ORIGINAL ARTICLE

A User-centered Auditory Design in Interaction Between Electric Vehicle and Pedestrian

Han ZHANG* and Seung Hee LEE**

* Graduate School of Comprehensive Human Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-0006, Japan

** University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-0006, Japan

Abstract: The growing maturity of autonomous driving technology has enabled designers to study how to give car owners better driving experience. However, the study of the auditory interaction between the pedestrian and the autonomous car is still not known sufficiently. The goal of this research is to study the user experience, real emotional feedback of the pedestrian for alarm sounds from the autonomous car. The whole research included one online survey and two experiments. There were three samples of sound used in this study: engine sounds, bells, and soft alarms. Experiments were conducted with factors such as sound preferences, the risk of sound, and different road environments analyzed the differences in pedestrian intelligence processing. The results of the user preference showed participants preferred soft sounds and noticeable sounds. However, pedestrians can feel sounds convey degrees of risk, but sound conveys the level of risk that is not affected by the environment.

Keywords: *Electric vehicle, Warning sounds, User experience*

1. INTRODUCTION

The growing maturity of autonomous driving technology is convenient for people's daily life. Many car manufacturers have been researching sound interactions outside electric cars and designing smart alarm systems. However, elements surrounding the needs of the pedestrians, such as their reaction after hearing the alarm, whether they can judge the risk of the sound, or the ideal volume level, are not satisfied.

To solve the lack of auditory information problem, this study aims to redesign proven sound quality parameters and sources in existing studies, present these warning signals from the car to pedestrians with a side-by-side simulated realistic roadside scenario, and explore an effective method to communicate the level of risk.

The research hypothesis is that participants can judge the information conveyed by different frequencies and types of warning sounds. Having a new sound design for warnings will improve the interaction between pedestrians and electric vehicles by conveying more accurate information of risk.

2. STATE OF THE ART

2.1 The electric vehicle problem

Electric vehicles (EV) are widely used in many countries because they are recognized as technologies to reduce air pollution in the region while increasing the safety of transportation energy [1]. EV sales grew by 41% globally

in 2016 and their prominence on roads is expected to increase dramatically over the next decades [2]. Moreover, most autonomous vehicles (AV) are based on EVs [3]. The potential impact of AVs on future personal mobility and freight has received considerable attention [4]. For example, a significant contribution to the reduction of Greenhouse Gas emissions in the transport section is part of the many benefits EVs and AVs bring to society [5]. However, a concerning problem of EVs is the lack of the discernable sound characteristic of the combustion engine. In that regard, pedestrians have an increasing difficulty to judge the distance of EVs in their vicinity, which leads to a failure in recognizing danger in time. A study by NHTSA showed that hybrid and electric vehicles are 37% more likely to cause accidents involving pedestrians, and 57% more likely to cause accidents involving cyclists in 2011 [8,9].

The environmental safety-related stimuli of pedestrians during walking are mainly visual, but it is also essential to understand safety-related auditory information [6,7]. Vehicles emit engine and tire noise which pedestrians can use to identify safe crossing opportunities. Indeed, pedestrians use changes in sound signals at the roadside to assess risk and also experience physiological reactions to vehicle noises [6,8].

In that sense, one of the most realistic and reasonable solutions is to provide auditory information to solve the quietness problem that automobile manufacturers and governments are currently facing. Using warning sounds emitted from the vehicle would be a type of sensory interaction between car and pedestrian.

2.2 Regulation

To solve this problem, some countries have adopted guidelines requiring all EVs to emit sound at low speeds.

For instance, in Japan, the “Committee for examining quietness of hybrid vehicles” consisting of academics, visually impaired people, and automobile manufacturers was established within the Ministry of Land, Infrastructure and Transport in July 2009 [9]. Besides, the “Guidelines for silent noise control of hybrid vehicles” was announced in January 2010 [10]. This guideline presents the requirements for an “Acoustic Vehicle Alerting System (AVAS)” (vehicle approach notification device) in “electric vehicles and fuel cell vehicles” [10,11]. In the guideline, the following requirements are indicated for this equipment:

- In the speed range from the start of the vehicle until 20 km/h, a sound must be automatically generated.
- The sound must be a continuous sound that conveys the running state of the vehicle.
- The sound should change depending on the speed of the vehicle, such as the volume or pitch being automatically changed, to make it easy to recognize.

