Investigation of Frequency-Selective Loudness Reduction and Its Recovery Method in Hearables

Watanabe, Hiroki
Kanemoto, Sota
Terada, Tsutomu
Tsukamoto, Masahiko

(Citation)
IEEE Access, 12:49916-49926

(Issue Date)
2024-04-04

(Resource Type)
journal article

(Version)
Version of Record

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 30 January 2024, accepted 29 March 2024, date of publication 4 April 2024, date of current version 12 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3385123

RESEARCH ARTICLE

Investigation of Frequency-Selective Loudness Reduction and Its Recovery Method in Hearables

HIROKI WATANABE[1], SOTA KANEMOTO[2], TSUTOMU TERADA[2], (Member, IEEE), AND MASAHIKO TSUKAMOTO[2], (Member, IEEE)

[1]Graduate School of Information Science and Technology, Hokkaido University, Sapporo 060-0814, Japan
[2]Graduate School of Engineering, Kobe University, Kobe 657-8501, Japan

Corresponding author: Hiroki Watanabe (hiroki.watanabe@ist.hokudai.ac.jp)

This work was supported in part by Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (KAKENHI) under Grant JP21K11973, in part by Japan Science and Technology Agency (JST) Precursory Research for Embryonic Science and Technology (PRESTO) under Grant JPMJPR2138, and in part by the JST Core Research for Evolutional Science and Technology (CREST) under Grant JPMJCR18A3.

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Human Ethics Committee of Graduate School of Engineering, Kobe University under Permission Nos. 04-41 and 04-42, and performed in line with the Declaration of Helsinki.

ABSTRACT With the ongoing spread and functional improvement of hearables, we may soon find ourselves in a society where users are wearing hearables at all times. In a hearable environment of this kind, the constant presentation of aural information to users may impede their ability to hear external noises that require their attention. For example, suppose the constant presentation of information in a particular frequency band causes a reduction in the subjective perception of sound pressure (loudness) of the corresponding frequency band. In such a case, the response to environmental sounds that indicate danger (e.g., the sound of an approaching car or an emergency alarm) may be delayed, leading to potential disaster. In this study, we investigated 1) how the presentation of a sound of a specific frequency through a hearable affects the loudness; and 2) which stimulus sound is most effective for recovering the decrease in loudness. In the first investigation, a loudspeaker presented the sound of a specific frequency that imitates environmental sound, and a hearable gave a stimulus sound of a particular frequency based on the frequency of the loudspeaker sound. The results showed that the loudness decreased by more than 10.0% in all stimulus sounds listened to with hearables, and the amount of the decrease tended to be larger the closer the frequency of the loudspeaker sound was to that of the hearable sound. In the second investigation, we hypothesized that the presence of specific recovery stimulus sounds would be effective in quickly restoring any loudness that had decreased, and the results showed that the amount of recovery was greater for all the recovery stimulus sounds we used compared to when the stimulus sounds were not presented.

INDEX TERMS Hearables, human–computer interaction, loudness recovery, loudness reduction, wearable computers.

I. INTRODUCTION

In recent years, the demand for earphone-type wearable computers called "hearables" has been increasing [1], [2]. Hearables are connected to devices such as smartphones to enable device operation by voice and the acquisition of information by audio. Many hearables feature a transparency mode, which captures the external sound by means of an outside microphone and plays it back from the inside speaker, as well as various noise-canceling functions. Further capabilities are expected in the future, such as manipulating human auditory perception [3] and acquiring user's biometric information [4]. Therefore, we may soon find ourselves in a hearable environment in which users wear hearables at all

The associate editor coordinating the review of this manuscript and approving it for publication was Alessio Vecchio.

times. In light of such a hearable environment, this study focuses on human auditory perception. Auditory information is essential sensory information because it not only enables us to hear music and announcements but also helps us identify danger by sound. Auditory information is transmitted from the outer to the middle and inner ear in the auditory peripheral system and converted into electrical signals in the inner ear. For the perception of sound direction, research has shown that the head-related transfer function, a characteristic of the change in the frequency response of sound depending on the auricle, can be expressed as a physical quantity [5]. Although various auditory characteristics have been clarified, we focus in the current study on changes in the subjective perception of sound pressure (loudness) due to auditory habituation. One example of loudness reduction is the phenomenon where the sound after listening to loud music seems quieter than before listening to music [6]. While loudness reduction typically occurs across frequencies for general sounds, when we consider the presentation of information in a hearable, there may be cases where sounds are presented only at a specific frequency/frequency band, e.g., a person may only hear certain sounds through manipulating auditory perception [3] or by listening to music that has been excessively equalized. The human auditory nerve responds differently to different frequencies [7], [8], so listening to a particular frequency could cause a loudness change at certain frequencies. If this phenomenon occurs when, for example, a user is listening to sounds in the same frequency band as a car horn or emergency alarm, which are environmental sounds that represent danger, the user's response might be delayed, which could lead to unexpected peril (Fig. 1(a)). In a pervasive environment where audio is constantly presented through a hearable, a user wearing a hearable may experience a reduction in loudness without being aware of it. In other words, although it is possible to be aware of the difficulty in hearing sounds when the loudness reduction in all frequency bands, it is difficult to notice a decrease in the loudness only in specific frequency bands because the user can hear sounds other than the specific frequencies as normal. Studies have shown that the loudness returns to the normal state after maintaining silence for a certain period after the loudness reduction [9]; however, it is not realistic to enforce silence for extended periods during the daily use of a hearable. On the other hand, since certain natural sounds, such as birdsong and running water, can possibly contribute to the recovery of attention and stress, the presentation of certain sounds may be effective in the quick recovery of loudness reduction [10] (Fig. 1(b)). Since it is easier to present specific sounds to the device wearer in a hearable environment, effective recovery sounds could be utilized immediately after a decrease in loudness.

In light of this background, we conducted two investigations (Fig. 1) to examine (i) how the presentation of a sound of a specific frequency through a hearable affects the user's perception of loudness, and (ii) which stimulus sound would be most effective for recovering the decrease in loudness. In the first investigation, a sound of a specific frequency is presented by a loudspeaker to imitate environmental sound, and a stimulus sound of a specific frequency based on the frequency of the loudspeaker sound is presented through a hearable. Participants periodically estimated the loudness of the presented loudspeaker sound, and the effect of listening to the stimulus sound on the loudness was determined from the change in the estimated value. The results showed that the loudness decreased by more than 10.0% for all stimulus sounds listened to with hearables, and the amount of the decrease tended to be larger the closer the frequencies of the loudspeaker sound were to those of the hearable sound. In the second investigation, we hypothesized that the presence of specific recovery stimulus sounds would be effective in quickly restoring the decreased loudness. We investigated the effect on recovery by presenting recovery stimulus sounds with multiple sound effects after the loudness had decreased and then comparing the loudness before and after the presentation. The results showed that the amount of recovery was greater for all the recovery stimulus sounds compared to when no stimulus sounds were presented. Our findings also indicated that listening to a sound with a band-stop filter applied to a specific frequency of white noise and a fade-out effect was the most effective in restoring loudness reduction.

The main contributions of this paper are as follows.

- We clarified the issue of loudness reduction in a hearable environment.
- We confirmed that a reduction in loudness at certain frequencies occurs in a hearable environment.
- We investigated the recovery stimulus sound for frequency-selective loudness reduction and demonstrated that it effectively restores loudness.

Section II of this paper describes related research. In Sections III and IV, we present our experiments on the loudness change and effective recovery sound, respectively. Section V discusses the application of loudness recovery to hearables. We conclude in Section VI with a brief summary and mention of future work.

