inter.noise 2000
The 29th International Congress and Exhibition on Noise Control Engineering
27-30 August 2000, Nice, FRANCE

I-INCE Classification: 6.3

COMPARISON BETWEEN THE EFFECTS OF ADDITIONAL VERBAL AND VISUAL INFORMATION ON THE PERCEPTION OF ENVIRONMENTAL SOUNDS

Y. Suzuki*, K. Abe, K. Ozawa***, T. Sone****

* Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ku, 980-8577, Sendai, Japan

** Faculty of Systems Science and Technology, Akita Prefectural University, 84-4 Ebinokuchi, Tsuchiya, 015-0055, Honjo, Japan

*** Faculty of Engineering, Yamanashi University, 4-3-11 Takeda, 400-8511, Kofu, Japan

Tel.: +81-22-217-5461 / Fax: +81-22-217-5535 / Email: yoh@ais.riec.tohoku.ac.jp

Keywords:
ENVIRONMENTAL SOUNDS, PERCEPTION OF SOUND, VERBAL AND VISUAL INFORMATION, CROSS-MODAL CONSISTENCY

ABSTRACT
Results of three experiments with the SD method, Experiments A (sound only), AL (sound with written verbal explanation) and AV (sound with moving picture), to study factors in evaluating environmental sounds were compared to examine the additional effect of verbal and visual information on the evaluation of environmental sounds. Sixty-six kinds of environmental sounds were rated on 38 scales expressing sound quality, characteristics of sound field, those of sound source and emotional aspects evoked by the sounds. The results show that the verbal and the visual information affect the evaluation of environmental sounds. The cross-modal consistency as to whether the source of sound is clearly perceptible in the picture and whether the movement of the source synchronizes with the sound also influences the evaluation of sound. In conclusion, it seems that an image intrinsic to a specific sound evoked by cross-modal information affects the evaluation of environmental sounds.

1 - INTRODUCTION

We understand the environment by using various types of information provided by sounds coming to our ears from surroundings including sound quality, the nature of sound fields, sound sources and so on. Naturally, when we listen to environmental sounds, some emotion is naturally evoked by the information they carry. Furthermore, to obtain information on sound sources and sound fields, we must use every sense, not just the auditory one. Bearing all of these aspects in mind, we have been investigating factors for evaluating and describing environmental sounds. Three types of psychoacoustical experiments, i.e. Experiments A (Audio), AL (Audio and Linguistic) and AV (Audio and Visual), have been conducted in our studies [1-4]. In Experiment A, the stimuli were composed of only sounds [1], [3]. In Experiments AL and AV, written explanations of sounds and moving pictures, corresponding to the sounds were presented in addition to the sounds respectively [2-4]. As a result of these investigations, seven factors were commonly extracted. Three of them are the factors of sound quality, i.e. "aesthetic state," "brightness" and "volume and spatiality." The other four factors were a factor concerning information on sound sources, a factor concerning sound localization, a factor concerning the significance of sound, and a factor expressing nostalgia. Comparison of Experiments A and AL showed that the verbal information affects the perception of the sound [2], [4]. As a result of the comparison of Experiments A and AV, addition of the moving picture affected all the factors except for the brightness [3,4]. In this paper, the results of the three experiments are compared particularly paying attention to the difference in the effects of verbal and visual information.

2 - OUTLINE OF THE EXPERIMENTS

Common condition: The SD (Semantic Differential) method with a seven-category scale was used for the evaluation. Sixty-six environmental sounds were adopted as stimuli [1-4]. Some examples of the stimuli are listed in Table 1. Thirty-eight pairs of adjectives in Japanese were used for rating the stimuli. The pairs can be divided into three groups: Group I consisted of 15 pairs of adjectives describing sound quality used in the previous studies, Group II consisted of 12 pairs of adjectives describing emotional state due to sounds, and Group III included 11 pairs of adjectives describing information concerning the kinds of sounds. Seven subjects, aged 22 to 24 years, all with normal hearing acuity, participated in the experiment.

