Auditory space perception during linearly self-motion

Shinichi Sakamoto, Zhenglie Cui, Yōiti Suzuki
Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ku, Sendai, 980-8577, Japan

Wataru Teramoto
Department of Computer Science and Systems Engineering, Muroran Institute of Technology, Muroran, 050-8585, Japan

Jiro Gyoba
Graduate School of Arts and Letters, Tohoku University, Sendai, 980-8576, Japan

Summary
Spatial inputs from the auditory periphery can be changed with listener’s various movements relative to the sound source. Nevertheless, humans can perceive a stable auditory environment and appropriately react to a sound source. This suggests that the inputs are reinterpreted in the brain, while being integrated with information on the movements. Little is known, however, about how these movements modulate auditory perceptual processing, especially under linearly moving environment. We investigate the effect of linearly self-motion on auditory space representation. Results of our experiments showed that the sound position judged as being located at the subjective coronal plane was displaced compared with the the listener’s physical coronal plane. This distortion was observed irrespective of the style of the self-motion (active or passive). Moreover, the distortion was observed both with and without vestibular stimulation. These results suggest that self-motion information, irrespective of its origin (vestibule or any other sources), is crucial to evoke this distortion of auditory space perception.

PACS no. 43.66.+y, 43.66.Qp

1. Introduction

We can perceive auditory spatial information stationarily, even when we are moving, which implies that inputs from the auditory periphery are interpreted in the brain by integrating them with information on the movements of the head and whole body. Such movement signals used for sound localization can be derived from vestibular information [1].

Several reports have described the influence of the vestibular semicircular canal signals on auditory localization. Although a few reports of some studies have described improvements in sound localization by active and passive head rotations with low angular displacement amplitudes [2, 3, 4, 5], most preceding studies have demonstrated large systematic errors, rather than improvements, especially during rapid head motion [6, 7]. These findings suggest that the vestibular semicircular-canal system plays an important role in space perception.

Aside from information originating in the semicircular-canal system, sensory information from the macular receptors of the otolith system (utricle and saccule) might also play a role in this respect. The otolith system can detect linear acceleration, whereas the semicircular-canal system can detect rotatory acceleration. Although some studies suggest that information from the otoliths, as well as the semicircular canals, influence the auditory localization/lateralization in azimuth, it is not clear how auditory space representation in depth is modulated by linear motion.

We have focused on the effect of linearly self-motion perception on auditory space representation. As the factor of the self-motion perception, vestibular information and the style of the motion (active/passive) are considered. In this paper, we reviewed our recent research results related to this topic and discuss the mechanism how people perceive outside sound environment during motion.
2. Distortion of auditory representation by passive movement [8]

This research was undertaken to investigate how the sound position aligned with the subjective coronal plane (SCP) was displaced, while manipulating the direction of self-motion (forward or backward) and its acceleration. The coronal plane divides a body vertically into anterior and posterior sections.

2.1. Experimental Apparatus

The experiments took place in a corridor in the Research Institute of Electrical Communication, Tohoku University (Fig. 1(a)). Sound-absorbing materials were placed on the sidewalls in the part of the corridor in which the experiments were conducted (about a 3-m section) to attenuate the sound reflections. The observers were transported using a robotic wheelchair (iXs Research Corp., Fig. 1(b)). The experimenters had exclusive wireless control over the movements of the wheelchair. The maximum A-weighted sound pressure level of ambient environmental noise, including noise from the wheelchair, was 60 dB while the wheelchair was in operation.

Auditory stimuli were presented using 17 full-range loudspeakers (30 mm, 0254-7N101; HOSIDEN Corp.) installed in small cylindrical plastic boxes (108 cm³). These loudspeakers were on the right-hand side, aligned with the direction of movement of the wheelchair at 10-cm intervals and at a height of 1.32 m (almost equivalent to the height of the seated participant’s ears). The auditory stimulus was presented at the moment the wheelchair intersected an orthogonal laser set beside the baseline (Fig. 1(a)). Specifically, analog signals from the laser were converted to digital signals using a data acquisition device (NI USB-6289, National Instruments Corp.) connected to a laptop computer. The inputs were processed using a LabVIEW program (National Instruments Corp.). The audio data were output through audio interfaces (UA-25EX and Marantz, PM-54DS; Roland Corp.). The system delay from sensing the position of the wheelchair to the onset of the auditory stimulus was 3 ms or less.

