# Measurement of attention to auditory signal in noisy environment

Hiroshi SATO[1]

[1] National Institute of Advanced Industrial Science and Technology

1-1-1 Higashi, Tsukuba, Ibaraki 305-8566, Japan

## ABSTRACT

When people find or localize a target sound in noisy environment, he/she employ his/her attention to the sound. If we can measure degree of attentions to a sound in daily life situation, the results can be used to control and/or design sound environment as well as auditory sign. However, it is difficult to measure auditory attention especially in real life situation.

The intention of this study is to develop a tool to be used in real public space for evaluating performance of acoustic guide signal to present target directions. The challenge of this study is to measure the degree of attention to a sound with gaze measurement and presenting the relation between S/N and gazing detection.

Keywords: sound localization, auditory attention, visual attention, gaze

## 1. INTRODUCTION

"Cocktail party effect," which is one of the most famous auditory event in the field of hearing and cognitive science (1), is closely related to attention of human being (2). People can distinguish target sound in noise, pay attention to the target sound and catch up the message of the sound. Once we can detect the target sound, we tend to rotate our head and to look at the sound source (e.g. face of a talker).

One of the important application of sound localization is an acoustic guide signal in public facilities for visually impaired person as well as normal person (3). Acoustic guide signals are designed to lead visually impaired persons to their respective destination points. It is often used at ticket gates in railway stations in Japan because the government encourages to use it by publishing the guideline (4) and Japanese Industrial Standard (5).

Authors revealed effects of noise and reverberation on sound localization of the Japanese standard signal in laboratory and suggests following three factors (3);

1) Effects of noise on sound localization can be presented as a function of the signal-to-noise ratio.
2) The initial delay time and reverberation energy are more important than the reverberation time.
3) A temporal pattern of a guide signal is an important factor for sound localization in reverberant sound fields.

Although these factors are controllable in laboratory studies, it is quite difficult to evaluate effectiveness of acoustic guide signal in real public spaces to grab pedestrians' attention and to guide them to target destinations.

The intention of this study is to develop a tool to be used in real public space for evaluating performance of acoustic guide signal to present target directions. This new study challenges to employ measurement of visual attention to evaluate auditory attention. The challenge also includes dynamic evaluation while walking in public space.

[1] sato.hiro@aist.go.jp

In this study, degree of attention to a sound in noise was measured in a laboratory by gaze measurement and the relationship between S/N and gaze detection will be discussed.

## 2. GAZE MEASUREMNT TO JUDGE POSITION OF SOUND SOURCE IN NOISY SOUND FIELDS

### 2.1 Apparatus and stimuli

A "Ping-Pong" sound, which is used in railway stations in Japan (4), and pink noise burst (duration: 1 s), which gives good sound localization to a listener, were used as a stimuli in the experiment. Figure 1 presents the temporal pattern of the sound pressure of the "Ping-Pong" signal. The signal duration was approximately 4 s. The first half part of signal was a 770 Hz triangle wave. The latter part was 660 Hz. A triangle wave has high-level harmonics components at higher frequencies.

Presentation levels of signals and noise were determined at the listener's position. Signals were presented 50 dB and 38 dB in A-weighted maximum sound pressure level with Fast response. Table 1 presents experimental conditions for each subject. Figure 2 presents the arrangement of loudspeakers for signals (TANNOY SYSTEM 600 in black) and for noises (Fujitsu ten TD-10 in white) used for this study. Listening tests were conducted in a semi anechoic chamber. All loudspeakers were set at 1.2 m height and 1.2 m distance from the center of the listening position. Each signal was presented from one of 4 loudspeakers for signals shown as black in Figure 2.

An ambient noise with Hoth spectrum was presented with 50 dB in A-weighted equivalent continuous sound pressure level at the listening position with 2 loudspeakers. Noise signal from both loudspeakers were incoherent. Gaps between signals was 2 seconds. Table 1 presents experimental conditions.

