Method for Selecting Appropriate Sounds to Convey Alerts to Train Drivers

Ayano SAITO

Yufuko ABE

Ayako SUZUKI

Ergonomics Laboratory, Human Science Division

Tomoharu TAKIMOTO

Hidefumi NISHIMOTO

West Japan Railway Company

The increasing number and variety of sounds emitted inside driving cabs has created the concern that this may distract drivers. A process has been developed to select sounds and information conveyed through spoken messages (speech information) to ensure they reflect a level of warning commensurate with the importance of their content. This selection method was developed on the basis of interviews conducted with people responsible for system design, and tests and surveys carried out with on-board staff. The importance of the information to be conveyed was split into four categories (hazard level) according to the likelihood of an accident and example sounds and speech information were given for each hazard level based on the degree of warning to be impressed on drivers.

Keywords: auditory warning, speech information, sound sign, train drivers

1. Introduction

Until now only a small range of in-cab auditory warnings and information existed, such as buzzers, chimes or bells. Today due to the development of sound producing devices and driver-support technology, there is a wider variety of sounds to choose from and the potential volume of information emitted in cabs has increased. While these changes do offer some advantages and a sense of safety, they could also become overwhelming and cause the driver confusion. In the medical field, these problems have already pointed out [1].

One route to avoiding these drawbacks could be to match sounds with the urgency of warnings to be made to drivers commensurate with the importance of the information to be conveyed, since this would allow the driver to recognize the importance of the signal intuitively from the sound itself.

To this end, this paper proposes a method for selecting sounds according to the level of the warning to be given to drivers corresponding to importance of the information being conveyed.

The demands on driver attention for warnings, differ according to signal system, safety equipment and the number of crew. This paper is based on a typical conventional line with wayside signaling, S-type ATS (automatic train stop) and P-type ATS (automatic train stop, with speed checking) and two operational staff. The expression “auditory warning and information” in this paper is used as a collective term referring to audio information given to the driver from train devices and is divided into “speech information” consisting of information given in spoken words and “sound signals” containing no speech.

2. Review and framework of the study

Regarding the design of “auditory warnings and information” : Wake et al. pointed out that there were three

view points as follows: the design of the information to be conveyed, the design of the timing of the emission and the design of the sounds themselves [2].

In aviation, for the design of information, alerts in the cockpit are divided into three categories: “warnings” which must attract immediate awareness and response; “cautions” which require immediate awareness and subsequent response; and “advisory” which require awareness and may require a subsequent response [3]. “Warnings” are divided into a further two groups, “time critical warnings” and “other warnings [4].” In the automobile sector, priority management of information is becoming more important because of the growing amount of information given to drivers. As such, a proposal was made to classify information by severity and urgency. One of the methods proposed to do this divides information into 9 categories according to combinations of warning levels and importance levels. The required criterion values for each warning level are defined for each category [5].

On the design of the sounds themselves: the impression given by a sound depends on a number of variables, such as volume, frequency, timbre, duration, envelope, temporal patterns, frequency shift, combination of sounds, and so on. There are many studies about the relationship between these variables and the sound impression they give [6], [7], [8]. In the method mentioned above, it is proposed that the sound should exceed the criterion value for the warning level defined for the category it belongs to by means of a parameter selection, based on the relationships [5]. For example, the criterion value for “K1” category is the highest, 7, in which case a higher frequency (value 3.2), a faster tempo (2.5) and repetition (1.6) should be added to make their sum exceed 7.

Although this method may be applied to the railway field, more studies are necessary because of differences between the two fields. For example, the aim of the method mentioned above is to make efficient use of frequencies between 1 kHz and 4 kHz. In contrast to this, recent train cabs have a wider variety of sounds, including some that

are lower than 1 kHz, high-low alternative sounds and miscellaneous tones like buzzers or bells. There are more parameters to be studied to understand their influence on the impression made on railway drivers.

In this paper, the information to be conveyed is classified first according to required warning level, and then examples of sounds commensurate with the urgency of the warning in each category are given. The following railway specific aspects were taken into account:

- Drivers must look towards the front to see signals and other alerts. Visual information that requires time to be recognized, such as long sentences should be avoided while the train is moving. So, auditory information should be a priority.
- Some alerts have already been in use for a long time. Therefore, consistency between these and new sounds should be taken into account.

With regards to “speech information,” a driver does not have to learn what it means, and it can convey complicated matters. At the same time, it may take drivers longer to listen until the end, or could lead to a misunderstanding if some words are unheard. So, a minimum amount of knowledge about speech information is desired. To this end, the content of the spoken message, and the order and tone in which it should be delivered, were studied. The result of the study is summarized later and the details can be found in reference [9].

