공학석사학위논문

Exploring the User Preference of Auditory Icons as In-Vehicle Signals Under Autonomous Driving Contexts

자율주행 맥락하에서 차량 내 청각 신호음에 대한 사용자 선호도 분석 연구

2022 년 8 월

서울대학교 대학원

산업공학과

ADRIANCE WILFRED ANGGANG

Exploring the User Preference of Auditory Icons as In-Vehicle Signals Under Autonomous Driving Contexts

자율주행 맥락하에서 차량 내 청각 신호음에 대한 사용자 선호도 분석 연구

지도교수 윤 명 환

이 논문을 공학석사 학위논문으로 제출함

2022 년 08 월

서울대학교 대학원

산업공학과

ADRIANCE WILFRED ANGGANG

ADRIANCE WILFRED의 공학석사 학위논문을 인준함

2022 년 08 월

위 원 장 ____이 덕 주____(인)

부위원장 ____윤 명 환____(인)

위 원 ____서 은 석____(인)

Abstract

Exploring the User Preference of Auditory Icons as In-Vehicle Signals Under Autonomous Driving Contexts

Adriance Wilfred
Department of Industrial Engineering
The Graduate School
Seoul National University

The rise of autonomous technology that has been incorporated into vehicles allows the autonomous vehicles to shifted its functionality as an interactive system where providing interaction and feedback between the user and system is essential. In addition, auditory user interface has been used in vehicle technology to reduce cognitive workload and provide information to the drivers. However, autonomous vehicle is still regarded as a new technology domain, and it is necessary to investigate what type of in-vehicle signals feedback that should be designed to the passenger depending on the context-of-use and scenarios involved. In this thesis, the three main research aims are; (1) to present a design proposal for in-vehicle signals feedback for autonomous vehicles based on passenger's perspective, (2) to explore the passenger's sound preference for in-vehicle signals feedback used in autonomous vehicle, and (3) to suggest a fully derived scenario when designing an in-vehicle signals feedback used in autonomous vehicles based on user-centered design process. To achieve the research aim, this thesis focuses on investigating whether the design of in-vehicle signal types such as earcon and auditory icon, and temporal pattern of information signal types would affect the passenger's preferences by measuring its perceivability, intuitiveness and consistency or appropriateness as an in-vehicle signal.

This thesis includes two experiments; a pilot test and a large scale online sound evaluation study. Prior to the sound set evaluation, a pilot test was conducted on a total of 13 participants with an average age of 27.23(±7.53) to investigate whether the auditory sound sample that was created for sound evaluation has the congruity that matches with the intended information (confirmatory, error, detection, in progress, alert and warning), and to further develop the scenario for passengers in autonomous vehicles context. There were two measures used for the pilot test, which is perceivability and intuitiveness to determine if the designed sound sample with temporal pattern matches with the intended information as this paper suggested. The pilot test was conducted in an acoustic chamber, and participants were asked to give their evaluation in a 7-points Likert scale for perceivability and intuitiveness of the sound samples, and conducted survey of multiple choices to select the appropriate scenarios for each sound. The data obtained for perceivability and intuitiveness were analyzed using analysis of variance (ANOVA) and Bonferroni correction post-hoc test for multiple comparisons. Result of the pilot test shown that all sound samples are perceivable intuitively designed with the intended information, except for in progress type signal. Hence, in progress type signals will need to be re-created for this study. Also, out of the 27 scenarios that was developed prior to the pilot study, this study narrowed down 15 essential scenarios which in-vehicle signal feedbacks are imperative to autonomous vehicles based on passenger's context.

The sound set evaluation was conducted online with a total of 125 participants with an average age of 37.15(±11.4) to investigate which type of sounds (a mixture of earcons and auditory icons, or a set of earcon/auditory icon consecutively) they prefer by measuring consistency/appropriateness measure in 7-points Likert scale. In progress sounds were re-created in ascending, descending, variated and simple tone parameters, and were evaluated by its satisfaction measures. The data obtained for consistency/appropriateness were analyzed using pairwise *t*-test comparison for each sound sets. The in-progress sounds were analyzed using four-way analysis of variance (ANOVA). Lastly, all of the participants' opinions were collected for qualitative analysis by performing text network analysis for visualization. Results from the independent samples *t*-tests for each scenario shown that users or listeners prefer a consistent ‘family’ of sounds, rather than a mixture of earcons and auditory icons in a scenario. The result from the in-progress sounds also shows that a descending-simple tone melody sounds has high satisfaction level. In the discussion, this study discussed whether the research aim is fulfilled based on the results obtained and added implications for the sound design. In summary and conclusion, this study also discussed the limitation of this study and the future direction

Keywords: Auditory user interface, Auditory design, Autonomous vehicles, Human-vehicle interaction, User-centered design
Student Number: 2020-28244

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Research Background

Autonomous technology has brought substantial impact in our daily lives ranging from robotics, appliances and even in vehicles. The emergence of autonomous technology in vehicles over the past recent years has received attention from the public, and researchers have been discussing the domain widely in several different aspects of research. Prior to the incorporation of autonomous technology in vehicles, National Highway Traffic Safety Association (NHTSA) reported that an estimated of 31,720 accidental deaths for the first 9 months of 2021, and showed an increase of 12% compared to the previous year record, despite stay-at-home measures implemented on March 2021 (National Center for Statistics and Analysis, 2021). About 90% of many accident cases were caused due to human error during driving (Trucks, 2013; Bengler et al., 2014). Thus, many researchers would consider automation would resolve such human error issues related to traffic accidents. Many had expected that autonomous vehicles would ensure to bring positive impacts such as reducing car accidents, energy consumption, pollution, congestion, and increasing transportation accessibility (Bagloee et al., 2016).

To further understand what 'autonomous vehicle' means, we would have to explore the definition of the term 'autonomous'. Luck et al., (2003) suggested that 'autonomy' was defined as "an agent's ability to generate on its own goals" (Luck et al., 2003), which was different than 'automation' term where Groover (2007) stated that 'automation' is "a physical technology that reduces or minimizes the need of human process intervention" (Groover, 2007). However, Bradshaw et al., (2013) added on their perspective on 'autonomous' and defined it as self-sufficiency, where an entity is capable to take care of itself (Bradshaw et al., 2013; Payre et al., 2021). Kaber (2018) proposed a conceptualized framework set model for 'autonomous' should consist of self-governing, viable and independent (Kaber, 2018). Overall, 'autonomous vehicles' can be defined as self-driving vehicles that fulfills its main transportation capabilities similarly to traditional vehicles (Li et al., 2016; Gordon & Lidberg, 2015; Eskandarian, 2012). Ilková & Ilka (2017) added on by stating that 'autonomous vehicle' is "a vehicle that can guide itself without human conduction" (Ilková & Ilka, 2017).

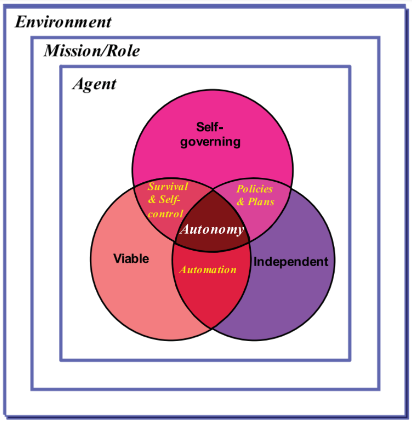


Figure 1. Set theory concept of automated and autonomous agents (Kaber, 2018)

Despite the many definitions of 'autonomous vehicles' suggested and proposed by many researchers, NHTSA and Society of Automotive Engineers (SAE) provided a more specific definition of autonomous vehicles based on its context through levels of automation. The National Highway Traffic Safety Administration (NHTSA) classified levels of vehicular autonomation into five levels, from Level 0 to Level 4 (NHTSA, 2013). On the other hand, SAE International redefines the level of automation taxonomy into six levels from Level 0 to Level 5 in standard J3016^{TM}, which distinguishes the high autonomation (Level 4) and full autonomation (Level 5) depending on driving scenarios (SAE, 2014). The SAE International standard J3016^{TM} was later adopted by the NHTSA and U.S. Department of Transportation (USDOT) in September 2016 (NHTSA, 2017; Ahmed et al., 2022; Ilková & Ilka, 2017).

SAE level	Name	Narrative Definition	Execution of Steering and Acceleration/ Deceleration	*Monitoring* of Driving Environment	Fallback Performance of *Dynamic Driving Task*	System Capability *(Driving Modes)*
***Human driver* monitors the driving environment**						
0	No Automation	the full-time performance by the *human driver* of all aspects of the *dynamic driving task*, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
1	Driver Assistance	the *driving mode*-specific execution by a driver assistance system of either steering or acceleration/deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	Human driver and system	Human driver	Human driver	Some driving modes
2	Partial Automation	the *driving mode*-specific execution by one or more driver assistance systems of both steering and acceleration/ deceleration using information about the driving environment and with the expectation that the *human driver* perform all remaining aspects of the *dynamic driving task*	**System**	Human driver	Human driver	Some driving modes
***Automated driving system* ("system") monitors the driving environment**						
3	Conditional Automation	the *driving mode*-specific performance by an *automated driving system* of all aspects of the dynamic driving task with the expectation that the *human driver* will respond appropriately to a *request to intervene*	System	**System**	Human driver	Some driving modes
4	High Automation	the *driving mode*-specific performance by an automated driving system of all aspects of the *dynamic driving task*, even if a *human driver* does not respond appropriately to a *request to intervene*	System	System	**System**	Some driving modes
5	Full Automation	the full-time performance by an *automated driving system* of all aspects of the *dynamic driving task* under all roadway and environmental conditions that can be managed by a *human driver*	System	System	System	**All driving modes**

Table 1. Taxonomy and definition for terms related to driving automation systems for on-road motor vehicles (SAE International J3016™, 2014)

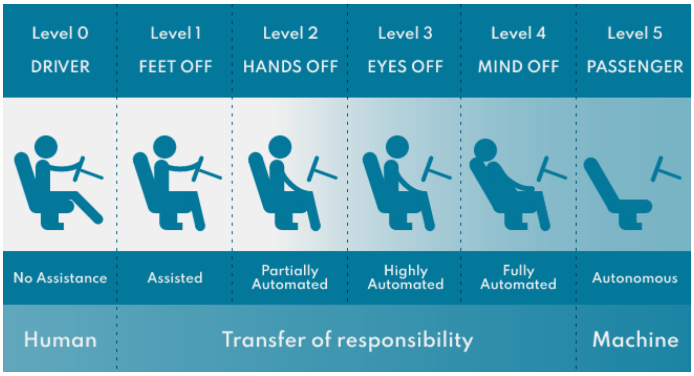


Figure 2. Levels of automation proposed by SAE International 2014

Based on the levels of automation proposed by SAE, most of the modern cars and vehicles have already incorporated low-level of autonomous (Level 1) technology such as Advanced Driver Assistance Systems (ADAS) and Adaptive Cruise Control (ACC) that aid drivers to drive more safely (Insurance Institute for Highway Safety (IIHS), 2020; Rukonić et al., 2021), but driver must be ready to resume takeover tasks and control when necessary. Camara et al., (2021) further explains the terms of each levels of automation by SAE. The Level 2 of the autonomous is considered as “hands-off”, where the automated system takes full control of the steering, and the driver must be ready to resume full control when needed. Level 2 technology can be regarded similarly to Lane Keeping Assistant (LKA) technology. Level 3 of the autonomous is “eyes-off”, where drivers can safely turn their attention away from the driving tasks, however drivers would still need to intervene with immediate response such as emergency braking. Level 4 is “mind-off” where driver’s attention is not required for safety except certain circumstances. Aside the limited circumstances, the vehicle is able to safely take over control from human. And lastly, as Figure 2 suggests, the Level 5 term is “passenger”, where not a single human intervention is required at all (Camara et al., 2021).

Future autonomous vehicles are expected to incorporate integrated intelligent technologies in the vehicle, and as the level of automation increases, it will result the drivers to shift their role as passengers (Šabić et al., 2021; SAE 2014). The integrated technologies adapted in the vehicle will also allow the current human-vehicle interaction (HVI) to have similar interaction as the human-robot interaction (Murali et al., 2021), where the vehicle would automatically perform driving tasks such as changing lanes while allowing the drivers to enjoy movies leisurely during mid-driving. However, in a fully autonomous vehicle system where the driver’s role has shifted as a passenger, these interactions would require feedback from the intelligent system itself to allow users or drivers to obtain information during the autonomous tasks. Furthermore, as the visual and attention of the driver (or in higher level of autonomation context, a ‘passenger’) were used for other activities instead of driving, the auditory aspect is the only medium to convey information of the system's progress to the passenger.

In traditional vehicles, most information that is conveyed to the drivers are in the form of auditory cues such as earcons and auditory icons, rather than text-to-speech type feedbacks (Nees et al., 2016). This leads to the issue that users may be easily misunderstood or misheard the auditory cues if sound designers failed to understand the significant factors when creating sound feedbacks for the autonomous vehicles. Hence, this prompts for a necessity to design for a more intuitive auditory feedback for autonomous vehicles.

Furthermore, numerous researchers and designers would resolve to evaluate auditory cues and sounds in products or systems using usability ratings such as System Usability Scale (SUS) (Brooke, J., 1996) or User Experience Questionnaire (UEQ) (Schrepp, M, 2015), however, the current scales provide only general usability of the product or system and there was no specific user experience evaluation related to auditory sounds, especially for non-speech type sounds (Tomlinson et al., 2018), which is widely used in vehicles. There were many heated discussions regarding the concerns of auditory-related research in terms of usability (B. F. G. Katz and G. Marentakis, 2016), and some claimed that the applications of auditory user experience are unrealistic (Goudarzi, 2016). As a result, the impracticality mentioned previously can be tackled by shifting the research focus towards approaching user-centric design especially in new domains (Barrass, 2012; Cornejo et al., 2018). However, there were several studies took affective engineering (or Kansei engineering) approach on the study of auditory user experience in vehicles (Kim et al, 2018; Park et al., 2019; Moon et al., 2019). Despite all the auditory related research on user experience aforementioned in this research's background, there were still lack of in-depth research on covering the auditory user experience in autonomous vehicles.

1.2 Research Objective

The main objectives of this study are as following: (1) Present a design proposal for in-vehicle's signals feedback for passenger-oriented autonomous vehicle, (2) explore the passenger's sound preference for in-vehicle's feedback used in autonomous vehicle, and (3) suggest a fully derived scenario when designing in-vehicle's signal feedback in future autonomous vehicles based on user-centered design process.

1.3 Organization of the Thesis

The thesis is composed of 5 chapters. The first chapter introduces the background and objective of this research. The second chapter provides a summary of findings from previous studies related to the in-vehicle signal types (earcons and auditory icons, auditory information types, and acoustic parameters), auditory user interface (AUI) used in vehicles, and the development of scenario based on the autonomous vehicles passenger's context-of-use. The third chapter contains the details of the pilot study and sound evaluation that was conducted to achieve the aim of this thesis. This chapter includes the aim of pilot study, procedures, methodology and result which leads to the necessity of sound evaluation. Similarly, in the same chapter, it includes sound evaluation's procedure, methodology and results. The fourth chapter presents discussions of the results of the experiment. Finally, the fifth chapter includes concluding remarks, limitation of the study and possible future research directions of this thesis.

Chapter 2

Literature Review

2.1 Auditory Types

2.1.1 Earcon and Auditory Icon

Earcons and auditory icons have been used widely in many fields of researches ranging from mobile phones, home appliances and vehicles which relates to the performance, situational awareness and user's efficiency when navigating menus or interfaces (Garzonis et al., 2009; Walker et al., 2013; Roginska, A., 2013; Larsson et al., 2009). To simply put, earcons are basic, structured, abstract and simple non-speech melody consisting a few musical notes (Walker et al., 2006; Hoggan et al., 2009; Oswald, D., 2012), whereas auditory icons are an alternative to earcon which the non-musical sounds conveys a resemblance thing they represent in our everyday life (Gaver, W., 1994). A well-known example of an auditory icon is the crumbling sound of paper trash when we perform deleting tasks on our computer or smartphones. Hoggan et al. (2009) added that auditory icons are semantically linked to the natural, everyday sounds they represent, and the meaning should be easy to understand and remember, similar to pictorial approach (Hoggan et al., 2009). In previous research, users perceive errors quickly and intuitively when an earcon is applied during performing their task (Brewster & Crease, 1999). Brewster (1999) created a menu-based graphical interface, and provided three selection where one of the selections is a wrong choice. Participants who clicked the wrong choice will be notified their error and as a result, participants tend to correct their wrong choices when earcon feedback is given. There were also many studies that compares the types of auditory cue (earcon and auditory icon) effectiveness in vehicles (Bonebright et al., 2007; Bussemakers et al., 2007; Šabić et al., 2021). Šabić investigated how well drivers respond to the types of auditory warnings while performing a stimulated driving task under various noise conditions (Šabić et al., 2021). Therefore, this study considered and selected the design guideline in composing for earcon and auditory icon as stated for standardization of our study with the other researches on auditory displays.

2.1.2 Auditory Information Types

One of the most important aspects of auditory feedback is the ability to convey intended information to the listener. As discussed in 2.1.1, earcons and auditory icons are non-verbal auditory cue that is often used as a means of feedback for operations or conditions of mobile phones, home appliances and even vehicles. However, the auditory feedback that is conveyed need to have information congruency in order for the users to intuitively perceive the feedback's meaning (Hoggan et al., 2009). Early studies approached information mapping used simple rhythmic sounds. Patterson (1999) conducted early research to investigate the evidence that rhythm, tempo, and speed of auditory signals induce significant effect on the perceived urgency of the warning alarm (Edworthy et al., 1995; Patterson, 1999). Palomäki further solidifies the evidence by using sound samples that has different variation in tempo, number of beats and rate of predictability to associate these rhythms to adjectives (Palomäki, 2006). The approach indicated that auditory rhythm can encode information to the listener. Walker also added on that appropriate mappings, polarities and scaling for auditory information display is necessary (Walker, 2002). Despite many researches on tempo pattern and rhythm that relates to information mapping, Hoggan et al., suggested a more accurate study related to the congruency of information as an auditory feedback. Hoggan et al. explored the auditory parameters such as tempo and rhythm to be mapped with four information types; confirmations, errors, progress updates and warnings (Hoggan et al., 2009). Based on Hoggan's findings, this study took similar approach of the four information types that were suggested and incorporated it in an autonomous driving passenger's context for design and evaluation.

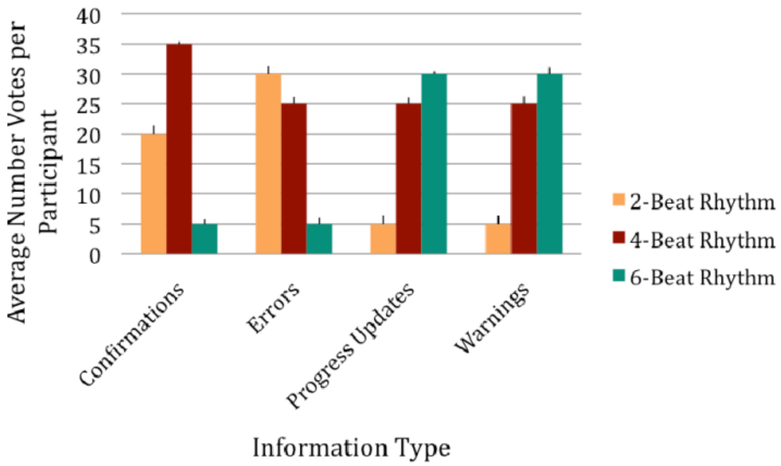
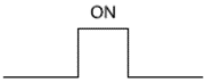
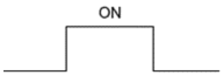
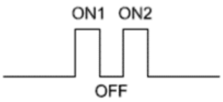
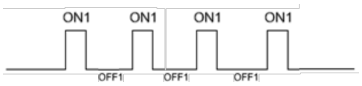
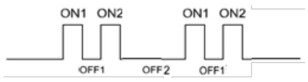
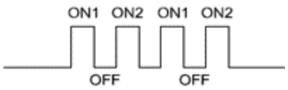


Figure 3. The average number of votes for each rhythm according to information type (Hoggan et al., 2009)

In designing a more accurate information type mapping that has the congruity with the acoustic parameter, especially tempo pattern and rhythm, the International Standard ISO-24500 (Ergonomics – Accessible design – Auditory signals for consumer products) provides a complete standardized guideline for auditory signals of products to ensure the listener to be able to hear the signals and understand the objective and meaning of signaling (International Organization for Standardization, 2010). The International Standard ISO-24500-2010 defines auditory signal as "sound emitted from a product for the purpose of conveying information to help the user to use the product correctly" (International Organization for Standardization, 2010). Furthermore, auditory signals convey information in an abstract manner than spoken instructions, thus, temporal patterns (hereinafter will have its meaning as 'tempo and rhythm') should be designed as such to allow users to understand without having the need of further instruction (International Organization for Standardization, 2010).

Hence, this study incorporates both Hoggan et al.'s findings and taking the ISO-24500-2010 as a reference to suggest a better design of information types of auditory signals which will be used in autonomous vehicles by using temporal patterns. This study proposes a better representation of temporal pattern and will be used for the sound evaluation and research. The in-depth details of information type, its pattern and explanation are shown in Table 2 below.

Table 2. The information type and temporal pattern for auditory signals suggested.

Information type	Information pattern (ON/OFF)	Explanation
Confirmatory		An audible tone indicating that the operation entered by the user has been performed correctly.
Error		An audible tone that indicates that the operation entered by the user has not been performed correctly.
Detection		An audible tone that informs the user that the system has detected a particular behavior of the user, a biometric signal, or an object of interest to the user.
In Progress		An audible tone that informs the user that the system is working on a particular task (for the purpose of preventing user interference)
Alert		An audible tone that requires attention from the user to be notified.
Warning		An audible tone that requires the user's immediate intervention by the system.

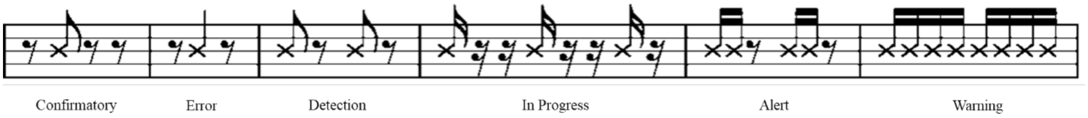



Figure 4. The beat notation for sound signal information type

2.1.3 Acoustic Parameters

Previous literature had composed earcons and auditory icons according to a set of parameters such as pitch, frequency, duration, tempo and tone (Orzessek & Falkner, 2006; Hoggan et al., 2009; Foley et al., 2020). In Geldard's research on the minimum distinguishable duration for in-vehicle signals feedback is that it should not be shorter than 0.1 seconds or longer than 2 seconds (Geldard, 1960). Walker (2006) added on that through his research on the improvement of navigation performance in auditory display menus. The guideline that was proposed in his research when composing both earcons and auditory icons should have the duration of the cue that last on average 1.26 seconds, which is in a range of 0.31 seconds until to 1.67 seconds (Walker et al., 2006). In investigating auditory acoustic parameters in terms of tempo, Yu et al. took an approach to sonification design in their research by maintaining a similar tempo for all auditory cue variations (Yu et al., 2015). Furthermore, in a study of investigating user preference for vehicle warning sounds among gender and age groups conducted by An et al., sound pitch that ranges between 400 Hz to 500 Hz (in a range of C and B♭ on the 5th octave) shown significant positive results compared to 1000 Hz and above (An et al., 2020).

2.2 Auditory User Interface (AUI)

2.2.1 Auditory User Interface (AUI) in Autonomous Vehicles

Auditory user interface (AUI), or auditory display has always been important in our daily lives. From our smartphones, home IoT and in-vehicles, auditory interfaces have been fulfilling its role in alerting and notifying users. Auditory user interface has been widely used in vehicles for drivers to reduce secondary task mental workload during driving (Seagull et al., 2001) before autonomous technology was incorporated in vehicles. Over the past years, auditory interfaces played significant part in vehicles by giving feedback and convey additional information to drivers intuitively without imposing cognitive workload (Baldwin & Struckman-Johnson, 2002; Forlizzi & Battarbee, 2004). As the rise of autonomous technology being incorporated into vehicles, there were several existing works related to the usage of in-vehicle signals or auditory cues as user interface to provide information or alarm the drivers about situations on the road (Merat et al., 2009; Merat et al., 2014). The types of in-vehicle signals or auditory cues often incorporated in vehicles are basic tones and chimes (Nees & Walker, 2011). However, the auditory user interface used in autonomous vehicles, for instance, Google's self-driving cars integrated chime sound to let human driver take over the task manually (Bilger, 2013), and basic tone was used in most semi-autonomous vehicles which requires the driver's attention to perform manual take over the driving task during hazardous situations (Huang & Pitts, 2022).

2.2.2 Auditory User Experience Measurement

Understanding user experience factors has becoming essential in research and in development process of a certain product or service. The interaction between users and a product or service are complex, dynamic and yet, subjective (Schneider et al., 2018). Also, investigating user experience will provide methods to understand user's behavior, needs and emotions during these product interactions (Jodi & Katja, 2004; Hassenzahl & Tractinsky, 2006). Due to this, it is important to investigate the user experience overall (including auditory user experience) to determine whether the users will continue or stop using the product or service (Kahneman et al., 1999).

Hassenzahl et. al., modelized a theoretical model which differentiates user's attraction towards a user interface by two distinctive qualities: (1) pragmatic quality, and (2) hedonic quality (Hassenzahl et al., 2003; Hassenzahl & Tractinsky, 2006). Both qualities are subjective aspects of a user interface that were used to measure user experience. In later literature, Schrepp defines 'pragmatic quality' as traditional usability aspects, such as efficiency, effectiveness, and learnability, which focuses on task related design aspects. Whereas, 'hedonic quality' is defined as quality aspects, such as originality and beauty, which are not directly related to the tasks the user wants to accomplish (Schrepp, 2015).

From these two qualities, this study will mainly investigate the perceivability (pragmatic quality), intuitiveness (pragmatic quality), and consistency/appropriateness (hedonic quality) of the in-vehicle signals feedback which will be used in future autonomous vehicles based on passenger's perspective. The 'perceivability' will measure the degree to which a user can accurately extract information from a given auditory stimuli (Smith & Walker, 2002). Garzonis et al., defines 'intuitiveness' as the immediacy of recognition of notifications and their relation to the service (Garzonis et al., 2009). However, to align with this study's context, this study measures 'intuitiveness' as the degree to which a user can immediately recognize the intended information from a given auditory stimuli. Green stated that the goal of psychoacoustics is to understand the relation between auditory stimulus and observer's reaction or response (Green, 1964), hence brought to the need of mapping between the sound stimuli with intended information conveyed to the listener. To show that the auditory stimuli is acceptable or preferred, the 'consistency or appropriateness' measures were added in this study. The 'consistency' or 'appropriateness' in this study will measure the degree on how the congruency or consistent the auditory stimuli as a series set of sounds without inducing the feeling of awkwardness to the user (Marshall et al., 2007).

2.3 Ideation for Scenario Development for Autonomous Vehicles

2.3.1 Context-of-use of Autonomous Vehicles

In order to develop the auditory feedback design that is specifically for passenger-oriented autonomous vehicles, this study incorporates user-centered design (UCD) process which were widely used in the design of products or interactive system where the user is the main focus. In this study, user-centered design process benchmarking will be taken from the International Standard ISO-9241-210 (Ergonomics of human-system interaction – Part 210: Human-centered design for interactive systems) as reference. The International Standard ISO-9241-210 defines the term usability as “the extent to which a product can be used by specified users to achieve specific goals with effectiveness, efficiency and satisfaction in a specific context of use” (International Organization for Standardization, 2019). The aforementioned definition of the term, usability, are highly dependent on context-of-use, which covers users (in this study, passengers), tasks, equipment and environment (International Organization for Standardization, 2019; Alonso-Ríos et al., 2010). Additionally, the advantage of incorporating user-centered design in this study is to improve the passengers’ auditory user experience, reducing discomfort and stress, and also, prompting the usability for passengers to comprehend the meaning of the auditory feedbacks. According to ISO-9241-210, there is a need to understand and specify the context-of-use of the system in order to plan for the user-centered design process.