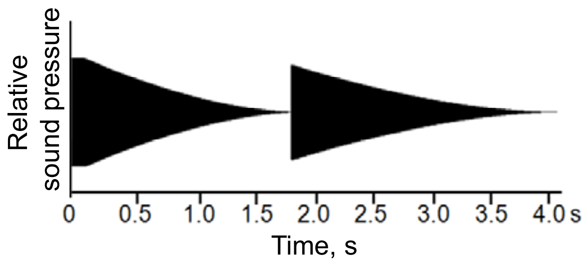


Figure 1 – Temporal pattern of the "Ping-Pong" sound.

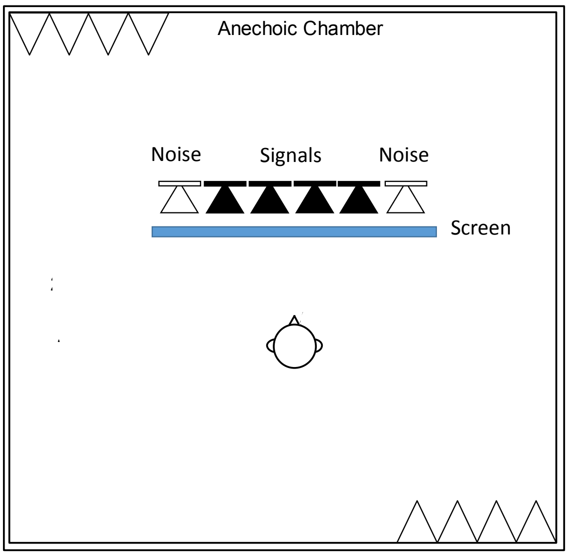


Figure 2 – Arrangement of loudspeakers in the semi anechoic chamber. Occupied marks indicate location of loudspeakers for signal.

Table 1 – Experimental conditions

| Condition | WN50 | WN38 | WN50N | WN38N | P50 | P38 | P50N | P38N |
|---|---|---|---|---|---|---|---|---|
| Signal | White Noise, | | | | "Ping-pong" sound | | | |
| Signal level, dB | 50 | 38 | 50 | 38 | 50 | 38 | 50 | 38 |
| Noise | No additional noise | | Hoth Spectrum Noise, 50 dB(A) | | No additional noise | | Hoth Spectrum Noise, 50 dB(A) | |
| Repeats | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

### 2.2 Subjects and procedure

Because this test was challenging and difficult to control in detail, three of researchers who were very familiar with sound and visual localization. They didn't report any vision and hearing problems. Subjects were asked to wear an eye-tracking glass (Tobii glass). They were instructed to look at where sound presents on the screen and to look at the center position, where a blue dot was put on the screen, when waiting for the next signal. Eye-tracking recordings were done all the time of the experiment and were edited for each condition individually. At the end of the all sessions, subjects were asked to look at center location of each loudspeaker, where the instructor put an IR marker only this timing, and center position to match the gaze position and positions of loudspeakers. Four of IR markers were used to wider the range of view area to be analyzed. Sampling rate of the eye-tracker is 30 ms.

### 2.3 Results and discussions

View area of the eye-tracking glass was 640 x 480 pixels. Only gaze position in horizontal axis were analyzed. First and second trial in the condition WN50 were omitted from analysis as training sessions.

Figure 3 presents the results of third trial of condition WN50. This condition is noise free condition and easy to localize presented sound. Subject A and subject C didn't look at center position but subject B did as instructed. There are difference between sound localization and visual localization for all subjects. It seems difficult to distinguish center position and both side of loudspeakers beside the center visually. This is because of distance and accuracy of gaze measurement. Response distribution of subject C would be ideal but still sound localization is shifted from visual localization. The instruction to ask subjects looking at the center gather data near to center and make it difficult to analyze around center position.

Figure 4 compares WN50 and WN38 to find the effect of presentation level. All subjects presents more less the same tendency to both conditions but weaker presentation level seems to make more scatter than stronger.