2.2. Experimental Procedure

This experiment examined data of eight participants (21–38 years; one woman, seven men). All participants had normal hearing with no history of vestibular deficiency.

The five sessions comprised two with forward motion (0.2 m/s² and 0.4 m/s²), two with backward motion (0.2 m/s² and 0.4 m/s²), and one with no motion (baseline). The sound was presented when the chair moved 2.0 m and 1.0 m, respectively, with acceleration of 0.2 m/s² and 0.4 m/s². Consequently, the wheelchair velocity when the sound was presented was 0.9 m/s, irrespective of the acceleration. The auditory stimulus consisted of 30 ms of pink noise modulated by 5 ms raised-cosine onset and offset windows at an average sound pressure level of 80 dB (44.1 kHz sampling frequency).

Each blindfolded participant sat on the wheelchair with the head fixed to the wheelchair with an elastic band. They were asked to answer whether they perceived the sound “backward” or “forward” relative to their coronal plane (i.e., a two-alternative forced-choice task). A test sound was presented from one of the loudspeakers when the chair reached a particular location (baseline). The baseline was a point aligned with the physical coronal plane (i.e., the interaural axis) at the moment a test sound was delivered. The distance was defined as the physical distance between the baseline and test sound. The physical test sound position varied among trials according to a staircase method. The test sound position ranged from −80 cm to 80 cm in 10-cm intervals (see Fig. 1; the negative and positive values respectively denote the rear and frontal spaces). In one sequence, the initial position of the sound was 80 cm from the baseline (descending series). In another sequence, the initial position was −80 cm (ascending series). These two staircase sequences were intermixed randomly. The step size of the staircase was 10 cm. Each staircase sequence was terminated after five reversals of the response sequence. Consequently, 10 reversals were obtained from these two staircase sequences in each session. They were averaged to obtain the alignment of the sound position with the SCP.

2.3. Results

The mean sound positions aligned with the participants’ SCPs are presented as a function of acceleration in Fig. 2. Negative and positive values respectively indicate the rear and frontal spaces. A repeated-measures analysis of variance (ANOVA) with one within-participant factor (two forward and two backward motions (±0.2 m/s² and ±0.4 m/s²), and no motion conditions) revealed a significant effect of the experimental condition (F(4, 28) = 9.88, p < .01). A
multiple comparison (Tukey’s HSD method, $a < .05$) revealed that the mean sound positions aligned with the participants’ SCPs moved forward considerably in the direction of self-motion with an increase in acceleration for the forward motion conditions, although no effect was observed for the backward motion conditions.

3. Effects of active and passive movement on the distortion of auditory representation [9]

In the research written in section 2, participants’ movement was controlled by the experimenter. Therefore, participants were unable to know when and where they were moving. This movement is regarded as “passive” movement. However, it is unknown whether the difference between active and passive movement affects observed phenomena. Several studies have suggested that whether the listener’s motion is active or passive affects sound localization [10]. In this research, perceived auditory space during walking was asked to the participant and the effects of active and passive movement on the distortion of auditory representation were investigated.

3.1. Experimental Apparatus

The same wheelchair and speaker array in the previous research were used to provide acceleration and auditory information to participants. To reduce the reflection from a seat back of the chair, the seatback was removed. A participant’s head was fixed by a small and thin headrest. To record the movement during walking, position sensor (TX-4; FASTRAK) was set on the participant in active motion condition.

3.2. Experimental Procedure

Six participants participated in this experiment. All had normal hearing with no history of vestibular deficiency.

Three conditions were no motion (reference), active motion, and passive motion conditions. In the active motion condition, blindfolded participants were asked to walk straight ahead at the velocity of 0.3 m/s. The movement was recorded at the sampling rate of 16 Hz. Then, recorded movements was approximated by logistic function. In the passive motion condition, participants were transported forward by a robotic chair. One of the recorded movements was selected randomly and used as the passive motion. The auditory stimulus consisted of 30 ms of pink noise modulated by 5-ms Hanning onset and offset windows at an average sound pressure level of 80 dB (sampling frequency: 44.1 kHz). In each condition, they were asked to answer whether they perceived the sound “backward” or “forward” relative to their coronal plane (i.e., a two-alternative forced-choice task). A test sound was presented from one of the loudspeakers when the body (active condition) or chair (passive condition) reached a particular location (baseline). The distance was defined as the physical distance between the baseline and test sound. The test sound position ranged from $-80$ cm to $80$ cm in 10-cm intervals. The actual sound position varied from trial to trial according to a randomized maximum likelihood sequential procedure.