2.3 The current soundscape

Trucks and public service vehicles with their verbal reversing sound warning, or ambulances with their sirens, have used bespoke sounds to warn of their proximity for decades. These are effective for their own needs, but less so for commercial vehicles. Manufacturers will have to install at least one external speaker on electric and hybrid vehicles to emit a constant sound like an internal combustion engine vehicle in operation at low speeds [12].

Some manufacturers have developed warning sound systems for EVs according to regulatory requirements. Since 2011, electric vehicles using manual switching warning sound systems have appeared on the market, including Honda FCX Clarity [13], Chevrolet Volt [14], Hybrid M35 [15], Nissan Leaf [16]. However, Tesla Motors and Volkswagen only added artificial sounds after the law was established [17, 18].

The internal combustion engine has been powering cars for more than a century, thus engine noises have become deeply ingrained in people’s everyday lives and modern culture [19]. Therefore, most manufacturers choose to restore engine sounds to the EV due to the familiarity with pedestrians [20]. As there is no new sound to adapt to, it is possible to simulate an approaching vehicle just by increasing the volume of the sound. This has become the most effective risk assessment tool used in vehicles and is one of the main ways that pedestrians measure safety.

So far, engine and horn sounds are the most common sources of alarm in current vehicles. The engine sound allows pedestrians to estimate the speed of the car and their relative distance to the vehicle, and the horn cannot convey relevant information [21]. Therefore, the question arises if using traditional warning sounds is the best method to convey auditory information and if there is a more user-centered approach to express risk factor.

2.4 Evaluation in acoustic research

A test frequently used in acoustic research to evaluate user’s preferences of sounds is pairwise comparison [22]. This test is used to compare the degree of sensation between two sound samples that are emitted in pairs. The advantage of pairwise comparisons over the sequencing method is that the experimental process is simpler. Also, respondents are more accessible to make objective decisions about the perception of two sound samples. Chi et al., indicated that the preference rating results of the two methods of comparison test were comparable [23].

All pair-comparison results constitute a complete evaluating matrix. For each participant, the answer of each paired comparison completes the result P_{ij} where i and j stand for the order in row and column, respectively [23]. If the former sound was preferred to the latter one, P_{ij} equals to 1; if the former sound was less preferred, P_{ij} equals -1; if the participant could not decide a preference, P_{ij} equals 0 [24]. Each pair of comparison results constitute an evaluating matrix [25].

In addition, Parizet found a listener can make some mistakes on some trials of sounds in the tests of comparisons by pairs. These mistakes are caused by circular error [24]. This can cause to overlook the question, forget the answer or skip to the next question.

Parizet poses that it can consider a circular error under the following conditions:

$$\begin{aligned} P_{12} \geq A \text{ and } P_{23} \geq A \text{ and } P_{13} \leq -A \\ \text{or} \\ P_{12} \leq A \text{ and } P_{23} \leq A \text{ and } P_{13} \geq A \end{aligned} \quad (1)$$

Given that the number of possible triads $A_t^3 = \frac{t!}{3!}$ where t is the number of sound sounds, the circular error rate can be derived as in Equation 1 [24].

In conclusion, for each response, if there was a circular error, $\delta = 1$; if not, then $\delta = 0$. For each warning sounds, the circular error rate was calculated for all of the pair-comparison responses. The criterion of $C = 0.25$ circular error rate was screen out of the distraction of participants [26].

3. METHODOLOGY

The whole research includes 1 online survey and 2 experiments, as shown in Figure 1. A total of 221 respondents (129 males and 92 females) aged between 18 and 65 years old took part in the online survey to collect data about user preference of sound (pair-comparison test) and sound risk (risk assessment test). After that, 20 students for the University of Tsukuba (10 males and 10 females / mean age = 23.4) participated in Experiment 1 and 2. Experiment 1 was conducted in a semi-anechoic room. Participants were asked to do a pair-comparison test and risk assessment test from the online survey, to compare whether the environment affects sound preferences and degree of risk.

For Experiment 2, participants were required to stand in front of a line to watch a video and try to imagine walking on the road. When presenting the approach warning sound, they needed to adjust the volume of the sound, evaluate the sound and the risk of the environment.

A detailed explanation of each survey and experiment is presented in Section 3.2.

3.1 Sound design

In this study, three types of warning sounds were used: 3 engine sounds, 4 bell sounds, and 4 soft alarm sounds (vibration sounds). The volume of the sounds measured with a Sound Level Meter ranged between 47 and 72 dB(A). The audios were converted to 16 bit, 44.1 kHz, 1411 kbps stereo Mp3 files using Adobe Audition CC (a digital audio workstation).