There are three phases which will represent the context-of-use for autonomous vehicle passenger; (i) pre-usage, (ii) usage, and (iii) post-usage. The ‘pre-usage’ phase includes the ingress and pre-driving activity, ‘usage’ phase will include mid-driving activities, and ‘post-usage’ will include egress and post-driving activities. The context-of-use for this study will be represented in a timeline of these activities, and each roles and tasks from the user (passenger) and the interactive system (autonomous vehicle) will be represented according to the Figure 4 below. In this study, the target user group will be 'passengers', and the target interactive system would be 'auditory feedback of the autonomous vehicles'. The target group and interactive system will be assigned its role/goal and tasks of which the entities perform.

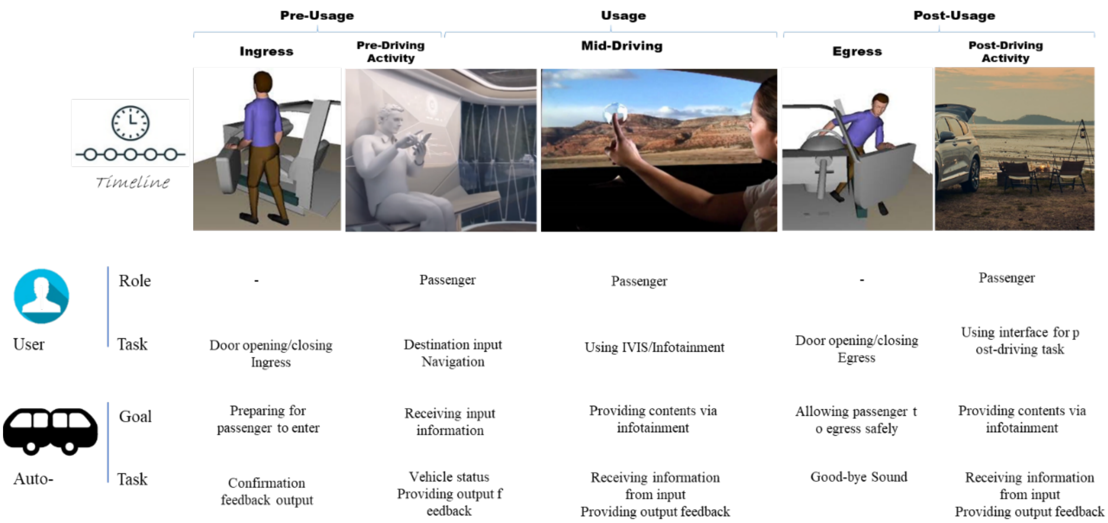


Figure 5. The role and tasks for user and the system

2.3.2 Scenario Development

Scenario in terms of usability and design refers to a description of a set of users, a work context and a set of tasks that users perform (Nardi, 1992). Furthermore, Nardi added that the purpose of a scenario is to provide explicit concrete vision of how some human activity could be supported by technology, in this study's focus, passengers and autonomous vehicle system (Nardi, 1992). Scenario development, or scenario-based design (SBD) were also widely used in human-computer interaction field where it acts as a tool in various stages of system development from problem define to solution envision, also helping stake holders to contribute to the analysis, design and evaluation of the systems (Robins et al., 2010). Scenarios can be in a form of textual narratives, video mock ups, storyboards of annotated panels or physical situations that contrive to support certain user activities (Carroll, 1997; Robins et al., 2010). Hence, in this study, a timeline form of scenario will be presented as the approach for this scenario development. From the context-of-use in 2.1, the goal and task of the autonomous vehicles is extended based on the necessity for auditory feedback to be given to the passenger, and the three phases of the timeline is further extended to five phases; (i) ingress, (ii) pre-driving, (iii) mid-driving, (iv) egress, and (v) post-driving.

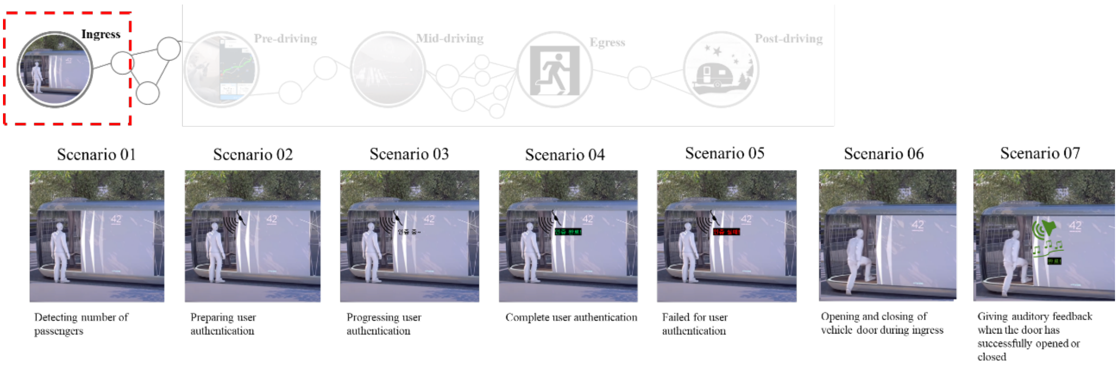


Figure 6. Scenarios for (i) ingress phase

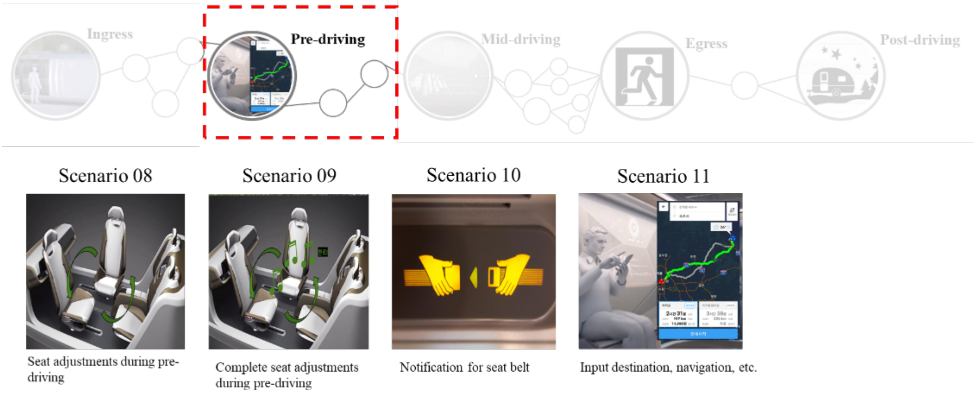


Figure 7. Scenarios for (ii) pre-driving phase

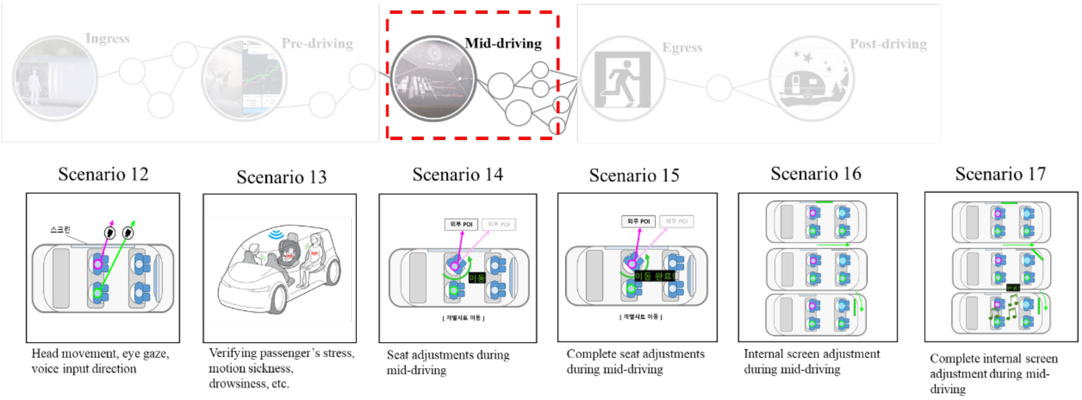


Figure 8. Scenarios for (iii) mid-driving phase

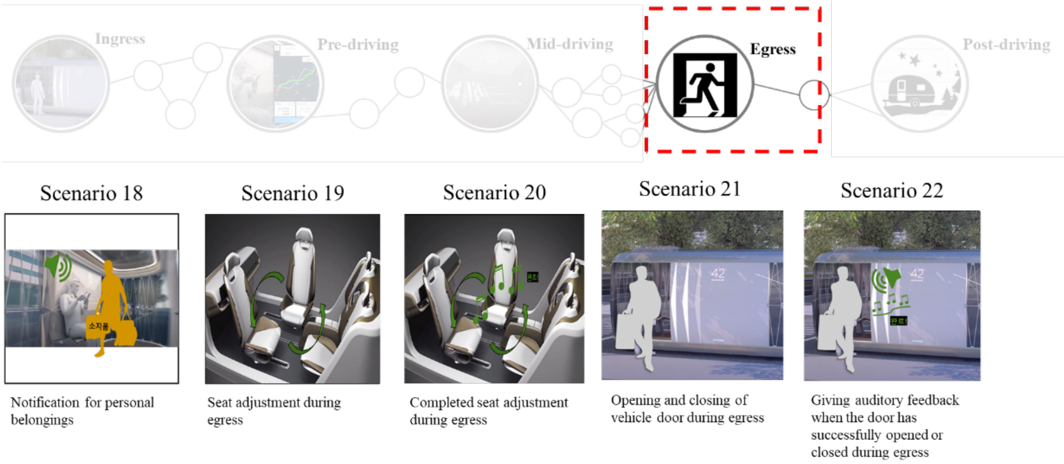


Figure 9. Scenarios for (iv) egress phase

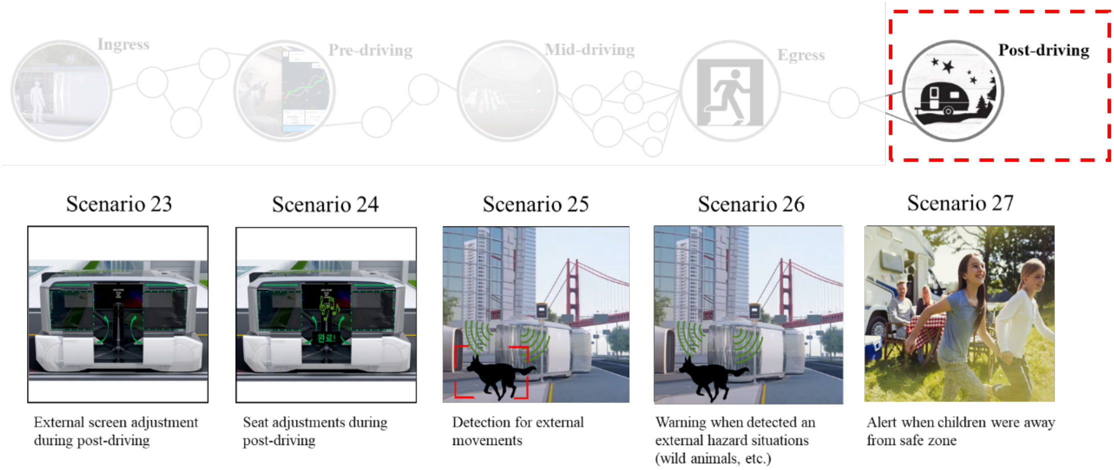


Figure 10. Scenarios for (v) post-driving phase

Chapter 3

Sound Experiment and Evaluation

3.1 Pilot Test

3.1.1 Overview and aim for pilot test

In Chapter 2, a review of the literature showed that there are different types of auditory types, information types and several acoustic parameters which relates to this study. However, there is a need to investigate whether the auditory sample created for the sound evaluation has the congruity that matches with the intended information based on the temporal pattern, and also to further develop the scenario of riding an autonomous vehicle based on passenger's context-of-use. Hence, in this study, this pilot test aims to investigate the intuitiveness and perceivability of each auditory sample created according to the developed scenario.

3.1.2 Participants

A total of 13 participants (6 males, 7 females) with an average age of 27.23(±7.53) participated in the pilot test. A majority of the participants about 53.9% (7 participants) had experienced being a passenger in any form of transportation means for about 6~15 times per week, whereas the remaining had at least ridden any transportation once in a month. The participants were screened through hearing test using the DB-23000 Audiometer. Participants were asked to raise their hand according to a randomized sheet of frequency 500Hz, 1000Hz, 2000H and 4000Hz at a threshold of 5dB. All participants had no abnormalities from the hearing test, and reported that they had no hearing disabilities.

3.1.3 Stimuli

Thirty-six sound samples (18 earcons, and 18 auditory icons) created using the Adobe Audition software by varying different information types and acoustic parameters were used for the pilot test. Each sound sample lasted approximately between 0.1 seconds to 2.5 seconds. The sound sample stimuli acoustic parameters are created as follows. All sound samples' loudness was remained constant at 10dB to avoid misconceptions during the evaluation.

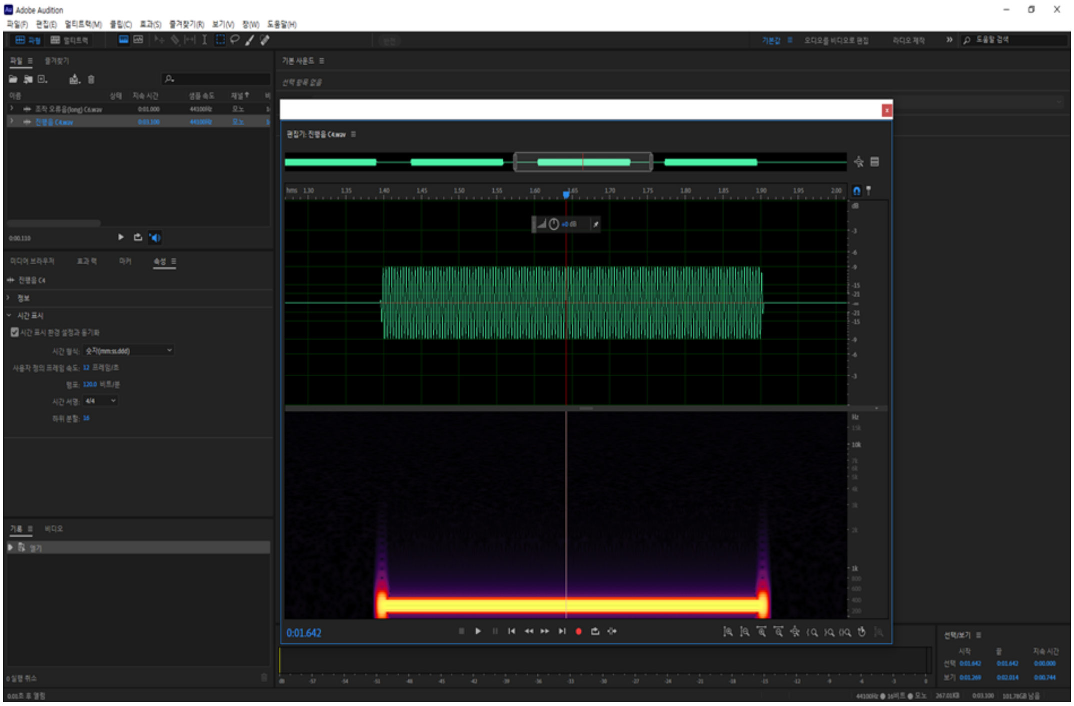
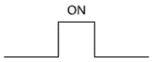
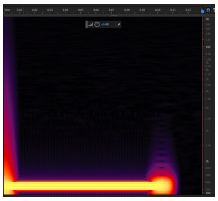
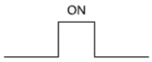
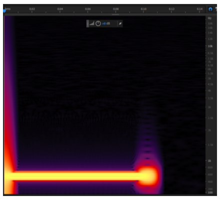
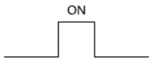
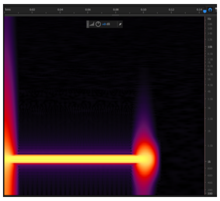
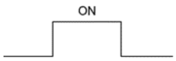
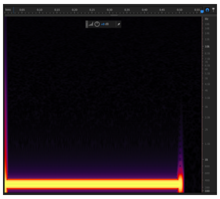
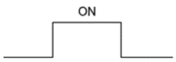
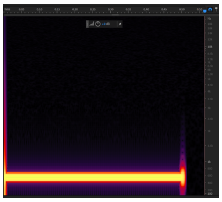
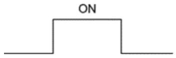
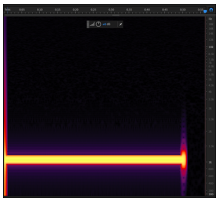
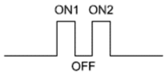
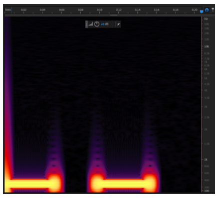
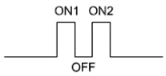
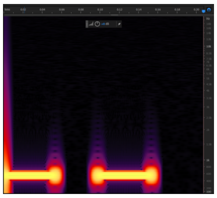
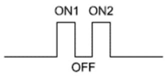
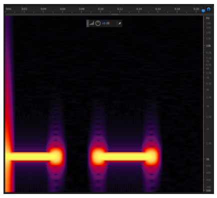
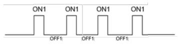
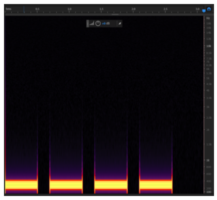
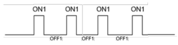
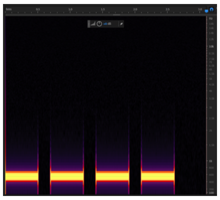
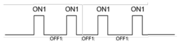
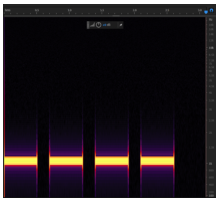
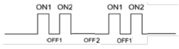
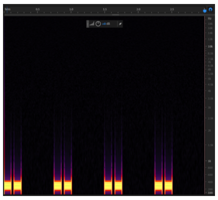
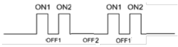
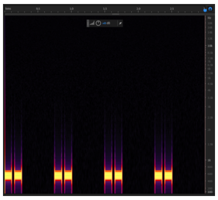
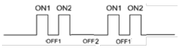
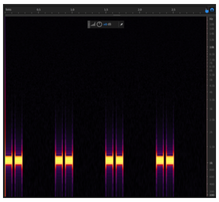
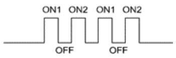
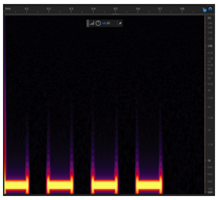
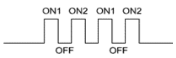
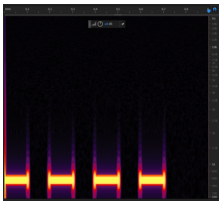


Figure 11. One of the sound sample stimuli created using Adobe Audition software

Table 3. The sound sample stimuli for earcons

Sound Type	Code	Pitch	ON time (s)	OFF time (s)	Duration (s)	Sound Pattern	Spectrograph of Stimuli
Earcon	A1	250Hz	0.1	-	0.1		
Earcon	B1	500Hz	0.1	-	0.1		
Earcon	C1	1000Hz	0.1	-	0.1		
Earcon	A2	250Hz	0.5	-	0.5		
Earcon	B2	500Hz	0.5	-	0.5		

Earcon	C2	1000Hz	0.5	-	0.5		
Earcon	A3	250Hz	0.05	0.05	0.15		
Earcon	B3	500Hz	0.05	0.05	0.15		
Earcon	C3	1000Hz	0.05	0.05	0.15		
Earcon	A4	250Hz	0.5	0.2	2.6		
Earcon	B4	500Hz	0.5	0.2	2.6		

Earcon	C4	1000Hz	0.5	0.2	2.6		
Earcon	A5	250Hz	0.1	0.05, 0.5	2.5		
Earcon	B5	500Hz	0.1	0.05, 0.5	2.5		
Earcon	C5	1000Hz	0.1	0.05, 0.5	2.5		
Earcon	A6	250Hz	0.1	0.1	0.7		
Earcon	B6	500Hz	0.1	0.1	0.7		

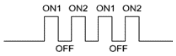
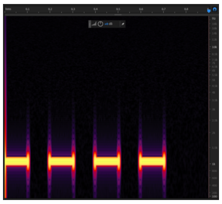
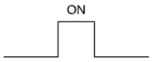
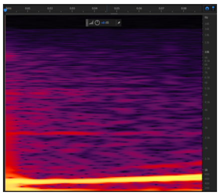
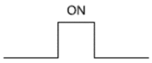
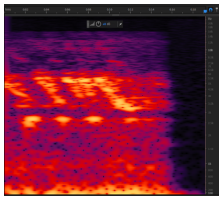
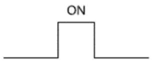
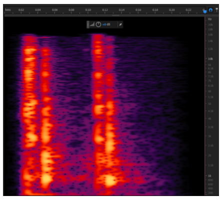
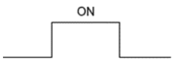
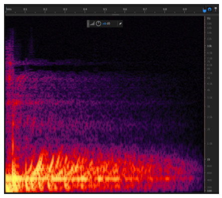
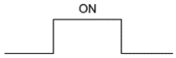
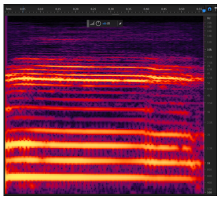
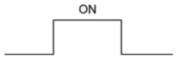
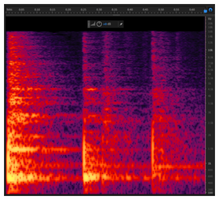
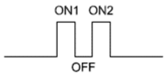
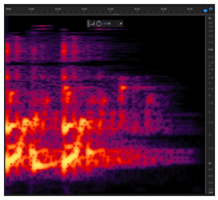
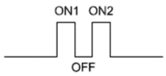
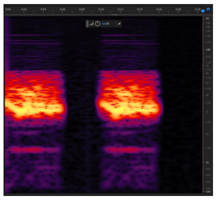
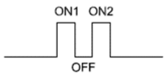
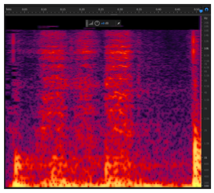
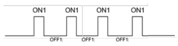
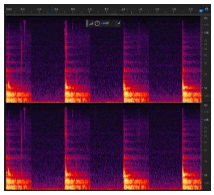
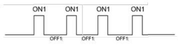
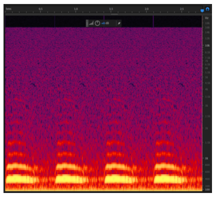
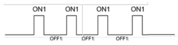
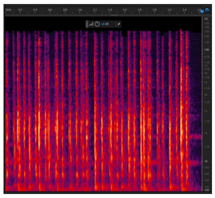
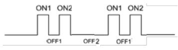
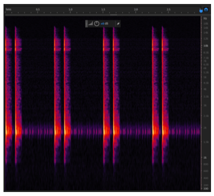
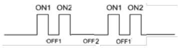
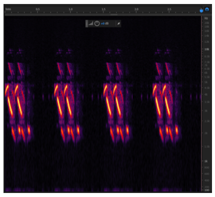
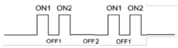
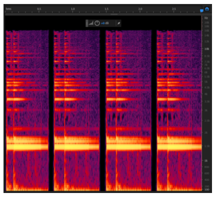
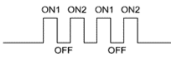
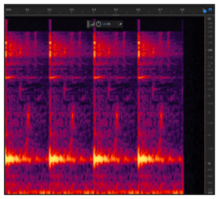
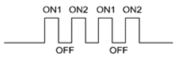
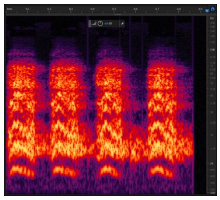
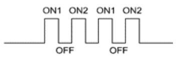
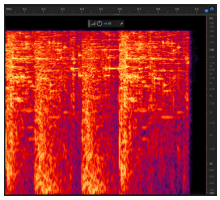
Earcon	C6	1000Hz	0.1	0.1	0.7		

Table 4. The sound sample stimuli for auditory icons

Sound Type	Code	Pitch (Fund. Frequency)	ON time (s)	OFF time (s)	Duration (s)	Sound Pattern	Spectrograph of Stimuli
Auditory Icon	D1	500 ~800Hz (Fund. Frequency)	0.1	-	0.1		
Auditory Icon	E1	2500 ~2800Hz (Fund. Frequency)	0.1	-	0.1		
Auditory Icon	F1	600 ~1000Hz (Fund. Frequency)	0.1	-	0.1		
Auditory Icon	D2	400 ~500Hz (Fund. Frequency)	0.5	-	0.5		

Auditory Icon	E2	5200 ~5500Hz (Fund. Frequency)	0.5	-	0.5		
Auditory Icon	F2	400 ~600Hz (Fund. Frequency)	0.5	-	0.5		
Auditory Icon	D3	800 ~1000Hz (Fund. Frequency)	0.05	0.05	0.15		
Auditory Icon	E3	2800 ~3000Hz (Fund. Frequency)	0.05	0.05	0.15		
Auditory Icon	F3	100 ~200Hz (Fund. Frequency)	0.05	0.05	0.15		
Auditory Icon	D4	1000 ~1300Hz (Fund. Frequency)	0.5	0.2	2.6		

Auditory Icon	E4	300 ~400Hz (Fund. Frequency)	0.5	0.2	2.6		
Auditory Icon	F4	800 ~900Hz (Fund. Frequency)	0.5	0.2	2.6		
Auditory Icon	D2	1800 ~2000Hz (Fund. Frequency)	0.1	0.05, 0.5	2.5		
Auditory Icon	E2	3200 ~3500Hz (Fund. Frequency)	0.1	0.05, 0.5	2.5		
Auditory Icon	F2	1400 ~1500Hz (Fund. Frequency)	0.1	0.05, 0.5	2.5		
Auditory Icon	D6	1000 ~1200Hz (Fund. Frequency)	0.1	0.1	0.7		

Auditory Icon	E6	1500~1800Hz (Fund. Frequency)	0.1	0.1	0.7		
Auditory Icon	F6	3500~4000Hz (Fund. Frequency)	0.1	0.1	0.7		

3.1.4 Measurement

A 7-point Likert scale were used for this pilot test to measure the participants perceivability to hear the sounds and whether each intended information is intuitive to the participants. Similar measurement methods were used in sound evaluation assessment and auditory-related researches (Amann & Anderson, 2014; Chi et al., 2017; Lazaro et al., 2022). The evaluated measures for intuitiveness of the sound samples in this pilot test will determine if the sound samples created matches with the intended information. The scenario evaluation is a multiple-choice questionnaire where it requires participants to select the appropriate scenario for each of the sound sample given.

3.1.5 Experiment Procedure

The pilot test is divided into three parts; (i) pre-evaluation, (ii) main evaluation, and (iii) post-evaluation. Prior to the pilot test, explanation was given to all participants regarding the overall study aim for the pilot test and they were asked for their consent to participate in the pilot test. After obtaining their consent, all participants were asked to complete a set of short demographic survey form and fill in their basic information. The pilot test procedure was explained in details. All participants were ensured that they understood the pilot test procedure and questionnaire meanings. The participants were also screened through hearing test using the DB-23000 Audiometer by asking them to raise one of their hand if they were able to hear a set of randomized auditory signal set with a frequency of 500Hz, 1000Hz, 2000Hz and 4000Hz at 5dB loudness.

The main pilot test evaluation was conducted in the experiment room located in Seoul National University Engineering building 39 in a closed acoustic chamber for sound evaluation. Previous studies have confirmed that sound evaluation or jury testing experiment should be conducted in a closed room where there is no sound interference during evaluation or experiment (Otto et al., 2001; Brizon & Medeiros, 2012). During the main pilot test evaluation, the participants listened to each sound sample and were instructed immediately to evaluate its perceivability and intuitiveness. Participants were also instructed to choose the appropriate scenario from the multiple choice for each sound sample heard. All of the sound samples were in random order to eliminate order effects during the evaluation. The main pilot evaluation was divided into two sections with a 10 minutes break to avoid fatigue effects.