3. Classification of information [10]

Based on interviews conducted with people from railway companies and responsible for system design and the answers from questionnaires sent to over 500 crew members, a system for classifying information, suited to the railway environment, was proposed. The upper part of table 2 provides a summary of this work. The information was split into four categories (hazard levels) according to the likelihood of an accident, and selection criteria were devised for each category.

“Hazard level IV” indicates a situation which is likely to lead to an accident and where driver action is very important to avoid the danger because there is no automated system backup. “Hazard level III” is for a situation which is likely to lead to an accident but where the protection system will make the train stop or slow down automatically, without wholly recovering the situation to normal. “Hazard level II” indicates a situation where the protection system not only resolves the hazard but also returns the situation to normal, automatically; alternatively, it indicates a situation which is unlikely to lead to an accident. “Hazard level I” is a situation in which information is conveyed for reference.

4. The sound character assignment to each hazard level

In this study it is important for the “Hazard level” and the perceived sound impression to be consistent in each category. An index was used to determine the feeling of urgency in a situation, called “perceived warning level.”

On the basis of other studies, a sound set consisting of 33 sounds was compiled to cover the broadest range of “perceived warning levels” possible. Driver impressions for each of these sounds were collected.

The sound set was compiled by funneling over 150 sounds in preliminary experiments; the set contains sounds similar to those which already exist. They were roughly divided into three groups: continuous sounds, short sounds and spoken words. Table 1 includes a description of the sounds.

The experiment was conducted in a meeting room near some train tracks. The background noise was approximately 37 dB (LAeq, 5) with no trains passing and 45 dB with trains passing. Figure 1 shows the devices used in the experiment. Sounds were emitted from a speaker (YAMAHA MSR100). The participants were seated on chairs arranged in an arc-like form at a distance of 2000 mm from the speaker to achieve a uniform hearing dis-

Table 1 Description of sounds tested

Broad category	Sub-category	Description
Continuous sounds	Continued sounds (7)	• 3 seconds saw-tooth wave or triangular wave Frequency (440 Hz, 700 Hz, 2100 Hz) • 6 seconds of bell rolling
	Intermittent sounds (6)	• 6 seconds saw-tooth wave or triangular wave Frequency (440 Hz, 700 Hz, 2100 Hz)
	Sweeping sounds (2)	• ascending from 300 Hz to 900 Hz within 0.5 seconds Codal saw-tooth wave With or without intermission
	High-low alternating sounds (5)	• 6 seconds saw-tooth wave or triangular wave Musical interval, alternating tempo, intermission and degree of decay were varied
Short sounds	Single sounds (4)	• decayed sounds Triangular wave, saw-tooth wave, bell, code
	More than one sound (6)	• Triangular wave, saw-tooth wave, bell Musical interval and number of sounds were varied
Spoken word (3)		Male voice. Two kinds of tone. Two kinds of words.

Fig. 1 Equipment used in the experiment

tance. The height of the center of the speaker was 1400 mm for close to head level. A mic used to record the sounds was placed at a distance of 1000 mm from the speaker.

44 active drivers participated in the experiment. Their average age was 36.6 and the average number of years of experience driving was 11.2 years.

The order of the sounds was randomized in each group. The sound levels were 73 dB and 67 dB at the mic, to produce realistic conditions. Half the participants were exposed to 73 dB sounds first while the other half were first exposed to 67 dB sounds.

For each sound, the participants were requested to rate on a scale of 1 to 5, as indicated on the questionnaire sheets, the “perceived warning level,” the “urgency” they felt, the “restlessness” they experienced and “the possibility of the emitted sound being confused with another existing sound.”

Figure 2 shows the average score for “perceived warning levels.” As there were no significant differences among sound levels, excluding 4 sounds which had low scores, each plot includes both sound levels.

As shown in Fig. 2, the following acquired knowledge was confirmed: sweeping sounds created a strong feeling of warning, this sensation intensified as the frequency of the intermittent sound increased and intervals between high and low alternating sounds shortened; saw-tooth waves created a stronger impression than triangular waves. The warning level perceived from spoken messages varied.

The correlation between “perceived warning level” and “urgency” and between “perceived warning level” and “restlessness” was relatively high ($r=0.90$, $r=0.78$ respectively).

For each “Hazard level,” the reference values required for the “perceived warning level” and caused by the sounds, were defined with reference to existing warnings which, after evaluation by drivers, were considered to match in terms of sound and degree of importance. For example, if an existing warning was deemed to belong to “Hazard level III,” the “perceived warning level” of the sound was 3 as well, this would show a good match at the preliminary survey stage, and the subsequent reference value for the “perceived warning level” would be set to 3, corresponding to “Hazard level III.”