	Place where recorded	**Primary sounds**
c	Arcade	Music, hum of voices
e	Office room	sound of broom sweeping a floor
h	Office room	sound of typing on a computer keyboard
i	Skating center	cheering
j	Tunnel	sound of passing cars
n	Entrance hall of a railway station	hum of voices
o	Platform at a railway station	sound of a train starting
q	Underground passage at a railway station	hum of voices, footsteps
r	open space at a railway station	Sound of cars, signals, and footsteps
s	Pedestrian crossing	sound of signals
t	Pool	splash, hum of voices
u	top of a cliff facing the sea	roaring of waves
w	near an airport	birds singing
x	Mountain torrent	sound of a torrent
y	Crossing	Sound of a passing car in the rain
z	Park	sound of a fountain
C	Harbor	Sound of cranes in operation
E	Seashore	roaring of waves
I	Buddhist temple	sound of a temple bell
J	Conference room	Sound of glass being scratched
L	Conference room	Sound of a blackboard being scratched
M	Conference room	sound of raindrops
O	Primary school	sound of chimes
P	under a railway bridge of the Shinkansen	Sound of a passing Shinkansen
Q	Gymnasium	sound of footsteps
R	Waterfall	Roaring of a nearby (30 m) waterfall
S	Waterfall	Roaring of a distant (200 m) waterfall
T	Laboratory	sound of workstation fans
V	Lounge	sound of hammering
W	Lounge	Sound of tableware being washed
Y	Lounge	ticking of a clock
0	Subway platform	sound of a subway stopping
2	Parking lot	birds singing
8	Zoo	Sound of swimming hippopotami
α	Zoo	cawing
β	Farm road	Sound of an engine being revved up

Table 1: Some examples of the recorded environmental sounds used as stimuli.

Experiment AL: Explanation of the place where the sound had been recorded and the sound source(s) of the primary sound(s) were written at the top of the evaluation sheet for each sound stimulus. Subjects were requested to read the written information prior to the presentation of each sound stimulus.
Experiment AV: Moving pictures corresponding to the sound stimuli were synchronously given to subjects. Subjects were asked, as in Experiments A and AL, to evaluate the sound.

3 - RESULT OF EXPERIMENTS A, AL AND AV

To investigate the difference in the effects of verbal and visual information on factor scores to be assigned to the stimuli, the results from the three experiments were merged and analyzed with the method of factor analysis. Applying the criterion for eigen-values (> 1.0), seven factors were extracted. The first three factors, i.e., F1, F2 and F3, were regarded as the three factors of sound quality: "aesthetic state," "brightness" and "volume and spatiality," respectively. Other than these, four factors were derived. F4 is a factor concerning information on sound sources, F5 is a factor concerning sound localization, F6 is a factor concerning the significance of sound, and F7 is a factor expressing nostalgia. In the results of analysis, each stimulus has three factor scores relating to experimental conditions. A two-way analysis of variance in which the experimental condition and the sound stimuli were adopted as factors were executed for each of the seven factor scores. The result showed that the main effect of the experimental condition was not significant on all factors. The interaction between stimulus sound and experimental condition was significant in all the factors except for F2. This suggests that the brightness factor is hardly affected by the verbal and visual information. For the other factors showing the significant interaction between experimental condition and sound stimulus, the difference in the effects of verbal and visual information on factor scores of the stimuli were examined. The differences in some factors exhibiting characteristic changes are shown in Figs. 1, 2, 3 and 4. Symbols in the figures indicate the stimuli. The abscissa shows factor scores. The starting point of each arrow corresponds to the factor score of the sound in Experiment A. For each stimulus sound, on the other hand, the end points of the upper arrow shows factor scores of Experiment AV, and the end point of the lower arrow shows factor scores of Experiment AL. The arrows are arranged from the top to the bottom in the order of magnitude of the difference in factor scores between AL and AV. The stimuli, for which the differences of Experiments A vs AL, A vs AV and AL vs AV are significant beyond the 5% level by a multiple comparison test (LSD) are marked by \$, # and *, respectively.

4 - DISCUSSION

These figures show that the verbal and the visual information can affect the auditory evaluation of the environmental sounds. Moreover, it is seen that the effects of the verbal and the visual information are similar in general since the directions of the arrows corresponding to Experiments AL and AV are identical for most of the sound stimuli. However, if the effects of the addition of verbal and visual information are examined in detail for each sound stimulus, some characteristic differences are found as described below.