II. RELATED WORK

A. HEARING ASSISTANCE AND ENHANCEMENT IN HEARABLES

There are many products and studies related to hearables designed to assist or extend the functionality of the user's hearing. NuraTrue is a wireless earbud that calculates auditory sensitivity by playing various sounds into the ear and measuring the otoacoustic emissions produced and then automatically adjusts the sound to the user's ear characteristics [11]. LinkBuds are a type of always-worn earphones that, in contrast to earplug-type earphones, allow users to listen to their surroundings and conversations without blocking the ear canals [12]. Jabra sells earphones with a feature that optimizes call volume according to ambient noise [13]. Wei et al. have developed earphones with a better frequency response in the frequency band below 1000 Hz than ordinary earphones by using graphene, a material with excellent

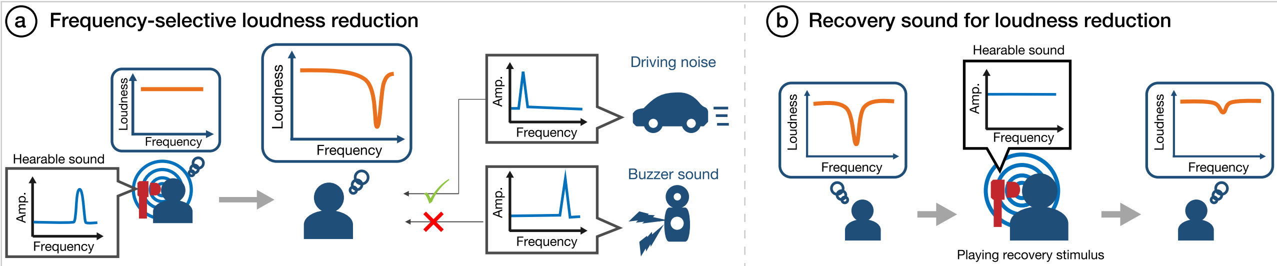


FIGURE 1. **Overview of (a) frequency-selective loudness reduction and (b) recovery stimulus for loudness reduction.**

electrical conductivity [14]. Hoshina et al. investigated the effect of the active noise control (ANC) function on the output level from earphones in noisy environments and confirmed that it can mitigate the risk of hearing loss [15]. Yamanobe et al. measured the balance function using earphones with a built-in accelerometer and identified a positive correlation with results using a stadiometer, indicating the usefulness of measuring balance with a non-medical device [16]. Ando et al. showed that it is possible to recognize face-related movements such as opening and closing the mouth and changing the direction of the face by acquiring changes in air pressure inside the ear canal using a barometric sensor [17]. Taniguchi et al. confirmed that an earphone-type sensor equipped with a light-emitting diode and a phototransistor can be utilized to measure the shape change of the ear canal caused by tongue movement, which can be used in applications such as music players [18]. Choi et al. utilized earphones with a built-in optical heart rate sensor and an accelerometer to acquire changes in facial muscle movement and confirmed that the earphones could detect and recognize seven facial expressions of the user, including anger and surprise [19].

Although a wide range of research has investigated how to improve the comfort and sound quality of hearables and of user-device interaction, to the best of our knowledge, none of these studies have considered the effect on hearing caused by the specific frequency of sound emitted by the hearable.

B. LOUDNESS CHANGE

Several studies have examined changes in loudness, especially those changes caused by listening to earphone sound alone. Carterette investigated the effect of noise on the loudness of the comparison sound by presenting both a noise stimulus with multiple frequency bandwidths centered at 1500 Hz and a comparison tone at 1500 Hz through earphones, and adjusting the volume of the comparison sound every minute. The results showed that the loudness reduction was the largest when the stimulus sound was the same as the comparison sound, and that the reduction decreased when the noise had a wide bandwidth [20]. Wagner et al. utilized magnitude estimation to examine the effect on hearing of a comparison sound in which six specific frequency stimulus sounds ranging from 500 to 8000 Hz were presented to one ear with earphones at 80-dB SPL and then presented to either ear at the same frequency at 70-dB SPL. Their findings showed that the loudness was reduced by 28% to 40%, with the amount of the reduction increasing for lower-frequency sounds [21]. Hellman et al. continuously presented four different sound volumes ranging from 5-dB SPL to 40-dB SPL to the right ear via earphones and examined the change in loudness of these sounds using magnitude estimation. The results showed that the loudness of the 40-dB SPL sound decreased by approximately 20% and that of the 5-dB SPL sound by 70% to 100% [22]. Botte et al. presented intermittent tones to one ear and continuous comparison tones to the opposite ear using earphones and examined the change in loudness in comparison listening when the speed of intermittency was varied under intermittent tone stimulation [23]. They confirmed that the amount of decrease in loudness tended to increase as the speed of intermittency increased. Charron et al. presented both intermittent stimulus tones and continuous comparison tones to one ear using an earphone and examined the change in loudness while listening to comparison tones of 21 different specific frequencies in response to an intermittent tone stimulus of one specific frequency [24]. They found that the loudness decreased the most when the stimulus and comparison tones had the same frequency, and that the amount of the decrease lessened as the frequency difference between the stimulus and test tones increased.

To the best of our knowledge, the above studies on changes in loudness have only verified the loudness reduction using sound heard through earphones; in other words, the effects of the sound stimulation of a hearable on the hearing of environmental sounds have not been investigated. In a hearable environment, it is assumed that the user always receives information from the hearable and also hears environmental sounds through transparency mode. Therefore, in this study, we investigate the effects of sound stimulation through a hearable on listening to environmental sounds.

C. RECOVERY OF LOUDNESS REDUCTION

Research on the recovery of loudness has been conducted mainly by examining recovery times and trends due to silence. Hirsh et al. presented a series of low-frequency sounds at 120-dB SPL for 3 min and periodically investigated changes in the minimum audible threshold [9]. The results showed that the thresholds temporarily returned to normal or recovered to above normal within about 1 minute after the presentation of the continuous sound, and that the thresholds that returned to normal rose again when a 500-Hz sound was used. Ward et al. showed that the recovery time for a threshold loss of 14.0 dB is approximately 16 hours [25]. They also investigated the trend of hearing recovery by periodically measuring the threshold after a minimum audible threshold loss of approximately 50 dB [26]. The results showed that although recovery occurred as a function of the logarithm of time from 200 to 500 minutes after stimulation, it was a linear function of time thereafter. Mills et al. investigated recovery times by periodically examining subsequent changes in minimum audible thresholds under two conditions: 24 hours of low-frequency octave-band noise at 84-dB SPL or 8 hours at 90-dB SPL [27]. Their findings showed that recovery took approximately 48 hours under the former condition and approximately 12 hours under the latter. Bell et al. investigated the effects of three variables on threshold lowering and recovery: stimulus sound frequency, loudness, and silence time to minimum audible threshold verification [28]. They found that the thresholds increased in proportion to the loudness of the stimulus sound and that continuous measurement of thresholds resulted in slower recovery than when measurements were made after a silent condition. Arieh et al. tested the time required for recovery by comparing the volume of the target sound and the comparison sound and having the participants select the louder one after multiple bursts of 80-dB SPL sound [29]. The recovery time was 70.5 s after five presentations and 129 s after 40 presentations, indicating that the recovery time increased monotonically with the number of times the stimulus sound was presented. Scharf et al. examined the recovery from loudness reduction induced by 160 seconds of intermittent sound stimulation to one ear at 60-dB SPL and found that it occurred after 30 seconds or more of silence [30].