After the main pilot test evaluation, the participants were asked to write their opinions regarding the sound stimuli's intended information and the necessity of sound among the choices of scenarios. Lastly, participants were asked about their overall evaluation of the pilot test. Overall pilot test took approximately 120 minutes.

3.1.6 Data Analysis Approach

The data obtained from the dependent variables, which are perceivability and intuitiveness of the sound were analyzed using one-way analysis of variance (ANOVA) for each of the sound information types (confirmatory, error, detection, in progress, alert and warning) and set the significance level at p-value of 0.05. Post-hoc test was also conducted for multiple comparisons using Bonferroni correction at $\alpha = 0.05$ to reduce the instance of a false positive. Meanwhile, for the multiple-choice questionnaire, the data were analyzed using frequency analysis in order to figure out the cumulative frequency of the appropriate scenario for each sound sample chosen by the participants.

3.1.7 Results

The result of the analyses is presented based on the tables and graph below. In this pilot test, the result of perceivability of the intended sound information type will be presented first, then followed by the intuitiveness of each sound information types. The results will include descriptive statistic tables, ANOVA summary tables and graphs for perceivability and intuitiveness for each sound information types. Also, results for the frequency analysis for appropriate scenario under the multiple-choice questionnaire will be presented according to the scenario phases (from ingress to egress) in a table and histogram graphs.

Perceivability

The perceivability of the sound samples result data for pilot test is presented in Table 5, 6 and 7 below. Based on the data in Table 5, the perceivability mean value for confirmatory auditory signals which has a value of 3.513±(0.166) is the lowest and followed by detection signal with 3.551±(0.149), error signal with 4.769±(0.163), in progress signal with 4.872±(0.185), alert signal with 4.974±(0.171) and the highest perceivability mean value is warning signal with 5.308±(0.153).

Table 5. Descriptive statistics for Perceivability

	Mean	Std. Error	95% Confidence Interval for Mean	
			Lower Bound	Upper Bound
Confirmatory	3.513	0.166	3.182	3.843
Error	4.769	0.163	4.444	5.094
Detection	3.551	0.149	3.255	3.848
In Progress	4.872	0.185	4.504	5.239
Alert	4.974	0.171	4.634	5.315
Warning	5.308	0.153	5.003	5.612

Table 6 and 7 shows a summary of one-way ANOVA with Bonferroni's Post Hoc Test for perceivability. From Table 6, the result shows that there were significant differences $p < 0.001$ between each groups of sound information. Post-hoc analysis with Bonferroni correction indicated that among the sound information types, confirmatory signal and detection signal has no significance among each type. Similarly, error signal, in progress signal, alert signal and warning signal also has no significance (however, error signal and warning signal has slightly significant difference, $p < 0.05$).

Table 6. ANOVA summary for Perceivability

	Sum of Squares	df	Mean Square	F	*p*-value
Between groups	9468.002	1	9468.002	2708.515	0.00
Within group	269.165	77	3.496		
Total	9737.167				

Table 7. One-way ANOVA (Bonferroni's Post-hoc Test) for Perceivability

		Mean Difference	Std. Error	*p*-value
Confirmatory	Error	-1.256**	0.234	0.00
	Detection	-0.038	0.252	1.00
	In Progress	-1.359**	0.257	0.00
	Alert	-1.462**	0.245	0.00
	Warning	-1.795**	0.225	0.00
Error	Confirmatory	1.256**	0.234	0.00
	Detection	1.218**	0.231	0.00
	In Progress	-0.103	0.195	1.00
	Alert	-0.205	0.203	1.00
	Warning	-0.538*	0.176	0.05
Detection	Confirmatory	0.038	0.252	1.00
	Error	-1.218**	0.231	0.00
	In Progress	-1.321**	0.242	0.00
	Alert	-1.423**	0.221	0.00
	Warning	-1.756**	0.226	0.00
In Progress	Confirmatory	1.359**	0.257	0.00
	Error	0.103	0.195	1.00
	Detection	1.321**	0.242	0.00
	Alert	-0.103	0.179	1.00
	Warning	-0.436	0.185	0.31
Alert	Confirmatory	1.462**	0.245	0.00
	Error	0.205	0.203	1.00
	Detection	1.423**	0.221	0.00
	In Progress	0.103	0.179	1.00
	Warning	-0.333	0.164	0.68
Warning	Confirmatory	1.795**	0.225	0.00
	Error	0.538*	0.176	0.05
	Detection	1.756**	0.226	0.00
	In Progress	0.436	0.185	0.31
	Alert	0.333	0.164	0.68

* $p < 0.05$, ** $p < 0.001$

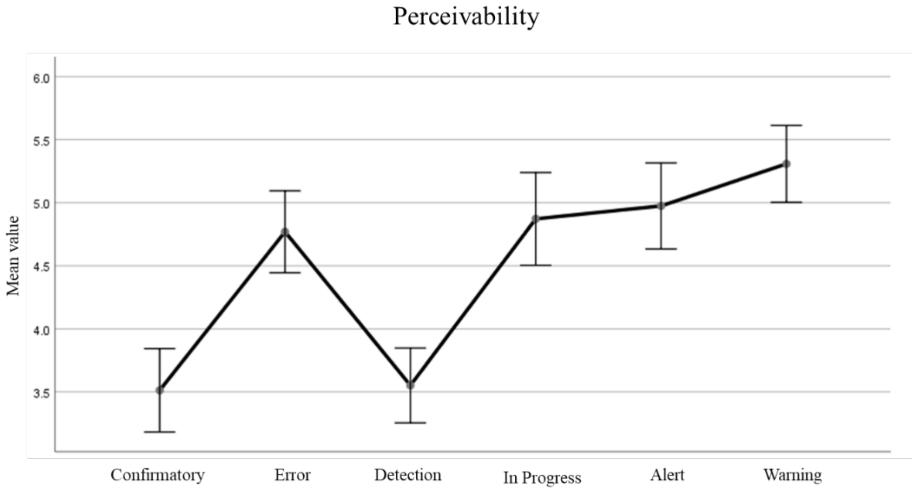


Figure 12. Graph for Perceivability

Intuitiveness (Confirmatory type signal)

The intuitiveness of the confirmatory type signal is measured by asking the participants to evaluate in a 7-point Likert scale based on how intuitive they think that the sound samples they heard would convey its message or information as a 'confirmatory signal'. Thus, the intuitiveness of the intended sounds as confirmatory type is presented in the tables below.

The intuitiveness of the sound samples as a confirmatory type signal result data for pilot test is presented in Table 8, 9 and 10 below. Based on the data in Table 8, the intuitiveness (confirmatory) mean value for the intended confirmatory signals has the highest value of 4.244±(0.215) which is the highest as expected. This means that the listeners perceived the confirmatory type signals hold the intuitive meaning of 'confirmatory' as this study intended. Meanwhile, in-progress type signals has the lowest mean value of 2.346±(0.157) which means that the in-progress type signal does not give any information about 'confirmation' in the context.

Table 8. Descriptive statistics for Intuitiveness (as 'Confirmatory' type signal)

	Mean	Std. Error	95% Confidence Interval for Mean	
			Lower Bound	Upper Bound
Confirmatory	4.244	0.215	3.815	4.672
Error	3.167	0.193	2.781	3.552
Detection	3.808	0.197	3.416	4.199
In Progress	2.346	0.157	2.034	2.658
Alert	2.590	0.157	2.276	2.903
Warning	2.744	0.179	2.387	3.100

Table 8 and 9 shows a summary of one-way ANOVA with Bonferroni's Post Hoc Test for intuitiveness (as 'confirmatory' type). From Table 8, the overall result shows that there were significant differences $p < 0.001$ between the intuitiveness of confirmatory type signal among each other groups of sound information. Post-hoc analysis with Bonferroni correction, on the other hand, indicated that only the intended confirmatory type signal has significant difference compared to error, in progress, alert and warning at the significance level of $p < 0.05$ and $p < 0.001$, except for detection type signal. This can be inferred that both confirmatory and detection type signal has similar information pattern type which unable the listener to distinguish the conveyed information. Thus, from this result, it is secure to assume that the current sound temporal pattern (tempo and beat) with ON(0.1s)/OFF(0s) time that is intended can be used as a 'confirmatory' type signal.

Table 9. ANOVA summary for Intuitiveness (as 'Confirmatory' type signal)

	Sum of Squares	df	Mean Square	F	*p*-value
Between groups	4642.470	1	4642.470	1076.080	0.00
Within group	332.197	77	4.314		
Total	4974.667				

Table 10. One-way ANOVA (Bonferroni's Post-hoc Test) for Intuitiveness (as 'Confirmatory' type signal)

Dependent variable			Mean Difference	Std. Error	*p*-value
(Intuitiveness) How intuitive do you think this sound would be as a "confirmatory signal"?	Confirmatory	Error	1.077*	0.293	0.01
		Detection	0.436	0.300	1.00
		In Progress	1.897**	0.261	0.00
		Alert	1.654**	0.266	0.00
		Warning	1.500**	0.282	0.00

* $p < 0.05$, ** $p < 0.001$

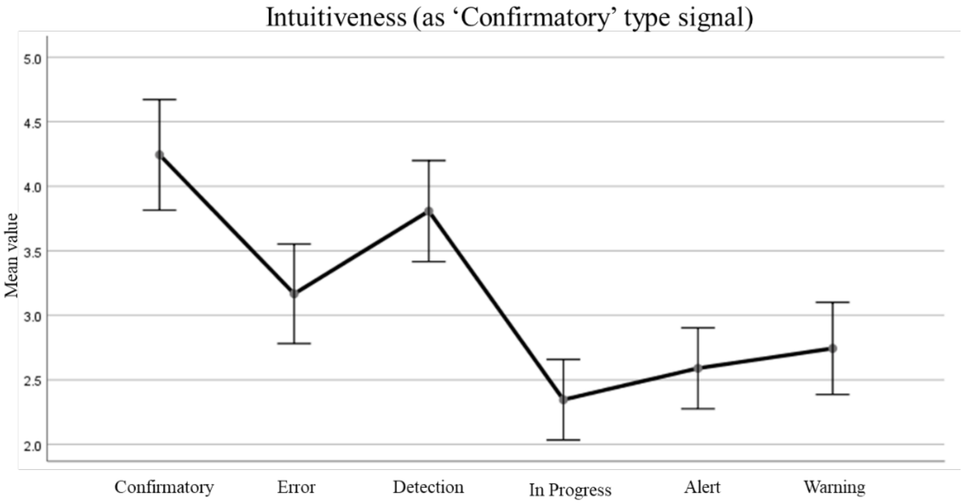


Figure 13. Graph of Intuitiveness (as 'Confirmatory' type signal)

Intuitiveness (Error type signal)

Similar with the above-mentioned intuitiveness measure, the intuitiveness of error type signal is measured by asking the participants to evaluate in a 7-point Likert scale based on how intuitive they think that the sound samples they heard would convey its message or information as an 'error signal'. Thus, the intuitiveness of the intended sounds as error type is presented in the tables below.

The intuitiveness of the sound samples as a confirmatory type signal result data for pilot test is presented in Table 11, 12 and 13 below. Based on the data in Table 11, the intuitiveness mean value for the intended error signal has the value of 3.744±(0.199). Though the intended error type signal has similar mean value with in-progress, 3.590±(0.222) and alert, 3.859±(0.183), it is confirmed that confirmatory type signal has the lowest intuitive mean value for error type, 2.859±(0.170) compared to the intended error type signal.

Table 11. Descriptive statistics for Intuitiveness (as 'Error' type signal)

			95% Confidence Interval for Mean	
	Mean	Std. Error	Lower Bound	Upper Bound
Confirmatory	2.859	0.170	2.521	3.197
Error	3.744	0.199	3.348	4.139
Detection	3.231	0.206	2.821	3.640
In Progress	3.590	0.222	3.147	4.032
Alert	3.859	0.183	3.494	4.224
Warning	4.077	0.187	3.705	4.449

Table 12 and 13 shows a summary of one-way ANOVA with Bonferroni's Post Hoc Test for intuitiveness (as 'error' type). From Table 12, it shows that there were significant differences $p < 0.001$ between the intuitiveness of error type signal among each other groups of sound information. Also, from the post-hoc analysis with Bonferroni correction, it was found that only confirmatory type signal has significance effect with error type signal at $p < 0.05$. Other information types, such as detection, in-progress, alert and warning show no significance which indicated that there is a need for listener to have attention to the sound whether it is in an error situation or progress notification. Also, as long as the intended error type signal has significance difference from confirmatory type signal, the intended temporal pattern (tempo and beat) with ON(0.5s)/OFF(0s) time can be used as a 'confirmatory' type signal for the main evaluation.

Table 12. ANOVA summary for Intuitiveness (as 'Error' type signal)

	Sum of Squares	df	Mean Square	F	*p*-value
Between groups	5930.675	1	5930.675	1060.382	0.00
Within group	430.658	77	5.593		
Total	6361.333				

Table 13. One-way ANOVA (Bonferroni's Post-hoc Test) for Intuitiveness (as 'Error' type signal)

Dependent variable			Mean Difference	Std. Error	*p*-value
(Intuitiveness) How intuitive do you think this sound would be as an "error signal"?	Error	Confirmatory	0.885*	0.293	0.05
		Detection	0.513	0.224	0.37
		In Progress	0.154	0.265	1.00
		Alert	-0.115	0.231	1.00
		Warning	-0.333	0.254	1.00

* $p < 0.05$, ** $p < 0.001$

Intuitiveness (as 'Error' type signal)

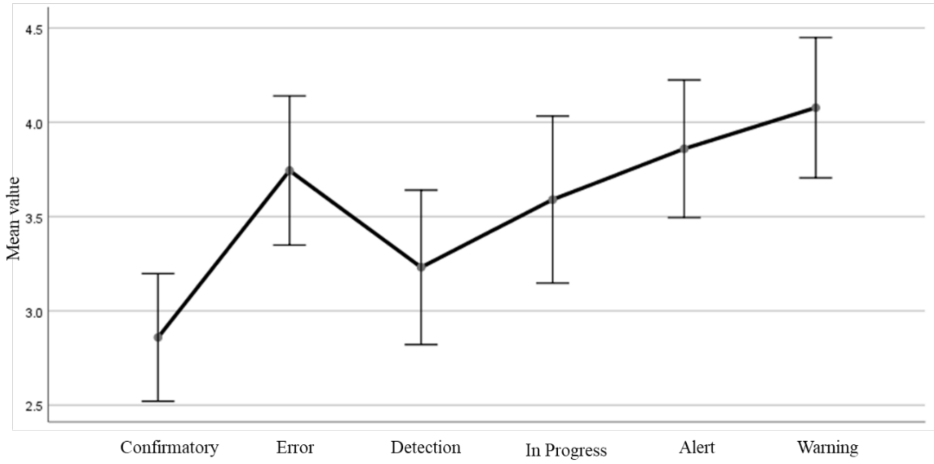


Figure 14. Graph of Intuitiveness (as 'Error' type signal)

Intuitiveness (Detection type signal)

The intuitiveness of detection type signal is measured by asking the participants to evaluate in a 7-point Likert scale based on how intuitive they think that the sound samples they heard would convey its message or information as a ‘detection signal’. Thus, the intuitiveness of the intended sounds as detection type is presented in Table 14, 15 and 16 below.

Based on the data in Table 14, the intuitiveness mean value for the intended detection type signal has the value of 3.885±(0.211). Although the intuitiveness mean value for the intended detection type signal is not the highest, it can be seen that alert type signal recorded the highest mean value among other types of sound information at 4.564±(0.166).

Table 14. Descriptive statistics for Intuitiveness (as ‘Detection’ type signal)

			95% Confidence Interval for Mean	
	Mean	Std. Error	Lower Bound	Upper Bound
Confirmatory	3.359	0.191	2.978	3.740
Error	3.410	0.223	2.966	3.854
Detection	3.885	0.211	3.464	4.305
In Progress	3.731	0.245	3.244	4.218
Alert	4.564	0.166	4.234	4.894
Warning	3.936	0.215	3.509	4.363

Table 15 and 16 shows a summary of one-way ANOVA with Bonferroni’s Post Hoc Test for intuitiveness (as ‘detection’ type). From Table 15, in overall, there are significant differences between sound type groups. Post-hoc Bonferroni correct shows that the intended detection type has distinguish significance with alert type signal at $p < 0.05$. Meanwhile, other sound types, such as confirmatory, error, in-progress and warning does not show any significance. Based on the post-hoc analysis also, the detection type signal and warning type signal both has similar temporal pattern (tempo and beat) of ON1=ON2, with OFF1, thus, allowing the listeners to perceived both detection type signal and warning signal as the same. In addition, the detection and confirmatory also shows no significance difference among both types due to the fact that both sound types are applicable in a certain scenario such as 'detecting passenger before entering the vehicle' and 'confirming the passenger's presence before entering the vehicle'.

Table 15. ANOVA summary for Intuitiveness (as 'Detection' type signal)

	Sum of Squares	df	Mean Square	F	*p*-value
Between groups	6808.173	1	6808.173	807.343	0.00
Within group	649.327	77	8.433		
Total	7457.500				

Table 16. One-way ANOVA (Bonferroni's Post-hoc Test) for Intuitiveness (as 'Detection' type signal)

Dependent variable			Mean Difference	Std. Error	*p*-value
(Intuitiveness) How intuitive do you think this sound would be as a "detection signal"?	Detection	Confirmatory	0.526	0.294	1.00
		Error	0.474	0.201	0.32
		In Progress	0.154	0.272	1.00
		Alert	-0.679*	0.216	0.04
		Warning	-0.051	0.221	1.00

* $p < 0.05$, ** $p < 0.001$

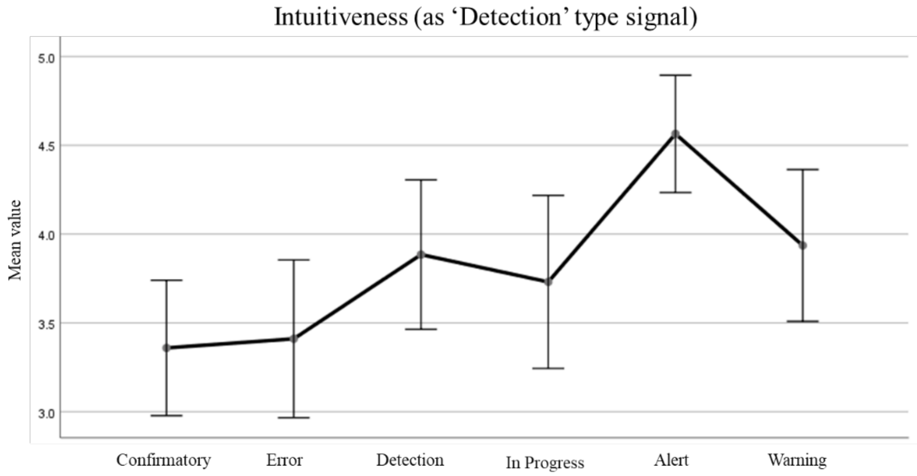


Figure 15. Graph of Intuitiveness (as 'Detection' type signal)

Intuitiveness (In Progress type signal)

The intuitiveness of in-progress type signal is measured by asking the participants to evaluate in a 7-point Likert scale based on how intuitive they think that the sound samples they heard would convey its message or information as an 'in-progress signal'. Thus, the intuitiveness of the intended sounds as in-progress type is presented in Table 17, 18 and 19 below.

Based on the data in Table 17, the intuitiveness mean value for the intended in-progress type signal has the value of 3.256±(0.201). Although confirmatory type signal has the lowest intuitive mean value for the intended in-progress signal with 2.949±(0.192), the other information type, such as error 3.103±(0.212), detection 3.744±(0.193), alert 3.615±(0.191), and warning 3.013±(0.194) has similar mean value with the intended in-progress sound.

Table 17. Descriptive statistics for Intuitiveness (as 'In Progress' type signal)

	Mean	Std. Error	95% Confidence Interval for Mean	
			Lower Bound	Upper Bound
Confirmatory	2.949	0.192	2.566	3.331
Error	3.103	0.212	2.679	3.526
Detection	3.744	0.193	3.360	4.127
In Progress	3.256	0.201	2.856	3.657
Alert	3.615	0.191	3.236	3.995
Warning	3.013	0.194	2.627	3.398

Table 18 and 19 shows a summary of one-way ANOVA with Bonferroni's Post Hoc Test for intuitiveness (as 'in-progress' type). From Table 18, in overall, there are significant differences between sound type groups. However, based from the post-hoc Bonferroni correct, the result indicated each group of sound information types has no significance at all with the intended in-progress type signal at $p < 0.05$ or $p < 0.001$. This can be inferred that the in-progress type signal can be confused with either confirmatory, error, detection, alert or warning as the information or message regarding the system's status of 'in-progress' does not conveyed effectively to the listeners or users. Hence, from this data obtained in pilot test, the intended 'in-progress' information type of signal should be re-defined or re-designed to match its congruity purpose accordingly.

Table 18. ANOVA summary for Intuitiveness (as 'In Progress' type signal)

	Sum of Squares	df	Mean Square	F	*p*-value
Between groups	5034.669	1	5034.669	945.925	0.00
Within group	409.831	77	5.322		
Total	5444.500				

Table 19. One-way ANOVA (Bonferroni's Post-hoc Test) for Intuitiveness
(as 'In Progress' type signal)

Dependent variable			Mean Difference	Std. Error	*p*-value
(Intuitiveness) How intuitive do you think this sound would be as an "in-progress signal"?	In Progress	Confirmatory	0.308	0.297	1.00
		Error	0.154	0.248	1.00
		Detection	-0.487	0.228	0.53
		Alert	-0.359	0.251	1.00
		Warning	0.244	0.277	1.00

* $p < 0.05$, ** $p < 0.001$

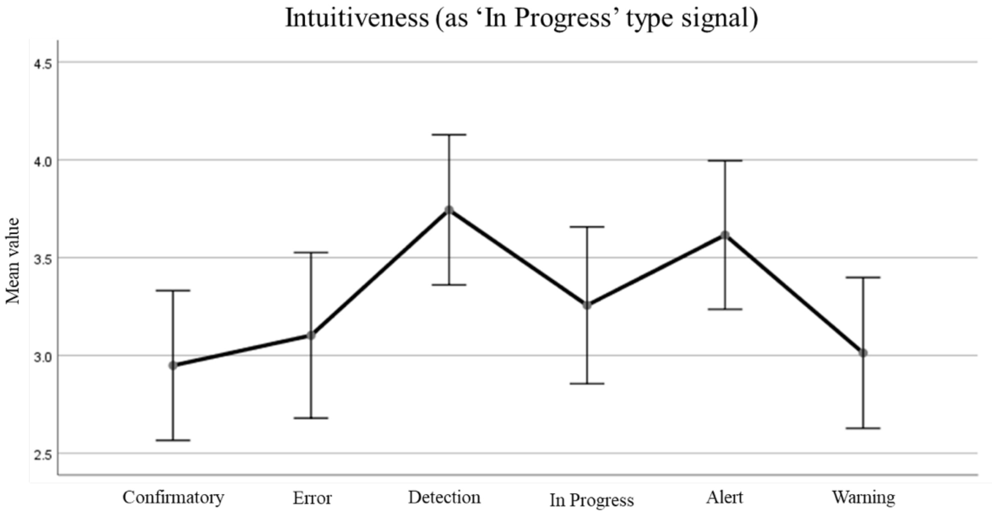


Figure 16. Graph of Intuitiveness (as 'In Progress' type signal)

Intuitiveness (Alert type signal)

The intuitiveness of alert type signal is measured by asking the participants to evaluate in a 7-point Likert scale based on how intuitive they think that the sound samples they heard would convey its message or information as an 'alert signal'. Thus, the intuitiveness of the intended sounds as alert type is presented in Table 20, 21 and 22 below.

Based on the data in Table 20, the intuitiveness mean value for the intended alert type signal has the value of 4.282±(0.206) which has the highest mean value among other information types, followed by warning type signal with the mean value of 4.179±(0.195). Both alert and warning types have similar intended information but with different levels of perceived urgency. On the other hand, confirmatory type signal has the lowest mean value of 2.436±(0.158) among all other information types.

Table 20. Descriptive statistics for Intuitiveness (as 'Alert' type signal)

			95% Confidence Interval for Mean	
	Mean	Std. Error	Lower Bound	Upper Bound
Confirmatory	2.436	0.158	2.122	2.750
Error	3.077	0.196	2.686	3.468
Detection	2.731	0.178	2.377	3.084
In Progress	4.026	0.228	3.572	4.479
Alert	4.282	0.206	3.871	4.693
Warning	4.179	0.195	3.792	4.567

Table 21 and 22 shows a summary of one-way ANOVA with Bonferroni's Post Hoc Test for intuitiveness (as 'alert' type). From Table 21, the result shows that there were significant differences $p < 0.001$ between each groups of sound information. Post-hoc analysis with Bonferroni correction indicated that among the sound information types, confirmatory, error and detection type signals have great significant value at $p < 0.001$ compared to alert type. However, the in-progress and warning type of signal shows no significance with the intended signal type as the intended information that is conveyed through in-progress and warning is undistinguishable from the listener's perspective. Hence, it is also true that both alert and warning types of signal are designed to prompt the listener's attention for intervention. Alert type signal also can be viewed as a 'soft warning', and there are no significant difference between warning type signal.

Table 21. ANOVA summary for Intuitiveness (as 'Alert' type signal)

	Sum of Squares	df	Mean Square	F	*p*-value
Between groups	5586.942	1	5586.942	888.420	0.00
Within group	484.224	77	6.289		
Total	6071.167				

Table 22. One-way ANOVA (Bonferroni's Post-hoc Test) for Intuitiveness (as 'Alert' type signal)

Dependent variable			Mean Difference	Std. Error	*p*-value
(Intuitiveness)	Alert	Confirmatory	1.846**	0.255	0.00
How intuitive do		Error	1.205**	0.244	0.00
you think this		Detection	1.551**	0.223	0.00
sound would be as		In Progress	0.256	0.263	1.00
an "alert signal"?		Warning	0.103	0.264	1.00

* $p < 0.05$, ** $p < 0.001$

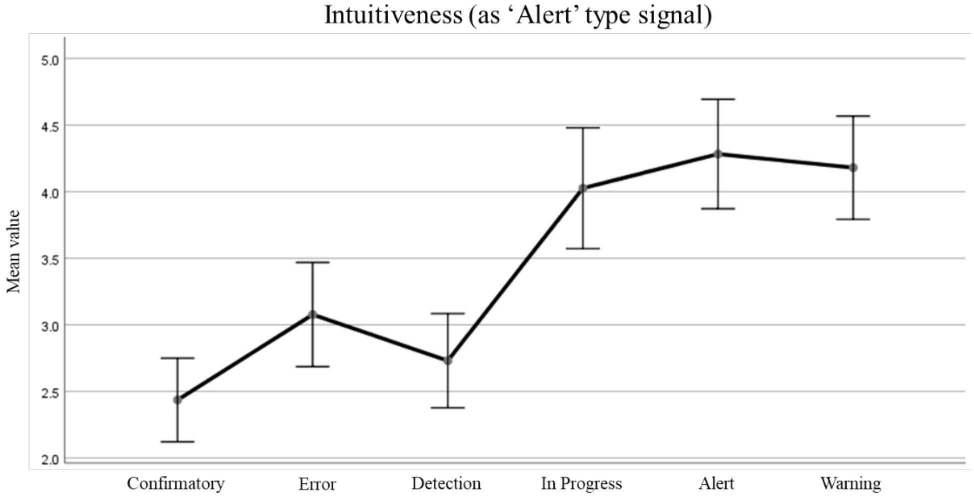


Figure 17. Graph of Intuitiveness (as 'Alert' type signal)

Intuitiveness (Warning type signal)

The intuitiveness of warning type signal is measured by asking the participants to evaluate in a 7-point Likert scale based on how intuitive they think that the sound samples they heard would convey its message or information as a 'warning signal'. Thus, the intuitiveness of the intended sounds as warning type is presented in Table 23, 24 and 25 below.