These sounds were used as stimuli for the paired comparison test in the survey and Experiment 1 and played on the speakers in the semi-anechoic room for the volume adjustment test.

A detailed information about the interval of each sound sample and the frequency of sound are described in the next section.

Figure 1: Experimental structure

3.1.1 Bell and soft alarm

Otto suggested that tested sound signals should be between 2–5 s for continuous and transient sounds [21]. Each sound should be presented at least four times. Bell and soft alarm sounds tested in this study were edited based on the above suggestions.

Bell sounds and soft alarm sounds were obtained from a commercially available CD. The duration of the bell (Figure 2) was approximately 4 s while the soft alarm (Figure 3) was edited into a 2 s interval. For the sound of the bell and soft alarm, four different frequencies were made with a cutoff frequency of 12 kHz and a slope of 15 dB/octave.

3.1.2 Engine sound

As presented in Chapter 2, engine sounds have been engrained into the soundscape of our society. Based on that, engine sounds were recorded from a Tesla S P90D. Three different kinds of engine sounds (simulated V8 engine sounds) were used. As shown in Figure 4, the first row of tracks shows the sound of a car driving at a constant speed of 20 km/h, the second row shows the sound from an approaching car running at a speed of 20 km/h, and the last row represents a car slowly stopping.

3.1.3 Road environment conditions

To test whether sounds always convey the same level of risk or if they vary based on the environment, four different road environment conditions used (Table 1).

Figure 2: Bell alarm waveform

Figure 3: Soft alarm waveform

Figure 4: Engine sound waveform

Table 1: Road environment conditions

Environment	L _{Aeq} (dB)	Environment Condition
1	68.3	Downtown area The road is divided into two paths and pedestrian can also walk on the driveway. Pedestrians often hear street propaganda broadcasts from various home appliance stores while walking on this road. The main sound noise are vocals and traffic.
2	71.8	Residential area On the two-lane road of the residential street, there is no visible dividing line between the side walk and the lane. The main environmental sound comes from traffic noise when vehicles pass.
3	78.5	Heavy traffic area The sidewalks and driveway are completely separated and pedestrian only can walk on the walking road. The main noise comes from the sound of the vehicle and the sound of traffic signals.
4	64.2	Vehicle-accessible shopping area This is a narrow shopping street with densely populated roads, cars and pedestrian on the same road. The sound of footsteps and voice is the main noise of the environment.

Environment 1 is a 2-lane street in a downtown area with an average sound level of 68.3 dB, including street propaganda (music from shops and sales announcement) and the sound of the crowd. Environment 2 is a 2-lane street in a residential area with an average sound level of 71.8 dB. Road traffic is the main source of noise and there is a significant difference in noise level when a vehicle is passing. Environment 3 is a 4-lane street with a lot of traffic and an average sound level of 78.5 dB. There are few pedestrians, road traffic noise is constantly heard, and the fluctuation of the noise level is relatively small. Environment 4 is a vehicle accessible shopping area and the average sound level is around 64.2 dB. Traffic volume is low, with the sound of the crowd and speaking noise as the main source of noise.

3.2 Survey design

A survey containing 68 questions was designed in Python and uploaded to GitHub. The online survey used for this study is located at <http://sounds.ziggear.us/>. Sequel Pro (Structured English Query Language) was used to manage the MAC database stored in MySQL (Structured Query Language).

3.2.1 Survey structure

The 221 respondents of the online survey were gathered from the 2nd of September 2018 at 15:00 until the 15th of September at 22:00 (JST). The respondents took an average of 31.2 minutes ($SD = 29.3$ min) to complete the survey. None of the respondent had obvious hearing abnormalities.

Before the survey began, the respondents were instructed to “compare three kinds of sounds based on their preference and evaluate the risk degree of those sounds”. The respondents were told to imagine themselves as pedestrians walking on the road.

The survey was divided into three parts. First, general demographic questions about age and gender (Q1-Q2). In the second part, respondents were required to complete

a pairwise comparison test composing of 55 questions (Q3-Q57). Each question used two random sounds for comparison while two sound icons were displayed on the left and right sides on the interface. Each respondent clicked on these icons to play a sound and then selected one of the icons to express their preference. However, if the respondent thought that there were no difference between the two sounds, they could click on “no specific preferences”. Once the respondent has chosen, they pressed the ok button to jump to the next question.