For “Hazard level IV,” based on references to existing warnings belonging to this level, a “perceived warning level” of about 4 was thought to be appropriate, generated, for example, by intermittent high frequency sounds, emergency bells and sweeping sounds. For “Hazard level III,” based on references to “ATS brake actuation,” which is

Score of 1 for “no feeling,” 2 for “slight feeling,” 3 for “feel something” 4 for “strong feeling” and 5 for “urgent feeling”

Fig. 2 Average score for perceived warning level

considered to be a typical warning belonging to this level, a “perceived warning level” of about 3 was thought to be appropriate, generated, for example, by a saw-tooth wave continuous sound, intermittent sounds and high-low alternating sound. For “Hazard level II,” based on references to “changes in speed limits,” which is thought to be typical warning belonging to this level, a “perceived warning level” of about 2 was thought to be appropriate, generated, for example, by a triangular wave continuous sound or a combination of two sounds. For “Hazard level I,” sounds with a “perceived warning level” less than 2 were thought to be appropriate, such as short sounds regardless of timbre or number of sounds emitted.

It should be emphasized that the sound examples assigned to each “Hazard level” are by definition only examples and that assignment of these characteristics to each “hazard level” is not strict. For example, sweeping sounds used in the experiment showed a stronger impression of a warning because of the fast tempo and wide ranging variation in frequency. However, if the tempo was slowed with narrower frequency variation, the impression generated would be different. It is expected that perceived urgency of intermittent sounds or high-low alternating sounds would decrease if the tempo slowed down or if intervals were extended.

5. Method Proposed

Table 2 shows the proposed method for selecting sounds based on the study described above.

When a sound has to be found to convey new information the first step should be to determine the “Hazard level.” For example, a situation where a door indicator lamp fails while the train is running, and there is no automated system to slow the train down would be classified as “Hazard level IV” because the situation needs an immediate response from the crew due to the high possibility of a door opening and a passenger falling out.

The second step is to determine whether the sound signal should be a sound or a spoken message. This decision will depend on the occurrence frequency of the hazard in question. For infrequent hazards, a spoken message is recommended because no effort is required to interpret the sound. When drivers are regularly exposed to a sound, such as during training or device tests, sound signals may be acceptable. For frequent situations, sound signals are advised because frequent spoken messages would soon become annoying and speech may take longer to assimilate. When the information to be conveyed is too complicated or if drivers prefer speech, spoken messages are acceptable. Sound signals should be chosen more carefully, because many sounds already exist. As in the door indicator case, spoken information is recommended because it is infre-

Table 2 Proposed method (summary)

	Hazard level			
	I	II	III	IV
Example situations ※	Reminder of next stop	Approaching to or exceeding the speed limit	Entering within the protection area of red signal	Door open while train running, Fire alarm,
Lists of features	Reminder Information to assist prediction Complements of visual display Notification Feedback to driver's action	Not necessarily to notify danger Train is stopped or speed downed by automatic train protection system The danger disappears after activation of the ATP system Low possibility of accident	Presence of danger Train is stopped by the ATP system Normal situation is not resumed after activation of the ATP system Relatively high possibility of accident	Presence of danger There is no backup system High possibility of accident, malfunction or failure of the ATP system to activate
Sound signals (frequently)	Short sound, decayed sound	Continuous sound with Triangular wave Combination of two sounds	Continuous sound with saw-tooth wave Intermittent sound Alternative sound Bell (existing)	Relatively strong sounds like faster tempo, frequency shift, higher frequency intermittent sound and bell rolling
Speech information (infrequently)	Normally pitched voice announcing “situational information” after preceding sound indicates hazard level. The sound is optional. Or announcing “action to be taken” after “situational information.”		Voice at command pitch stating “action to be taken” and “situational information” in this order. A sound indicates Hazard level may replace the action when if it is difficult to display it.	

※ Example situations will change according to actual train protection devices or safety policies. For example, non-activation of train protection system is considered to be level I if the train does not move until the protection system is activated, but is deemed to be level IV if the train moves without activation of the protection system.

quent.

When using sound signals, according to Fig.2, sound should be selected ensuring that the “perceived warning level” corresponds to the “Hazard level.” Additional factors to taken into account include: hearing ability of elderly people, total number of auditory signals and distinctiveness from background noise or other auditory information. To ensure distinguishability of one sound from another, it is desirable to make a difference based on at least two parameters such as frequency, duration, intensity and temporal patterns between the two sounds [4]. Sounds from the same device need a similar timbre.

When using spoken information, as for “Hazard levels” III or IV, the recommended action to be taken and situational information should be conveyed assertively using expressions such as, “Stop! Door indicator off!” When there is more than one possible action, or there is no need to describe the action to be taken because it is obvious, the spoken advice maybe replaced with a sound indicating the appropriate Hazard level. For example, “Pip, pip! Fire!” For “Hazard levels” I or II, situational information should be spoken in a neutral tone, such as “Next is a stop”; or with an added recommended action, such as “Next is a stop. Check the timetable.” Another measure which is thought to be effective, is to add a preceding sound to indicate the “Hazard level,” such as “Ding, Next is a stop” to reduce the possibility of not being heard. The words should be selected on the basis of distinguishability, i.e. they should be concrete, concise, clear and unambiguous, and should not easily be confused with other terms.