3.3. Results

The mean sound positions aligned with the participants’ SCPs are presented in Fig. 3. The baseline denotes a sound position aligned with the participants’ physical coronal plane. Negative and positive values respectively denote the rear and frontal spaces. Repeated-measures analysis of variance (ANOVA) with one within-participant factor (no motion, active motion and passive motion conditions) revealed a significant effect of the experimental condition ($F(2, 10) = 17.62, p < .05$). A multiple comparison (Ryan’s method, $p < .05$) revealed that the mean sound positions aligned with the participants’ SCPs significantly moved backward in the opposite direction of self-motion for the active and passive motion conditions compared with SCP in the no motion condition. No significant difference was found between active and passive motion conditions.

4. Effects of visually-induced self-motion on the distortion of auditory representation

Aside from information originating in the vestibular system, the visual system could also play an important role in self-motion perception. Large-field visual
motion can induce the sensation of self-motion (vection). Previous studies reported that rotation of a visual environment around the vertical axis caused a displacement of sound source in the direction of visual motion (i.e., in the opposite direction of induced self-motion) [11]. This is in line with the results from the semicircular stimulation.

The present research investigated whether visual stimulation that simulates linear self-motion affects sound localization in depth. It is possible that a difference in the origin of self-motion information could produce different results.

### 4.1. Experimental Apparatus

The experiment was conducted in a sound attenuated room in the Research Institute of Electrical Communication, Tohoku University. All visual stimuli were projected with a projector (P’DG-DHT100JL; SANYO; refresh rate: 60 Hz; resolution: 1,280×1,024) on a 150-inch projection screen. Participants were seated 1.48 m from the center of the screen (field of view: 90°×78°) with their head fixed to the chair with a headrest. An array of 11 full-range loudspeakers (30 mm, 0254-7N101; HOSIDEN) was placed perpendicular to the screen (i.e., parallel to the anterior-posterior axis of the human body) along the wall to the left of the participants at a height of 1.35 m (almost equivalent to the height of the seated participant’s ears). The central speaker of the array was placed 1.48 m lateral to the participants on their physical coronal plane. This speaker was defined as the baseline (0°).

The angles between the remaining speakers and the center of participants’ head were ±4°, ±8°, ±16°, ±24°, and ±32°. Negative and positive values indicate the rear and frontal space, respectively. Audio data were output through audio interfaces (HDSP MADI & M32-DA; RME) using a power amplifier (MP-3016; Mishima Planning).

### 4.2. Experimental Procedure

Twelve participants (21–40 years; three females, nine males) participated in the experiment. All participants had normal or corrected-to-normal vision, normal hearing, and no vestibular dysfunction.

A random-dot pattern simulating either forward or backward linear self-motion or no motion was displayed on the screen as the visual stimulus. Auditory stimuli consisted of 30 ms of pink noise modulated by 5-ms raised-cosine onset and offset windows at an average sound pressure level (A-weighted) of 54 dB (sampling frequency: 44.1 kHz).

There were two directions (forward and backward) of simulated self-motion and two accelerations (0.15 and 0.3 m/s²) for each direction. The two self-motion directions were tested in different blocks. In each block, the no motion session, where the static dot pattern was presented, was conducted first, followed by the two acceleration sessions (0.15 and 0.3 m/s²). The order of acceleration conditions was randomized. Two forward and two backward self-motion blocks were conducted in random order. Each experimental session consisted of a number of trials for the staircase procedure.