The third part (Q58-68) consisted of questions on the risk assessment test. Degree of risk was divided into four stages: low, medium, high, extreme. Respondents clicked on the speaker icon to listen at least once, and evaluated the sound conveying the highest degree of risk in which stage.

3.3 Experiment

20 participants (10 males and 10 females) aged between 20 and 27 years old (mean = 23.4 years, $SD = 4.41$) took part in the experiment. Participants were all undergraduate and graduate students from different departments of the University of Tsukuba. The experimental data were collected from 16th September 2018 8:00 to 24th September 20:00 (JST). The respondents took an average of 34.3 min to complete the survey ($SD = 30.6$ min). All 20 participants could understand simple English words, and none had any obvious hearing abnormalities.

The experiment was performed in a semi-anechoic room at the University of Tsukuba in Japan. The experiment was divided into three parts. First, in Experiment 1, participants were asked to do a pair-comparison test and risk assessment test, to collect information about user preference of sound and sound risk. Then, for Experiment 2, participants were required to stand in front of a line to watch a video and try to imagine walking on the road. When the approach warning sound was presented, they needed to adjust the volume of the sound, evaluate the sound and the environmental risk. Before the end of all experiments, participants were asked to do a simple interview. The detailed flow of the experiment is presented in Figure 5.

3.3.1 Apparatus

The experimental setup is shown in Figure 6. PC 1 synchronizes the large-screen playing the environment video, while the participants use a Bluetooth mouse to

*SV1 = sound volume 1; SR = sound risk; ER = environment risk

Figure 5: The procedure of the experiment

Figure 6: Setup of the experiment

control a slider visible on PC 2 to change the volume of the sound in order to take the volume test when the warning sound was emitted. At the same time, the sound level and pressure was measured and recorded using a sound level meter. The PC used for the sound value adjustment test was also used to record all experiment data. The environment videos were run on PC 1.

3.3.2 Experimental environment

In Experiment 2, the participants were asked to face the large screen from 1.5 meters away and to imagine that they were on a road. As shown in Figure 7, PC 2 provided a warning sound located 2 meters diagonally behind the participants that simulated a vehicle passing by. The background sound was played 3.2 meters away from them. The participants' field of view was horizontal to the screen. The brightness of the display was 80% of the highest brightness (about 160 Lumen). The intensity levels of environment sound ranged between 55–75 dB measured by the sound level meter depending on the comfortable threshold of the participants. When the background stimuli were presented, 11 kinds of warning sounds were presented in random order with random intervals.

There were four tests in each environmental session. The first test was to adjust the volume of the sound so it could be clearly recognized regardless of environmental noise. Next was to adjust the volume of the warning sound so it could just be audible. The third test was to calculate the sound's risk level, and the last test was to calculate the environmental risk.

Figure 7: The experimental environment

3.4 Interview

After Experiment 1 and Experiment 2 had been finished, a short post-experiment interview was conducted with the participants to know their preferences and attitudes regarding the three kinds of warning stimuli. Also, they were asked to explain the reason for their preferences and gave some suggestions to improve the experiment. Four questions were being asked as follows:

- Q1 What are your preferences from the 11 kinds of warning sounds? Why is that? Are there any sounds that impressed you? If there were, what are the reasons?
- Q2 From the three kinds of sound stimuli (bell, soft alarm, engine) what is the risk level order from high to low?
- Q3 In the four kinds of environment, what kind of warning sound did you think was necessary and which one did you feel strange?
- Q4 Do you have any suggestions about the experiment? Do you have any recommended sound type according to your life experience?

In the process of talking to the participants, the experiment investigator recorded keywords and important information on paper. At the end of the experiment, the investigator bounded together interview papers with their respective consent form and experiment guiding document.

4. RESULTS

4.1 Online Survey

4.1.1 Paired-comparison test

Based on a complete evaluating matrix calculation, 132 of the 221 participants were excluded from the comparative analysis because their circle error rate was greater than 0.25. There were no obvious individual elements from the 132 participants (e.g., age, gender). Analysis of variance (ANOVA) indicated that different sound frequencies have significantly different preference result ($p < 0.01$).