Once the new sound or speech information is decided, training is necessary.

Following these steps should make it possible to associate the appropriate level of warning to the new sound commensurate with its importance, whilst ensuring consistency with other existing sound alerts.

The following points should be taken into account:

- Visual signs such as lights should be provided in association with the sound for important information.
- Information belonging to each “Hazard level” should be customized according to safety policies.
- To avoid excessive perception of urgency, exceptional situations should not be considered to determine the “Hazard level.”

This study did not examine cases where two or more sounds are emitted at the same time. For this type of situation, it is thought that displays aimed at different sensory organs should be used, or that low priority sounds should be eliminated, however, further studies are necessary.

6. Conclusions

A process has been developed to select sounds and speech information that match with the level of severity of the danger they are signaling, based on interviews with people responsible for designing systems, and tests and surveys conducted with train crews. At first, the information was split into four categories (hazard levels) according

to the likelihood of an accident. Then, 33 sounds including sounds similar to existing sounds were emitted in front of 44 actual train drivers and their impressions were collected. Based on the outcome, reference values for the perceived warning level thought to match each hazard level were set, with reference to existing warnings, and example sounds, deemed to correlated with each hazard level were given.

When a sound has to be found to convey new information the first step is to determine the “Hazard level.” The next step is to decide whether a sound or voice signal should be used, depending mainly on frequency of occurrence of the event. Deciding sounds or speech information based on the example sounds or speech composition elements is expected to help listeners recognize the importance of each warning intuitively, from the sound itself.

References

- [1] Edworthy, J. & Stanton, N., “A user-centered approach to the design and evaluation of auditory warning signals. 1. Methodology,” *Ergonomics*, Vol. 38, No. 11, pp.2262-2280, 1995.
- [2] Wake, S., Okada, Y., Asahi, T., “Designing Auditory Signals on Human Interface. Proposal of “Matrix Design Method” and Designing Auditory Signals of Windows Access Tool for Visually Impaired Users,” *Bulletin of Japanese Society for the Science of Design*, Vol.49, No.5, pp.41-50, 2003.
- [3] 14CFR Part 25: Airworthiness Standards -Transport Category Airplanes.
- [4] Federal Aviation Administration (U.S.A.): Flightcrew Alerting, *Advisory circular* No.25, 1322-1, 2010.
- [5] Patent Publication information 2007-128267 : A method of alarm sounds setting and alarming devices in vehicles, 2007 (in Japanese).
- [6] Ebe, K., Kimura, K., Inagaki, H., Doi, S., “Characteristics of Alarms and Perceived Urgency : A proposal of urgency-map,” *Journal of Society of Automotive Engineers of Japan*, Vol. 56, No. 3, pp. 27-32, 2002 (in Japanese).
- [7] Acoustical Society of Japan, *Acoustical Science Series 1, Sensibility of sound tone -Evaluation and Creation of Timbre and Sound Quality-*. Corona Publishing Co.,Ltd. Tokyo, Japan, pp. 173-189, 2010 (in Japanese).
- [8] Acoustical Society of Japan, *Acoustics Science Series 5, Science of sound signs -Design Methodology of Sounds Convey Message-*. Corona Publishing Co.,Ltd. Tokyo, Japan, 2012 (in Japanese).
- [9] Saito, A., Abe, Y., Suzuki, A., Mizukami, N., Fujinami, K., Takimoto, T., “Presentation Style of Voice Messages as Alerts for Train Drivers : from the points of the Alert Levels,” *Journal of Human Life Engineering*, Vol. 17, No. 2, pp. 33-39, 2016 (in Japanese).
- [10] Saito, A., Abe, Y., Suzuki, A., Takimoto, T., Nishimoto, H., “Process to Select Appropriate Sounds to Convey In-cab Driver Alerts,” *RTRI Report*, Vol. 30, No. 9, pp. 17-22, 2016 (in Japanese).

Authors

Ayano SAITO, Dr. Eng.
Senior Researcher, Ergonomics Laboratory,
Human Science Division
Research Areas: Body Size Variation,
Usability,

Tomoharu TAKIMOTO
Researcher, Safety Research Institute, West
Japan Railway Company

Yufuko ABE
Senior Researcher, Ergonomics Laboratory,
Human Science Division
Research Areas: Noise Evaluation

Hidefumi NISHIMOTO
Railway Operations Headquarters Technical
Research & Development Department, West
Japan Railway Company

Ayako SUZUKI, Dr. Lit.
Assistant Senior Researcher, Ergonomics
Laboratory, Human Science Division
Research Areas: Social Psychology,
Occupational Health Psychology