Figure 5 compares WN38 and WN38N to find the effect of ambient noise. All subjects present more scatter with noise case (Wn38N) than without noise case (WN38). Furthermore, Subject A and Subject C presents wider range of gaze point with noise case than without noise case. Subject A looked at more left side and Subject C did more right side and presents clear phantom image.

Figure 6 compares WN50 and P50 to find the difference of signal. All subjects present more scatter in P50 than WN50. There is no difference of distribution of frequency between signals.

Figure 7 compares P50 and P50N to find the effect of ambient noise for "Ping-Pong" sound. All subjects present more scatter with noise than without noise as same as WN signals.

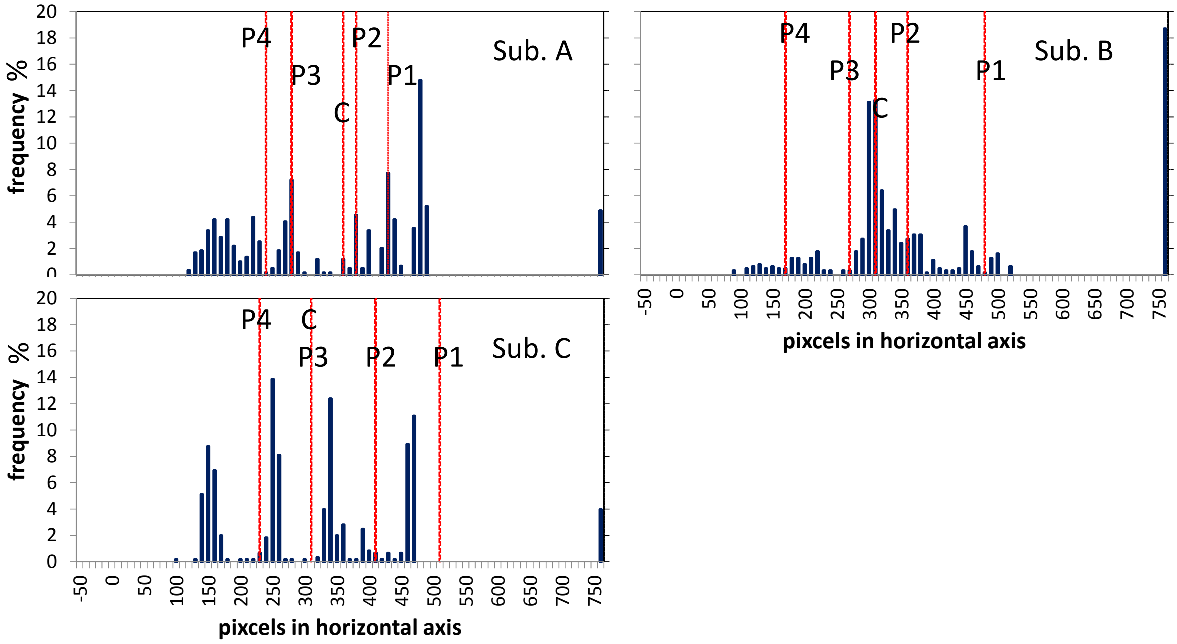


Figure 3 – Results of third trial of Condition WN50. Red lines presents visual localization of each loudspeakers (P1-P4) and center position.

