Other's gaze direction affects sound localization

Haru Sukemiya*, Takahiro Kawabe, Yuki Yamada and Kayo Miura

**Faculty of Human-Environment Studies, Kyushu University, 6-19-1, Hakozaki, Higashi ward, Fukuoka city, 8128581, Japan*

E-mail:sukemiya@psycho.hes.kyushu-u.ac.jp

This study investigated the effect of visual cue of other's gaze on sound localization in two experiments. In Experiment 1, by using the manual pointing task to a sound source, we examined the accuracy of sound localization when target sound and gaze cue were presented simultaneously. Results showed that perceived location of sound source was displaced in the gaze direction, when target sound was presented in the gazed spatial field. In Experiment 2 manipulating stimulus onset asynchrony (SOA) between onsets of target sound and gaze, we showed that perceived location of sound source was biased toward gaze direction at short SOA of less of 300 ms, regardless of whether or not the gaze cue appeared ahead of target sound. The results in Experiment 1 and 2 indicate that the shift of visuospatial attention caused other's gaze cue affects sound localization.

Key words : *Sound localization, Other's gaze direction, Visuospatial attention, Visuo-audio interaction*

Introduction

In social communication, we can share an object of attention together with other person by utilizing other's or mutual gaze signals.

Recently, some behavioral studies have indicated that the visuospatial attention can shift to the position (or direction) where someone else is looking, even when gaze direction is not a good predictor as to where a visual target may be present (e.g., Driver et al., 1999; Friesen et al., 2004). In addition, this attentional shift seems to affect detection, localization and identification of visual target. For instance, Driver et al. (1999) measured reaction time (RT) needed to detect a visual target, when a real face looking to the left or right and visual target were presented successively with some target stimulus onset asynchrony (SOA). There were 2 trial types: first was that visual target was presented at gaze direction, and second was the opposite case. In results sampled RT performance at 100, 300, and 700 ms cue-target SOAs, RT was longer and there was no effect of gaze direction between trial types at short, 100 ms SOA. At 300 ms SOA, RT was shorter when target was presented in the gazed visual field, however, RT increased at 700 ms SOA when visual target was presented in the ungazed visual field, in the situation that subjects was preliminarily informed where target was presented. This effect of gaze direction seems to appear even at longer SOA of 1200 and 1800 ms (Friesen et al., 2004). They suggested that attention shifted reflexively to gazed position at short SOA of 300 ms SOA, while it shifted voluntary at long SOA of over 700 ms.

Although these studies have discussed the effect of gaze direction on detection speed for the target, it has not been investigated whether or not other's gaze affects audiospatial localization of target yet. In this study, we investigated the effect of gaze cue on sound localization and its temporal characteristic, by using manual pointing task to a sound source.

Experiment1

In this Experiment, we examined the accuracy of sound localization by using the manual pointing task to target sound which was presented simultaneously with gaze cue.

Method

Subjects. Six subjects (4 female and 2 male) participated in, those who were all university undergraduate and graduate students, and had normal hearing as well as vision and right handedness.

Apparatus and Stimuli. Fig. 1 illustrates the layout of apparatus. The auditory stimulus was white noise sound (60dB) and presented from one of four loudspeakers, located at 15° and 30° to the left or the right with respect to subject's median plane. The face display was presented on screen of personal computer

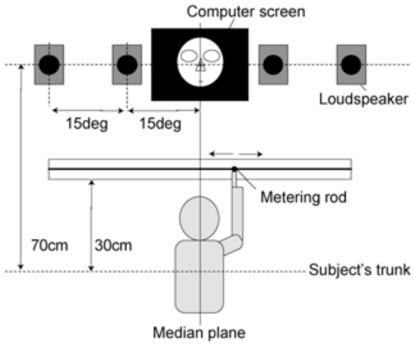

Figure 1. Set of experimental apparatus in top view.

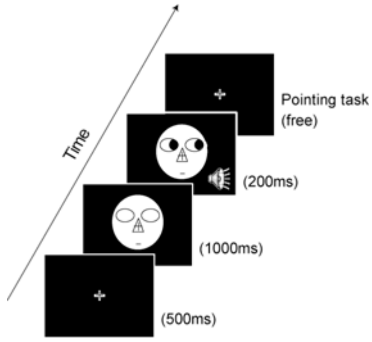

Figure 2. Illustration of the trial sequence in Experiment 1.