Although these studies have clarified various aspects of the recovery of loudness, they examined the transition of recovery during a state of silence in which no sound is presented after the loudness has decreased. In the actual use of a hearable, a recovery method that forces silence every time loudness decreases is impractical because it would cause undue inconvenience to the wearer of the device. Since loudness reduction is caused by the sounds heard, loudness can be quickly restored by providing specific sound stimuli. In a hearable environment, it is possible to utilize sound stimuli for loudness recovery and effectively prevent danger caused by loudness reduction because sound information can be presented at all times. Therefore, we investigate which sounds are most effective in restoring the loudness reduction presented by a hearable.

III. INVESTIGATION OF FREQUENCY-SELECTIVE LOUDNESS REDUCTION

We first investigate the effect of a specific frequency sound stimulus given through a hearable on the loudness when listening to loudspeaker sounds that imitate environmental sounds. Fig. 2 shows the overview of the experiment in this section, where participants were asked to quantify the perceived loudness of the loudspeaker sounds presented to them, and the change in loudness was evaluated by comparing the values before and after listening to the specific frequency sound.

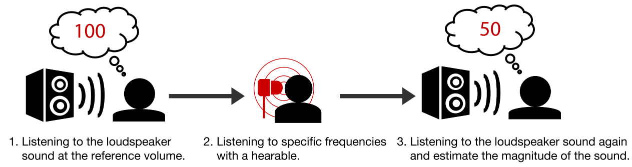


FIGURE 2. Overview of experiment 1.

A. EXPERIMENTAL SETUP

This experiment was conducted in the environment shown in Fig. 3 using the settings in Fig. 4. Participants wore hearables (Anker Liberty Air 2 pro) with transparency mode enabled on both ears. A loudspeaker (Fostex PM0.4C) located in front of the participants played a sinusoidal sound of a specific frequency for 84 seconds continuously at 60-dB SPL. Participants estimated the magnitude of the loudspeaker sound with a positive integer at a specific time indicated by a PC placed at hand and input the value using a keyboard. The estimation was done at the timing denoted by the black circle in Fig. 4, i.e., for the first time at five seconds after the start of the loudspeaker sound, for the second time at 15 seconds, for the third time at 22 seconds, and every eight seconds thereafter for a total of ten estimations.

The earphone sound was played at 70-dB SPL for 24 seconds, beginning 20 seconds after the start of loudspeaker sound playback. Table 1 lists the frequencies of the speaker and earphone we used: 500 Hz, 1000 Hz, and 3000 Hz for the loudspeaker and seven types for the earphone based

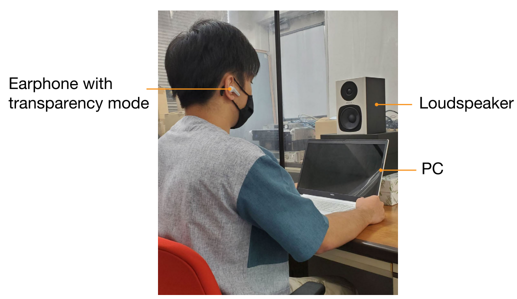


FIGURE 3. Participant during experiment.

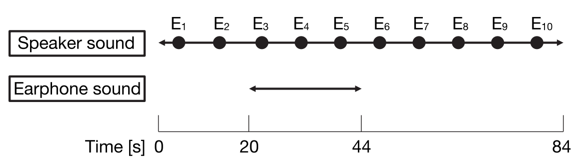


FIGURE 4. Sound playback time and estimated timing (black circle).

TABLE 1. Speaker and earphone frequencies.

Speaker [Hz]	500	1000	3000
Earphone [Hz]	None	None	None
	250	500	1500
	400	840	2520
	450	920	2760
	500	1000	3000
	550	1080	3240
	600	1160	3480
	1000	2000	6000

on the frequency of each loudspeaker sound. The human auditory nerve responds to different parts at different frequencies and is thought to have auditory filters with different center frequencies and bandwidths at different locations [7], [8], so considering the bandwidth of the auditory filter with the frequency of the loudspeaker sound, we decided to investigate the sound of five specific frequencies within the bandwidth and two specific frequencies outside the bandwidth. For an initial control experiment in which no earphone sound was played, participants performed 21 trials (3 loudspeaker sounds × 7 earphone sounds) and three trials (3 loudspeaker sounds × no stimulus sound), for a total of 24 trials. A 90-second interval was permitted between trials to avoid any effect from the previous trial. The participants were ten males in their 20s. This experiment was conducted with the approval of the Human Ethics Committee of Kobe University (04-41).

B. RESULTS

We obtained the loudness change AC_t by

$$AC_t = 100(E_1 - E_t)/E_1, \quad (1)$$

where t is time and E_t is the estimated value at time t. AC_t takes values ranging from 0% to 100%. The larger the value, the more difficult the participant perceives the loudspeaker sound to hear, indicating the amount of loudness reduction.

Figs. 5, 6, and 7 show the average decrease in loudness for all participants when listening to speaker sounds at 500 Hz, 1000 Hz, and 3000 Hz, respectively. The horizontal axis indicates the time from the start of the loudspeaker sound presentation, the vertical axis indicates the amount of the loudness reduction, and the light-brown area indicates the interval of the earphone sound presentation. The 24-second period during which the earphone sound was presented was excluded from the estimation because the participants heard both the earphone sound and the loudspeaker sound at the same time, which made it difficult to estimate the loudspeaker sound alone. In Fig. 5, the earphone sound that decreased the loudness the most for the 500-Hz loudspeaker sound was the stimulus sound at 500 Hz (i.e., the same frequency as the loudspeaker sound) and showed a maximum decrease of 50.6% at E_6. The minimum decrease in loudness when the earphone sound was presented was 17.4% of the stimulus with a 1000-Hz sound at E_{10}. In Fig. 6, the earphone sound with the greatest loudness reduction for

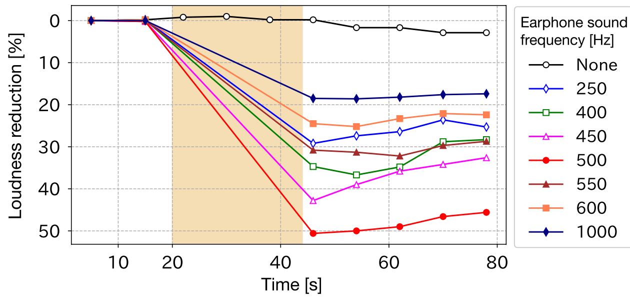


FIGURE 5. Loudness reduction at 500-Hz loudspeaker sound.

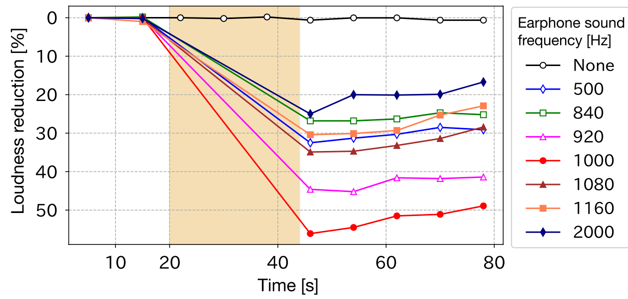


FIGURE 6. Loudness reduction at 1000-Hz loudspeaker sound.