Figure 1 corresponds to the aesthetic factor (F1). Large differences in the effects of the verbal and the visual information are found for sounds such as "sound of a torrent (x)," "sound at a harbor (C)," "sound of cars, traffic signals, and footsteps (r)," "sound of a fountain (z)" and "roaring of waves (u)." For these stimuli, the factor scores increase more or decrease less by adding the visual information than by adding the verbal information. It should be noted that most of those stimuli relate to sounds of water. Reminding that F1 is the aesthetic factor, subjects' image that the sound of natural water is pleasant and beautiful could be enhanced by the moving pictures of a torrent, a harbor, a fountain and the seashore. This effect is typical in "sound at a harbor (C)." The score of this stimulus is clearly increased by the moving picture while it is hardly changed by the verbal information.

Figure 2 corresponds to the volume and spatiality factor (F3). The difference between Experiments AL and AV is significant in "sound of cars, traffic signals, and footsteps (r)," "sound of workstation fans (T)" and "sound of chimes (O)." For most of these stimuli, the factor scores in Experiment AV highly decreased. Similar tendency can be seen for other stimuli such as "ticking of a clock (Y)" and "sound of a passing Shinkansen (P)," though the changes are not statistically significant. The large decrease in scores by the moving pictures may be attributable to the immobility of the pictures. The pictures of "sound of workstation fans (T)" and "ticking of a clock (Y)" are obviously (almost) static and no Shinkansen train appeared in the picture of (P) because it was recorded under an elevated railroad. This suggests that static or invisible sound sources in moving pictures arouse only small attention to the sound sources, i.e. the main objects for the evaluation, while written verbal information may call clear attentions to the sound sources, even if the sound sources are static or invisible. On the other hand, for some of the stimuli, the factor scores were increased by the moving pictures while they were decreased by the written explanation. This phenomenon is found in "sound of glass being scratched (J)," "sound of hammering (V)" and "roaring of a nearby waterfall (R)." In these stimuli, the sound sources are dynamic, well synchronized with the sound stimuli and clearly appeared in the moving pictures. We speculate that this consistency between a given sound and a moving picture enables subjects to focus on the sound sources and sounds, resulting in affecting the auditory evaluation of sound.

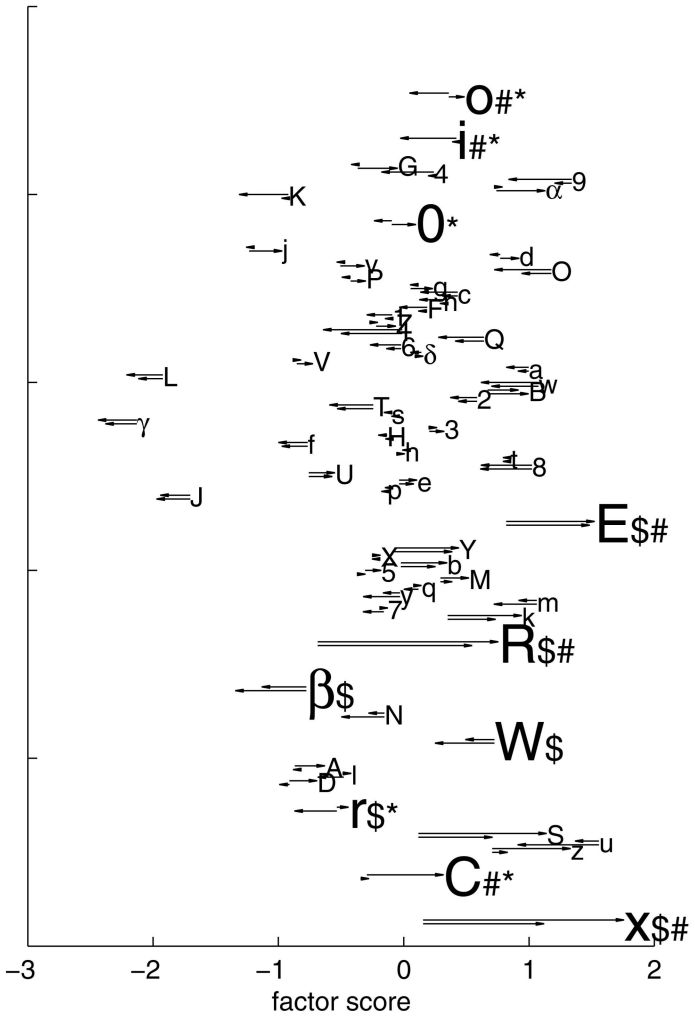

Figure 1: Factor scores of Exp. A, AL and AV aesthetic state (F1).