Based on the data in Table 23, the intuitiveness mean value of warning type signal is 2.949±(0.210) which is lower than the highest mean value of in-progress 3.167±(0.255). Meanwhile, alert type has similar mean value compared to warning type at 2.846±(0.226), whereas, confirmatory, error and detection has the lower mean value of 1.141±(0.044), 1.705±(0.124) and 1.423±(0.109) respectively.

Table 23. Descriptive statistics for Intuitiveness (as 'Warning' type signal)

			95% Confidence Interval for Mean	
	Mean	Std. Error	Lower Bound	Upper Bound
Confirmatory	1.141	0.044	1.054	1.228
Error	1.705	0.124	1.458	1.952
Detection	1.423	0.109	1.207	1.640
In Progress	3.167	0.255	2.659	3.675
Alert	2.846	0.226	2.397	3.296
Warning	2.949	0.210	2.530	3.368

Table 24 and 25 shows a summary of one-way ANOVA with Bonferroni's Post Hoc Test for intuitiveness (as 'warning' type). Table 24 shows that the overall information types of signal between groups has significance at $p < 0.001$. From Table 25, results of the post-hoc analysis with Bonferroni correction revealed that the confirmatory, error and detection types of signals were significantly lower than the in-progress and alert types of signals at $p < 0.001$. Except for in-progress information type signal, the warning type and alert type both has no significant differences which aligns to the intended sound information congruity that these both signals should have the ability to convey alert notifications or warning situations to the listener when intervention is needed in autonomous vehicle.

Table 24. ANOVA summary for Intuitiveness (as 'Warning' type signal)

	Sum of Squares	df	Mean Square	F	*p*-value
Between groups	2275.692	1	2275.692	598.101	0.00
Within group	292.974	77	3.805		
Total	2568.667				

Table 25. One-way ANOVA (Bonferroni's Post-hoc Test) for Intuitiveness (as 'Warning' type signal)

Dependent variable			Mean Difference	Std. Error	*p*-value
(Intuitiveness)	Warning	Confirmatory	1.808**	0.213	0.00
How intuitive do		Error	1.244**	0.223	0.00
you think this		Detection	1.526**	0.238	0.00
sound would be as		In Progress	-0.218	0.319	1.00
a "warning signal"?		Alert	0.103	0.277	1.00

* $p < 0.05$, ** $p < 0.001$

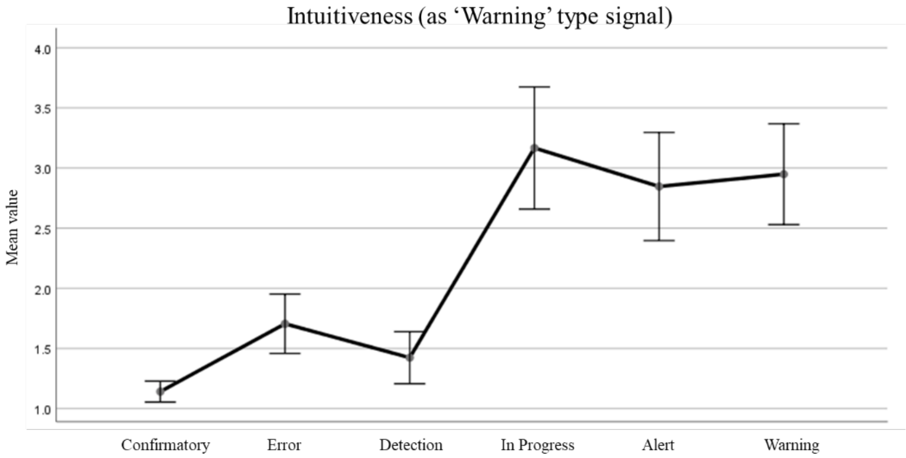


Figure 18. Graph of Intuitiveness (as 'Warning' type signal)

Appropriateness between intended auditory information types and scenarios

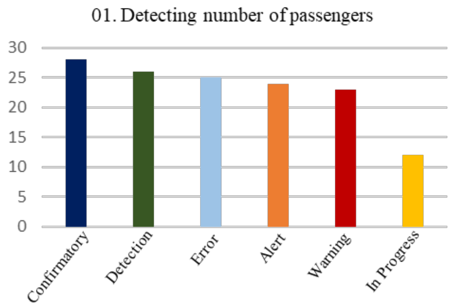
To evaluate which scenarios would be the best fit for the intended auditory information type signals, this study incorporates multiple-choice questionnaires for the participants to select all the appropriate scenarios for the respective information type signals (confirmatory, error, detection, in-progress, alert and warning). Prior to this pilot test, a timeline of passenger-oriented context of riding an autonomous vehicle has been developed. Through this pilot test evaluation, this study will add appropriate scenarios or remove any unnecessary scenarios which will be used for the main sound set evaluation.

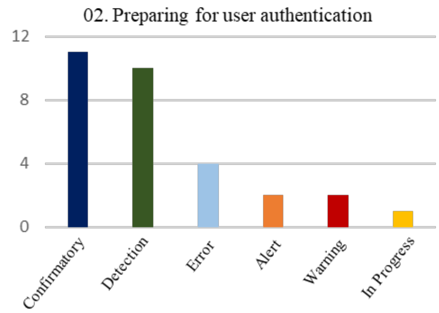
In Chapter 2, the developed scenarios consist of five phases; (i) ingress phase, (ii) pre-driving phase, (iii) mid-driving phase, (iv) egress phase, and (iv) post-driving phase. The result for this pilot test will include these phases in an orderly manner and the result of frequency analysis will be represented in tables and histograms.

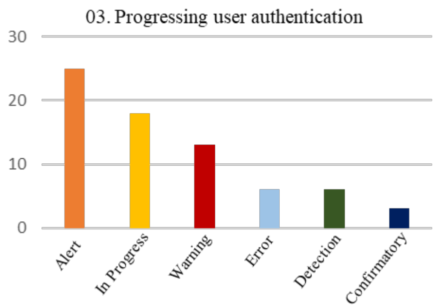
Table 26 shows the overall result for the frequency analysis based on the multiple-choice questionnaire that the participants answered for the pilot test. For the (i) ingress phase, a total of 138 frequency counts where the participants accumulatively consider the scenario if 'detecting number of passengers' is appropriate for the in-vehicle signals feedback, followed by the second highest total of frequency count of 113, which is the 'failed for user authentication' scenario. For the (ii) pre-driving phase, 'notification for seat belt' scenario records the highest total frequency (120 counts) followed by seat adjustments during pre-driving (101 counts). In (iii) mid-driving phase, a total frequency of 92 counts accumulated for 'head movement, eye gaze and voice input direction' scenario, and then followed by 'verifying passenger's stress, motion sickness, drowsiness, etc.' scenario with 84 counts. For (iv) egress phase, the highest total frequency is recorded for 'notification for personal belongings' scenario with 140 counts, and followed 'seat adjustments during egress' with 71 counts. Finally, for (v) post-driving phase, 'detection for external movements (such as wild animals, etc.)' records at 151 counts, and 'external screen adjustment during post-driving' with 103 counts.

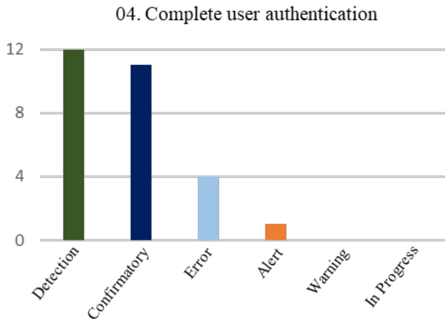
Table 26. Frequency analysis for the appropriateness of scenario and intended information types

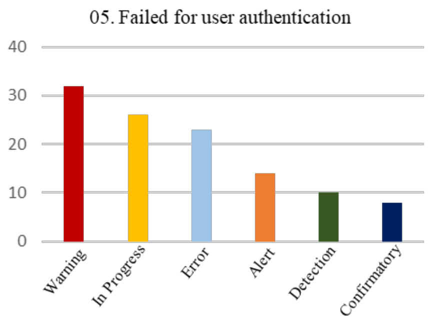
		Information Types						Σ(Total frequency)
		Confirmatory	Error	Detection	In Progress	Alert	Warning	
(i) Ingress Phase	01. Detecting number of passengers	28	25	26	12	24	23	138
	02. Preparing for user authentication	11	4	10	1	2	2	30
	03. Progressing user authentication	3	6	6	18	25	13	71
	04. Complete user authentication	11	4	12	0	1	0	28
	05. Failed for user authentication	8	23	10	26	14	32	113
	06. Opening and closing of vehicle doors during ingress	4	2	3	5	1	0	15
	07. Giving auditory feedback when the door has successfully opened or closed	2	0	2	1	1	0	6
(ii) Pre-driving Phase	08. Seat adjustments during pre-driving	5	11	14	18	32	21	101
	09. Complete seat adjustments during pre-driving	21	15	19	0	3	7	65
	10. Notification for seat belt	7	16	14	30	27	26	120
	11. Input destination, navigation, etc.	29	6	9	7	0	1	52
(iii) Mid-driving Phase	12. Head movement, eye gaze, voice input direction	14	11	23	8	18	18	92
	13. Verifying passenger's stress, motion sickness, drowsiness, etc.	5	12	14	15	21	17	84
	14. Seat adjustments during mid-driving	4	11	3	11	19	6	54
	15. Complete seat adjustments during mid-driving	17	8	14	0	1	5	45
	16. Internal screen adjustment during mid-driving	3	2	3	1	3	3	15
	17. Completed internal screen adjustment during mid-driving	3	4	2	1	1	2	13
(iv) Egress Phase	18. Notification for personal belongings	5	21	13	30	33	38	140
	19. Seat adjustments during egress	2	8	13	13	21	14	71
	20. Completed seat adjustments during egress	25	9	14	0	2	5	55
	21. Opening and closing of vehicle doors during egress	10	7	8	6	6	0	37
	22. Giving auditory feedback when the door has successfully opened or closed during egress	1	2	2	0	2	0	7
(v) Post-driving Phase	23. External screen adjustment during post-driving	13	14	19	17	24	16	103
	24. Seat adjustments during post-driving	24	15	20	0	2	5	66
	25. Detection for external movements	19	20	21	29	32	30	151
	26. Warning when detected an external hazard situations (wild animals, etc.)	0	6	2	18	10	19	55
	27. Alert when children were away from safe zone	1	2	1	1	1	1	7

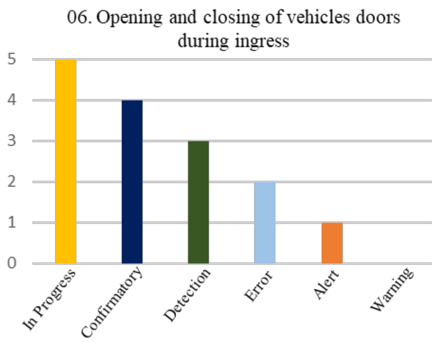

01. Detecting number of passengers
0
5
10
15
20
25
30
Confirmatory
Detection
Error
Alert
Warning
In Progress

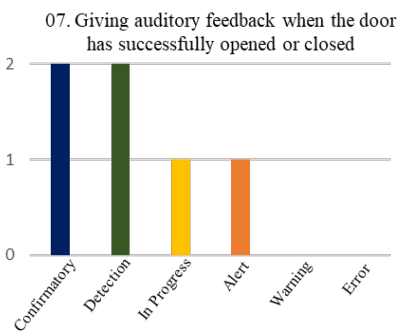

02. Preparing for user authentication
0
4
8
12
Confirmatory
Detection
Error
Alert
Warning
In Progress

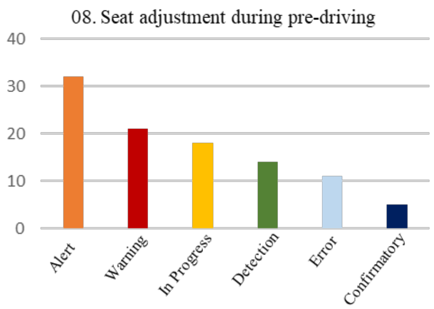

03. Progressing user authentication
0
10
20
30
Alert
In Progress
Warning
Error
Detection
Confirmatory

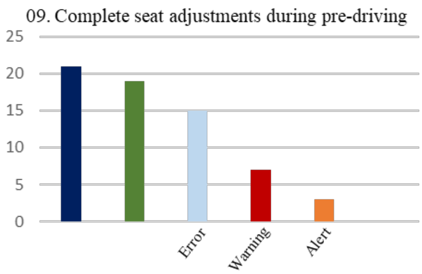

04. Complete user authentication
0
4
8
12
Detection
Confirmatory
Error
Alert
Warning
In Progress

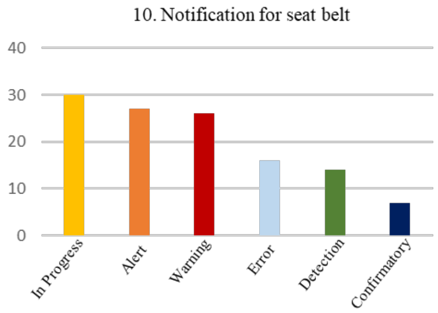

05. Failed for user authentication
0
10
20
30
40
Warning
In Progress
Error
Alert
Detection
Confirmatory

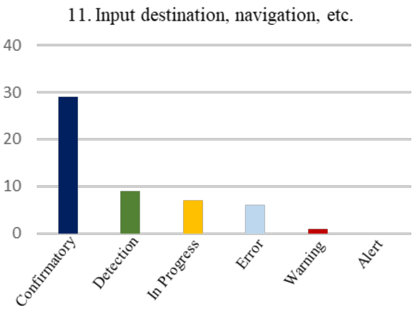

06. Opening and closing of vehicles doors during ingress
0
1
2
3
4
5
In Progress
Confirmatory
Detection
Error
Alert
Warning

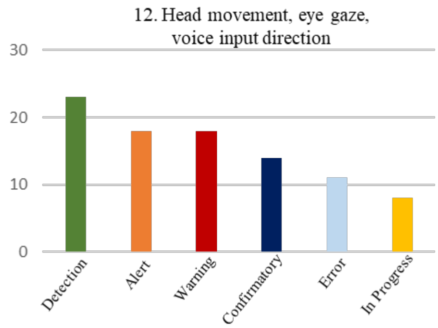

07. Giving auditory feedback when the door has successfully opened or closed
0
1
2
Confirmatory
Detection
In Progress
Alert
Warning
Error

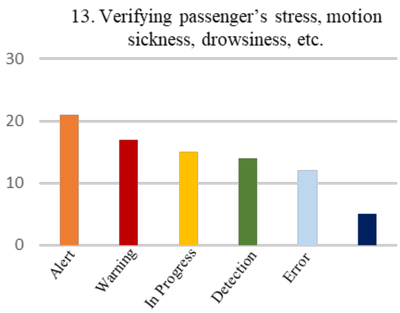

08. Seat adjustment during pre-driving
0
10
20
30
40
Alert
Warning
In Progress
Detection
Error
Confirmatory

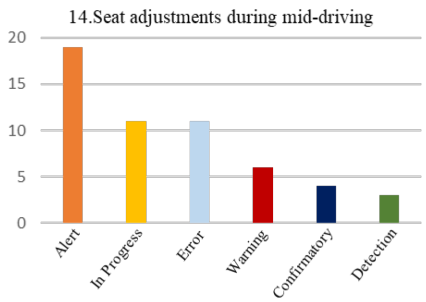

09. Complete seat adjustments during pre-driving
25
20
15
10
5
0
Error
Warning
Alert

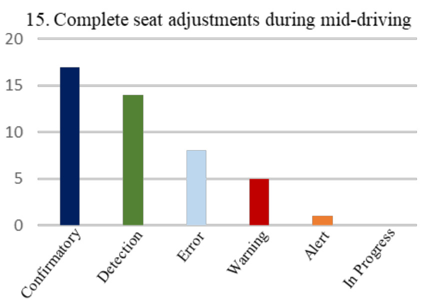

10. Notification for seat belt
40
30
20
10
0
In Progress
Alert
Warning
Error
Detection
Confirmatory

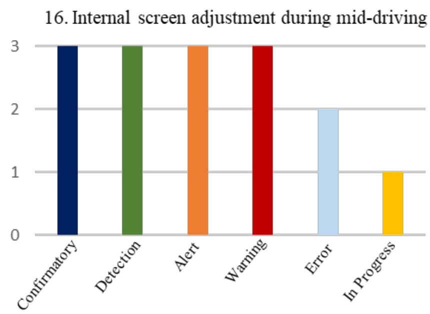

11. Input destination, navigation, etc.
40
30
20
10
0
Confirmatory
Detection
In Progress
Error
Warning
Alert

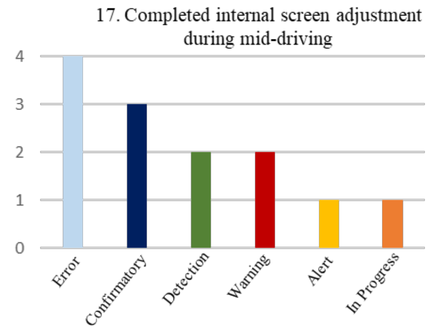

12. Head movement, eye gaze, voice input direction
30
20
10
0
Detection
Alert
Warning
Confirmatory
Error
In Progress

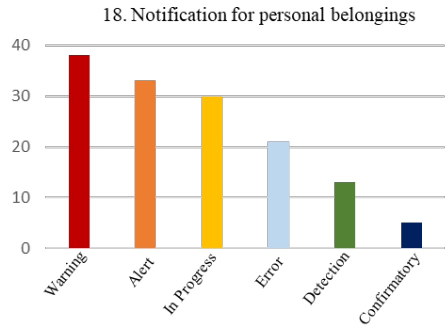

13. Verifying passenger's stress, motion sickness, drowsiness, etc.
30
20
10
0
Alert
Warning
In Progress
Detection
Error

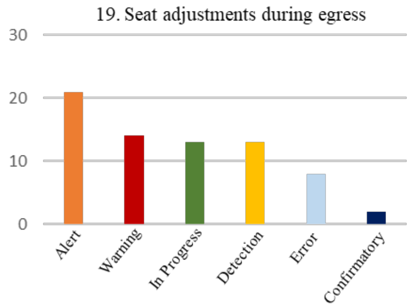

14.Seat adjustments during mid-driving
20
15
10
5
0
Alert
In Progress
Error
Warning
Confirmatory
Detection

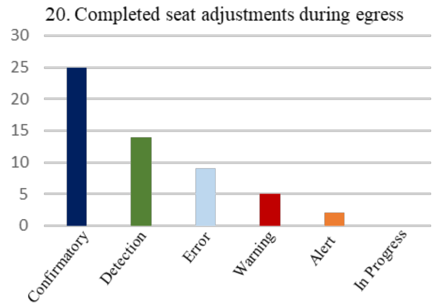

15. Complete seat adjustments during mid-driving
20
15
10
5
0
Confirmatory
Detection
Error
Warning
Alert
In Progress

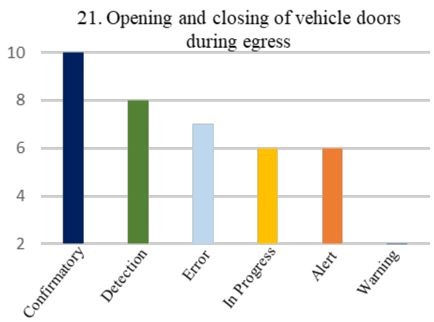

16. Internal screen adjustment during mid-driving
3
2
1
0
Confirmatory
Detection
Alert
Warning
Error
In Progress

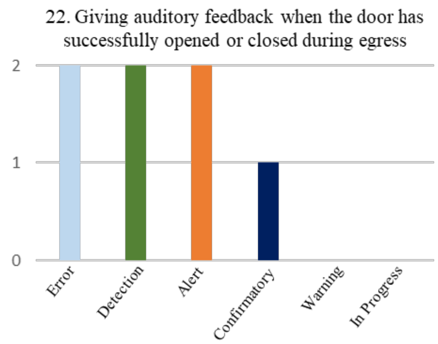

17. Completed internal screen adjustment during mid-driving
4
3
2
1
0
Error
Confirmatory
Detection
Warning
Alert
In Progress

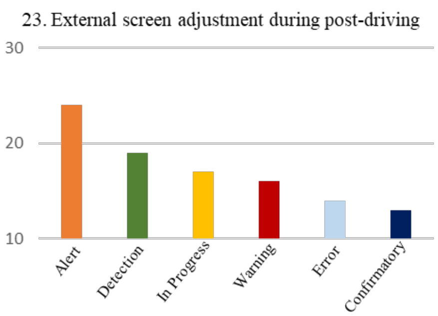

18. Notification for personal belongings
40
30
20
10
0
Warning
Alert
In Progress
Error
Detection
Confirmatory

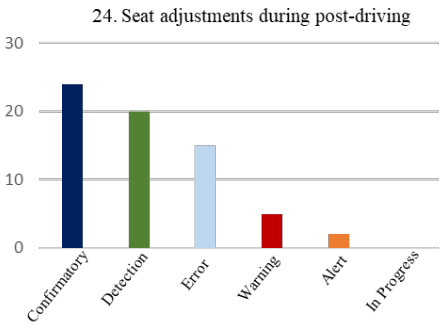

19. Seat adjustments during egress
30
20
10
0
Alert
Warning
In Progress
Detection
Error
Confirmatory


20. Completed seat adjustments during egress
30
25
20
15
10
5
0
Confirmatory
Detection
Error
Warning
Alert
In Progress


21. Opening and closing of vehicle doors during egress
10
8
6
4
2
Confirmatory
Detection
Error
In Progress
Alert
Warning


22. Giving auditory feedback when the door has successfully opened or closed during egress
2
1
0
Error
Detection
Alert
Confirmatory
Warning
In Progress


23. External screen adjustment during post-driving
30
20
10
Alert
Detection
In Progress
Warning
Error
Confirmatory


24. Seat adjustments during post-driving
30
20
10
0
Confirmatory
Detection
Error
Warning
Alert
In Progress

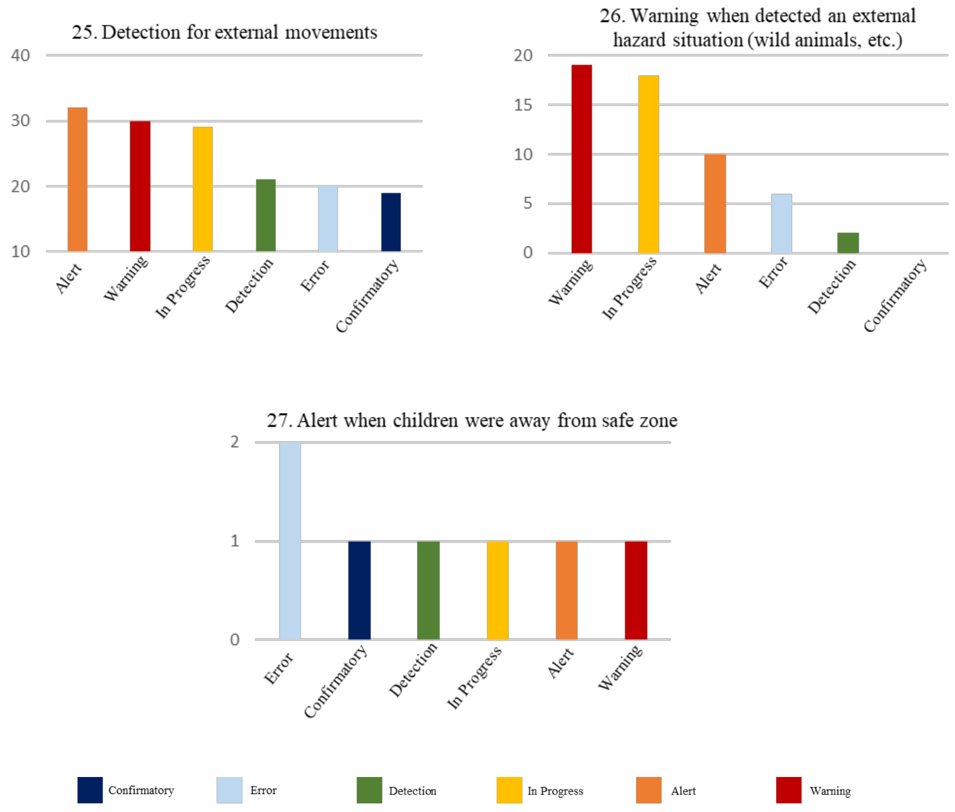


Figure 19. Histograms of frequency analysis for appropriateness between intended auditory information types and 27 scenarios

Based on the results of the frequency analysis, repetitive scenarios during pre-driving, mid-driving and post-driving such as ‘seat adjustments’, and ‘complete seat adjustments’ should be evaluated once rather than dividing these into pre-, mid-, and post-phases. Also, for the main sound set evaluation, ‘internal and external screen adjustments’ scenario will be removed, as it is not confirmed that future technology of the autonomous vehicle will incorporate such screen movements technologies. The top three scenarios (repetitive scenarios exempted) with the highest frequency value will be selected for the main sound set evaluation.

3.2 Sound Set Evaluation

3.2.1 Overview of sound set evaluation

Subsequent to the pilot test conducted in 3.1, this sound set evaluation will be the main study of this paper. Based on the data obtained from the pilot test, all sound sample stimuli indicated that there was significant difference among information type groups through the intuitiveness measure, except for 'in-progress' type sound. Also, it is confirmed that the in-vehicle signal with the temporal pattern that was hypothesized did match with the sound's intended information. Hence, in this sound set evaluation, 'in-progress' type sounds will be re-created and re-evaluated according to the appropriate scenario. Among all 27 scenarios that was evaluated in the pilot test, 15 scenarios were selected for this main evaluation. In addition, the condition for the scenario selection was based on the highest total value of frequency count, and repetitive scenarios, such as 'seats adjustments during pre-driving' and 'seat adjustments during mid-driving' is combined as one scenario. Furthermore, this main sound set evaluation aims to determine the appropriateness or consistency of the sound as a set of family of earcons and auditory icons based on a more specified version of scenarios. To fulfill the purpose of this study, the scenarios were divided into 10 scenarios regardless the timeline for passenger's context, but rather a specified version of scenario that is needed for appropriateness or consistency evaluation.

3.2.2 Participants

A total of 125 participants (58 males accounted for 46%, 67 females accounted for 54%) with an average age of 37.15(±11.4) participated in the main sound set evaluation. About 42 (33%) of the participants aged in their 20s, 30 (24%) aged in their 30s, 27 (22%) are at the age of their 40s, and 26 (21%) are aged above 50s. The participants on this study was recruited via an online survey platform service, and were requested to perform this sound set evaluation by using earphones or headsets to remove environmental effects that would affect their evaluation. Also, prior to the evaluation, participants were screened to have no issue with hearing disabilities through a screening question before participating.

3.2.3 Stimuli

The stimuli which will be used for this main sound set evaluation is similar with the created 36 sound samples, however, only a partial selection of the sound sample stimuli will be used according to the number of highest numbers of frequency, and highest total frequency for a sound type (earcon or auditory icon) based on the pilot test's result. Also, the sound sample stimuli which will be selected will also depending on the specified scenario. Based on Figure 20, the three highest value of total frequency scenarios will be the selection criterion for this main evaluation. Thus, from 27 scenarios, 15 were selected.