(28.5 cm width × 21.5 cm height). The face display consisted of a black line drawing of a face presented on a skin color background. The round face outline subtended 15° height and 10° width, and the fixation point was presented at the center of display. Loudspeakers and screen of personal computer were located along horizontal line with 70 cm away from subject's trunk, with their physical center positions coincided with eye height of subjects. The perceived location of target sound was record by the metering rod, which allows subjects to move by their index finger along the horizontal line, arranged at 30 cm away from their trunk. Because this and next experiment were conducted in complete dark laboratory, with controlling acoustic echo, subjects could not utilize any visual information except the fixation and the face display.

Procedure. Subjects sat on a comfortable chair, and their head were fixed by chin-head rest. Before starting each trial, subjects put their right index finger on the metering rod and closed their eyes. After subjects opened their eyes and looked at the fixation point, the face display without black eyes was presented behind the fixation point for 1000 ms. And then, black eyes as the gaze cue and the target sound were presented simultaneously for 200 ms (see Fig. 2). Subject's task was to move the metering rod in the direction of target sound with respect to their body align with that of the metering rod, without removing their eyes from the fixation point, after the face display and target sound were disappeared.

Design. Two experimental conditions were employed: target sound locations (L30° ,L15° ,R15° and R30°) and trial types. There were three trial types: consistent or inconsistent gaze cues with the target and no gaze cue as control. Pointing responses toward each target location was repeated 6 times for each of trial types, and a total of 72 trials were performed in pseudo-random order.

Results and Discussion

We computed the magnitude and direction of displacement between physical target sound direction and pointed direction with respect to subject's trunk as visual angle. For each of consistent/inconsistent conditions, difference from error for control condition was calculated. To indicate the direction of pointing error, we used the sign "-" and "+" for leftward and rightward displacements, respectively. Mean pointing errors and standard deviations at every target sound locations for each of trial types are shown in Fig. 3.

A 2-way ANOVA was conducted with trial types (target consistent/inconsistent with gaze direction) and target sound locations (L30° , L15° , R15° and R30°) as within subject variables. The ANOVA revealed that the trial types × target sound locations interaction was

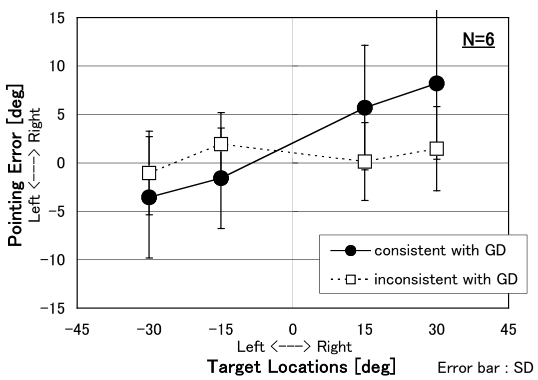

Figure 3. Mean pointing error and standard deviation at each target locations, for the consistent and inconsistent with gaze direction condition. Error bars show standard deviation.

significant (F (3, 15) = 3.90, $p < .05$). Specifically, there were significant tendency and difference between the consistent and the inconsistent condition in R15° (p = 0.07) and R30° ($p < .05$). Additionally, there were significant differences between L30° and R15°, R30°, and between L15° and R15°, R30° in the consistent condition ($p < .05$ in all), though there was no significant difference in the inconsistent condition. These results show that the perceived location of sound shifts toward the other's gaze, when a target sound presents on an extension of gaze, though this effect disappears on the opposite side of gaze direction. It indicates that the shift of visuospatial attention caused by observation other's gaze affects sound localization and that the shift of visual spatial attention can affect auditory as well as vision localization performance.

Experiment2

In this Experiment, we manipulated SOA and examined the temporal characteristic of the effect of gaze cue on sound localization.

Method

Subjects. Four university undergraduate and graduate students (3 female and 1 male), with normal hearing and vision and with right handedness, participated this experiment.