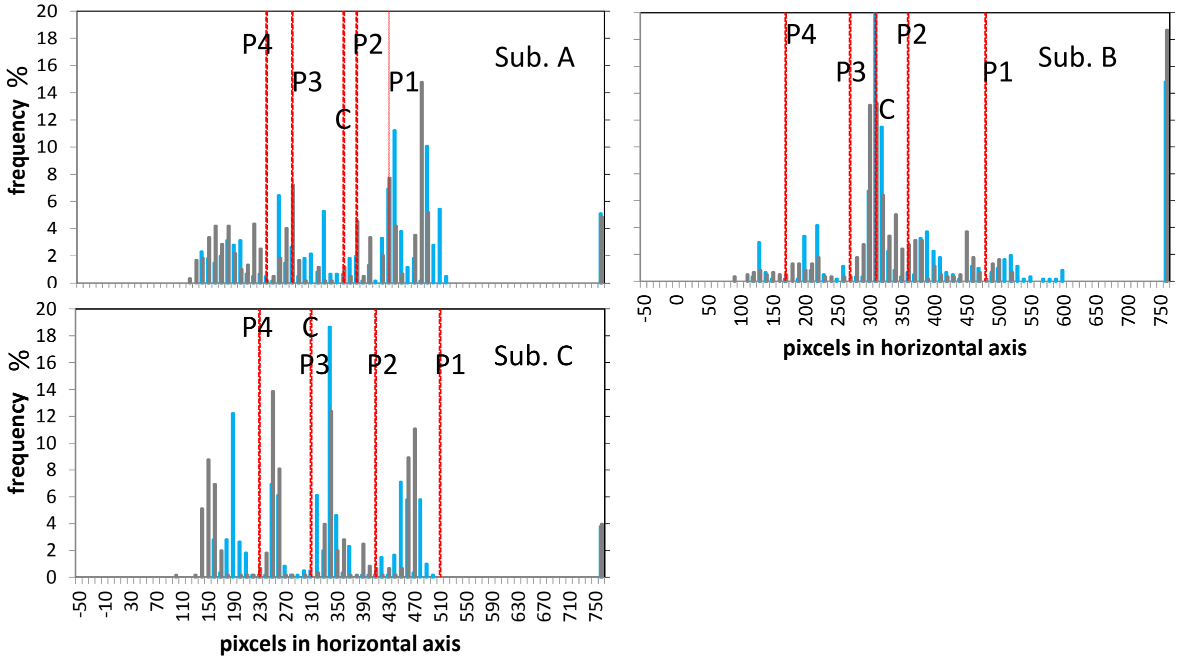


Figure 4 – Comparison of the results between Condition WN50(gray bars) and WN38(blue bars). Red lines presents visual localization of each loudspeakers (P1-P4) and center position.

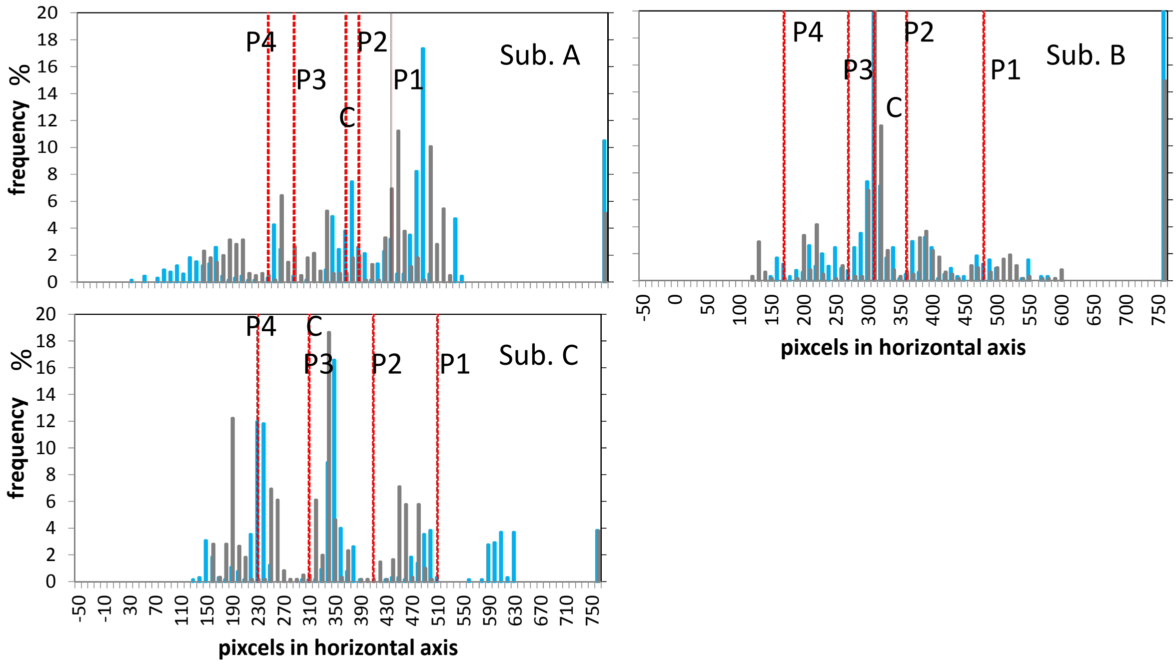


Figure 5 – Comparison of the results between Condition WN38(gray bars) and WN38N50(blue bars). Red lines presents visual localization of each loudspeakers (P1-P4) and center position.