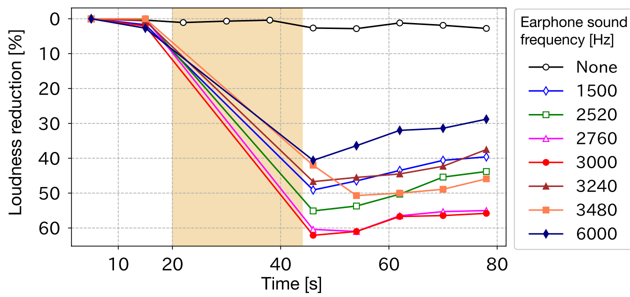


FIGURE 7. Loudness reduction at 3000-Hz loudspeaker sound.

the 1000-Hz loudspeaker sound was the 1000-Hz stimulus sound, with a maximum reduction of 56.1 % at E_6. The minimum decrease in loudness when the earphone sound was presented was 16.7% of the stimulus with the 2000-Hz sound at E_{10}. In Fig. 7, the earphone sound with the greatest loudness reduction for the 3000-Hz loudspeaker sound was the 3000-Hz stimulus sound, with a maximum reduction of 62.1 % at E_6. The minimum decrease in loudness when the earphone sound was presented was 28.8% of the stimulus with the 6000-Hz sound at E_{10}.

For all loudspeaker sounds, the loudness remained almost unchanged with a decrease of approximately 3% when the earphone sounds were not presented. In contrast, when the earphone sound was presented, the loudness decreased rapidly. The amount of the decrease was greatest when the frequencies of the loudspeaker sound and the earphone sound were the same, and we can confirm that the loudness hardly recovered in the 40 seconds after the decrease.

Figs. 8, 9, and 10 show the average decrease in loudness for all participants when listening to speaker sounds of 500 Hz, 1000 Hz, and 3000 Hz, respectively, at E_{10}. The horizontal axis shows the frequency of the earphone sound, and the

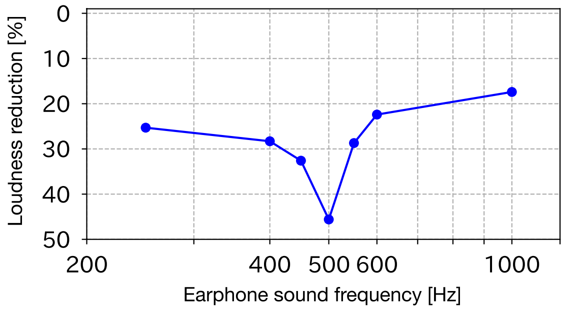


FIGURE 8. Loudness reduction at E_{10} in 500-Hz loudspeaker sound.

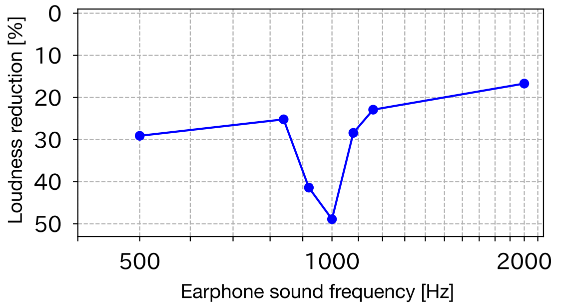


FIGURE 9. Loudness reduction at E_{10} in 1000-Hz loudspeaker sound.

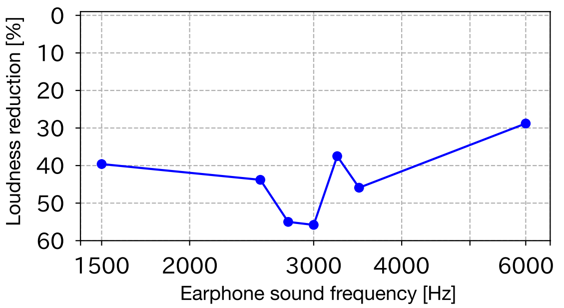


FIGURE 10. Loudness reduction at E_{10} in 3000-Hz loudspeaker sound.

vertical axis shows the amount of the loudness reduction. We can confirm here that the loudness decreased the most when the frequencies of the loudspeaker sound and the earphone sound were the same, and that the amount of the decrease in loudness tended to be lower when the frequency difference between the loudspeaker sound and the earphone sound increased. Also, the results were roughly symmetrical on the logarithmic frequency axis for the amount of reduction in loudness. Comparing the amount of loudness reduction when the frequencies of the loudspeaker sound and the earphone sound were the same, we found that the amount of the reduction increased as the frequency increased: by 45.6% when the loudspeaker sound was 500 Hz, by 48.9% when it was 1000 Hz, and by 55.8% when it was 3000 Hz. For all loudspeaker sounds, the loudness did not decrease the most when the frequency of the earphone sound was twice the frequency of the loudspeaker sound.

Tables 2, 3, and 4 summarize the minimum, maximum, and mean loudness reduction for all participants for each speaker sound (500 Hz, 1000 Hz, and 3000 Hz, respectively) at E_{10}. As we can see, for both the minimum and maximum

TABLE 2. Minimum, maximum, and mean loudness reduction at E_{10} for 500-Hz loudspeaker sound.

Speaker [Hz]		500						
Earphone [Hz]		250	400	450	500	550	600	1000
AC_{10} [%]	Min.	2.0	6.0	6.0	6.0	4.0	4.0	4.0
	Max.	60.0	40.0	50.0	64.0	40.0	50.0	28.6
	Mean	25.3	28.3	32.6	45.6	28.7	22.4	17.4

TABLE 3. Minimum, maximum, and mean loudness reduction at E_{10} for 1000-Hz loudspeaker sound.

Speaker [Hz]		1000						
Earphone [Hz]		500	840	920	1000	1080	1160	2000
AC_{10} [%]	Min.	4.0	4.0	14.0	10.0	6.0	0.0	0.0
	Max.	60.0	40.0	80.0	70.0	50.0	60.0	40.0
	Mean	29.1	25.2	41.4	48.9	28.4	22.9	16.7

TABLE 4. Minimum, maximum, and mean loudness reduction at E_{10} for 3000-Hz loudspeaker sound.

Speaker [Hz]		3000						
Earphone [Hz]		1500	2520	2760	3000	3240	3480	6000
AC_{10} [%]	Min.	10.0	12.0	10.0	10.0	12.0	8.0	0.0
	Max.	90.0	84.0	85.7	90.0	66.7	80.0	68.0
	Mean	39.6	43.8	55.0	55.8	37.5	45.9	28.9

values, the loudness reduction tended to be larger when the frequencies of the loudspeaker sound and the earphone sound were the same, and the reduction became smaller as the frequency difference increased. Focusing on the minimum value, we can see that the amount of decrease in loudness remained around 10% and that there were stimulus sounds with a 0 % decrease in loudness. In contrast, focusing on the maximum value, some stimulus sounds had a 90% reduction in loudness when the loudspeaker sound was 3000 Hz. We can also see that the larger the frequency of the loudspeaker sound, the larger the maximum amount of the decrease tended to be.

C. DISCUSSION

In the above experiment, we assumed that loudness would change with the stimulation of a specific frequency of sound presented through earphones, and that the amount of change would differ for each frequency of sound. As hypothesized, the loudness of the loudspeaker sound was decreased by the stimulation of the earphone sound, and the amount of decrease in loudness was larger when the frequency of the earphone sound was similar to that of the loudspeaker sound. We therefore conclude that, in a hearable environment, the loudness of the ambient sound obtained by the transparency mode changes depending on the sound emitted by the hearable.