Figure 8 shows the preference value of 11 warning sounds of all 89 participants (132 out of 221 because of their concentration), all pair-comparison results constitute a complete evaluating matrix (see Section 2.4). E1 (a car

Figure 8: Pair comparison result of all 89 participants

Table 2: Sound risk rating of 11 warning sounds by percentage

	Low	Medium	High	Extreme	Total
B4	3.1	9	86	1.9	100
B3	9	3	74	14	100
B2	3.8	76	10	10.2	100
B1	75	11	10	4	100
SA4	3.7	5.3	15	76	100
SA3	5.3	5.3	85	4.4	100
SA2	4.8	73	7.3	14.9	100
SA1	77	6.3	11.2	5.5	100
E3	76	8	11	5	100
E2	2	73	20	5	100
E1	4.8	8.7	81	5.5	100

driving at a constant speed of 20 km/h) was perceived to be the worst, while B2 (medium frequency of bell) and B3 (high frequency of bell) were the preferred ones. Also, an ANOVA comparison test found a significant difference within the types of sounds ($F=88$, $p<0.0001$). A post hoc analysis using paired t -test with Bonferroni adjustment found significant differences for the sounds B1, E2, and SA3. Those sounds were the highest among their categories.

4.1.2 Risk assessment test

Risk assessment test was performed to analyze each warning sounds conveying with the warning levels. The degree of risk that the sound can convey is divided into the low, medium, high, and extreme. Table 2 shows among the options that convey low risk, more respondents chose engine sounds. While soft warnings are considered to convey extreme danger best.

The calculated risk rate based on the 11 different types of warning sounds is illustrated in Table 2. It showed that more respondents chose low frequency warning sounds to convey less dangerous risk. For example, 75% of respondents chose low frequency bell sound (B1) as “low risk”, while 77% of respondents believe that a low frequency soft alarm (SA1) seemed risk level was low.

Besides, 76% of respondents judged that the highest frequency soft alarm warning sound (SA4) conveyed extreme risk. However, only 1.9% of respondents believed that the highest frequency bell warning sound (B4) conveyed a level of risk that was extreme.

4.2 Experiment 1

4.2.1 Paired-comparison test

According to the statistical analysis in the survey results, calculate whether there had a circular error at first. Base on a complete evaluating matrix calculation, 1 out of 20 participants from pair comparison analysis because they had a greater than 0.25 circular error rate in the soft alarm sounds evaluation. There were no obvious individual elements from the rest of participants (e.g., age, gender). Analysis of variance (ANOVA) indicated that different sound frequencies had significantly different preference results ($p < 0.01$).

In the results of the survey, the respondent's rate of circle errors was 59.7% (132 out of 221). But in Experiment 1, the rate at which the circle error occurred was reduced to 5% (1 out of 20). Based on the explanation for the circle error in Section 2.4, the result shows that the respondents more focused to the test in a controlled environment. Because in online survey, the respondent completed the test in uncertain environments. However, in the Experiment 1 participants were in a semi-anechoic room that they would pay more attention to the test and each warning sounds could be clearly distinguished.

Figure 9 shows the preference value of the 11 warning sounds) base on an evaluating matrix (see Section 2.4). SA1 (low frequency of soft alarm) was perceived to be the worst, while B2 (medium frequency of bell) was the most preferred. For most of the preferred warning sounds, the results of the survey are the same as those in experiment.

However, for the worst warning sound, E1 (a car driving at a constant speed of 20 km/h) changed to SA1 (low frequency of soft alarm). Also, an ANOVA comparison test found a significant difference within the types of sounds. For 20 people the ANOVA is $F = 9.89$, $p < 0.0001$. A post hoc analysis using paired t -test with Bonferroni adjustment found significant differences for the sounds B1 and E2. Those sounds were the highest among their categories.

4.2.2 Risk assessment test

Risk assessment test was performed to analyze how each warning sounds conveyed a warning level. Table 3 shows among the options that convey low risk, more respondents

Figure 9: Pair comparison result of all 19 participants**Table 3:** Sound risk rating of 11 warning sounds by percentage in Experiment 1

	Low	Medium	High	Extreme	Total
B4	5	5	55	35	100
B3	5	35	25	35	100
B2	20	60	20	0	100
B1	40	55	5	0	100
SA4	5	45	40	10	100
SA3	30	55	25	0	100
SA2	75	10	15	0	100
SA1	85	10	5	0	100
E3	50	30	10	10	100
E2	0	30	40	30	100
E1	65	30	0	5	100

chose soft alarm. While Bell is considered to convey extreme danger best.

According to calculate the risk level based on 11 different types of warning sounds. The risk rate is illustrated clearly in Table 3. It shows that more participants chose low frequency warning sounds that can convey less dangerous risk. For example, 40% of participants chose B1 (low frequency bell) warning sound as “low risk”, while 85% of participants believe that SA1 (low frequency soft alarm) risk level was low.