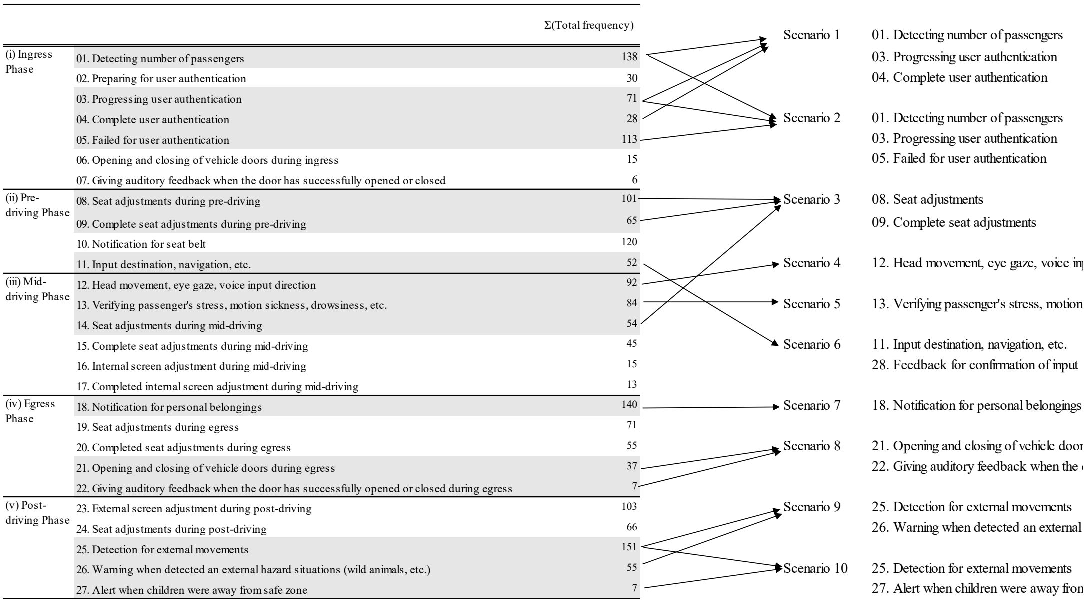


Figure 20. The selection criteria for the specified scenario for main sound set evaluation

Table 27. Freqeuncy analysis for each earcon, auditory icon that is appropriate for each sceanrio (Scenario 1 to Scenario 5)

Sound Sample Code		Scenario 1				Scenario 2				Scenario 3			Scenario 4		Sceanrio 5	
		01. Detecting number of passengers	03. Progressing user authentication	04. Complete user authentication	Σ(Total frequency)	01. Detecting number of passengers	03. Progressing user authentication	05. Failed for user authentication	Σ(Total frequency)	08. Seat adjustments	09. Complete seat adjustments	Σ(Total frequency)	12. Head movement, eye gaze, voice input direction	Σ(Total frequency)	13. Verifying passenger's stress, motion sickness.	Σ(Total frequency)
Earcon	A1	6	0	1	43	6	0	1	57	2	6	40	2	22	0	23
	A2	5	3	1		5	3	3		3	3		2		4	
	A3	2	0	1		2	0	5		3	4		4		3	
	A4	3	3	0		3	3	4		5	0		3		6	
	A5	5	4	0		5	4	2		6	1		4		6	
	A6	4	5	0		4	5	2		6	1		7		4	
	B1	6	1	2	49	6	1	0	59	2	2	30	2	11	1	36
	B2	7	3	0		7	3	0		3	4		3		3	
	B3	6	0	3		6	0	1		4	2		3		4	
	B4	3	1	0		3	1	7		3	0		0		9	
	B5	5	5	0		5	5	2		7	0		1		9	
	B6	6	1	0		6	1	5		2	1		2		10	
	C1	7	0	2	48	7	0	0	62	2	3	32	2	17	1	38
	C2	6	2	0		6	2	4		4	1		2		6	
	C3	7	1	1		7	1	1		3	4		4		4	
	C4	3	1	0		3	1	8		2	0		1		11	
	C5	5	4	0		5	4	2		5	1		4		6	
	C6	8	1	0		8	1	2		5	2		4		10	
Auditory Icon	D1	3	1	2	46	3	1	1	47	0	4	33	0	10	1	13
	D2	4	1	3		4	1	1		0	5		0		2	
	D3	5	1	5		5	1	0		1	5		4		2	
	D4	0	3	0		0	3	4		2	0		2		3	
	D5	2	7	0		2	7	2		9	0		4		2	
	D6	4	5	0		4	5	3		6	1		0		3	
	E1	2	0	0	23	2	0	6	54	0	0	12	7	31	0	18
	E2	2	0	0		2	0	7		1	0		4		4	
	E3	4	1	0		4	1	4		2	0		9		2	
	E4	2	2	0		2	2	2		3	0		2		2	
	E5	5	3	0		5	3	3		4	0		4		5	
	E6	1	1	0		1	1	9		2	0		5		5	
	F1	3	1	4	28	3	1	1	43	0	6	19	0	1	1	12
	F2	1	0	0		1	0	8		0	2		0		3	
	F3	2	1	2		2	1	0		3	4		1		0	
	F4	2	7	0		2	7	0		1	1		0		0	
	F5	2	1	1		2	1	5		0	1		0		5	
	F6	0	1	0		0	1	8		0	1		0		3	

Table 28. Freqeuncy analysis for each earcon, auditory icon that is appropriate for each sceanrio (Scenario 6 to Scenario 10)

Sound Sample Code		Scenario 6		Scenario 7		Scenario 8			Scenario 9			Sceanrio 10		
		11. Input destination, navigation, etc.	Σ(Total frequency)	18. Notification for personal belongings	Σ(Total frequency)	21. Opening and closing of vehicle doors during egress	22. Giving auditory feedback when the door has successfully opened or closed during egress	Σ(Total frequency)	25. Detection for external movements	26. Warning when detected an external hazard situations (wild animals, etc.)	Σ(Total frequency)	25. Detection for external movements	27. Alert when children were away from safe zone	Σ(Total frequency)
Earcon	A1	6	43	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	A2	5		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	A3	2		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	A4	3		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	A5	5		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	A6	4		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	B1	6	48	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	B2	7		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	B3	6		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	B4	3		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	B5	5		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	B6	6		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	C1	7	23	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	C2	6		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	C3	7		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	C4	3		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	C5	5		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	C6	8		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
Auditory Icon	D1	3	49	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	D2	4		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	D3	5		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	D4	0		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	D5	2		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	D6	4		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	E1	2	46	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	E2	2		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	E3	4		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	E4	2		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	E5	5		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	E6	1		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	F1	3	28	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
	F2	1		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	F3	2		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	F4	2		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	F5	2		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	
	F6	0		[illegible]		[illegible]	[illegible]		[illegible]	[illegible]		[illegible]	[illegible]	

Based on Table 27 and 28, the criterion for two evaluate sound set are based on the number of highest total frequency per scenario, and the highest number of total frequencies accumulated for the overall sound type (earcon or auditory icon). For instance, in the first scenario ‘01. Detecting number of passengers’, sound sample with a code, C6 records as the highest total frequency for the scenario, F4 sound sample code as the highest total frequency for ‘03. Progressing user authentication’, and D3 sound sample code as the highest total frequency for ‘04. Complete user authentication’. Thus, sound sample code C6, F4 and D3 will be selected as the first ‘sound set 1’ for this evaluation. Meanwhile, to select the ‘sound set 2’ for pairwise comparison, ‘sound set 2’ will be selected based on the highest total accumulated frequency among all the other sound sample types (earcons and auditory icons), which in this case, will be the B-type sound sample. Hence, ‘sound set 2’ sounds selection will be B2, B5 and B3 which are the highest number of counts in the scenario among all other B-type sounds. For each scenario, the sound selection will be undergone such criteria for sound set evaluation.

The sound sample for ‘in-progress’ was created using Adobe Audition software with the ascending and descending melody for each simple tone and variated tone. Hence, there were 4 sound sample stimuli which will be used for ‘in-progress’ sound evaluation.

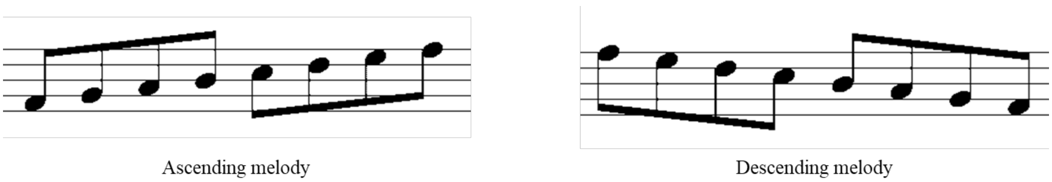
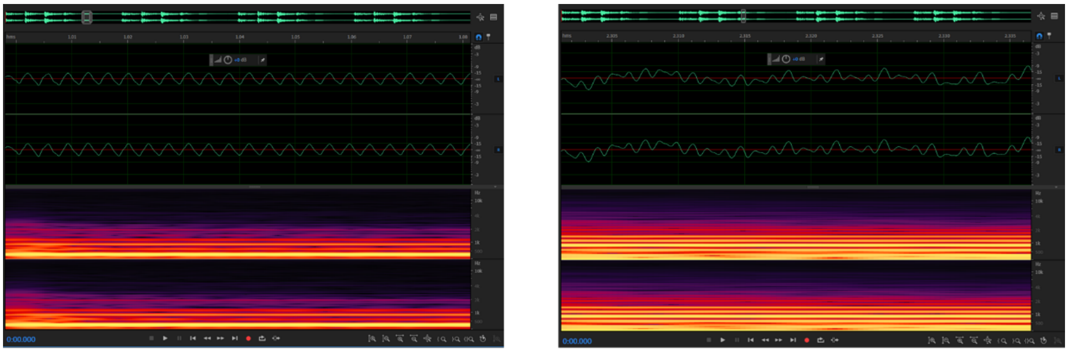


Figure 21. The musical notation for ascending melody and descending melody



Simple tone Variated tone

Figure 22. The sound frequency spectrograph for simple tone and variated tone.

3.2.4 Measurement

The measurement which will be used in this sound set evaluation is a 7-point Likert scale to measure the consistency of the sound as a set of family. Participants will give their evaluation based on the questionnaire of "Does this sound set has consistency to each other appropriately without feeling awkward?" for each sound set 1 and sound set 2. The intention of this measurement is to investigate whether consistency or appropriateness of a set of in-vehicle signals auditory feedback is important to increase the auditory user experience. Also, in the same scenario evaluation, participants were asked to choose which of the sound set they prefer. Thus, this sound set evaluation attempts to investigate whether passengers in an autonomous vehicle prefers a sound set of only earcons, only auditory icons, or a mixture of earcons and auditory icons.

For 'in-progress' sound evaluation, participants were asked to evaluate based on their level of satisfaction in a 7-point Likert scale for an overall scenario that involves an autonomous vehicle to convey information of 'in-progress' to the passenger or listener.

3.2.5 Experiment Procedure

This main sound set evaluation is divided into three parts; (i) pre-evaluation and screening, (ii) main evaluation, and (iii) post-evaluation. This paper considers the sound set evaluation be done via online for several reasons. Due to the coronavirus-19 pandemic, a large-scale sound evaluation and assessment had to be performed online according to the government standard measures of social distance. Also, previous studies have confirmed that online sound evaluation or assessment allows participant to be more independent and having a more sense of control (Shafiro et al., 2020). The online sound set evaluation was conducted in a similar manner with jury testing method, as it is commonly conducted when evaluating auditory related experience with regards to human perception (Rossi et al., 2005). Participants that were recruited via an online survey platform and were requested to read the research aim and instructions prior to the evaluation session. Participants were ensured that they understood the instructions well and received their consent before participating in the evaluation. After obtaining their consent, all participants were asked to complete a set of short demographic survey form and fill in their basic information. The evaluation procedure was explained in details. The participants online were also encouraged and advised that the use of earphones or headsets are necessary to perform their evaluation to avoid noise interferences from the background. All participants were also screened prior to the evaluation by using a simple question whether they are experiencing hearing disabilities.

The main evaluation was also conducted online by playing a set of sounds (sound set 1 and sound set 2) for each scenario (scenario 1 to scenario 10). After listening to both sound sets, participants were instructed to immediately evaluate each sound set's consistency according to 7-points Likert scale. For each scenario section, participants were asked to choose their preferred sound set, and were asked to write down their opinions for qualitative data exploration. Participants were also asked to evaluate their level of satisfaction and write their opinion for the 'in-progress' sound sample stimuli.

Lastly, for the post-evaluation, participants were instructed to write their overall opinion about the whole sound set evaluation. Overall sound set evaluation took approximately 60 minutes.

3.2.6 Data Analysis Approach

For sound set evaluation based on scenarios, the data obtained from the dependent variable, which is consistency/ appropriateness of the sound were analyzed using pairwise *t*-test comparison for each of the sound sets (sound set 1 and sound set 2) at the significance level *p*-value of 0.05. Meanwhile, for the 'in-progress' sound evaluation, the data were analyzed using four-way analysis of variance (ANOVA) for each of the in-progress sound parameters (ascending-simple tone, descending-simple tone, ascending-variated tone, descending-variated tone) and set the significance level at *p*-value of 0.05.

For qualitative data analysis, a Text Network Analysis is performed using UCINET to investigate the associated clustered word related to each sound set, and visualization per each scenario. The qualitative data was pre-processed using AutoMap by performing text preparation and text cleaning. After the text preparation and text cleaning, the qualitative texts were then visualized its network via NetDraw in UCINET by assigning appropriate nodes and centrality.

3.2.7 Results

The result of the analyses is presented based on the tables and graph below. In this main sound set evaluation, the result of the consistency of the sound set from scenario 1 to scenario 10 will be presented first, then followed by the result of satisfaction of 'in-progress' sound. The results will include descriptive statistic tables, independent sample *t*-test summary tables and graphs for the dependent variable, consistency. Meanwhile, the result of the analysis for the 'in-progress' will include descriptive statistic table, four-way ANOVA summary table, and graphs to indicate satisfaction mean score among parameters.

Scenario 1

The consistency or appropriateness of the sound set 1 and 2 data is presented in Table 29, and 30 below. An independent-sample *t*-test was conducted to compare the consistency or appropriateness for sound set 1 and sound set 2. Based on Table 30, there were significance differences (t(df) = 248, $p < 0.001$) in consistency/appropriateness scores for sound set 1 (M = 3.376, SD = 0.150) and sound set 2 (M = 5.416, SD = 0.108). Based on the result, it can be inferred that the sound set 2 is more appropriate as a set of sound, and has consistency to each auditory feedback sound within the set. Also, about 71% (89 of the participants) selected sound set 2 as more appropriate and preferable than sound set 1 (14%, 18 of the participants) for scenario 1.

Table 29. Descriptive statistics for sound set in Scenario 1

	N	Mean	Std. Error
Sound set 1	125	3.376	0.150
Sound set 2	125	5.416	0.108

Table 30. Independent samples *t*-test for Scenario 1

		Levene's Test for Equality		T-Test for equality means				
		F	*p*-value	*t*	df	*p*-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	25.334	0.00	-11.04**	248	0.00	-2.04	0.185
	Sound set 2			-11.04**	226.07	0.00	-2.04	0.185

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

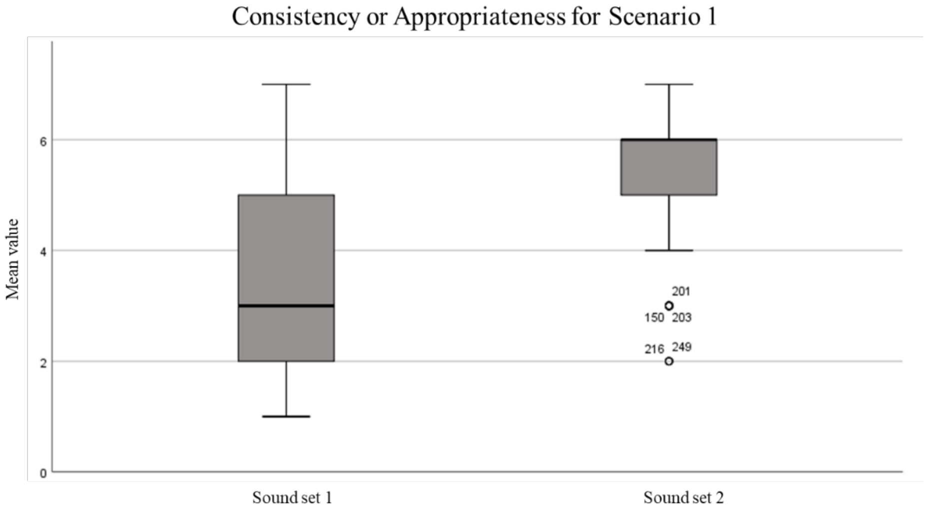


Figure 23. Graph of Consistency or Appropriateness for Scenario 1

Scenario 2

The consistency or appropriateness of the sound set 1 and 2 data for scenario 2 is presented in Table 31 and 32 below. Based on Table 32, there were significance differences (t(df) = 248, $p < 0.001$) in consistency/appropriateness scores for sound set 1 (M = 3.144, SD = 0.147) and sound set 2 (M = 5.352, SD = 0.110). Similar to scenario 1, sound set 2 is more appropriate as a set of sound, and consistent to each auditory feedback sound within the set. Also, about 71% (89 of the participants) selected sound set 2 as more appropriate and preferable than sound set 1 (13%, 16 of the participants).

Table 31. Descriptive statistics for sound set in Scenario 2

	N	Mean	Std. Error
Sound set 1	125	3.144	0.147
Sound set 2	125	5.352	0.110

Table 32. Independent samples *t*-test for Scenario 2

		Levene's Test for Equality		T-Test for equality means				
		F	*p*-value	*t*	df	*p*-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	18.063	0.00	-12.06**	248	0.00	-2.21	0.183
	Sound set 2			-12.06**	229.67	0.00	-2.21	0.183

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

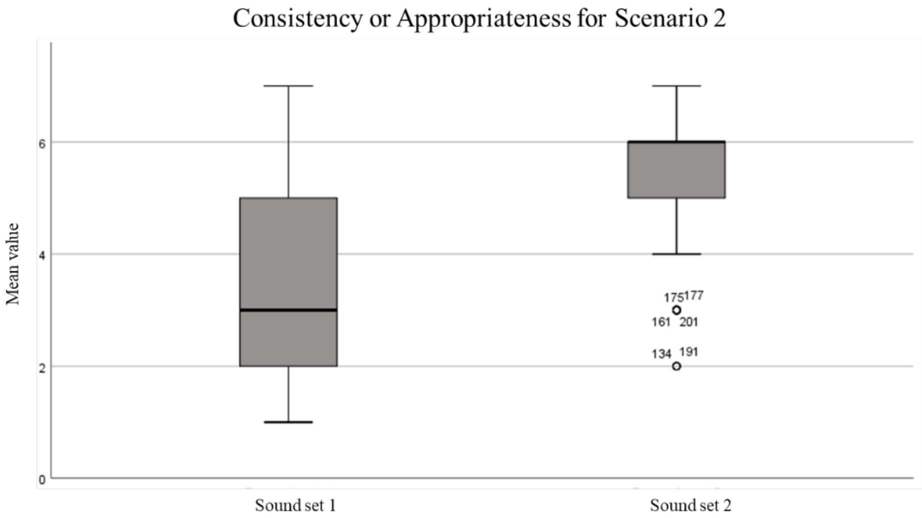


Figure 24. Graph of Consistency or Appropriateness for Scenario 2

Scenario 3

The consistency or appropriateness of the sound set 1 and 2 data for scenario 3 is presented in Table 33 and 34 below. Based on Table 34, there were significance differences (t(df) = 248, $p < 0.001$) in consistency/appropriateness scores for sound set 1 ($M = 4.200$, $SD = 0.143$) and sound set 2 ($M = 5.256$, $SD = 0.117$). This can be interpreted as that sound set 2 is more appropriate as a set of sound, and consistent to each auditory feedback sound within the set. Also, about 55% (69 of the participants) selected sound set 2 as more appropriate and preferable than sound set 1 (23%, 29 of the participants).

Table 33. Descriptive statistics for sound set in Scenario 3

	N	Mean	Std. Error
Sound set 1	125	4.200	0.143
Sound set 2	125	5.256	0.117

Table 34. Independent samples *t*-test for Scenario 3

		Levene's Test for Equality		T-Test for equality means				
		F	*p*-value	*t*	df	*p*-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	13.043	0.00	-5.70**	248	0.00	-1.06	0.185
	Sound set 2			-5.70**	238.83	0.00	-1.06	0.185

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

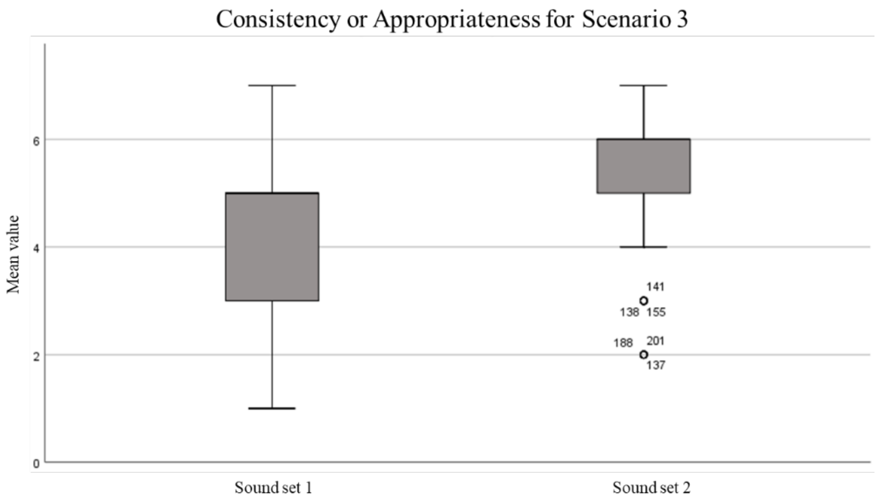


Figure 25. Graph of Consistency or Appropriateness for Scenario 3

Scenario 4

The consistency or appropriateness of the sound set 1 and 2 data for scenario 4 is presented in Table 35 and 36 below. Based on Table 36, there were no significance at $p > 0.05$, ($t(\text{df}) = 248$, $p = 0.69$) in consistency/appropriateness scores for sound set 1 ($M = 4.656$, $SD = 0.115$) and sound set 2 ($M = 4.720$, $SD = 0.119$). Also, about 32% (40 of the participants) selected sound set 2 as more appropriate and preferable than sound set 1 (26%, 33 of the participants).

Table 35. Descriptive statistics for sound set in Scenario 4

	N	Mean	Std. Error
Sound set 1	125	4.656	0.115
Sound set 2	125	4.720	0.119

Table 36. Independent samples *t*-test for Scenario 4

		Levene's Test for Equality		T-Test for equality means				
		F	*p*-value	*t*	df	*p*-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	0.162	0.69	-0.39	248	0.70	-0.06	0.165
	Sound set 2			-0.39	247.69	0.70	-0.06	0.165

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

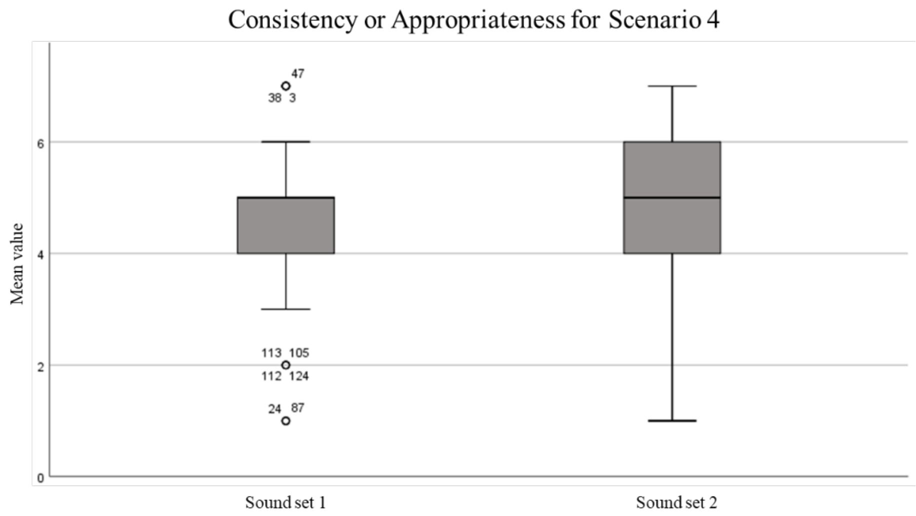


Figure 26. Graph of Consistency or Appropriateness for Scenario 4

Scenario 5

The consistency or appropriateness of the sound set 1 and 2 data for scenario 5 is presented in Table 37 and 38 below. Based on Table 38, there were also no significance at $p > 0.05$, (t(df) = 248, p = 0.60) in consistency/appropriateness scores for sound set 1 (M = 5.152, SD = 0.129) and sound set 2 (M = 4.080, SD = 0.129). Also, about 54% (68 of the participants) selected sound set 1 as more appropriate and preferable than sound set 2 (19%, 24 of the participants).

Table 37. Descriptive statistics for sound set in Scenario 5

	N	Mean	Std. Error
Sound set 1	125	5.152	0.129
Sound set 2	125	4.080	0.129

Table 38. Independent samples *t*-test for Scenario 5

		Levene's Test for Equality		T-Test for equality means				
		F	*p*-value	*t*	df	*p*-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	0.279	0.60	5.89	248	0.00	1.07	0.182
	Sound set 2			5.89	248.00	0.00	1.07	0.182

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

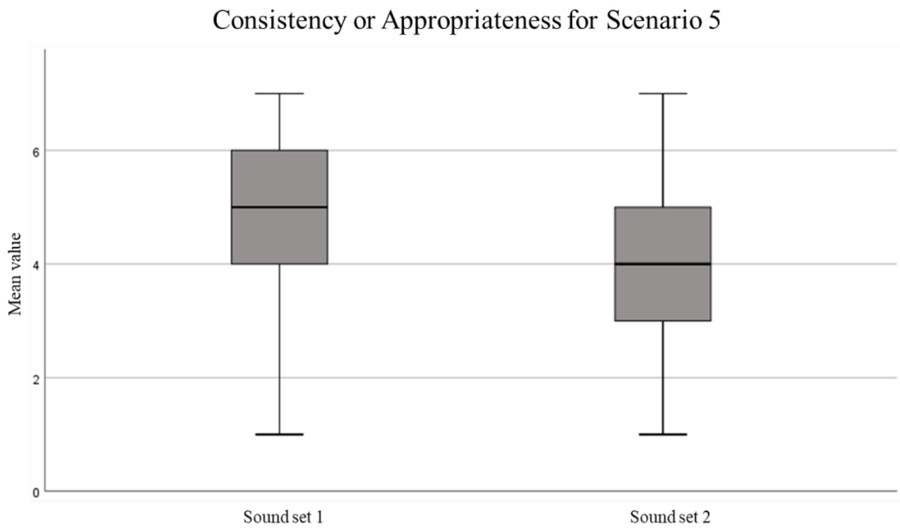


Figure 27. Graph of Consistency or Appropriateness for Scenario 5

Scenario 6

The consistency or appropriateness of the sound set 1 and 2 data for scenario 6 is presented in Table 39 and 40 below. Based on Table 40, there were significance differences (t(df) = 248, $p < 0.001$) in consistency/appropriateness scores for sound set 1 (M = 5.328, SD = 0.106) and sound set 2 (M = 3.312, SD = 0.141). From the mean score itself, it can be interpreted as that sound set 1 is more appropriate as a set of sound, and has consistency to each auditory feedback sound within the set than sound set 2. Also, about 67% (84 of the participants) selected sound set 1 as more appropriate and preferable than sound set 2 (15%, 19 of the participants).