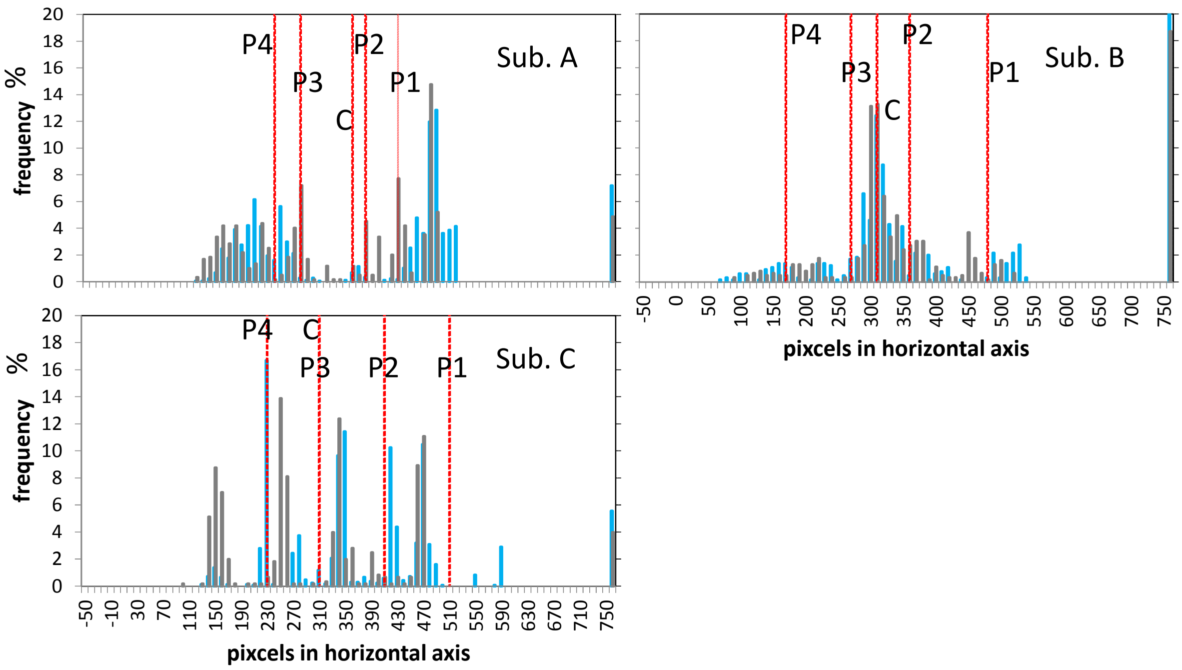


Figure 6 – Comparison of the results between Condition WN50 (gray bars) and P50(blue bars). Red lines presents visual localization of each loudspeakers (P1-P4) and center position.