The mean reduction in loudness for all stimulus sounds with earphones was greater than 10.0%. This is because the earphone sound was presented at 10-dB SPL louder than the loudspeaker sound. Our findings also revealed that the loudness decreased with higher-frequency sounds.

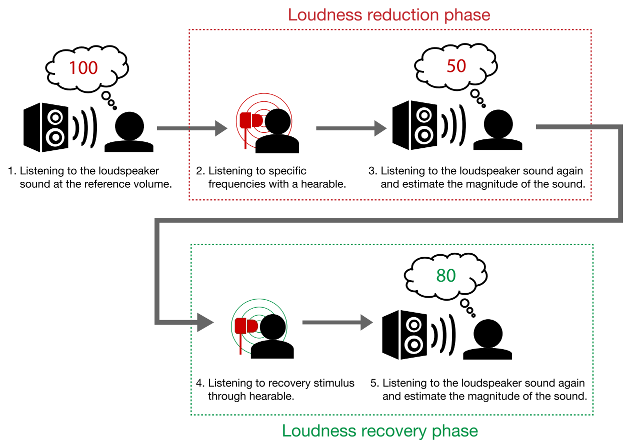


FIGURE 11. Overview of experiment 2.

Although the stimulus duration here was short (24 seconds), the earphone sound stimulus may have a greater effect on the loudness of the high-frequency band in environmental sound listening. These results suggest that presenting the stimulus sound with earphones at a higher volume and for a longer period may cause a further decrease in loudness. Tables 2, 3, and 4 show that there were individual differences in the amount of loudness reduction. When the frequencies of the loudspeaker and earphone sounds were equal, some participants showed a reduction in the loudness of 90% or more, while others showed a maximum reduction of 6%. However, we confirmed that the loudness reduction was greater when the frequency of the earphone sound and the loudspeaker sound were similar, indicating that the stimulation by the sound of a specific frequency affects the hearing of the user wearing the hearable.

IV. RECOVERY METHOD FOR LOUDNESS REDUCTION

While certain sounds can decrease loudness, as described in the previous section, it is also possible that the loudness change can be restored by presenting sounds that include all of the frequency band components that are expected to activate the entire hair cells in the ear, or by presenting sounds that affect the psychological rather than the physical aspect of the ear. Therefore, in this section, we investigate which stimulus sounds are effective in restoring reduced loudness. Fig. 11 shows an overview of this experiment. We used the same procedure as in the previous section to decrease loudness, and then presented the recovery stimulus again to investigate whether it is effective in restoring the loudness decrease.

A. EXPERIMENTAL SETUP

We conducted this experiment in the environment shown in Fig. 3. The presentation times of the loudspeaker sound and the earphone sound are shown in Fig. 12. Participants wore earphones with the transparency mode enabled in both ears and listened to a loudspeaker sound of a specific frequency created by a sine wave continuously at 60-dB SPL for

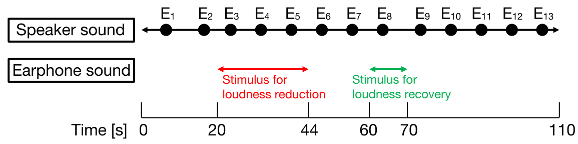


FIGURE 12. Sound playback time and estimated timing (black circles).

110 seconds from a loudspeaker placed in front of them. They then estimated the perceived magnitude of the loudspeaker sound at a specific time indicated by the PC placed in front of them using a positive integer, and input the value using the keyboard. The estimation was done at the timing denoted by the black circle in Fig. 12, i.e., for the first time at five seconds after the start of the loudspeaker sound, for the second time at 15 seconds, for the third time at 22 seconds, and every eight seconds thereafter until the eighth time. Furthermore, 72 seconds after the start of the loudspeaker sound, a ninth estimation was performed, and another one thereafter every eight seconds for a total of 13 estimations. As in the previous section, the first earphone sound was played at 70-dB SPL for 24 seconds, beginning 20 seconds after the start of the loudspeaker sound playback to produce a loudness reduction. The second earphone sound was a stimulus sound intended to restore loudness reduction and was played at 60-dB SPL for ten seconds, beginning 60 seconds after the start of loudspeaker sound playback.

The frequency of the loudspeaker sound we utilized was 1000 Hz, and the stimulus for loudness reduction was the same as the frequency of the loudspeaker sound (1000 Hz) that reduced the loudness the most in the previous section. As the second earphone sounds (recovery stimulus sounds), we selected the 11 types listed in Table 5 for a comprehensive survey. These recovery sounds were selected because pink noise is expected to have a relaxing effect thanks to its 1/f fluctuation, the sweep sound can activate the entire hair cells in the ear, and the white noise with a band-stop filter (BSF) applied contains all the frequency components that are not included in the first stimulus sound. The audio effects (fade-in, fade-out, fade-in/out, and intermittent) were also selected based on the assumption that they would positively affect the recovery. Participants performed a total of 12 trials, including a control experiment in which the recovery stimulus was not played. A 90-second interval was set between each trial to avoid any effects from the previous trial. The participants were 20 males in their 20s, eight of whom had participated in the experiment described in the previous section. This experiment was conducted with the approval of the Ethics Committee of Kobe University (04-42).

B. RESULTS

The results are shown in Fig. 13. As we can see, even when the same 1000-Hz stimulus sound was used to decrease the loudness, the amount of decrease varied depending on the trial. Therefore, we focus only on the amount of change before and after the recovery stimulus. The metric of recovery

TABLE 5. Stimulus sounds and effects used for recovering loudness reduction (BSF: band-stop filter, FI: fade-in, FO: fade-out, FIO: fade-in/out, and IEN: intermittence).

Recovery stimulus	Effect
Pink noise	None, FI, FO, FIO, IEN
White noise w/ BSF	None, FI, FO, FIO, IEN
Sweep signal	None

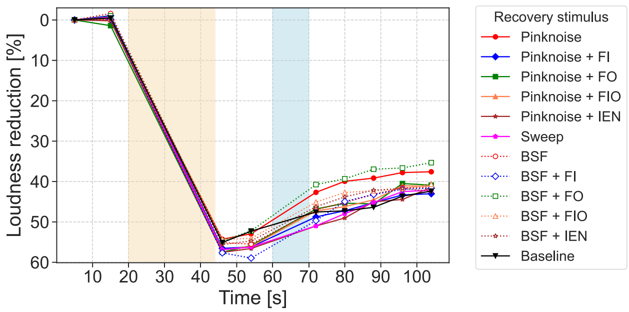


FIGURE 13. Loudness reduction rate.

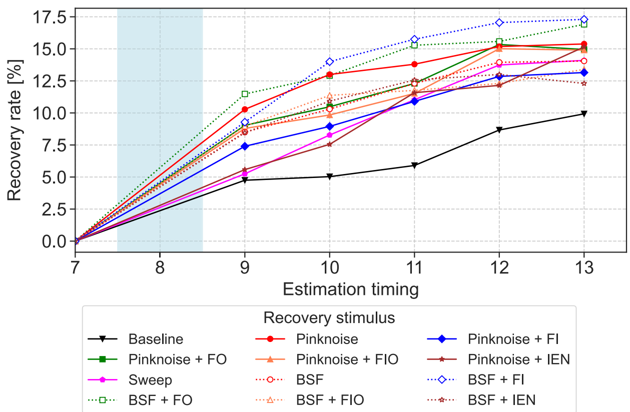


FIGURE 14. Amount of recovery for each stimulus sound.

from loudness reduction is obtained by

$$Rec_t = AC_7 - AC_t \ (t \geq 7). \tag{2}$$

Based on Eq. 1, the loudness reduction rate at the seventh presentation (just before the recovery stimulus sound) and the loudness reduction rate at the tth presentation are calculated. The difference between the two is used to determine the amount of recovery of loudness before and after the recovery stimulus.