Apparatus, Stimuli and Procedures. Apparatus, stimuli and procedure were almost identical to those in Experiment 1 (see Fig. 1, 2) except following: In this Experiment, a face display without eyes was presented in advance. Depending on SOAs, either of gaze cue or tone target was presented first. The eyes were disappeared simultaneously with the face display. Subject's task was to point to the direction of target sound immediately after target sound was disappeared.

Design. As like in Experiment 1, target sound locations (L30°, L15°, R15° and R30°) and trial type (consistent/inconsistent with gaze direction and no gaze cue as control) were employed as experimental conditions. Additionally, there were 6 types of SOA conditions: in half conditions, gaze cue was presented before target sound (-700, -300 and -100 ms SOAs) and in the half, opposite was true (100, 300 and 700 ms SOAs). For each of 6 SOA conditions, pointing responses toward target sound locations were repeated 6 times for each of trial types, and a total of 288 trials were performed in pseudo-random order.

Results and Discussion

The calculation procedure of pointing error was identical to Experiment 1. Mean pointing errors and

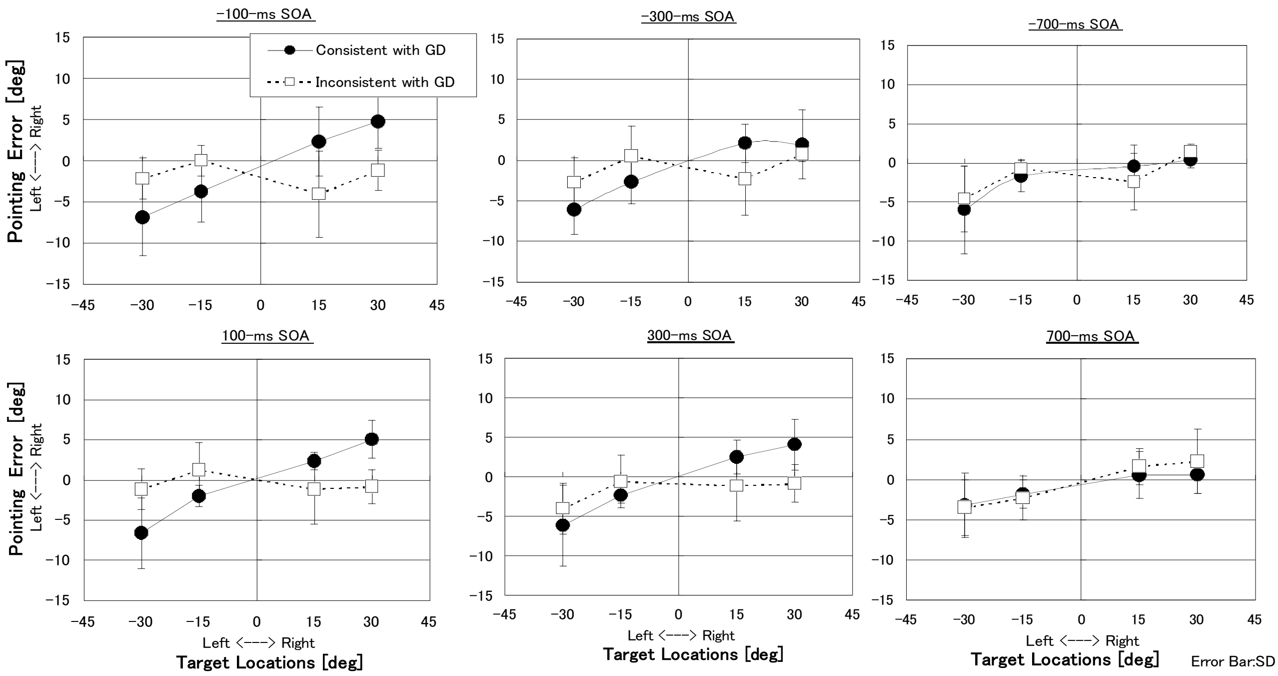

Figure 4. Mean pointing error and standard deviation at each of target locations, for each of SOA conditions. Block and white circles show pointing error in the condition where the visual field of target was consistent and inconsistent with gaze direction. Error bars show standard deviation.

standard deviations for each of SOA conditions are presented in Fig. 4.