Table 39. Descriptive statistics for sound set in Scenario 6

	N	Mean	Std. Error
Sound set 1	125	5.328	0.106
Sound set 2	125	3.312	0.141

Table 40. Independent samples *t*-test for Scenario 6

		Levene's Test for Equality		T-Test for equality means				
		F	*p*-value	*t*	df	*p*-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	15.955	0.00	11.45**	248	0.00	2.02	0.176
	Sound set 2			11.45**	230.27	0.00	2.02	0.176

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

Consistency or Appropriateness for Scenario 6

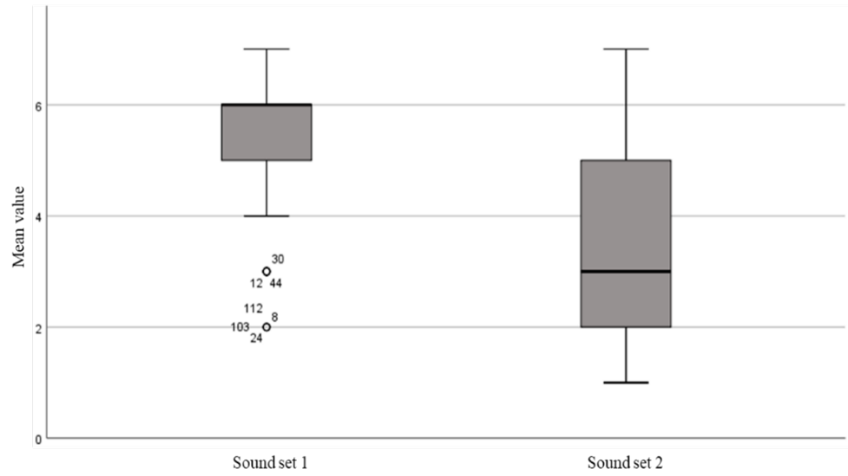


Figure 28. Graph of Consistency or Appropriateness for Scenario 6

Scenario 7

The consistency or appropriateness of the sound set 1 and 2 data for scenario 7 is presented in Table 41 and 42 below. Based on Table 42, there were also no significance at $p < 0.05$, (t(df) = 248, $p = 0.68$) in consistency/appropriateness scores for sound set 1 (M = 4.912, SD = 0.119) and sound set 2 (M = 4.784, SD = 0.117). Also, about 30% (38 of the participants) selected sound set 1 as more appropriate and preferable than sound set 2 (20%, 25 of the participants).

Table 41. Descriptive statistics for sound set in Scenario 7

	N	Mean	Std. Error
Sound set 1	125	4.912	0.119
Sound set 2	125	4.784	0.117

Table 42. Independent samples *t*-test for Scenario 7

		Levene's Test for Equality				T-Test for equality means		
		F	*p*-value	*t*	df	*p*-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	0.174	0.68	0.77	248	0.44	0.13	0.167
	Sound set 2			0.77	247.90	0.44	0.13	0.167

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

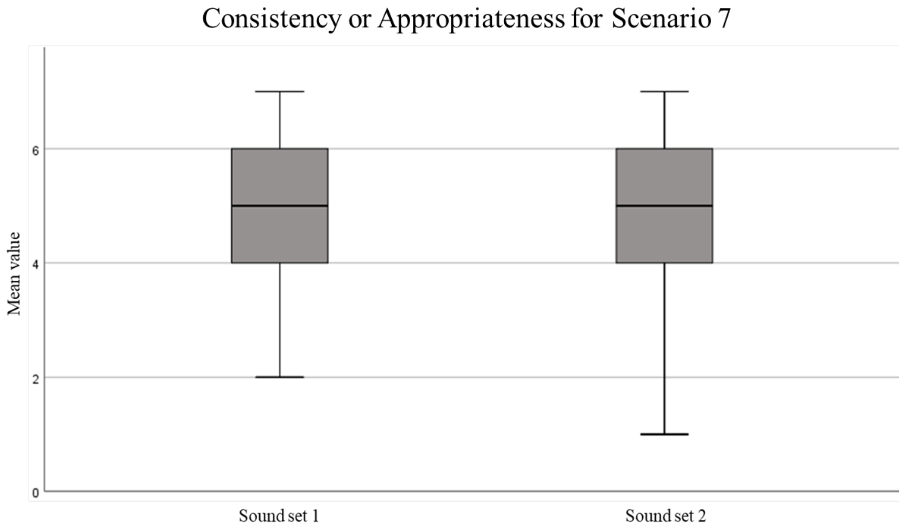


Figure 29. Graph of Consistency or Appropriateness for Scenario 7

Scenario 8

The consistency or appropriateness of the sound set 1 and 2 data for scenario 8 is presented in Table 43 and 44 below. Based on Table 44, there were significance differences (t(df) = 248, $p < 0.05$) in consistency/appropriateness scores for sound set 1 (M = 4.208, SD = 0.132) and sound set 2 (M = 5.072, SD = 0.123). Based on this, it is inferred as that sound set 2 is more appropriate as a set of sound, and has consistency to each auditory feedback sound within the set than sound set 1. Also, about 60% (75 of the participants) selected sound set 2 as more appropriate and preferable than sound set 1 (23%, 29 of the participants).

Table 43. Descriptive statistics for sound set in Scenario 8

	N	Mean	Std. Error
Sound set 1	125	4.208	0.132
Sound set 2	125	5.072	0.123

Table 44. Independent samples *t*-test for Scenario 8

		Levene's Test for Equality		T-Test for equality means				
		F	*p*-value	*t*	df	*p*-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	4.376	0.04	-4.78*	248	0.00	-0.86	0.181
	Sound set 2			-4.78*	246.87	0.00	-0.86	0.181

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

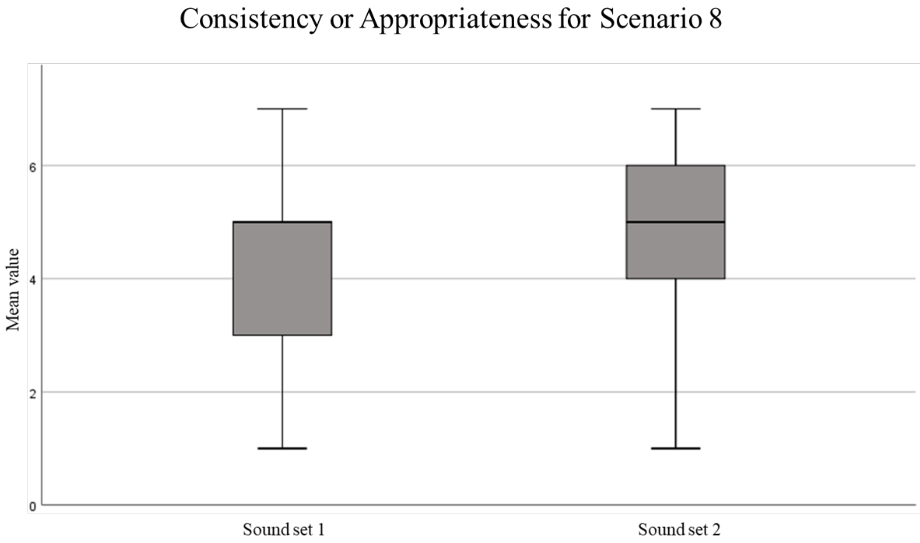


Figure 30. Graph of Consistency or Appropriateness for Scenario 8

Scenario 9

The consistency or appropriateness of the sound set 1 and 2 data for scenario 9 is presented in Table 45 and 46 below. Based on Table 46, there were significance differences (t(df) = 248, $p < 0.001$) in consistency/ appropriateness scores for sound set 1 (M = 3.816, SD = 0.152) and sound set 2 (M = 4.736, SD = 0.109). Based on this, it is inferred as that sound set 2 is more appropriate as a set of sound, and has consistency to each auditory feedback sound within the set than sound set 1. Also, about 38% (48 of the participants) selected sound set 2 as more appropriate and preferable than sound set 1 (27%, 34 of the participants).

Table 45. Descriptive statistics for sound set in Scenario 9

	N	Mean	Std. Error
Sound set 1	125	3.816	0.152
Sound set 2	125	4.736	0.109

Table 46. Independent samples *t*-test for Scenario 9

		Levene's Test for Equality				T-Test for equality means		
		F	*p*-value	*t*	df	*p*-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	25.771	0.00	-4.92**	248	0.00	-0.92	0.187
	Sound set 2			-4.92**	225.12	0.00	-0.92	0.187

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

Consistency or Appropriateness for Scenario 9

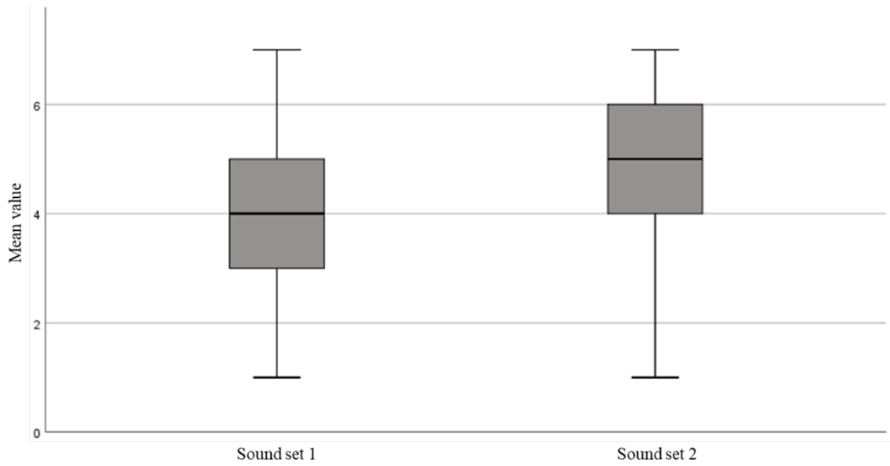


Figure 31. Graph of Consistency or Appropriateness for Scenario 9

Scenario 10

The consistency or appropriateness of the sound set 1 and 2 data for scenario 10 is presented in Table 47 and 48 below. Based on Table 48, there were significance differences (t(df) = 248, $p < 0.001$) in consistency/appropriateness scores for sound set 1 (M = 3.344, SD = 0.148) and sound set 2 (M = 5.304, SD = 0.113). It is also inferred as that sound set 2 has more appropriate as a set of sound, and has consistency to each auditory feedback sound within the set than sound set 1. Also, about 60% (75 of the participants) selected sound set 2 as more appropriate and preferable than sound set 1 (18%, 23 of the participants).

Table 47. Descriptive statistics for sound set in Scenario 10

	N	Mean	Std. Error
Sound set 1	125	3.344	0.148
Sound set 2	125	5.304	0.113

Table 48. Independent samples t-test for Scenario 10

		Levene's Test for Equality		T-Test for equality means				
		F	p-value	t	df	p-value (2-tail)	Mean Diff.	Std. Error Diff.
DV	Sound set 1	21.313	0.00	-10.53**	248	0.00	-1.96	0.186
	Sound set 2			-10.53**	232.25	0.00	-1.96	0.186

DV = Dependent variable, * $p < 0.05$, ** $p < 0.001$

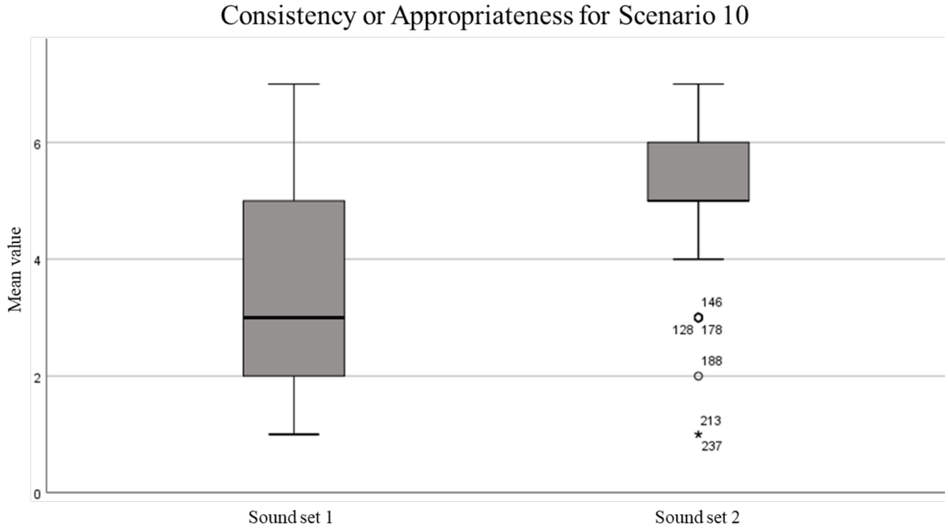


Figure 32. Graph of Consistency or Appropriateness for Scenario 10

In Progress Sound

To investigate the passenger's satisfaction of 'in-progress' sounds, the level of satisfaction result data for pilot test is presented in Table 49 and 50 below. Based on the data in Table 49, the satisfaction mean value for In-Progress Sound 1 (Descending-Simple melody) has the highest mean value of 4.472±(0.104) and followed by In-Progress Sound 4 (Ascending-Variated melody) with 4.416±(0.098), In-Progress Sound 2 (Descending-Variated melody) with 4.400±(0.098), and the lowest mean value is In-Progress 3 (Ascending-Simple melody) with 4.344±(0.106).

Table 49. Descriptive statistics for Satisfaction

			95% Confidence Interval for Mean	
	Mean	Std. Error	Lower Bound	Upper Bound
In-progress Sound 1	4.472	0.104	4.266	4.678
In-progress Sound 2	4.400	0.098	4.205	4.595
In-progress Sound 3	4.344	0.106	4.134	4.554
In-progress Sound 4	4.416	0.098	4.221	4.611

Table 50 shows a summary of ANOVA for the passenger's level of satisfaction. From Table 50, the result shows that there were no significant differences $p < 0.05$ between each in-progress sounds. This shows that all of the in-progress sounds are similar and acceptable for scenarios or situations that needed to let the listener knows that the system is in-progress of performing a task. Hence, In-Progress Sound 1 with a descending and simple tone melody is considered to be preferable among the other in-progress sounds.

Table 50. ANOVA summary for Satisfaction

	Sum of Squares	df	Mean Square	F	*p*-value
Between groups	1.040	3	0.347	0.268	0.85
Within group	641.728	496	1.294		
Total	642.768				

* $p < 0.05$, ** $p < 0.001$

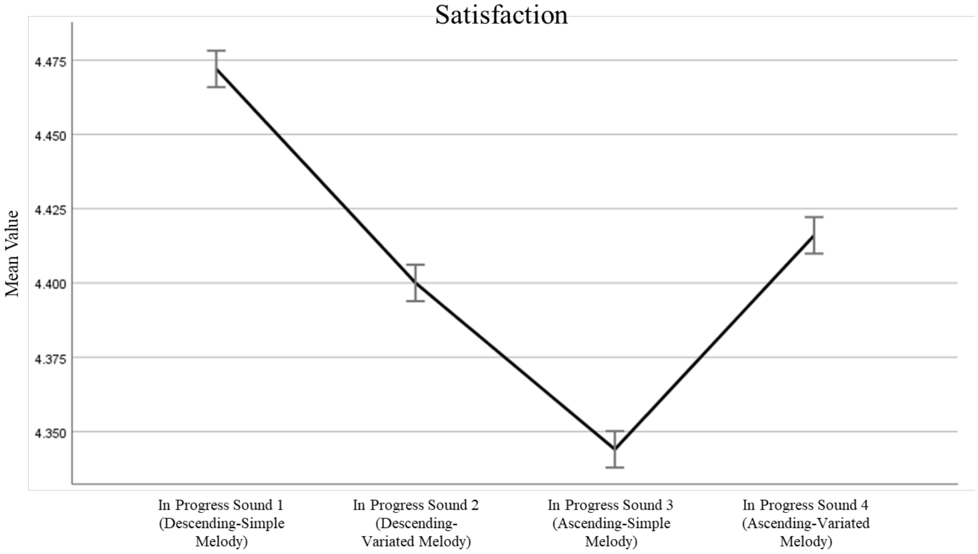


Figure 33. Graph for In-Progress Sounds Satisfaction

3.3 Text Network Visualization on User's Opinions

The qualitative data of the participants' opinions were collected and visualized using Text Network Analysis by using UCINET and NetDraw. First and foremost, the text opinions were collected according to each scenario (from scenario 1 to scenario 10), and the texts were translated from Korean language to English language via Google Translate. The translated texts were then prepared in Microsoft Excel and categorized according to scenarios. Words such as 'set 1', 'No.1', 'Sound 1' and etc. were unified and renamed as 'sound_set_1', and 'sound_set_2' respectively. The texts were double-checked to avoid error in the renaming of the 'sound_set'. The opinion texts were pre-processed using AutoMap by performing text cleaning, and followed by text preparation. Text cleaning included removal of extra spaces, conversion of British to American spellings, fix for common typos, expand common contractions and abbreviations, and resolute pronouns. Text preparation were done by removing single letters, conversion NGram, removal of pronouns, noise verbs, prepositions, day and month words, possessive words, complete numbers, and all noise words. Lastly, the texts were then refined by removing single symbols and converting all upper-case letters to lower-case for refinement. For visualization, the number of nodes were set in a range of 2 to 5 for each scenario. Text network analysis or semantic network analysis is an approach on a method for visualization of analytical reasoning and explorative analysis to gain qualitative insights (Drieger, 2013). Hence, in this study, the participants' opinions regarding the sound sets according to scenario were used for visualization purposes and to obtain qualitative insights only. The visualized text network for each scenario are presented as Figure 34 to 43 in the Appendix 1.

Chapter 4

Discussions

The main purpose of this study is to present a design proposal for in-vehicle signals feedback that is intuitive, to determine the sound preference for passenger in autonomous vehicle, and to suggest a fully derived scenario for the design of in-vehicle signal feedback in future autonomous vehicle based on passenger's context. In this chapter, discussions related to each purpose of this study will be presented and elaborated based on the pilot test and main sound set evaluation result from Chapter 3.

Design proposal for an in-vehicle signals feedback

Generally, from the pilot test, it is confirmed that the proposed design of signal information types based on temporal pattern as hypothesized in Table 2 does matches with the intended information type to be conveyed to the listener, except for in-progress type of sound. From perceivability result, all of the sound sample were perceivable to the listeners, and listeners were able to hear the sound samples played during the pilot test.

For the intuitiveness of all the sound types, the confirmatory type signal results shown that the listeners were able to identify the signal as confirmatory signal due to its short duration and one-beat tempo. As stated by Hoggan et al., study, confirmation or confirmatory signals should be designed in short rhythms (Hoggan et al., 2009).

For error type signal, if the length or duration of the signal was as short as the confirmatory signal, the users or listeners would not be able to obtain the information that the system is error. Hence, from this result, due to the tempo and duration of an error type signal is longer than the confirmatory signal, the participants in the pilot test were able to distinguish its difference and perceived that the intended error type signal is appropriate and matches with all of the scenarios that involves system error.

From the pilot test, detection type signals and warning type signal both were found out that there was no significant difference because both types of signal have similar temporal pattern (tempo and beat) of ON1=ON2, with OFF1, which resulting the listeners to perceived both detection type signal and warning signal as the same. However, considering the context scenario of passengers, there is a need to design detection-related scenario such as 'giving auditory feedback when the vehicle detects passenger approaching by the door for entrance' or 'detecting any movements of wild animals or external hazards to notify the passengers inside'. From this pilot test and sound evaluation, it is essential to re-consider a suitable temporal pattern or acoustic parameters when designing detection signal to avoid misconception with the alert or warning signals.

In-progress signals were found to have no significant difference between the other types of signal during the pilot test, hence, a re-design of in-progress signal as a melody with ascending, descending, simple tone and variated tone as parameter variables for the sound evaluation. Result from the sound evaluation shown that descending-simple tone melody has the highest satisfaction mean value compared to the others. This is because that when a system is performing a task that requires time to complete, it is important to let the user or listener know that the system should not be disturbed while performing the task until it is completed. Hence, the descending melody is appropriate in order to convey such meaning or information to the user, instead of an ascending melody.

For alert and warning type signals, the temporal patterns of the sounds shown that it matches with the intended information to be conveyed. repetitive temporal pattern and high frequency signals conveys the perceived of urgency to the drivers or listeners in order to take immediate intervention measures during hazard situations (Walker, B. N., 2007; Nees & Walker, 2011). It is also confirmed that repetitive tempo and beat increases annoyance to the listener and took initiative to turn off the signal (Adell et al., 2008).

Sound preference for passenger in an autonomous vehicle

Based from the sound set evaluation, results show that in most scenarios, similar type of sounds as a set has high consistency or appropriateness than a mixture of earcons and auditory icons. To put it simple, when a scenario requires in-vehicle signals feedback from one task to another, such as ‘detecting passenger’ and then ‘confirming the passenger’s identity’ and then ‘complete user authentication’ for egress, a set of earcon sounds only seems more preferable than a mixture set of earcon and auditory icons retrospectively. Also, in overall, a set of earcon sounds shows better results than a set of auditory icons. This result may cause when the auditory icon in the scenario was not appropriately designed. Unlike earcons, auditory icons would need to match the intended message to be conveyed to be consistent to its functionality, such as the sound of paper crumbling used to convey the ‘trash’ icon on our computer. However, in this study, these auditory icon samples were not appropriately designed according to the scenario, for example, the sound of water droplet does not convey the intended message for ‘completing user authentication’ scenario. Hence, this results that earcons are most preferable and its simple tone could convey a wide range of information, abstract or not, to the listeners (Walker et al., 2006; Hoggan et al., 2009; Oswald, D., 2012). Though earcons may require learnability to allow the listeners or users to fully understand its intended meaning (Dingler et al., 2008; Walker et al., 2006), but its vagueness or abstract characteristics would be appropriate and suitable to represent scenarios used in autonomous vehicles. Also, due to the earcons’ simple, basic and abstract characteristics of tone (Oswald D., 2012; Walker et al., 2013; Brewster et al., 1994), it is easily to associate earcons as a family of sounds with high consistency and appropriateness when designing a set of sounds used in a continuous scenario (Brewster et al., 1994). Furthermore, from this study, earcons can also be considered a universal auditory feedback, where a set of earcons shows that there is a high level of consistency with respect to most scenarios compared to the auditory icons. As the system's complexity increases, the requirements for information to be conveyed to the users also increases, hence, a simpler modality or feedback is required to avoid misconception or information loss.

Fully derived scenario for the design of in-vehicle signals feedback in future autonomous vehicles

Based on the pilot test and sound evaluation as well, this study was able to present a full context of scenarios from ingress to post-driving that can be used when designing for an in-vehicle signals feedback of autonomous vehicles. The pre-pilot test considers the time stamps and timeline for when auditory feedback will be needed when a passenger is riding an autonomous vehicle. The pre-pilot test scenario development considers three large categories of the passenger's context; (i) pre-usage, (ii) usage, and (iii) post-usage. The sub-categories include five activities; ingress, pre-driving activity, mid-driving, egress and post-driving activity. However, the scenario was then extended from five activities to 27 tasks/scenarios by considering all tasks required for an auditory signal feedback. Through pilot test, this study was able to identify which tasks or scenario are relevant to the usage of in-vehicle signals feedback by using frequency analysis, and then revamped the scenarios into a consecutive set of scenarios to investigate the consistency and appropriateness of sounds. 15 out of 27 scenarios were selected based on the highest number of counts, and as a result, the essential scenarios or situations where in-vehicle signals feedback is needed to be designed are 'detecting number of passengers before the passenger ingress', 'giving auditory signal feedback when the vehicle is performing user authentication process', ‘giving auditory signal feedback when user authentication has completed’, ‘giving auditory signal feedback when the user authentication has failed’, ‘performing seat adjustments’, ‘giving auditory signal feedback when the seat adjustments has completed’, ‘sound feedback when the vehicle system detects head movement, eye gaze, and voice input for gestures and voice recognition’, ‘allowing the passengers to be notified when in stress, motion sickness drowsiness, etc.’, ‘sending auditory signal feedback when the user input destination and navigation’, ‘giving a confirmatory feedback when the user has successfully input information’, ‘giving notification to the passengers when they forgotten to take their personal belongings during egress’, ‘opening and closing doors’, ‘giving auditory signal feedback when the door has completely opened or shut’, ‘giving auditory signal feedback when the vehicle senses external movements that could cause danger’, ‘giving warning signal when detected hazards’, and lastly, ‘to notify when small children are away from the safe zone during post-driving’.

Chapter 5

Conclusion

5.1 Conclusion

This study provides several insights related on the types of sound information and how to design it intuitively based on temporal patterns. It is important to understand that the temporal pattern also affects the users or listeners perceived information, for example, when a short beep was given, it represents confirmatory message, and if it beeps repetitively, it gives an alert or warning sign for users to take initiative to intervene. Overall, the pilot test result supported that the representation of the temporal pattern stated in Chapter 2 with the intended information type, excluding in-progress type sound are valid for future studies. Also, in summary, the sound set evaluation indicated that earcons are a considered universal auditory feedback or signal as it is simple, basic tone and abstract which is appropriate to be used in most vague scenario or situations, however, on the contrary, learnability is needed to fully convey its information. The scenario developed in this study can be considered as a fundamental basic scenario that are needed for future autonomous vehicles. Some future autonomous vehicles may incorporate certain technology which affects the usage scenario, but, in this study, the scenario of ingress to egress, and pre-driving to post-driving are fundamental, and can be used as a baseline for future study.

5.2 Contribution of the Thesis

This thesis presents two main contributions based on the result of the pilot test and sound evaluation study on the user preference of in-vehicle signal feedbacks between earcons and auditory icons under autonomous driving contexts. First, this study provides temporal patterns that was suggested and proposed as shown in Table 2 which can be used as a tool to evaluate any auditory-related situations that requires feedback. Second, the methodology used in this study provides a reference or benchmark when performing evaluation and analysis by focusing on user-centered design approach through the derivation of scenarios based on context-of-use analysis.

5.3 Limitation and future work

In this study, there were a few limitations that can be identified for future work. First limitation of this study is that the auditory icon sound stimuli which were used for the evaluation were designed solely based on the temporal pattern type and the theme used for the evaluation seemed to be less appropriate for the autonomous vehicle context. Secondly, the participants that were scouted for the pilot test were in the age of 20-30s where the participants represent a group of people with high adaptation of technology and not foreign to the concept of sound. We expect to recruit a more diverse group of people in the future work to represent the population for all age of groups to explore their auditory user experience for fully autonomous vehicles. Another limitation of this study is that subjective measurements for acoustic parameters with regards to the satisfaction of the sound samples should be included. Since, this study investigates on the intuitiveness and design proposal of what an auditory feedback would be based on the appropriateness of passenger riding an autonomous vehicle, thus, perceivability, intuitiveness, consistency/appropriateness were focused in this study.

Future studies may include the design and evaluation for other sensory feedbacks and interface such as visual and tactile in-vehicle feedbacks, specifically to provide alternative for passengers with hearing disabilities when riding a future autonomous vehicle. Furthermore, the future autonomous vehicle would require the ability to allow all passengers with and without physically challenged, thus a more in-depth context-of-use related to the passengers with a wheelchair or mobility aid devices should also be considered for future studies. To expand the subject, the preferences of sound signals according to gender or age. This study provides an overall appropriateness of the sound sample types and information types based on scenario, but future work could include the in-depth layer of whether gender or age affects the preference level or cognitive ability level.

Bibliography

[1] Adell, E., Várhelyi, A., & Hjälmdahl, M. (2008). Auditory and haptic systems for in-car speed management–A comparative real life study. Transportation research part F: traffic psychology and behaviour, 11(6), 445-458.

[2] Ahmed, H. U., Huang, Y., Lu, P., & Bridgelall, R. (2022). Technology Developments and Impacts of Connected and Autonomous Vehicles: An Overview. Smart Cities, 5(1), 382-404.

[3] Alonso-Ríos, D., Vázquez-García, A., Mosqueira-Rey, E., & Moret-Bonillo, V. (2010). A context-of-use taxonomy for usability studies. International Journal of Human-Computer Interaction, 26(10), 941-970.

[4] Amann, E., & Anderson, I. (2014). Development and validation of a questionnaire for hearing implant users to self-assess their auditory abilities in everyday communication situations: the Hearing Implant Sound Quality Index (HISQUI19). Acta oto-laryngologica, 134(9), 915-923.

[5] An, J. Y., Kim, Y. J., & Yoo, H. S. (2020, July). User Preference for Vehicle Warning Sounds to Develop AUI Guideline Focusing on Differences Between Sex and Among Age Groups. In International Conference on Human-Computer Interaction (pp. 3-13). Springer, Cham.

[6] Bagloee, S. A., Tavana, M., Asadi, M., & Oliver, T. (2016). Autonomous vehicles: challenges, opportunities, and future implications for transportation policies. Journal of modern transportation, 24(4), 284-303.