Figure 3 corresponds to the factor concerning sound localization (F5). For "sound of raindrops (M)," "splash and the hum of voice in a pool (t)," "sound of temple bell (I)," "a cry of crows (α)" and "sound of birds in parking area (2)," the factor scores was decreased by the moving pictures while they were increased by the verbal information. This difference seems to be caused by the vagueness of the sound sources in the pictures. This is supported by the contrary phenomenon that the factor scores were increased more by the moving pictures than by the written explanation in some stimuli in which the sound sources were clearly seen in the moving picture ("sound of an engine being revved up (β)," "sound of a floor being swept (e)", "sound of footsteps in gymnasium (Q)" and "sound of moving car in the rain (y)"). These suggest that evaluation of sounds concerning sound localization depends on the visibility of the sound sources. By the way, for "sound as cars pass in a tunnel (j)" and "the hum of voice and sound of footsteps in an underground passage at station (q)," the factor scores were decreased by the moving pictures even though the sound sources were clearly seen. The reason for it might be because the sounds in these stimuli could not be localized clearly at all.

Figure 4 corresponds to the factor concerning the significance of sound (F6). Large differences between the effects of the moving picture and the written verbal information are found for "sound of signals on a pedestrian crossing (s)," "sound of passing Shinkansen (P)," "the hum of voice and sound of footsteps in an underground passage at a station (q)" and "the hum of voice in a ticket barrier at a station (n)." In addition, the differences of factor scores from Experiment A to Experiment AV are generally larger than those from Experiment A to Experiment AL in these stimuli. This phenomenon suggests that the addition of a picture consistent to a specific sound enhances the *raison d'être* of the sound in the environment more effectively than in the case of written verbal information.

These results indicate that perception of sound changes if mental image that a listener has of the sound is

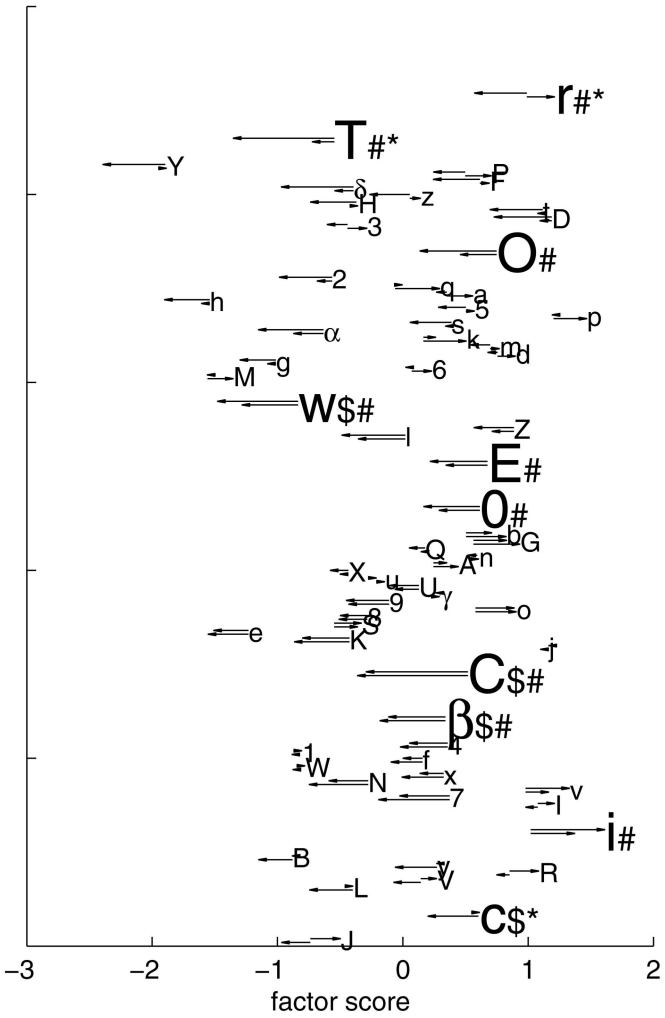

Figure 2: Factor scores of Exp. A, AL and AV volume and spatiality (F3).