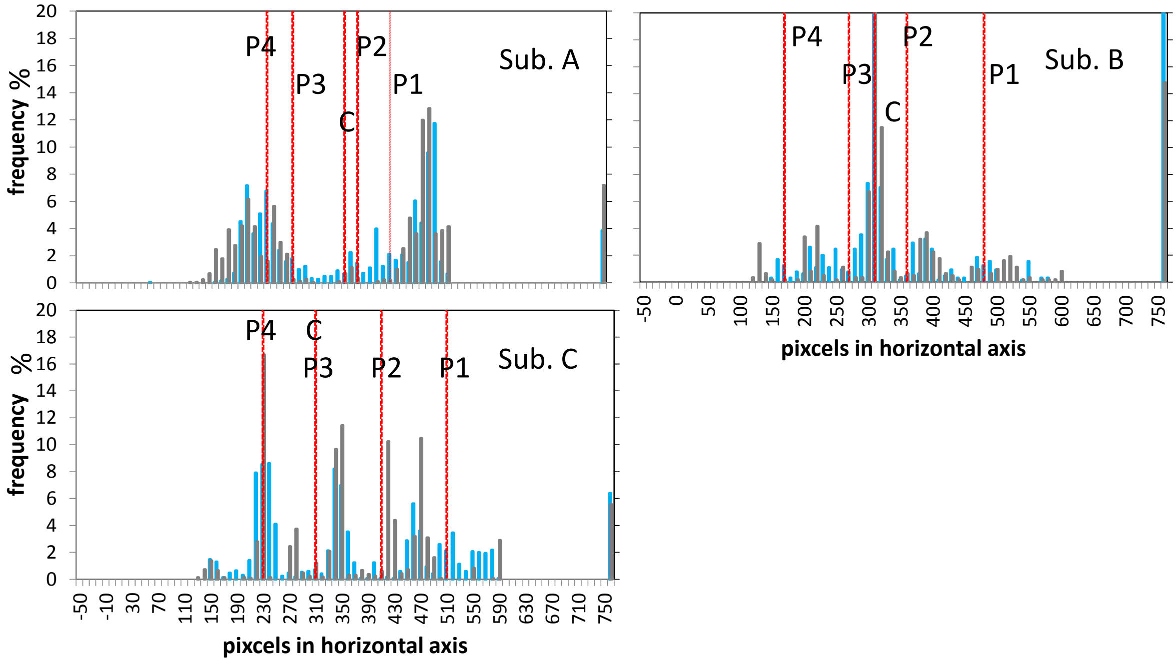


Figure 7 – Comparison of the results between Condition P50 (gray bars) and P50N (blue bars). Red lines presents visual localization of each loudspeakers (P1-P4) and center position.

# 3. CONCLUSIONS

## 3.1 Personal characteristics

As Figure 3 to Figure 7 suggests, gaze distribution patterns are different from subject to subject and personal patterns are same through conditions.

## 3.2 Difference of signal

As Figure 5 suggests, sound localization of white noise is more precise than Ping-Pong sound. However, this phenomenon is not quantified by this technique yet.

## 3.3 Effect of noise

As Figure 4 and Figure 7 suggests, adding ambient noise increases scattering of distribution.

## 3.4 Potential of gaze application to sound localization task in real world

The challenge is still undergoing and quantifying method is required to use gaze measurement technique. Calibration method to much the sound image and visual location shall be fixed. Using gaze measurement is still one of the strongest option to measure dynamic sound localization in real world.

# ACKNOWLEDGEMENTS

This work was supported by JSPS KAKENHI Grant Number 25282182.

# REFERENCES

1. E. C. Cherry and W. K. Taylor, "Some further experiments on the recognition of speech, with one and two ears," Journal of the Acoustic Society of America, 26, 554-559 (1954).
2. Barbara G. Shinn-Cunningham, "Object-based auditory and visual attention," Trends Cogn Sci., 12(5), 182–186. (2009).

3. Hiroshi Sato, Masayuki Morimoto and Hayato Sato, "Effects of noise and reverberation on sound localization of acoustic guide signals for visually impaired people in public spaces, "Noise Control Engr. J. 62 (1), January-February 2014.
4. Ministry of Land, Infrastructure, Transportation and Tourism, “Guidelines for Promoting Easily Accessible Public Transport Infrastructure (in Japanese),” published by Japanese government (2006).
5. JIS T0902, "Guidelines for older persons and persons with disabilities-Auditory guides in public space for mobility assist," Japanese Industrial Standards Committee, Tokyo, Japan (2014).
6. ITU-T Recommendation P.800, "Methods for subjective determination of transmission quality," ITU-T, Geneva, Switzerland (1996).