At the beginning of each trial, the static random-dot pattern was presented with a fixation point. When the participants pressed a button of the gamepad, the trial was started. In the no motion condition, a target sound was presented 1 s after the participants’ button press, and participants made a response. In the forward and backward conditions, after the button press, the random-dot pattern started to move. The initial velocity of simulated self-motion was 0.4 m/s (constant). As soon as participants reported self-motion perception, the velocity of simulated self-motion increased at a constant acceleration of either 0.15 m/s² or 0.3 m/s². A target sound was presented when the velocity of simulated self-motion reached 1.5 m/s (i.e., 6.0 s and 3.0 s after the acceleration in the 0.15 m/s² and 0.3 m/s² acceleration conditions, respectively). One second after the presentation of the target sound, the visual stimuli disappeared and participants indicated the direction in which the sound was perceived (front or back) relative to their coronal plane.

The test sound position varied from trial to trial according to a staircase method. In one staircase sequence, the initial position of the sound was 32° (descending series), and in another staircase sequence the initial position was −32° (ascending series). These two staircase sequences were randomly intermixed in a session. Each staircase sequence was terminated after 5 reversals of the response sequence. Thus, 10 reversals were obtained in a session. Because two sessions were conducted for each self-motion condition, 20 reversals were averaged to obtain the alignment of the sound position with the subjective coronal plane (SCP).
4.3. Results

Figure 4 shows mean sound positions aligned with participants’ SCP. The baseline point indicates a sound position aligned with participants’ physical coronal plane, and negative and positive values indicate the rear and frontal spaces, respectively. For each self-motion direction, a repeated-measures analysis of variance (ANOVA) with one within-participant factor (no motion, 0.15 m/s², and 0.3 m/s²) was performed for the sound localization data. For the backward condition, a significant effect of experimental condition ($F(2, 22) = 6.59, p < .006$) was observed. Multiple comparisons (Tukey’s HSD, $\alpha < .05$) revealed that the magnitude of mislocalization increased as acceleration increased. However, no effect of experimental condition was observed in the forward motion condition ($F(2, 22) = 2.58, p = .098$).

5. Discussion

In a series of these studies, we demonstrated that the sound position aligned with the SCP was displaced compared with the baseline irrespective of the style of the self-motion (active or passive); irrespective of the existence of vestibular stimulation. These results suggest that there is a certain link between self-motion perception and auditory space representation.

However, the direction of the self-motion was different among the researches when such SCP shift was observed. When vestibular information was presented via robotic wheel chair or under active motion condition, SCP shift was observed only during forward-motion. In contrast, when visual information was presented, such shift was observed only backward-motion. Such type of asymmetry is observed in other aspects of auditory perception [12, 13, 14]. Furthermore, human neuroimaging reports have described that auditory looming stimuli preferentially activate a neural network serving space recognition, auditory motion perception, and attention [14, 15]. Studies on linear vection have also reported an asymmetry in vection strength between forward and backward self-motion conditions: Stronger self-motion is induced by a contracting versus expanding flow patterns [16, 17]. We speculate that a closer link might be formed between the vestibular processing for forward self-motion and auditory space perception while between the visual processing for backward self-motion. The visual system has to discriminate between object motion and self-motion from retinal inputs. In particular, detecting approaching objects is essential for survival. Therefore, the link between visual information which is inducing forward motion might be connected to the object movement, not the self-motion. After detecting approaching objects, people would consider whether they have to avoid to obstacles, or not. This might be the reason why vestibular information have closer link to the forward motion.

6. Conclusion

In this study, we investigate the effect of linearly self-motion on auditory space representation. Results of a series of our studies showed that the sound position judged as being located at the subjective coronal plane was displaced compared with the the listener’s physical coronal plane. Moreover, this distortion was observed irrespective of the style of the self-motion (active or passive); irrespective of the existence of vestibular stimulation. These results suggest that self-motion information, irrespective of its origin (vestibule or any other sources), is crucial to evoke this distortion of auditory space perception.

Acknowledgement

We thank Jennifer Campos, William Martens, Denis Cabrera, Funimasa Furune, Hideaki Terashima,
Kazuki Moishi, Yukio Iwaya, Souta Hidaka, and Fumitaka Saito for their helpful comments and technical advice. This research was supported by a Grant-in-Aid for Specially Promoted Research (No. 19001004), Young Scientists (B) (No. 21730585) and Grant-in-Aid for Scientific Research (B) (No. 26280067) from the Ministry of Education, Culture, Sports, Science and Technology.

References