affected by other information. As for the olfactory sense, it was shown that verbal information given prior to an evaluation task influences subjective evaluation of the pleasantness of smell [5]. This phenomenon is called the labeling effect. Our experimental results show that a very similar labeling effect exists for the auditory sense. This labeling effect may be interpreted in the following way: In our previous report [3,4] we stated that the additional visual information may recollect the image intrinsic in a specific sound environment and the image may affect the perception of the sound environment. The present results indicate that such mental images intrinsic to specific sound environment may be recollected by additional cross-modal information, irrespective of visual or verbal. Moreover, whether the auditory information and cross-modal information are well consistent with each other seems to affect the way and the amount of the change of the evaluation when the cross-modal information is added. We could further speculate that these considerations might stand generally for any cross-modal and multi-modal evaluation of environment.

5 - CONCLUSION

Results of three experiments (Experiments A, AL and AV) were compared to examine the similarity and dissimilarity of the effects of additional cross-modal information on the auditory evaluation of environmental sounds. In Experiment A, the stimuli were composed of only sounds while in Experiments AL and AV, written verbal explanations and moving pictures, respectively, corresponding to the sounds were given to subjects in addition to the sounds. The results show that the verbal and the visual information similarly affect the auditory evaluation of the environmental sounds. For most of the stimuli, the changes in factor scores by the addition of the visual and verbal information seem to be similar. This means that additional cross-modal information, irrespective of visual or verbal, has essentially similar effects

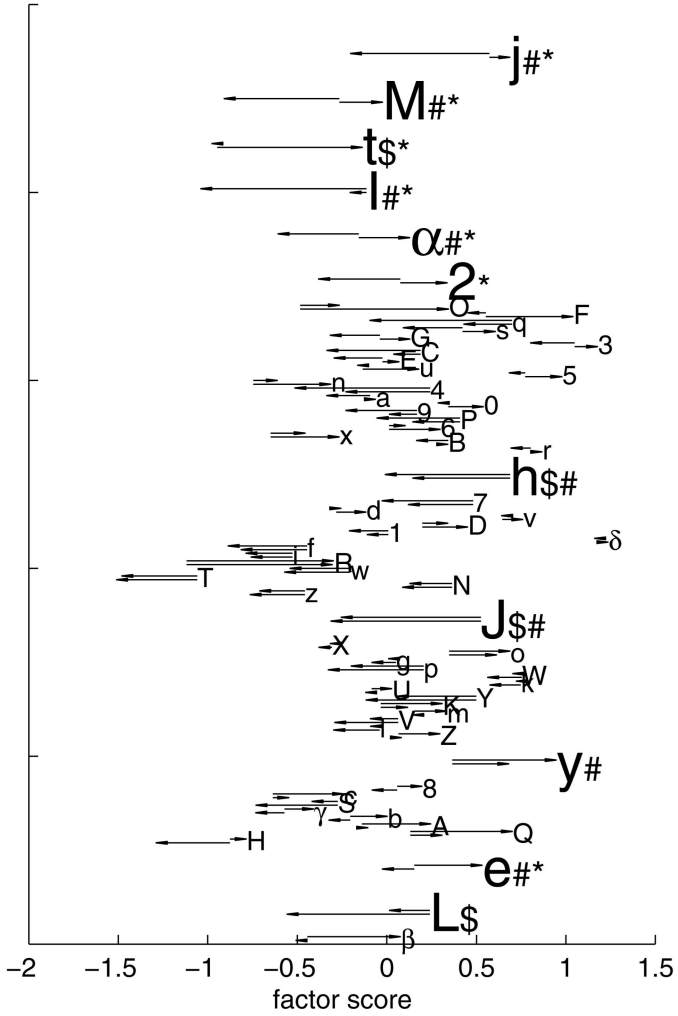

Figure 3: Factor scores of Exp. A, AL and AV sound localization (F5).

on the auditory evaluation of environmental sounds. For some of the stimuli, the changes caused by the additional visual and verbal information show different tendencies. Examination of such stimuli shows that whether the moving pictures are static or dynamic greatly affects the evaluation. The evaluation is also affected by the cross-modality consistency of stimuli such as whether sound sources is perceptible in the picture and whether movement of a sound source properly synchronizes with its sound. In summary, it seems that an image intrinsic to a specific sound evoked by additional cross-modal information affects the evaluation of the environmental sounds.