[7] Baldwin, C. L., & Struckman-Johnson, D. (2002). Impact of speech presentation level on cognitive task performance: Implications for auditory display design. Ergonomics, 45(1), 61-74.

[8] Barrass, S. (2012). The aesthetic turn in sonification towards a social and cultural medium. AI & society, 27(2), 177-181.

[9] Bengler, K., Dietmayer, K., Farber, B., Maurer, M., Stiller, C., & Winner, H. (2014).

Three decades of driver assistance systems: Review and future perspectives. IEEE Intelligent transportation systems magazine, 6(4), 6-22.

[10] Bilger, B. (2013). Auto correct: Has the self-driving car at last arrived. The New Yorker, 25.

[11] Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. Human–Computer Interaction, 4(1), 11-44.

[12] Bonebright, T. L., & Nees, M. A. (2007). Memory for auditory icons and earcons with localization cues. In Proceedings of the International Conference on Auditory Display (ICAD 07) (pp. 419–422). Atlanta: Georgia Institute of Technology.

[13] Bradshaw, J. M., Hoffman, R. R., Woods, D. D., & Johnson, M. (2013). The seven deadly myths of" autonomous systems". IEEE Intelligent Systems, 28(3), 54-61.

[14] Brewster, S. A., Wright, P. C., & Edwards, A. D. (1994). A detailed investigation into the effectiveness of earcons. In SANTA FE INSTITUTE STUDIES IN THE SCIENCES OF COMPLEXITY-PROCEEDINGS VOLUME- (Vol. 18, pp. 471-471). Addison-Wesley Publishing Co.

[15] Brewster, S. A., & Crease, M. G. (1999). Correcting menu usability problems with sound. Behaviour & Information Technology, 18(3), 165-177.

[16] Brooke, J. (1996). SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), 4-7.

[17] Bussemakers, M. P., & De Haan, A. (2000, April). When it sounds like a duck and it looks like a dog... auditory icons vs. earcons in multimedia environments. In Proceedings of the international conference on auditory display (pp. 184-189).

[18] Camara, F., Bellotto, N., Cosar, S., Nathanael, D., Althoff, M., Wu, J., ... & Fox, C. W. (2020). Pedestrian models for autonomous driving Part I: low-level models, from sensing to tracking. IEEE Transactions on Intelligent Transportation Systems, 22(10), 6131-6151.

[19] Carroll, J. M. (1997). Scenario-based design. In Handbook of human-computer interaction (pp. 383-406). North-Holland.

[20] Chi, C. F., Dewi, R. S., & Huang, M. H. (2017). Psychophysical evaluation of auditory signals in passenger vehicles. Applied ergonomics, 59, 153-164.

[21] Cornejo, S. D. H. (2018). Towards ecological and embodied design of auditory display. Georgia Institute of Technology

[22] da Silveira Brizon, C. J., & Medeiros, E. B. (2012). Combining subjective and objective assessments to improve acoustic comfort evaluation of motor cars. Applied Acoustics, 73(9), 913-920.

[23] Dingler, T., Lindsay, J., & Walker, B. N. (2008). Learnabiltiy of sound cues for environmental features: Auditory icons, earcons, spearcons, and speech. International Community for Auditory Display.

[24] Drieger, P. (2013). Semantic network analysis as a method for visual text analytics. Procedia-social and behavioral sciences, 79, 4-17.

[25] Edworthy, J., Hellier, E., & Hards, R. (1995). The semantic associations of acoustic parameters commonly used in the design of auditory information and warning signals. Ergonomics, 38(11), 2341-2361.

[26] Eskandarian, A. (Ed.). (2012). Handbook of intelligent vehicles (Vol. 2). London: Springer.

[27] Foley, L., Anderson, C. J., & Schutz, M. (2020, October). Re-sounding alarms: designing ergonomic auditory interfaces by embracing musical insights. In Healthcare (Vol. 8, No. 4, p. 389). MDPI.

[28] Forlizzi, J., & Battarbee, K. (2004, August). Understanding experience in interactive systems. In Proceedings of the 5th conference on Designing interactive systems: processes, practices, methods, and techniques (pp. 261-268).

[29] Garzonis, S., Jones, S., Jay, T., & O'Neill, E. (2009, April). Auditory icon and earcon mobile service notifications: intuitiveness, learnability, memorability and preference. In Proceedings of the SIGCHI conference on human factors in computing systems (pp. 1513-1522).

[30] Gaver, W. (1994). Using and creating auditory icons. Auditory display-Sonification, audification, and auditory interfaces, 417-446.

[31] Geldard, F. A. (1960). Some Neglected Possibilities of Communication: For some kinds of messages the skin offers a valuable supplement to ears and eyes. Science, 131(3413), 1583-1588.

[32] Gordon, T. J., & Lidberg, M. (2015). Automated driving and autonomous functions on road vehicles. Vehicle System Dynamics, 53(7), 958-994.

[33] Goudarzi, V. (2016). Exploration of sonification design process through an interdisciplinary workshop. In Proceedings of the Audio Mostly 2016 (pp. 147-153).

[34] Green, D. M. (1964). Consistency of auditory detection judgments. Psychological review, 71(5), 392.

[35] Groover, M. P. (2016). Automation, production systems, and computer-integrated manufacturing. Pearson Education India.

[36] Hassenzahl, M., Burmester, M., & Koller, F. (2003). AttrakDiff: A questionnaire to measure perceived hedonic and pragmatic quality. In Mensch & Computer (Vol. 57, pp. 187-196).

[37] Hassenzahl, M., & Tractinsky, N. (2006). User experience-a research agenda. Behaviour & information technology, 25(2), 91-97.

[38] Hoggan, E., Raisamo, R., & Brewster, S. A. (2009, November). Mapping information to audio and tactile icons. In Proceedings of the 2009 international conference on Multimodal interfaces (pp. 327-334).

[39] Huang, G., & Pitts, B. J. (2022). The effects of age and physical exercise on multimodal signal responses: Implications for semi-autonomous vehicle takeover requests. Applied Ergonomics, 98, 103595.

[40] Ilková, V., & Ilka, A. (2017, June). Legal aspects of autonomous vehicles—an overview. In 2017 21st international conference on process control (PC) (pp. 428-433). IEEE.

[41] Insurance Institute for Highway Safety (IIHS) (2020). Advanced driver assistance.

[42] International Organization for Standardization. (2010). Ergonomics — Accessible design — Auditory signals for consumer products (ISO Standard No. 24500:2010). https://www.iso.org/standard/46170.html

[43] International Organization for Standardization. (2019). Ergonomics of human-system

interaction — Part 210: Human-centered design for ineractive systems (ISO Standard 9241-210:2019). https://www.iso.org/standard/77520.html
[44] Kaber, D. B. (2018). A conceptual framework of autonomous and automated agents. Theoretical Issues in Ergonomics Science, 19(4), 406-430.
[45] Kahneman, D., Diener, E., & Schwarz, N. (Eds.). (1999). Well-being: Foundations of hedonic psychology. Russell Sage Foundation.
[46] Katz, B. F., & Marentakis, G. (2016). Advances in auditory display research. Journal on Multimodal User Interfaces, 10(3), 191-193.
[47] Kim, W., Park, D., Kim, Y. M., Ryu, T., & Yun, M. H. (2018). Sound quality evaluation for vehicle door opening sound using psychoacoustic parameters. Journal of Engineering Research, 6(2).
[48] Larsson, P., Opperud, A., Fredriksson, K., & Västfjäll, D. (2009, June). Emotional and behavioural response to auditory icons and earcons in driver-vehicle interfaces. In Proc. 21st International Technical Conference on Enhanced Safety of Vehicles, Germany.
[49] Lazaro, M. J., Kim, S., Choi, M., Kim, K., Park, D., Moon, S., & Yun, M. H. (2022). Design and Evaluation of Electric Vehicle Sound Using Granular Synthesis. Journal of the Audio Engineering Society, 70(4), 294-304.
[50] Li, L., Huang, W. L., Liu, Y., Zheng, N. N., & Wang, F. Y. (2016). Intelligence testing for autonomous vehicles: A new approach. IEEE Transactions on Intelligent Vehicles, 1(2), 158-166.
[51] Marshall, D. C., Lee, J. D., & Austria, P. A. (2007). Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness. Human factors, 49(1), 145-157
[52] Merat, N., & Jamson, A. H. (2009). How do drivers behave in a highly automated car?
[53] Merat, N., Jamson, H. A., Lai, F., & Carsten, O. (2014). Human factors of highly automated driving: results from the EASY and CityMobil projects. In Road vehicle automation (pp. 113-125). Springer, Cham
[54] Moon, S., Park, S., Park, D., Kim, W., Yun, M. H., & Park, D. (2019). A study on affective dimensions to engine acceleration sound quality using acoustic parameters. Applied Sciences, 9(3), 604.

[55] Murali, P. K., Kaboli, M., & Dahiya, R. (2021). Intelligent In-Vehicle Interaction Technologies. Advanced Intelligent Systems, 2100122.

[56] Nardi, B. A. (1992). The use of scenarios in design. ACM SIGCHI Bulletin, 24(4), 13-14.

[57] National Center for Statistics and Analysis. (2022, February). Early estimate of motor vehicle traffic fatalities for the first 9 months (January–September) of 2021 (Crash•Stats Brief Statistical Summary. Report No. DOT HS 813 240). National Highway Traffic Safety Administration.

[58] National Highway Traffic Safety Administration. (2017). Automated driving systems 2.0: A vision for safety. Washington, DC: US Department of Transportation, DOT HS, 812, 442.

[59] Nees, M. A., & Walker, B. N. (2011). Auditory displays for in-vehicle technologies. Reviews of human factors and ergonomics, 7(1), 58-99.

[60] Nees, M. A., Helbein, B., & Porter, A. (2016). Speech auditory alerts promote memory for alerted events in a video-simulated self-driving car ride. Human factors, 58(3), 416-426.

[61] Nees, M. A., Walker, B. N. (2011). Auditory Displays for In-Vehicle Technologies. Reviews of Human Factors and Ergonomics, 7(1), 58–99

[62] Orzessek, B., & Falkner, M. (2006). Sonification of autonomic rhythms in the frequency spectrum of heart rate variability. Georgia Institute of Technology.

[63] Oswald, D. (2012). Non-speech audio-semiotics: A review and revision of auditory icon and earcon theory. Georgia Institute of Technology

[64] Otto, N., Amman, S., Eaton, C., & Lake, S. (2001). Guidelines for jury evaluations of automotive sounds. Sound and Vibration, 35(4), 24-47.

[65] Palladino, D. K., & Walker, B. N. (2007). Learning rates for auditory menus enhanced with spearcons versus earcons. Georgia Institute of Technology.

[66] Palomaki, H. (2006). Meanings conveyed by simple auditory rhythms. Georgia Institute of Technology.

[67] Park, D., Park, S., Kim, W., Rhiu, I., & Yun, M. H. (2019). A comparative study on

subjective feeling of engine acceleration sound by automobile types. International Journal of Industrial Ergonomics, 74, 102843.

[68] Patterson, R. D. (1999). Guidelines for Auditory Warning Systems on Civil Aircraft, CAA paper 82017, London, Civil Aviation Authority, 1982. Cited in NA Stanton & J. Edworthy: Human Factors in Auditory warnings, 3-30. Hampshire, Ashgate.

[69] Payre, W., Birrell, S., & Parkes, A. M. (2021). Although autonomous cars are not yet manufactured, their acceptance already is. Theoretical Issues in Ergonomics Science, 22(5), 567-580.

[70] Robins, B., Ferrari, E., Dautenhahn, K., Kronreif, G., Prazak-Aram, B., Gelderblom, G. J., ... & Marti, P. (2010). Human-centred design methods: Developing scenarios for robot assisted play informed by user panels and field trials. International Journal of Human-Computer Studies, 68(12), 873-898.

[71] Roginska, A. (2013). Auditory icons, earcons, and displays: Information and expression through sound. The psychology of music in multimedia, 339.

[72] Rossi, G. B., Crenna, F., & Panero, M. (2005). Panel or jury testing methods in a metrological perspective. Metrologia, 42(2), 97.

[73] Rukonić, Luka & Pungu, Marie-Anne & Kieffer, Suzanne. (2021). UX Design and Evaluation of Warning Alerts for Semi-autonomous Cars with Elderly Drivers. 25-36. 10.5220/0010237000250036.

[74] Šabić, E., Chen, J., & MacDonald, J. A. (2021). Toward a better understanding of in-vehicle auditory warnings and background noise. Human factors, 63(2), 312-335.

[75] SAE International (2014) Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. J3016_201401.

[76] Schneider, T., Hois, J., Rosenstein, A., Ghellal, S., Theofanou-Fülbier, D., & Gerlicher, A. R. (2021, May). ExplAIn Yourself! Transparency for Positive UX in Autonomous Driving. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (pp. 1-12).

[77] Schrepp, M. (2015). User experience questionnaire handbook. All you need to know to apply the UEQ successfully in your project.

[78] Seagull, F. J., Wickens, C. D., & Loeb, R. G. (2001, October). When is less more? Attention and workload in auditory, visual, and redundant patient-monitoring conditions. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 45, No. 18, pp. 1395-1399). Sage CA: Los Angeles, CA: SAGE Publications.

[79] Shafiro, V., Hebb, M., Walker, C., Oh, J., Hsiao, Y., Brown, K., ... & Moberly, A. C. (2020). Development of the basic auditory skills evaluation battery for online testing of cochlear implant listeners. American Journal of Audiology, 29(3S), 577-590.

[80] Smith, D. R., & Walker, B. N. (2002). Tick-marks, axes, and labels: The effects of adding context to auditory graphs. Georgia Institute of Technology.

[81] Tomlinson, B. J., Noah, B. E., & Walker, B. N. (2018, April). Buzz: An auditory interface user experience scale. In Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1-6)

[82] Trucks, V. (2013). European accident research and safety report 2mi3. Volvo Trucks, Gothenburg, Sweden, Tech. Rep.

[83] Walker, B. N. (2002). Magnitude estimation of conceptual data dimensions for use in sonification. Journal of experimental psychology: Applied, 8(4), 211.

[84] Walker, B. N. (2007). Consistency of magnitude estimations with conceptual data dimensions used for sonification. Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, 21(5), 579-599

[85] Walker, B. N., Lindsay, J., Nance, A., Nakano, Y., Palladino, D. K., Dingler, T., & Jeon, M. (2013). Spearcons (speech-based earcons) improve navigation performance in advanced auditory menus. Human Factors, 55(1), 157-182.

[86] Walker, B. N., Nance, A., & Lindsay, J. (2006). Spearcons: Speech-based earcons improve navigation performance in auditory menus. Georgia Institute of Technology.

[87] Yu, B., Feijs, L., Funk, M., & Hu, J. (2015). Designing auditory display of heart rate variability in biofeedback context. Georgia Institute of Technology.