Besides, there were 75% of participants judged that the SA2 (medium frequency soft alarm) warning sound conveyed the low risk. And only 10% of participants believe that SA4 (the highest frequency soft alarm) warning sound conveyed a level of risk that is extreme.

For the engine sounds, 65% and 50% of participants chose E1 (20 km/h running sound) and E3 (driving came slowly stop) warning sound as “low risk”. However, there were no participants judged that the E2 (an approaching car running at a speed of 20 km/h) warning sound conveyed the low risk.

4.3 Experiment 2

4.3.1 Volume adjustment test

Figure 10 shows the interquartile and medians ranges of the averaged adjusted volume in environments. White symbols: warning sounds strongly audible. Black symbols: warning sounds to be able to hear. And horizontal dashed lines mean the average noise level of the background sound.

Four environment background condition with the adjusted sound levels (strongly audible and able to hear) were analyzed of variance (ANOVA).

The result of ANOVA showed that environmental background conditions and the volume of warning sound were statistically significant ($p < 0.001$). There were differences between adjustment levels of each warning sound, depending on the corresponding differences with the environmental noises.

It can also be seen from the results that the type of warning sounds plays an important role in a different environment. For example, in addition to the first environment downtown area, engine volume needs to be higher than the other sounds to be strongly audible while bell warning sound was more easily detected.

However, in a residential environment (Env.2), the sound of the bell is raised because there is a sound in the video that the tram passes through. The two sounds are similar, so the participants are difficult to distinguish, thus turning up the volume. However, in heavy traffic area (Env.3), the volume of strongly audible warning sounds

Figure 10: The interquartile and medians ranges of the averaged adjusted volume

was between 5 to 18 dB which higher than the subjects' ability to hear audible. The interquartile ranges reach to maximized 18 dB. The rather large variability is likely caused by the strong level fluctuations of the environmental background sounds (e.g. varying numbers of vehicle passing the road or pedestrian was through the streets of different environments).

4.4 Interview

After Experiment 1 and Experiment 2 had been finished, a short post-experiment interview was conducted with the participants to know their preferences and attitudes regarding three kinds of warning stimulus. The participants' answers were needs to be highlighted:

- In the third environment (traffic area), I feel no danger and I am more worried about bicycles.
- The last environment feels very dangerous. I feel that the car will suddenly drive from next to me.

- *The soft alarm sound is not harsh, very comfortable but does not play a warning role.*
- *The bell sound is like the sound of a tram, which is difficult to identify in the second environment (residential area)*

For the question: “What are your preferences in 11 kinds of warning sound? Why is that? Where there have been some sounds that impressed you? If there were, what are the reasons?” 80% of the participants thought that bell sounds were the most impressive and the alert function was strong, especially for the high frequency of bell.

Almost all the participants chose the sort order of bell > engine > soft alarm when participants were asked to sort the warning risks conveyed by three warning stimuli.

On the other hand, for the question: “In the four kinds of environment, what kind of warning sound did you think is necessary and which one are you feeling strange?” 60% of participants thought Soft alarms were not strange in any environment, but the lowest frequency bell and the highest frequency bell feel strange in the first and second environments (the downtown area and residential area). One of the participants presented such an emotional experience: “when I heard Low-frequency bells emitted in the shopping area, I thought the sound when the store checks out.”

In the end, participants recommended some sounds that they thought were a warning, such as a heartbeat, parking, S.O.S in a phone, etc.

5. DISCUSSION

5.1 Sound preference

Compared the results of survey and the experiment 1, for most preferred warning sounds, the results of the survey are the same as those in experiment 1. However, for the worst warning sound, E1 (driving at a constant speed of 20 km/h) was changed to the SA1 (low frequency). The results indicate that the low frequency of soft alarm was perceived to be worst while the medium frequency of bell was the preferred. For low-frequency soft warning sounds, participants felt uncomfortable and have no warning. However, for the sharp sound of the medium frequency, participants easy to recognize that dangerous is coming. Importantly, the results relate to preference evaluation rather than performance. There may be discrepancies between detection performance and taste assessment and preference opinions collected from various participants or tested in various contexts. Participants preferred soft sounds [27], noticeable sounds [28], intermittent and appropriate tempos and pitched sounds. However, taking

into consideration the signal detection, the preferences of the participants changed.