Fig. 14 shows the amount of recovery of loudness by listening to each recovery stimulus sound. The horizontal axis indicates the number of times the loudness was estimated, the vertical axis indicates the amount of loudness recovery, and the light-blue area indicates the interval of the recovery stimulus sound presentation. Note that the eighth estimation timing was excluded because it was difficult to estimate only the loudspeaker sound due to the challenge of simultaneously listening to the earphone sound and the loudspeaker sound. As we can see, the loudness tended to recover gradually after the recovery stimulus sound was presented. Compared

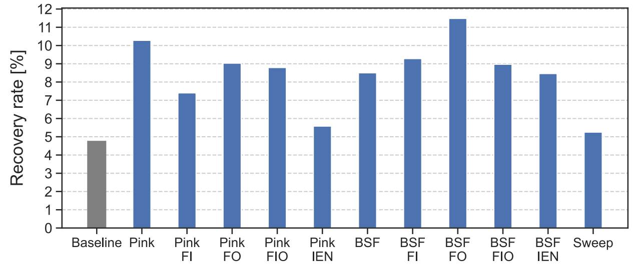


FIGURE 15. Recovery rate for each stimulus sound at the ninth evaluation.

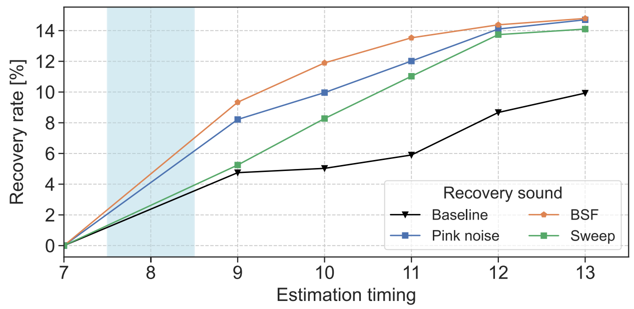


FIGURE 16. Recovery rate for each recovery stimulus sound.

to the case in which no stimulus sounds were presented (baseline), the recovery was greater for all stimulus sounds and estimation timings. Fig. 15 shows the recovery rate for each stimulus sound at the ninth estimation point, where we can see that, for all recovery stimuli, the amount of recovery was greater than the baseline. The stimulus sound with the highest recovery was the BSF-applied sound with a fade-out effect, and the stimulus sound with the lowest recovery was the sweep sound.

Table 6 lists the recovery rate at all recovery sounds. In the baseline, the amount of recovery from the ninth estimation to the thirteenth estimation was 5.2%. In contrast, eight out of 11 recovery stimulus sounds recovered more than 5.2% when the recovery stimulus was presented. In the ninth evaluation, which is immediately after the recovery stimulus, nine recovery stimuli showed a recovery rate more than 1.5 times higher than the case where no recovery stimulus was presented.

Comparing the recovery rate at the thirteenth time for the baseline and the recovery rate at the ninth time when the stimulus sound was presented, we can see that the recovery rate was higher than that of the baseline for the two types of recovery stimulus sound (pink noise and BSF + fade-out). In other words, the recovery stimulus sound causes faster recovery and greater recovery than the baseline.

Fig. 16 shows the recovery rate results summarized by recovery sound. As we can see, BSF had the highest recovery rate at the ninth estimation timing. The recovery rate for the sweep sound remained at the same level as the baseline; however, at the thirteenth estimation timing, the recovery rates converged at the same level for the three recovery sounds. These results suggest that BSF is better for faster and

TABLE 6. Recovery rate (FI: fade-in, FO: fade-out, and FIO: fade-in/out, and IEN: intermittence).

Recovery stimulus		None	Pink noise					BSF					Sweep
Effect		None	None	FI	FO	FIO	IEN	None	FI	FO	FIO	IEN	None
Recovery rate $Rec_t[\%]$	Rec_9	4.7	10.3	7.4	9.0	8.8	5.6	8.5	9.3	11.5	9.0	8.5	5.2
	Rec_{10}	5.2	13.0	9.0	10.5	9.8	7.6	10.3	14.0	12.9	11.4	10.9	8.3
	Rec_{11}	5.9	13.8	10.9	12.3	11.5	11.6	12.3	15.8	15.3	11.8	12.6	11.0
	Rec_{12}	8.7	15.2	12.8	15.3	15.0	12.1	14.0	17.1	15.6	12.3	13.0	13.7
	Rec_{13}	9.9	15.4	13.1	15.0	15.0	15.1	14.1	17.3	16.9	13.4	12.3	14.1
Rec_{13}-$Rec_9[\%]$		5.2	5.1	5.7	5.9	6.1	9.6	5.6	8.0	5.4	4.4	3.8	8.9

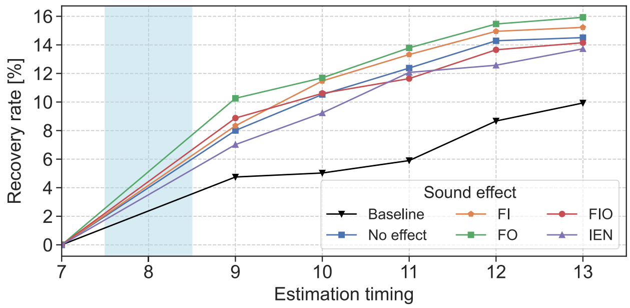


FIGURE 17. Recovery rate for each sound effect.

more effective recovery after loudness reduction, although the recovery rate is eventually comparable regardless of which recovery sound is selected.

Figure 17 shows the recovery rate results summarized by the audio effect. As we can see, while there were variations in the recovery rates for each effect, no significant differences could be confirmed. FO had the highest recovery rate throughout the ninth to thirteenth estimation timings, suggesting that the FO effect is most effective for recovery.

C. DISCUSSION

1) EFFECT OF RECOVERY STIMULUS SOUND

The experimental results in Fig. 15 indicate that, compared to the baseline, the recovery rate was greater for all recovery stimuli at the ninth estimation. This is presumably because the recovery stimulus mitigated the habituation to the first stimulus. In this experiment, since the sound pressure level of the loudspeaker sound and the recovery stimulus sound were the same, we conclude that the recovery was not due to the volume of the recovery stimulus but rather to the effect of hearing a recovery stimulus sound with a different frequency component than the first stimulus.

Although the recovery rate did not reach 100%, this is presumably due to the loudness reduction caused by the larger earphone sound (70-dB SPL) than the loudspeaker sound (60-dB SPL). The presentation of the recovery stimulus improved the loudness reduction between frequencies; however, it was difficult to improve the loudness reduction caused by the difference between the earphone and loudspeaker sound pressure levels themselves.

From Fig. 14, we can confirm that the recovery rate between the eleventh and thirteenth estimations did not change significantly depending on the presence or absence of the recovery stimulus; however, the recovery rate at the ninth and tenth estimations, which were immediately after the recovery stimulus sound was presented, was larger than that of the baseline. The mean recovery rate for the ninth estimation of all stimulus sounds was 8.5%, which is higher than the 4.7% recovery rate for the baseline. This suggests that the recovery stimulus presentation is particularly effective for recovery immediately after a decrease in loudness, and that the participant returns to normal hearing more quickly than if the recovery sound is not presented.

