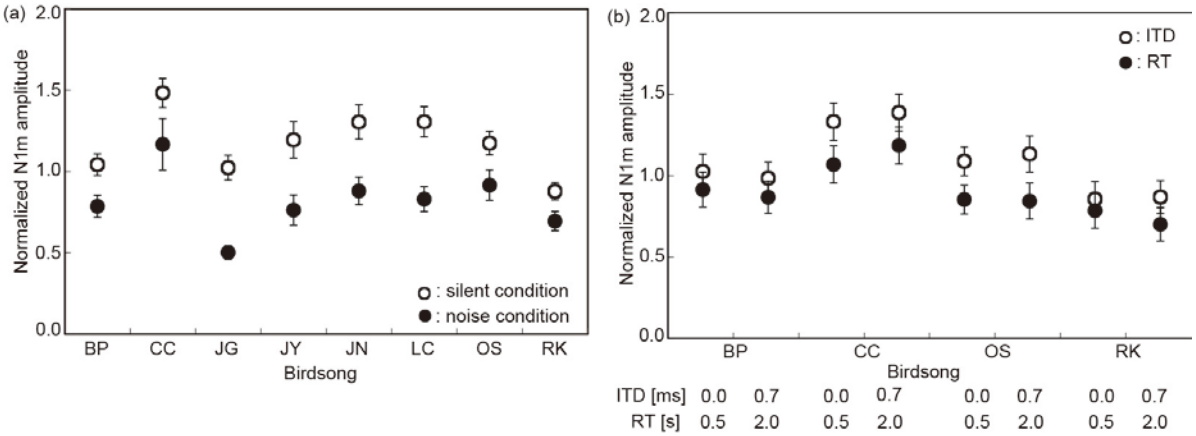

Figure 2: Normalized N1m amplitude for each songbird condition in (a) Experiment 1 and (b) 2.

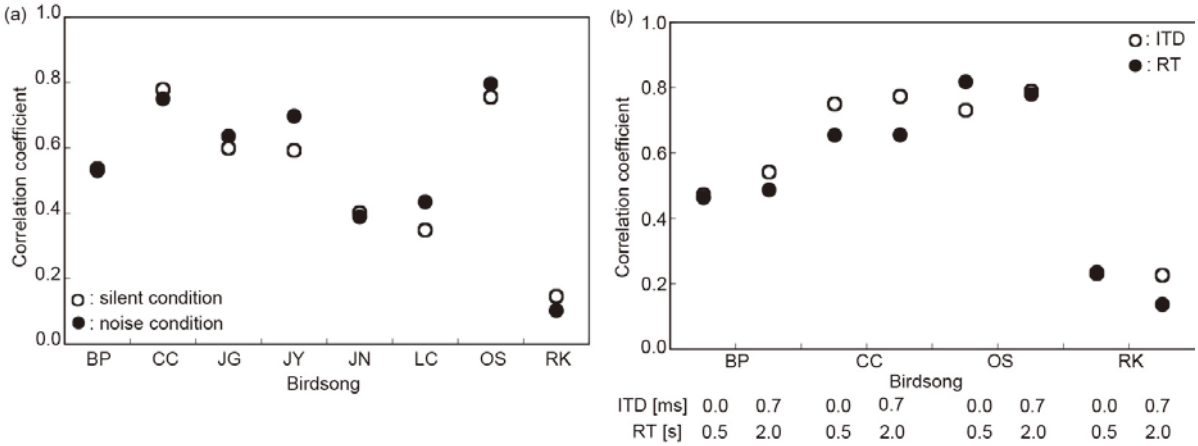

Figure 3: The maximum correlation coefficients between the waveform of the AEF response and envelope of the birdsong's waveform in (a) Experiment 1 and (b) 2.

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 more clearly localized sound (Soeta et al. 2005; Soeta and Nakagawa 2006; Soeta and Nakagawa 2008; Uppenkamp and Röhl 2013). 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. This 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 music (Smith et al. 2002) is transmitted to auditory cortex clearly (Aiken and Picton 2008). This suggests that auditory information of cuckoo and Oriental Scops Owl is well conveyed to human auditory cortex.

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

Aiken SJ, Picton TW (2008). Human cortical responses to the speech envelope. Ear Hear 29, 139-157.

Foundation for Promoting Personal Mobility and Ecological Transportation (2009). Basic case study report on movement support for traffic hub, (in Japanese).

Hämäläinen MS, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993). Magne-toencephalography – theory instrumentation and applications to noninvasive studies of the working human brain. Rev Mod Phys 65, 413–497.

Ministry of Land, Infrastructure, Transport and Tourism (2007). Guidelines for facilitating the move in public transportation system, (in Japanese).

Shimokura R, Soeta Y (2011). Characteristics of train noise in ground and underground stations with side and island platforms. J Sound Vib 330, 1621-1633.

Smith ZM, Delgutte B, Oxenham AJ (2002). Chimaeric sounds reveal dichotomies in auditory perception. Nature 416, 87-90.

Soeta Y, Nakagawa S, Tonoike M (2005). Auditory evoked magnetic fields in relation to the iterated rippled noise. Hear Res 205, 256-261.

Soeta Y, Nakagawa S (2006). Auditory evoked magnetic fields in relation to interaural time delay and interaural correlation. Hea Res 220, 106-115.

Soeta Y, Nakagawa S (2008). Relationship between loudness and auditory evoked N1m. Biomagnetism: Interdisciplinary Research and Exploration, 95-97.

Uppenkamp S, Röhl M (2013). Human auditory neuroimaging of intensity and loudness. Hear Res 307, 65-73.