5.2 Sound risk

The result of ANOVA showed that the frequency of the sounds and risk assessment are no statistically significant. Changing the frequency of the sound does not affect the impression of sound risk. It is possible that the participants paid attention to the characteristics conveyed by the sound traits rather than the frequency of the sound. As described in Section 2.3 the sound of the engine as a traditional warning sound has been more than a century. In comparison to convey the dangerous degree of sound frequency, the participants would be subconscious more sensitive and dependence to the engine, which is probably why most manufactures choose to reuse the engine sound in the electric vehicle.

5.3 Sound volume in different environment

The result of ANOVA showed that the warning sound conditions and the environment background conditions were statistically significant ($p < 0.001$). The result indicates that environment condition is a strong influence effect for warning sound volume adjustment. A volume of the sound level that is strongly audible in one environment (e.g., engine sound in the residential area) may be just audible in another background (e.g., bell sound in the shopping area).

The interdigital quartile ranges can reach 18 dB. This considerable change may be due to different noises in the ambient background sound (for example, the number of vehicles passing by on the road or pedestrians is through the streets in different environments).

For instance, the clearly audible volume of engine sound was higher than the average noise volume of the background sound about 10–15 dB in the downtown area; However, in residential area clearly audible volume of engine sound was lower than the average noise volume of the background sound about 10–18 dB.

It will be appreciated that providing only one emitted sound level may be defective when making guidance on the alarm sound level of the vehicle system.

For example, in Japan’s guidelines, it is recommended that the sound volume not to exceed the original driving sound at 20 km/h [28], which, according to the subsidiary documents, is below approximately 60 dB at a 2 m distance. In certain circumstances, such sound levels may be sufficient, but the sound might become difficult to audible and effective at higher background noise levels.

6. CONCLUSION

The current study tested warning sounds for engine sounds, bells, soft alarms from different functions and frequencies. Used the interview and the paired comparison test to collect user preference of sounds. Analyzing interview data and comparison results, designers could use soft sounds and noticeable sounds may improve pedestrian satisfaction when designing an alarm system.

By analyzing the risk test, the results show that pedestrians can feel sound conveys degrees of risk, low-frequency warning sounds can convey less risk. However, high frequency cannot convey the highest risk, the sound designers need to find other sound attributes to convey high-risk information.

Similarly, after cross-analysis of environmental factors and sound risks, the result indicates that sound conveys information that is not affected by the environment. Besides, in different environments, the volume value and the degree of risk of the warning sound have a strong influence. When designing an intelligent sound system, there is no necessity to consider that the same sound will convey different risks in different environments. At the same time, the need to issue various types of alarm sounds according to driving in different environments.

The present research focused on the pedestrian emotional feedback side of interaction with the autonomous driving car sound system. While gaining a lot of insights from results but only got some preliminary conclusions. The evaluation made in this study highlighted several limitations. Although four levels of risks were used for sound evaluation scale.

As results, participants were difficult to distinguish the difference between high and extreme level of sounds. Fewer evaluation scales may allow participants to assess sound risks more intuitively. Also, we only considered influence by factors of age and gender of the participants. Current experiments results show that gender and age do not affect sound risk and preference. Future work should also be accounting other influencing factors, such as background of education and nationality.

The next steps of this research aim to collect data more deeply in sound preferences, sound risk in a different environment by using existing technology, design an intelligent sound interaction system to improve the user experience of drivers and pedestrians.