2) LIMITATION

In this experiment, we set a 90-second interval between trials to avoid any potential influence from the previous trial; however, for some participants, the loudness did not return to normal even after 90 seconds, and this may have influenced the next trial. Also, participants had to estimate the magnitude of sound approximately every eight seconds; however, this period may be too short considering that they had to both manually enter a numerical value and correct any input errors. Moreover, focusing on this task may have distracted the participants from properly listening to the loudspeaker sounds. Therefore, more consideration should be given to the interval between trials and task content in the future.

The experiment was conducted assuming an environment in which the earphone sound was louder than the ambient sound. In future work, it will also be necessary to investigate the change in loudness in an environment where the environmental sound level is equal to or louder than the earphone sound level.

Participants in this experiment were asked to estimate the magnitude of the loudspeaker sounds five times after the presentation of the recovery sound; however, it is possible that they needed to estimate the loudness until it returned to normal. Table 6 shows that the BSF + FO sound had the largest amount of recovery immediately after the recovery sound; however, the pink noise + IEN and the sweep sound had large recovery amounts from the ninth to thirteenth recovery times. Therefore, if the estimation was continued after the thirteenth session, the recovery of these two sounds would be the largest and could return to normal in the shortest amount of time. Moreover, these two recovery sounds seemed as though they were presented multiple times to the participant, and this may have influenced the higher recovery between the ninth and thirteenth estimation timings.

The relationship between the amount of recovery and the recovery sound at longer periods should be investigated in the future, taking into account the multiple presentations of the recovery stimulus sound.

Finally, we used pure tones to conduct a basic investigation of loudness changes and the recovery of loudness at specific frequencies; however, in an actual environment, sounds containing multiple frequencies (e.g., voices and music) are likely to be included. In addition, the recovery stimulus sound should not be noise but rather a stimulus sound that provides information to the user while recovering the loudness. Therefore, the reduction and recovery of loudness involving multiple frequencies should be investigated in the future.

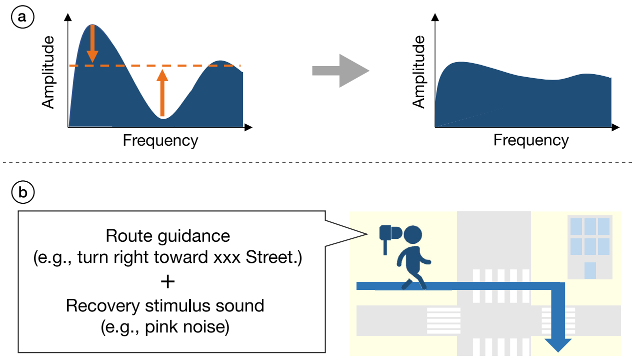


FIGURE 18. **Recovery stimulus sound presentation in applications.**

V. APPLICATION OF FREQUENCY-SELECTIVE LOUDNESS REDUCTION PREVENTION AND RECOVERY METHODS

In Sections III and IV, we examined the occurrence of loudness reduction at selective frequencies in hearables and clarified the effects of methods for recovering from these degradations. In this section, we discuss how to incorporate methods to prevent or recover from these problems into actual applications of hearables.

A. MANIPULATING FREQUENCY OF PLAYBACK SOUND

The results in Section III demonstrate that it is necessary to avoid listening to sounds that are biased only to specific frequencies to prevent loudness reduction at specific frequencies. In a hearable environment, each individual may listen to music with his/her preferred equalization, which may result in a bias toward a specific frequency or may present sounds at a specific biased frequency through manipulation of the external sound frequency [3]. We therefore feel it would be beneficial to de-emphasize only a specific band by using an equalizer (Fig. 18(a)) or to shift the frequency presentation band at regular intervals when a setting biased toward a specific frequency continues for a certain period. These measures will prevent the occurrence of loudness reduction in specific frequencies.

B. PRESENTING RECOVERY STIMULUS AS BACKGROUND SOUND

Although the technique discussed above can prevent loudness reduction in specific frequencies, loudness reduction may occur anyway because it is not always possible to cope with all potential loudness reductions. Therefore, when a loudness reduction occurs, we aim to recover the loudness reduction by playing a recovery sound. Assuming that the hearables are used in daily life, we will discuss the places that can be dangerous if the loudness is reduced. In particular, we consider two main locations where environmental sounds need to be heard.

The first location is in a train station, where the sounds of approaching trains, whistles, and announcements are used to warn people of danger. Therefore, when these sounds are detected by the hearables, the decreased loudness can be recovered by presenting the user with a recovery stimulus sound.

The second location is on roads, especially near intersections. As in train stations, the sounds of approaching cars, horns, and sirens are important environmental sounds. Walking on the side of these roads with a reduced loudness is dangerous and the loudness should always be restored to normal. In such an environment, the user needs to be presented with the recovery sound multiple times, so it is ideal to present the recovery stimulus sound with as little discomfort as possible. One approach would be to use the recovery stimulus sound as the background sound for notifications or announcements from the hearables. We feel this presentation method would prevent danger without inconvenience while simultaneously presenting peripheral information and route guidance (Fig. 18(b)).

VI. CONCLUSION

In this study, we investigated how listening to sounds of specific frequencies through a hearable affects loudness while listening to environmental sounds and explored which recovery sounds are most effective in restoring loudness reduction. Our findings showed that loudness could be reduced by more than 10% after stimulation with all earphone sounds. We also confirmed that the loudness decreased when the frequency of the loudspeaker sound and the earphone sound approximated each other, and that the presentation of the sound of a specific frequency affected the loudness of the specific frequency during environmental sound listening. After the loudness was decreased, we presented a recovery stimulus to participants wearing earphones to investigate its effects on the decreased loudness and confirmed that all tested recovery stimuli were effective in restoring loudness. In particular, the recovery rate immediately after the presentation of the BSF sound with a fade-out effect was approximately 7% greater than that without the recovery stimulus sound. Future studies will include the investigation of changes in loudness caused by speech or music that emphasize or attenuate specific frequency bands, as well as an exploration of the timing of the presentation of recovery sounds that is most effective in restoring loudness.