ACKNOWLEDGEMENTS

A part of this study was carried out under the Cooperative Research Project Program of the Research Institute of Electrical Communication, Tohoku University.

REFERENCES

1. **Abe. K., Ozawa, K., Suzuki, Y., & Sone T.**, Evaluation of environmental sounds using adjectives describing sound quality, emotional state, and information carried by sounds, *J. Acoust. Soc. Jpn.*, Vol. 54, pp. 343-350, 1998

2. **Abe, K., Ozawa, K., Suzuki, Y., & Sone T.**, The influence of verbal information on perception of environmental sounds, *J. Acoust. Soc. Jpn.*, Vol. 55, pp. 697-706, 1999

3. **Abe, K., Ozawa, K., Suzuki, Y., & Sone, T.**, The effects of visual information on the impression of environmental sounds, In *Proc. Internoise 99*, pp. 1177-1182, 1999

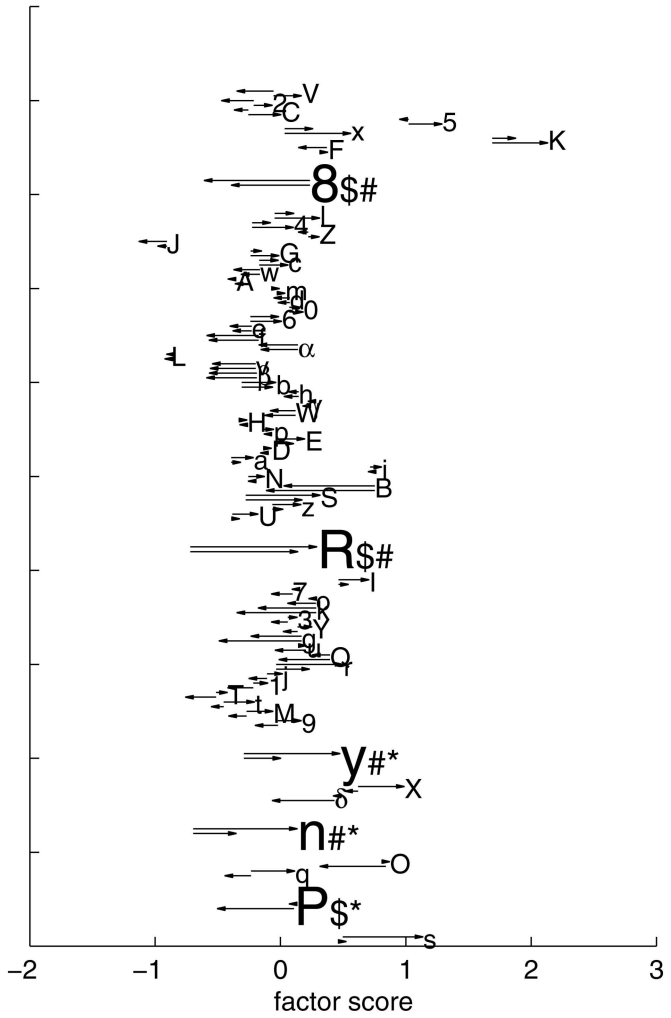

Figure 4: Factor scores of Exp. A, AL and AV significance of sound (F6).

4. **Suzuki, Y., Abe, K., Ozawa, K., & Sone, T.**, *Factors for perceiving sound environments and the effects of visual and verbal information on these factors*, Results of the 8th Oldenburg Symposium on Psychological Acoustics (A. Sshick, M. Meis & C. Reckhardt Eds.), pp. 209-233, 2000

5. **Ayabe-Kanamura, S., Kikuchi, T., & Saito, S.**, Effect of verbal cues on recognition memory and pleasantness evaluation of unfamiliar odors, *Perceptual & Motor Skills*, Vol. 85, pp. 275-285, 1997