For each of SOAs, 2-way ANOVAs were conducted with trial type (target consistent/inconsistent with gaze direction) and target sound locations (L30°, L15°, R15° and R30°) as within subject variables. The ANOVA revealed that similarly to Experiment 1, the trial types × target sound locations interactions were significant at the SOA of -100 ms ($F(3, 9) = 14.487$, $p < .001$), 100 ms ($F(3, 9) = 10.49$, $p < .005$) and 300 ms ($F(3, 9) = 9.05$, $p < .005$), and marginally significant at -300 ms SOA ($F(3, 9) = 3.04$, $p = .085$). In other words, there were significant differences between consistent and inconsistent conditions at the SOA of -100 and 100 ms in all target locations ($p < .05$ in all), and significant tendency at -300 ms SOA ($p < .10$). These results showed that perceived location of sound was displaced toward gaze direction in consistent condition at short SOAs, compared with inconsistent condition. In contrast, there were no significant interactions between trial types and target sound locations at the -700 and 700 ms SOAs ($p > .10$).

In sum, results of this experiment indicate that other's gaze as visual cue effectively work to displace the sound localization when temporal separation between visual and auditory events was shorter than 300 ms, regardless of whether or not visual gaze cue preceded the presentation of auditory target.

General Discussion

This study investigated the effect of other's gaze on sound localization in two Experiments. In Experiment 1, aiming to examine whether or not other's gaze affected the accuracy of sound localization, we showed that the perceived location of sound was displaced toward the other's gaze, when a sound and gaze cue were presented simultaneously (0 ms SOA). This indicates that the shift of visuospatial attention caused other's gaze cue affects sound localization. Driver et al. (1999), who reported that reaction time (RT) of detection of visual target were faster when visual target was presented at the gazed than at the non-gazed locations at short SOA of less than 600 ms, concluded that this could result from reflexive attentional shift along gaze direction. Although they discuss the effect of attentional shift by other's gaze on visual perception, our results presented new evidence that the visuospatial attentional shift affects the sound localization. This indicates that the spread of spatial attention can occur across auditory as well as visual spatial representation even when only vision was cued.

Additionally, Experiment 2 with the manipulation of SOAs revealed that the shift of perceived sound location arose only at short SOA of less than 300 ms, regardless of whether or not gaze cue preceded the presentation of sound. In contrast, this tendency seems to disappear in such long SOAs as 700 ms. These results indicate that gaze cue affecting sound localization can effectively work within a short temporal separation (i.e. within 300 ms) between visual and auditory events. This tendency is similar to the results of Driver et al. (1999) and Friesen et al., (2004). In conjunction with these studies and our results, temporal separation required for gaze cue to be active may be similar or common in both vision and auditory.

On the other hand, there was an alternative explanation with respect to results of Experiment 2. In this Experiment, subjects started the pointing movement after target sound disappeared. In other words, gaze cue appeared during pointing movement when followed the target sound. Some previous studies have suggested that the minimum delay needed for a visual signal to influence an ongoing movement in 80-100 ms, and that for the duration of visually-directed movement, as pointing and reaching movement, is typically 300-700 ms (e.g., Desmurget. & Grafton., 2000). Considered gaze cue affects during pointing movement, it may not be reflected in pointing performance when it presented after the onset of pointing movement, even if audiospatial representation was affected by gaze cue. To solve this problem, it may need to investigate the effect of the onset of pointing on localization performance.

References

Desmurget, M. & Grafton, S. (2000). Forward modeling allows feedback control of fast reaching movements. *Trends in Cognitive Science*, **4**(11), 423-431.

Driver, J., Davis, G., Ricciardelli, P., Kidd, P., Maxwell, E. & Baron-Cohen, S. (1999). Gaze perception triggers reflexive visuospatial orientating. *Visual Cognition*, **6**(5), 509-540.

Friesen, C.K., Ristic, J. & Kingstone, A. (2004). Attentional effects of counterpredictive gaze and arrow cues. *Journal of Experimental Psychology: Human Perception and Performance*, **30**(2), 319-329