REFERENCES

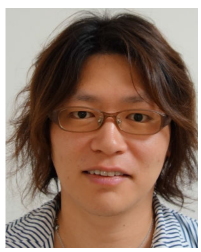
[1] J. Plazak and M. Kersten-Oertel, "A survey on the affordances of 'hearables,'" *Inventions*, vol. 3, no. 3, p. 48, 2018.
[2] P. Crum, "Hearables: Here come the: Technology tucked inside your ears will augment your daily life," *IEEE Spectr.*, vol. 56, no. 5, pp. 38–43, May 2019.
[3] H. Watanabe and T. Terada, "Manipulatable auditory perception in wearable computing," in *Proc. Augmented Hum. Int. Conf.*, Mar. 2020, pp. 1–12.
[4] Athavipach, Pan-ngum, and Israsena, "A wearable in-ear EEG device for emotion monitoring," *Sensors*, vol. 19, no. 18, p. 4014, Sep. 2019.
[5] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Localization*. Cambridge, MA, USA: MIT Press, 1997.
[6] C. G. Le Prell, S. Dell, B. Hensley, J. W. Hall, K. C. M. Campbell, P. J. Antonelli, G. E. Green, J. M. Miller, and K. Guire, "Digital music exposure reliably induces temporary threshold shift in normal-hearing human subjects," *Ear Hearing*, vol. 33, no. 6, pp. e44–e58, 2012.
[7] H. Fletcher, "Auditory patterns," *Rev. Modern Phys.*, vol. 12, no. 1, pp. 47–65, Jan. 1940.
[8] R. D. Patterson, "Auditory filter shape," *J. Acoust. Soc. Am.*, vol. 55, no. 4, pp. 802–809, Apr. 1974.
[9] I. J. Hirsh and W. D. Ward, "Recovery of the auditory threshold after strong acoustic stimulation," *J. Acoust. Soc. Amer.*, vol. 24, no. 2, pp. 131–141, Mar. 1952.
[10] T. O. Iyendo, "Exploring the effect of sound and music on health in hospital settings: A narrative review," *Int. J. Nursing Stud.*, vol. 63, pp. 82–100, Nov. 2016.
[11] Nura. (2023). *Nuratrue—Premium Wireless Earbuds With Our Award-Winning Personalised Sound*. [Online]. Available: https://www.nurasound.com/products/nuratrue
[12] Sony. (2023). *Sony Linkbuds True Wireless Open-Ear Earbuds, White*. [Online]. Available: https://electronics.sony.com/audio/headphones/truly-wireless-earbuds/p/wfl900-w
[13] Jabra. (2023). *Jabra Talk 45 Mono Bluetooth Headphones*. [Online]. Available: https://www.jabra.com/bluetooth-headsets/jabra-talk-45
[14] Y. Wei, G. Yu, X. Li, H. Shuai, W. Ye, Y. Qiao, H. Tian, Y. Yang, and T.-L. Ren, "High performance and wireless graphene earphone towards practical applications," in *Proc. 4th IEEE Electron Devices Technol. Manuf. Conf. (EDTM)*, Apr. 2020, pp. 1–4.
[15] T. Hoshina, D. Fujiyama, T. Koike, and K. Ikeda, "Effects of an active noise control technology applied to earphones on preferred listening levels in noisy environments," *J. Audiol. Otology*, vol. 26, no. 3, pp. 122–129, Jul. 2022.
[16] Y. Yamanobe, M. Fujioka, M. Ohashi, and H. Ozawa, "Potential usefulness of tracking head movement via a wearable device for equilibrium function testing at home," *J. Med. Syst.*, vol. 46, no. 11, p. 80, Oct. 2022.
[17] T. Ando, Y. Kubo, B. Shizuki, and S. Takahashi, "CanalSense: Face-related movement recognition system based on sensing air pressure in ear canals," in *Proc. 30th Annu. ACM Symp. User Interface Softw. Technol.*, New York, NY, USA: Association for Computing Machinery, Oct. 2017, pp. 679–689.
[18] K. Taniguchi, H. Kondo, M. Kurosawa, and A. Nishikawa, "Earable TEMPO: A novel, hands-free input device that uses the movement of the tongue measured with a wearable ear sensor," *Sensors*, vol. 18, no. 3, p. 733, Mar. 2018.
[19] S. Choi, Y. Gao, Y. Jin, S. J. Kim, J. Li, W. Xu, and Z. Jin, "PPGface: Like what you are watching? Earphones can 'feel,'" *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, vol. 6, no. 2, pp. 1–32, Jul. 2022.
[20] E. C. Carterette, "Loudness adaptation for bands of noise," *J. Acoust. Soc. Amer.*, vol. 28, no. 5, pp. 865–871, Sep. 1956, doi: 10.1121/1.1908497.
[21] E. Wagner and B. Scharf, "Induced loudness reduction as a function of exposure time and signal frequency," *J. Acoust. Soc. Amer.*, vol. 119, no. 2, pp. 1012–1020, Feb. 2006.
[22] R. Hellman, A. Miśkiewicz, and B. Scharf, "Loudness adaptation and excitation patterns: Effects of frequency and level," *J. Acoust. Soc. Amer.*, vol. 101, no. 4, pp. 2176–2185, Apr. 1997.
[23] M. C. Botte, G. Canévet, and B. Scharf, "Loudness adaptation induced by an intermittent tone," *J. Soc. Am.*, vol. 72, no. 3, pp. 727–739, Sep. 1982.
[24] S. Charron and M.-C. Botte, "Frequency selectivity in loudness adaptation and auditory fatigue," *J. Acoust. Soc. Amer.*, vol. 83, no. 1, pp. 178–187, Jan. 1988.
[25] W. D. Ward, A. Glorig, and W. Selters, "Temporary threshold shift in a changing noise level," *J. Acoust. Soc. Amer.*, vol. 32, no. 2, pp. 235–237, Feb. 1960.
[26] W. D. Ward, "Recovery from high values of temporary threshold shift," *J. Acoust. Soc. Amer.*, vol. 32, no. 4, pp. 497–500, Apr. 1960.
[27] J. H. Mills, J. D. Osguthorpe, C. K. Burdick, J. H. Patterson, and B. Mozo, "Temporary threshold shifts produced by exposure to low-frequency noises," *J. Acoust. Soc. Amer.*, vol. 73, no. 3, pp. 918–923, Mar. 1983.
[28] D. W. Bell and G. Fairbanks, "TTS produced by low-level tones and the effects of testing on recovery," *J. Acoust. Soc. Amer.*, vol. 35, no. 11, pp. 1725–1731, Nov. 1963.
[29] Y. Arieh, K. Kelly, and L. E. Marks, "Tracking the time to recovery after induced loudness reduction (L)," *J. Acoust. Soc. Amer.*, vol. 117, no. 6, pp. 3381–3384, Jun. 2005.
[30] B. Scharf, M.-C. Botte, and G. Canévet, "Récupération après adaptation induite de sonie," *Annee Psychol.*, vol. 83, no. 1, pp. 9–24, 1983.

HIROKI WATANABE received the B.E., M.E., and Ph.D. degrees in engineering from Kobe University, Hyogo, Japan, in 2012, 2014, and 2017, respectively.

From 2015 to 2017, he was a Research Fellow with Japan Society for the Promotion of Science (DC2), Kobe University. Since 2017, he has been an Assistant Professor with the Graduate School of Information Science and Technology, Hokkaido University, Hokkaido, Japan. His research interests include wearable computing and ubiquitous computing.

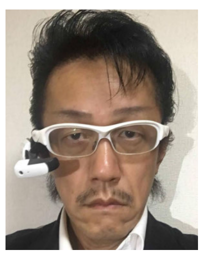
SOTA KANEMOTO received the B.E. degree in engineering from Ritsumeikan University, Shiga, Japan, in 2021. He is currently pursuing the master's degree with the Graduate School of Engineering, Kobe University, Hyogo, Japan.

His research interests include wearable computing and ubiquitous computing.



TSUTOMU TERADA (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in engineering from Osaka University, Osaka, Japan, in 1997, 1999, and 2003, respectively.

From 2000 to 2004, he was an Assistant Professor with the Cybermedia Center, Osaka University, where he was a Lecturer, from 2005 to 2007. From 2007 to 2018, he was an Associate Professor with the Graduate School of Engineering, Kobe University, Hyogo, Japan, where he has been a Professor, since 2018. His research interests include wearable computing, ubiquitous computing, and entertainment computing.



MASAHIKO TSUKAMOTO (Member, IEEE) received the B.E., M.E., and Ph.D. degrees in engineering from Kyoto University, Kyoto, Japan, in 1987, 1989, and 1994, respectively.

From 1995 to 1996, he was a Lecturer with the Graduate School of Engineering, Osaka University, Osaka, Japan, where he was an Associate Professor, from 1996 to 2004. Since 2004, he has been a Professor with the Graduate School of Engineering, Kobe University, Hyogo, Japan. His research interests include wearable computing and ubiquitous computing.