REFERENCES

1. Yamauchi, K.; A discussion on the problem of quietness of hybrid and electric vehicles and additional warning sound for these vehicles, *Proceedings of 7th Forum Acusticum*, Pj06-8, 2014.
2. REN21; *Renewables 2017 global status report*, 2017.
3. Jafary, B., Rabiei, E., Diaconeasa, M.A., Masoomi, H., Fiondella, L., and Mosleh, A.; A survey on autonomous vehicles interactions with human and other vehicles, *14th PSAM International Conference on Probabilistic Safety Assessment and Management*, 2018.
4. U.S. Emergency Information Administration; *Autonomous vehicles: Uncertainties and energy implications*, Issue in Focus from the *Annual Energy Outlook 2018*, pp.1-16, 2018.
5. Asian Development Bank; *Reducing carbon emissions from transport projects*, EKB: REG 2010-16, pp.1-96, 2010.
6. Barton, B.K., Lew, R., Kovesdi, C., Cottrell, N.D., and Ulrich, T.; Developmental differences in auditory detection and localization of approaching vehicles, *Accident Analysis and Prevention*, 53, pp.1-8, 2013.
7. Ashmead, D. H., Grantham, D. W., Maloff, E. S., Hornsby, B., Davis, T. J., and Rushing, E. G.; Auditory perception of motor vehicle travel paths, *Human Factors*, 54(3), pp.437-453, 2012.
8. Maffei, L., and Masullo, M.; Electric vehicles and urban noise control policies, *Archives of Acoustics*, 39(3), pp.333-341, 2014.
9. United Nations and I. T. Committee; *Economic and Social Council*, 2014.
10. UNECE; Proposal for guidelines on measures ensuring the audibility of hybrid and electric vehicles to be added to [R.E.3 and/or S.R.1], *Informal Documents for the 53rd GRB Session*, GRB-53-09, pp.1-3, 2011.
11. Chamard, J.C. and Roussarie, V.; Design of electric or hybrid vehicle alert sound system for pedestrian, *Acoustics 2012*, HAL ID: hal-00811353, 2012.
12. U.S. Department of Transportation, National Highway Traffic Safety Administration; *Minimum sound requirements for hybrid and electric vehicles - Draft environmental assessment*, 2013.
13. Honda Motor; *FCX Honda Clarity press*, 2007.
14. Chevrolet; *VOLT 2018*, 2018.
15. Nissan Motors; *2017 Nissan Rogue Hybrid, Owner's manual and maintenance information*, Nissan USA, 2017.
16. Nissan Motors; *Nissan Leaf introduction*, 2014.
17. Vynakov, O.F., Savolova, E.V., and Skrynnyk, A.I.; *Modern electric cars of Tesla motors company*,

- Automation of Technological and Business Processes, 8(2), pp.9-18, 2016.
18. De Bock, J.; Making the change happen, Volkswagen Group Investor Conference Presentation, 2018.
 19. ustwo-Auto; Humanising autonomy - Where are we going, 2018.
 20. Shoemaker, A.; Fear of change : Autonomous vehicle technology and the automobile as a cultural artifact, Senior Theses, Scripps College, pp.1-103, 2018.
 21. World Health Organization; A road safety manual for dicision-makers and practitioners, 2015.
 22. Budzynska, L., Jelonek, J., Lukasik, E., et al.; Multi-stimulus ranking versus pairwise comparison in assessing quality of musical instruments sounds, 118 AES Convention Paper, p.6482, 2005.
 23. Chi, C., Dewi, R. S., and Huang, M.; Psychophysical evaluation of auditory signals in passenger vehicles, Applied Ergonomics, 59, pp.153-164, 2017.
 24. Parizet, E.; Paired comparison listening tests and circular error rates, HAL Id: hal-00849430, Acta Acustica united with Acustica, 88(4), pp.594-598, 2002.
 25. Mao, D., Yu, W., Gao, Y., and Wang, Z.; Error rating of paired comparison listening test for sound quality assessment, 18th International Congress on Acoustics, pp.1799-1802, 2004.
 26. Meng, F., Li, S., Cao, L., Peng, Q., Li, M., and Wang, C.; Designing fatigue warning systems: The perspective of professional drivers, Applied Ergonomics, 53, pp.122-130, 2016.
 27. Genuit, K., and Fiebig, A.; Sound design of electric vehicles - Challenges and risks, 43rd International Congress on Noise Control Engineering, pp.1-10, 2014.
 28. Yamauchi, K., Menzel, D., Fastl, H., Takada, M., Nagahata, K., and Iwamiya, S.; Cross-cultural study on feasible sound levels of possible warning sounds for quiet vehicles, 40th International Congress and Exposition on Noise Control Engineering, pp.3236-3241, 2011.

Han ZHANG (Member)

Han Zhang is a PhD. Candidate at the Kansei Design Laboratory, Doctoral Program in Kansei, Behavioral, and Brain Sciences, Graduate School of Comprehensive Human Sciences, at the University of Tsukuba, Japan. She has a master degree in Kansei Science from University of Tsukuba and a bachelor degree in Industrial Design from the Huazhong University of Science and Technology in China. Her research theme is related to Kansei Design, Human-Machine Interaction and User Experience Design.

Seung Hee LEE (Member)

Seung Hee Lee, PhD. and Associate Professor, University of Tsukuba. Development of "Onigiri Machine-Device for safety of children" SCOPE, MIC 2010. Specialized in Kansei Interaction, Kansei Media. Creativity of designers using fMRI, interdisciplinary study with Brain Sciences. Now, has been leading Kansei Design Lab. at University of Tsukuba since 2001 and researching on human cognitive behavior to find original meaning of Kansei for global use.