Appendix 1. Text Network Visualization on User's Opinions

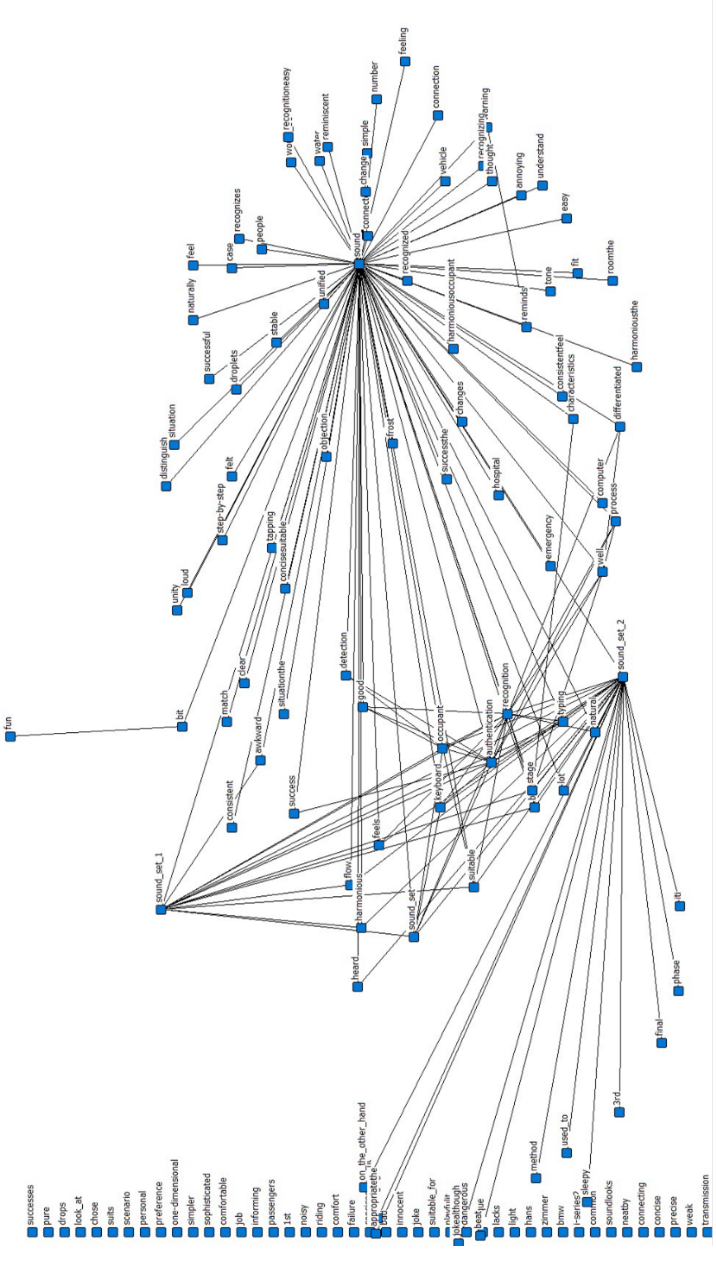


Figure 34. Text network visualization of participants' opinions for Scenario 1

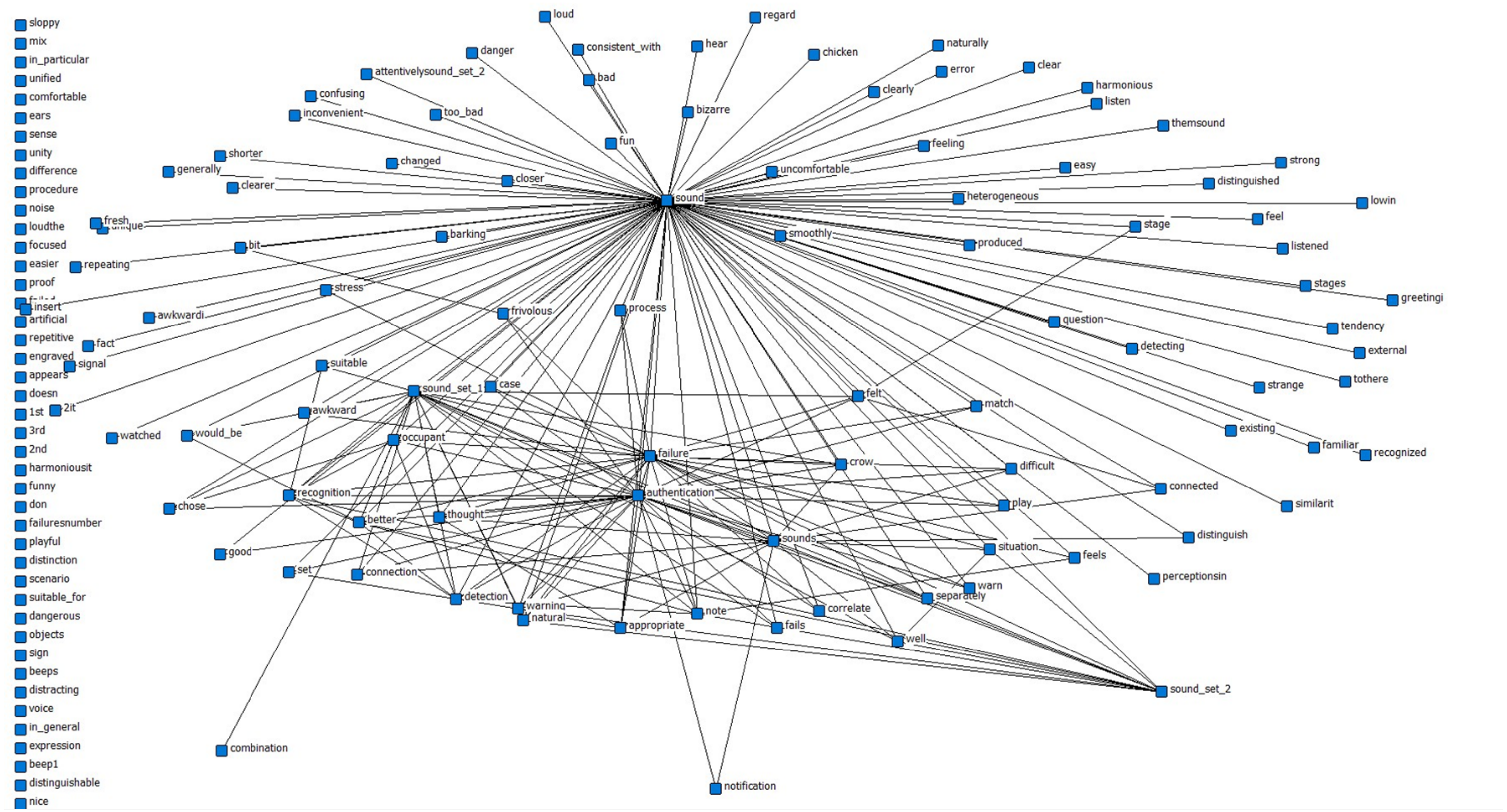


Figure 35. Text network visualization of participants' opinions for Scenario 2

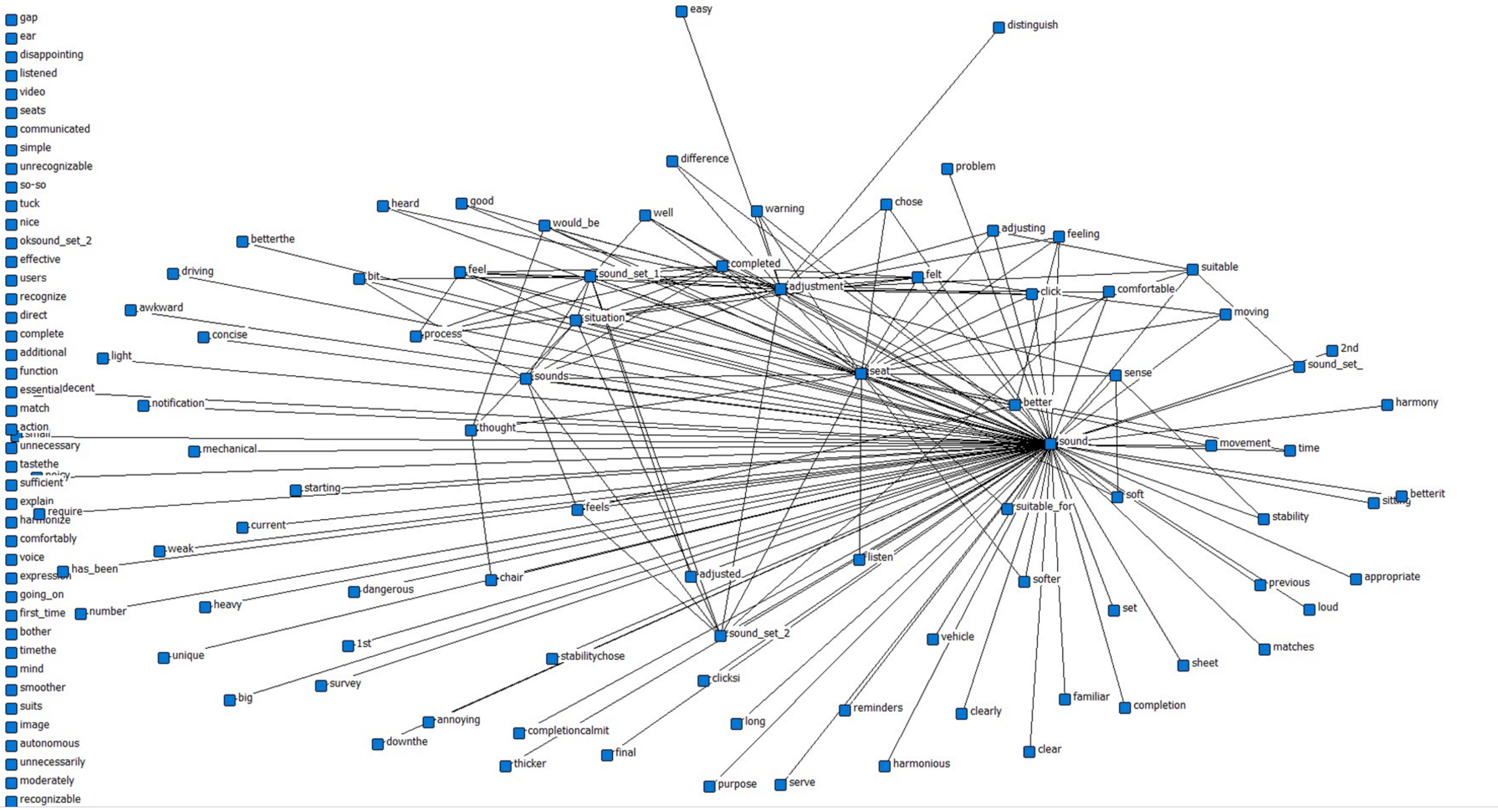


Figure 36. Text network visualization of participants' opinions for Scenario 3

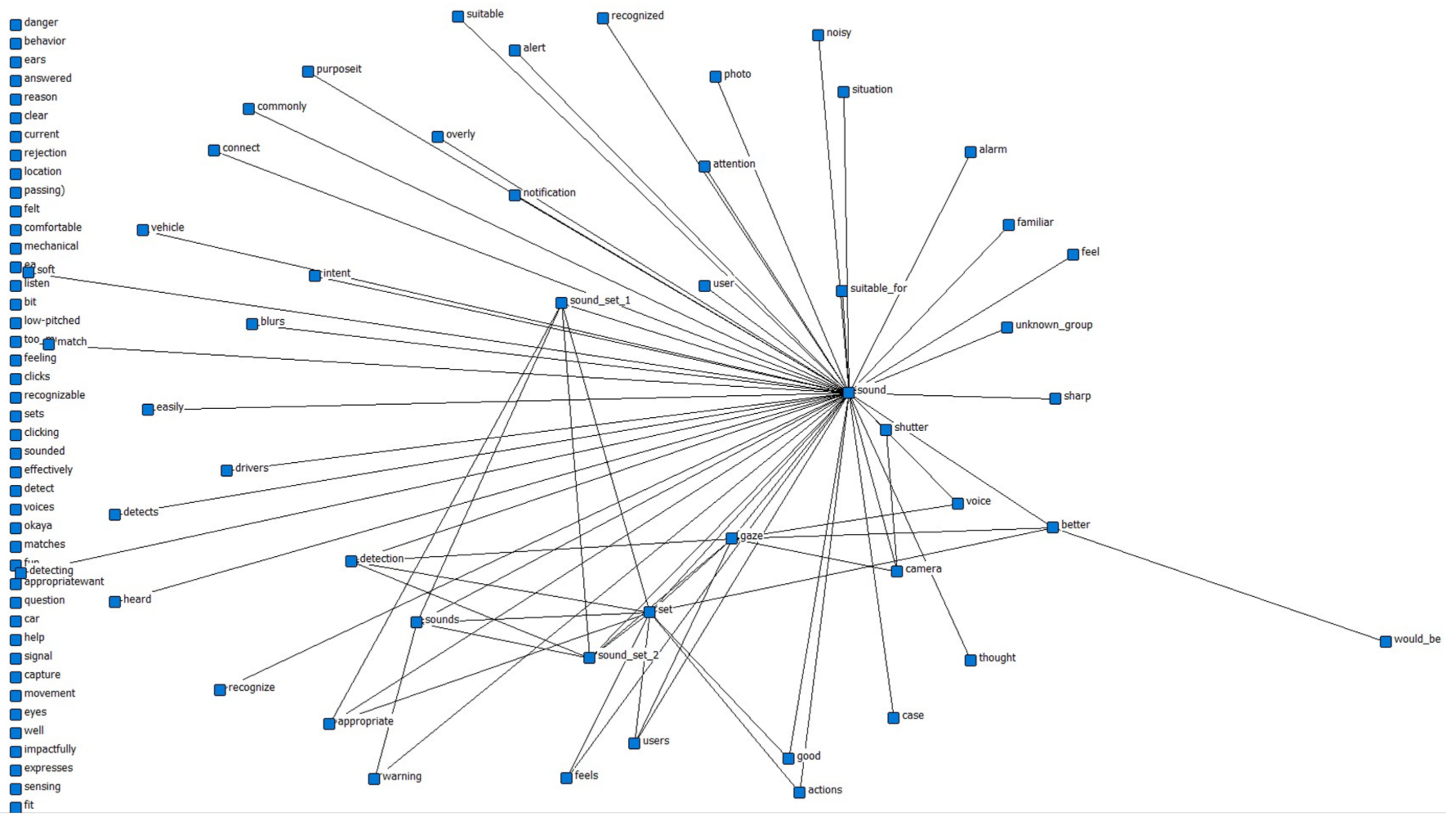


Figure 37. Text network visualization of participants' opinions for Scenario 4

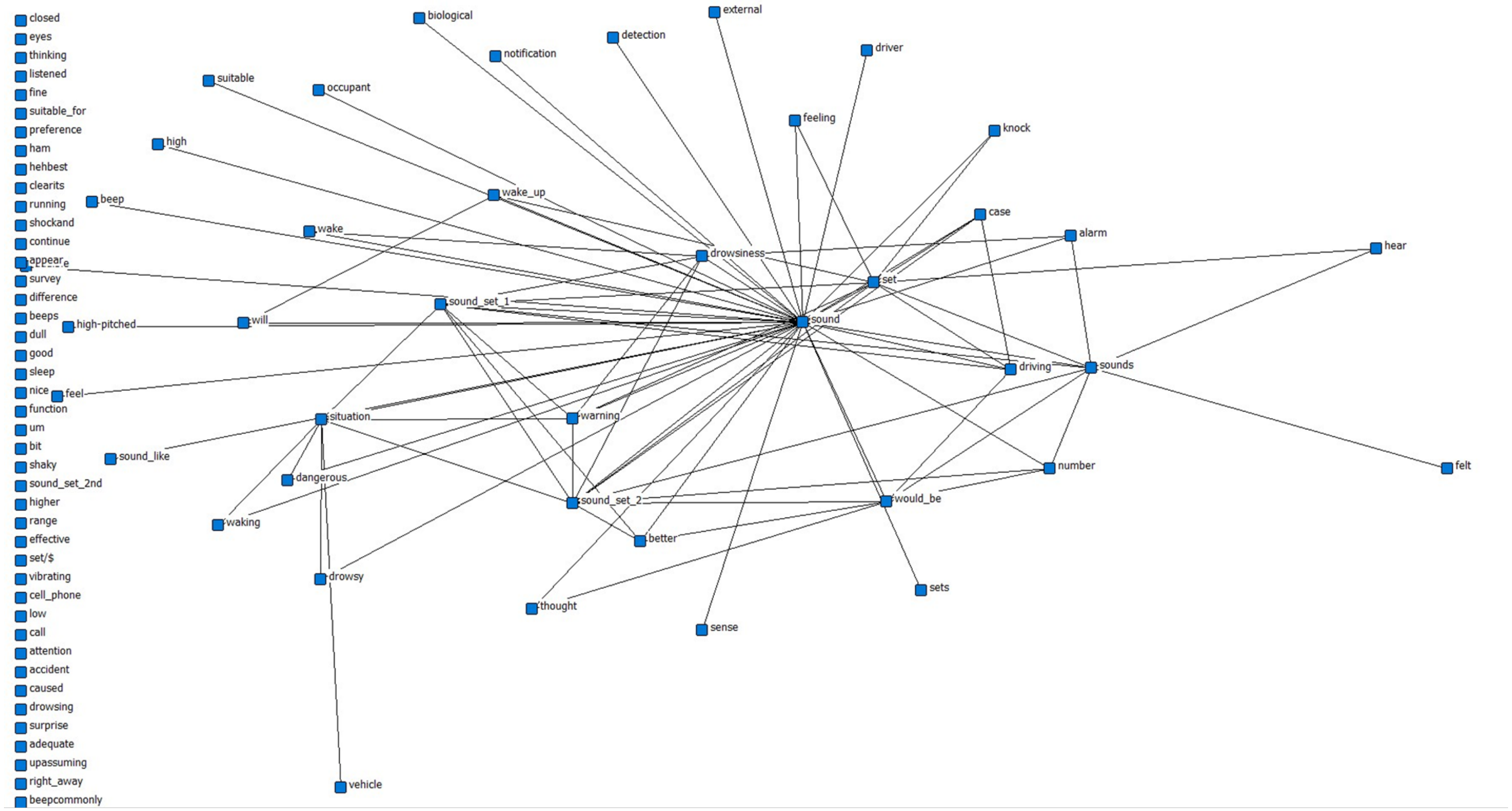


Figure 38. Text network visualization of participants' opinions for Scenario 5

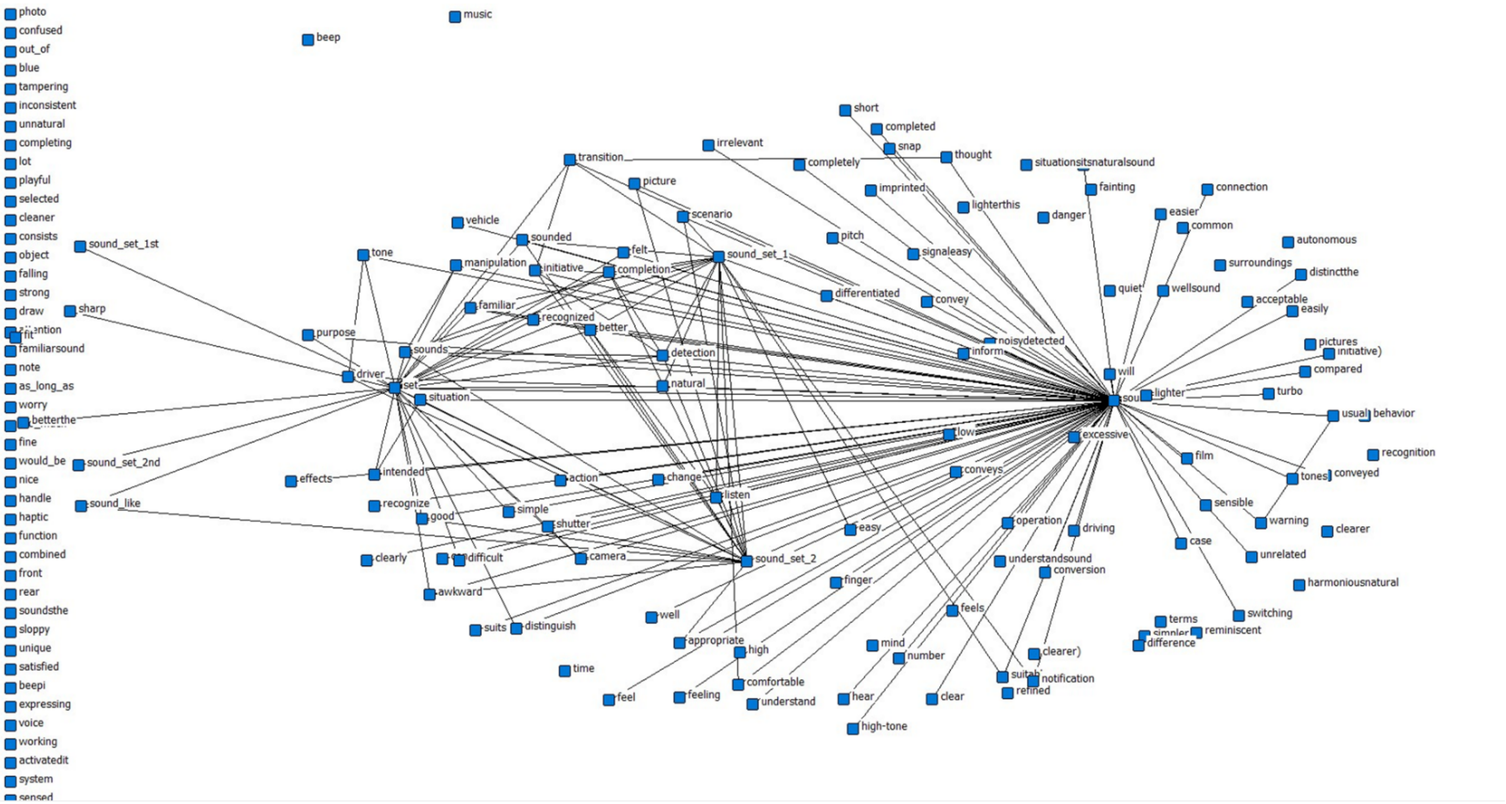


Figure 39. Text network visualization of participants' opinions for Scenario 6

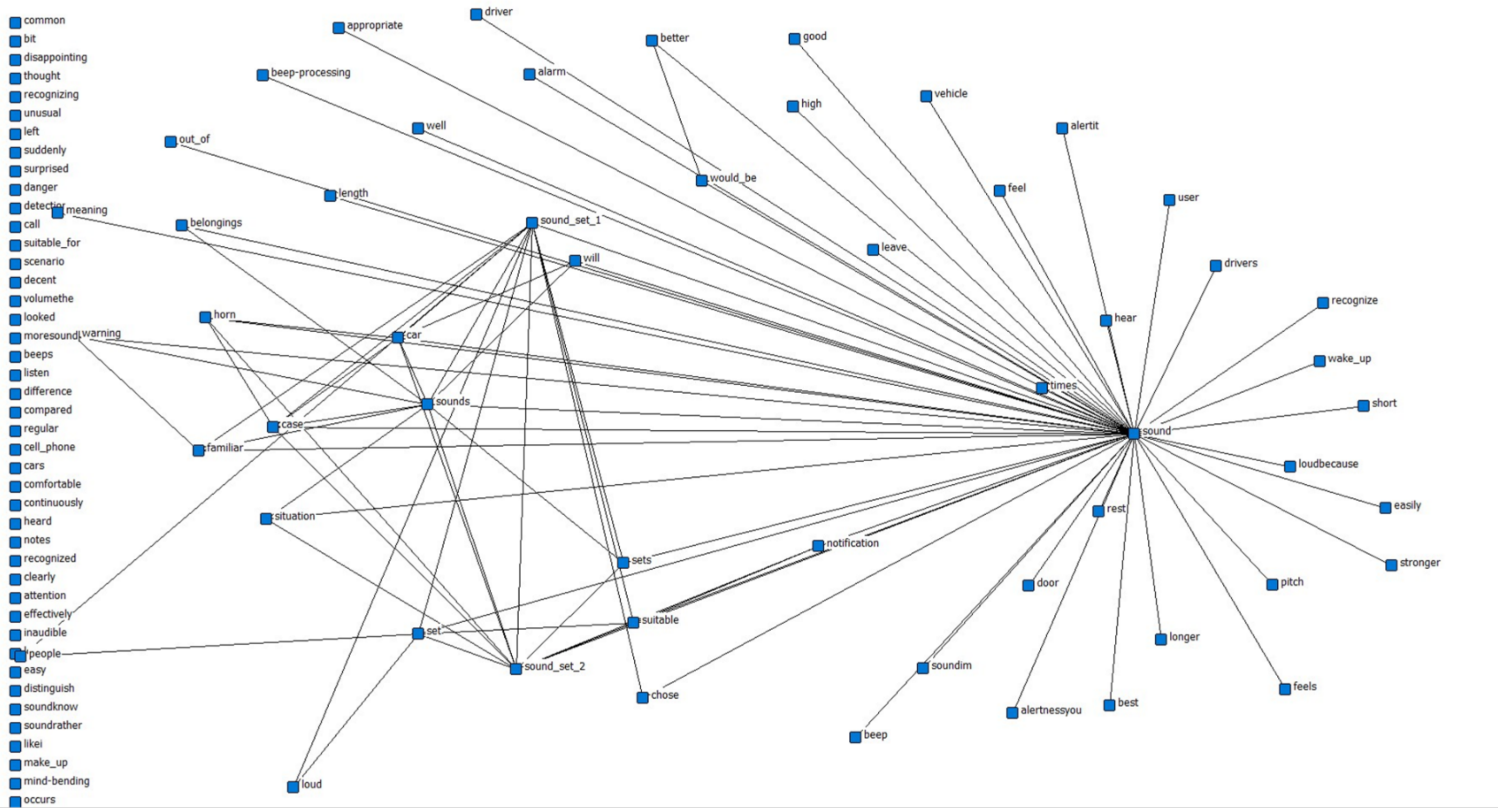


Figure 40. Text network visualization of participants’ opinions for Scenario 7

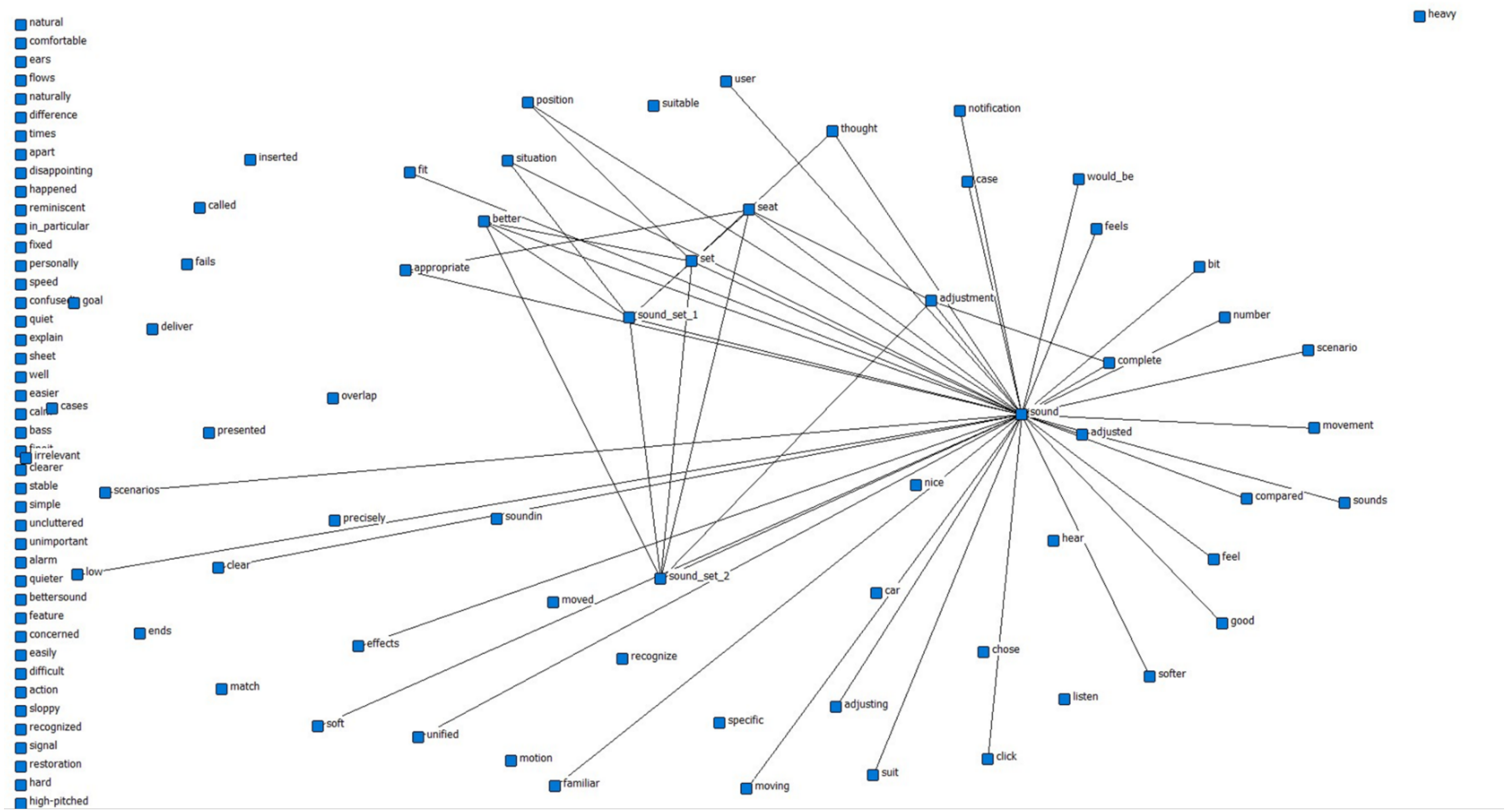


Figure 41. Text network visualization of participants' opinions for Scenario 8

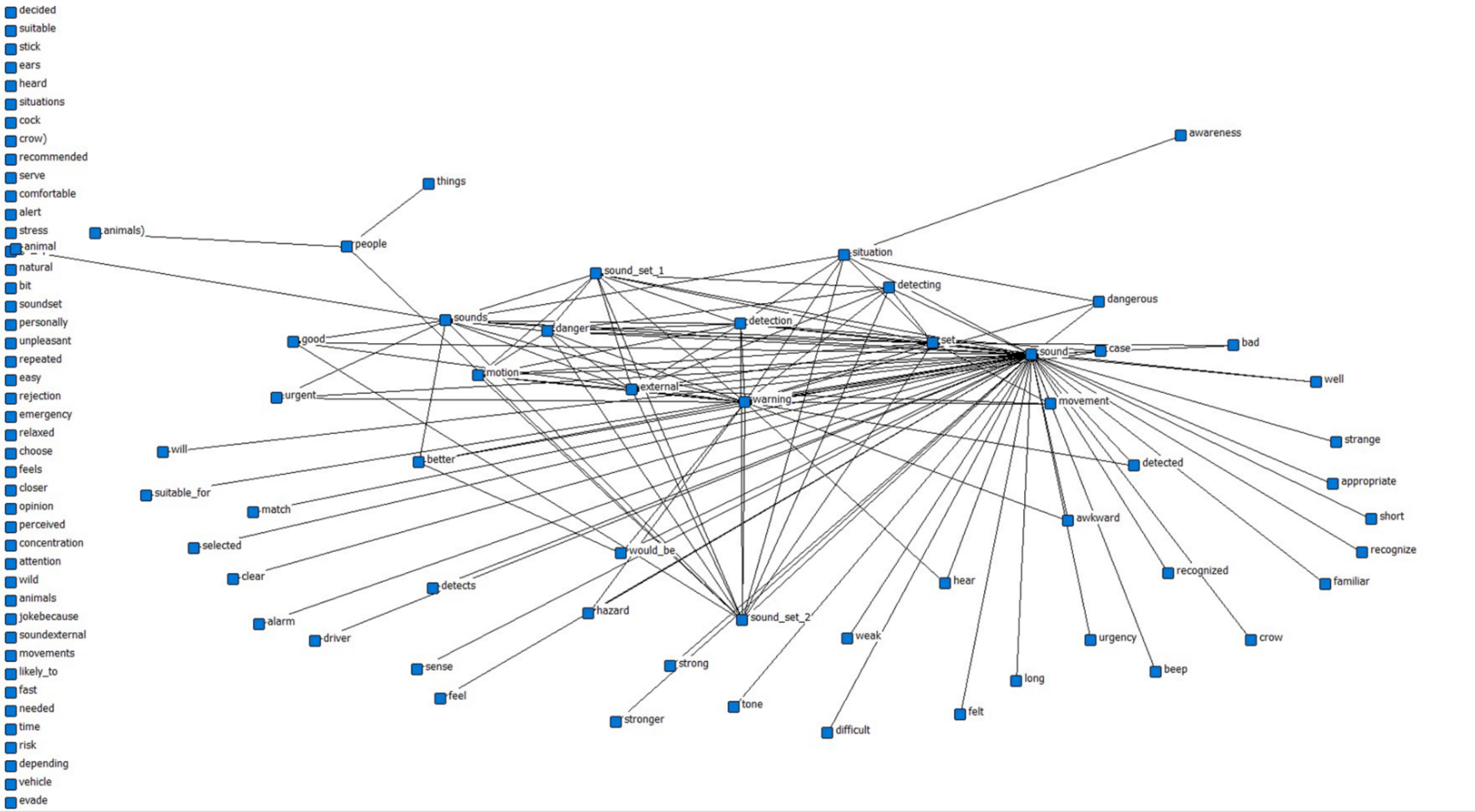


Figure 42. Text network visualization of participants’ opinions for Scenario 9

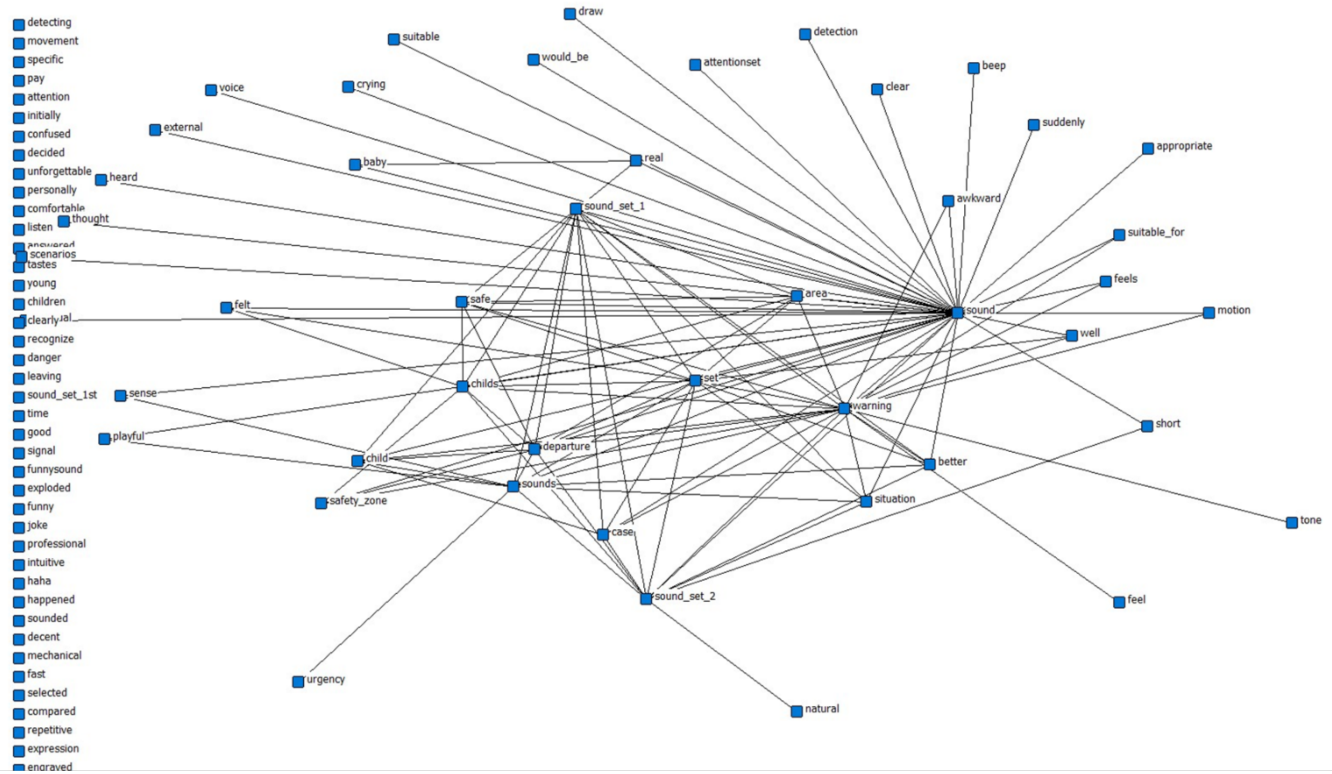


Figure 43. Text network visualization of participants' opinions for Scenario 10

국문초록

최근에는 차량 내에 자율 기술의 발달로 자율주행차의 기능이 고도화가 이루어졌고, 우리가 알고 있는 인간-차량 인터랙션은 점점 인간-로봇 인터랙션으로 패러다임이 변하고 있다. 차량 내에서 청각 유저 인터페이스는 운전자의 인지 부하를 줄이고, 운전자에게 정보를 제공하기 위해 차량 기술에 사용되고 있다. 그러나, 자율주행차는 새로운 기술 도메인으로 인해 사용자, 즉, 탑승자 사용맥락과 시나리오에 따라 청각 피드백 유형 설계가 필요하다. 본 논문에서는 세 가지 주요 연구 목표는 (1) 탑승자의 관점에 기반한 자율주행차량에 대한 직관적인 청각 피드백 설계 제안, (2) 자율주행차량에 적용된 청각 피드백에 대한 선호도, 그리고, (3) 자율주행차에서 필요한 청각 사용자 경험 시나리오 도출하는 것이다. 본 논문의 연구 목적을 달성하기 위해 제작된 청각적 피드백의 인지용이성, 직관성 일관성 또는 적절성을 측정하여 청각 피드백의 유형 및 정보 유형의 패턴이 탑승객의 선호에 영향을 미치는지 조사하는 방향을 잡았다.

본 논문은 파일럿 테스트와 대규모 온라인 사운드 평가로 실험을 진행하였다. 파일럿 테스트에는 총 13명 참가자 연령 27.23세(±7.53) 대상으로 실시하였고, 제작된 사운드 샘플의 의도된 정보(조작 확인음, 조작 오류음, 감지음, 진행음, 약경고음, 강경고음)와 인지적으로 용이성과 직관성이 있는지 진행하였다. 추가적으로, 청각 피드백이 필요한 자율주행차량의 시나리오를 도출하기 위해서도 빈도분석을 진행하였다. 인지용이성과 직관성 평가를 통해 얻은 데이터는 분산분석(ANOVA)과 다중 비교를 위한 본페로니 사후 검정을 사용하여

분석하였다. 파일럿 테스트 결과, 진행음 피드백을 제외하고 모든 사운드 샘플이 의도된 정보로 직관적으로 설계된 것으로 확인하였다. 따라서, 본 논문에서 사용되는 진행음을 다시 설계 및 제작해야 하는 것이다. 또한, 개발된 27가지 시나리오 중 탑승객의 사용맥락에 기반한 청각 피드백이 필요한 상황이 15가지 필수 시나리오로 도출하였다. 파일럿 테스트를 이어서, 사운드 평가는 평균 연령이 37.15세(±11.4)인 총 125명의 참가자를 대상으로 온라인으로 대규모 수행되었으며, 7점 척도로 일관성/적절성 측정을 통해서 어떤 사운드 유형 (이어콘과 오디토리 아이콘의 혼합 또는 일련의 이어콘/오디토리 아이콘)을 선호하는지 조사하였다. 진행음은 올라가는 멜로, 내려가는 멜로, 변형 및 단순 음색으로 4가지 파라미터로 재제작하였다. 본 평가에서 얻어낸 일관성/적절성 데이터는 각 사운드 세트에 대한 쌍별 *t*-테스트 비교를 사용하여 분석하였다. 진행음은 만족도로 측정하여 분산분석(ANOVA)을 사용하여 분석하였다. 마지막으로, 참가자들의 의견들을 정성적 분석을 위해 텍스트 네트워크 분석으로 시각화를 하였다. 각 시나리오에 독립적인 표번 t-테스트의 결과에 따르면, 사용자가 자율주행차량 탑승의 시나리오에서 이어콘과 오디토리 아이콘의 혼합보다는 일관된 사운드 세트를 선호한다는 결과가 나타났다. 또한, 진행음의 결과는 내려가는 멜로디와 단순 음색의 파라미터로 높은 만족도가 보였다. 마지막으로, 본 논문의 토의 부분에서는 파일럿 테스트와 온라인 대규모 사운드 평가 실험을 통해서 얻은 결과를 바탕으로 연구 목표의 달성에 대해 토의하였다. 결론 부분에서는 본 논문의 한계점과 향후 연구 방향에 대해서도 논의하였다.

주요어: 청각 유저 인터페이스, 청각 피드백 설계, 자율주행차, 인간-차량 인터랙션

학번: 2020-28244