Ergonomics In Design, Vol. 77, 2023, 605–613
https://doi.org/10.54941/ahfe1003420

Future Automobile Driving Space Voice Interaction: Adapt to the Driving Scenarios and User Personalities

Weitong Zhang and Jiaxin Chen

School of Design, Hunan University, Changsha 414000, China

ABSTRACT

With the continuous development of intelligent vehicles, Voice User Interface (VUI) has become the core interaction mode in the automobile driving space for its advantage of low interference to the driver. Although the in-vehicle VUI has been improved dramatically over the past, it is still mainly limited to functional aspects such as speech recognition accuracy and conversational fluency. In order to match users' requirements correctly, we conducted car driving simulation experiments and semi-structured interviews with 26 participants, who were asked to rate their experience of different styles of voice assistant persona in various driving scenarios and take an active part in discussions. Our research indicates that the preferred voice assistant personality exists difference in diverse driving scenarios. In addition, participants' gender and personality characteristics also have an influence on their preferences of voice assistant personality. We hope these early results open new research questions to improve in-vehicle VUI.

Keywords: Automotive ui, Voice user interface, Emotional voice assistant, User experience

INTRODUCTION

The eyes-free and hands-free characteristics of the voice user interface (VUI) benefit driving situations. Until now, however, it is common that users feel uncomfortable when using VUI in driving scenarios, since the current in-vehicle voice assistant has no capacity to provide reasonable feedback immediately based on the traffic scenes and driver status. However, the emotion of the voice actually changes according to the situations in human daily interactions. The monotonous emotion makes the in-vehicle voice assistant lack of naturalness, so that drivers feel like talking to a machine instead of a person, which shows there is still a lot of room for improving the in-vehicle VUI experience.

In-vehicle voice interaction has been explored from different aspects. Marie Jonsson's research found in-car voice assistants can influence driver attitudes and driving performance(Jonsson and Dahlbäck, 2009). This makes it important to find the most appropriate voice to suit each individual's preferences in the car driving scenarios (Strohmann et al., 2019). Braun et al. explored the relationship between voice personality and user experience and found that voice assistants that matched the user's personality had higher scores (Braun et al., 2019).

To explore the relationship between voice assistant emotional styles, driving scenarios and tasks, and users, we designed a series of voice assistants with different emotional styles (Joy, Relaxed, Urgent, and Neutral), recorded audio samples for five specified driving scenarios and tasks. Seven participants were invited to evaluate the voices, then based on the feedback, the voices were improved until they conformed to the set emotional styles. This was followed by a simulated driving experiment, in which several participants (N = 26) with driving experience were invited to test our voice assistants during five designated driving scenarios and tasks, by using questionnaire scoring and semi-structured interviews.

METHODS AND EXPERIMENTS

To investigate emotionally stylized voice assistants for driving scenarios, we first collected several typical emotional tones of in-car assistants. Then we designed a set of assistants which we finally evaluated in a car driving simulation study.

Emotional Voice Assistant Design

Our design of five assistants is based on Russell's two-dimensional emotional model (Posner et al., 2005). The model includes the dimensions Excited—Calm and Pleasant —Unpleasant as a basic emotional component (see Figure 1). We designed assistants to cover the different dimensions of the model. In particular, the emotions and their respective placement were: Joy (Pleasant, Excited), Relaxed (Pleasant, Calm), Urgent (Unpleasant, Excited), Depression (Unpleasant, Calm), Neutral (0, 0). Finally, after discussion, Depression (Unpleasant, Calm) was removed, which was generally considered inappropriate.

All emotional voice assistants were designed based on a two-dimensional emotion model. The traits are expressed through choice of words and intonation, while content and extent of speech output were identical for all

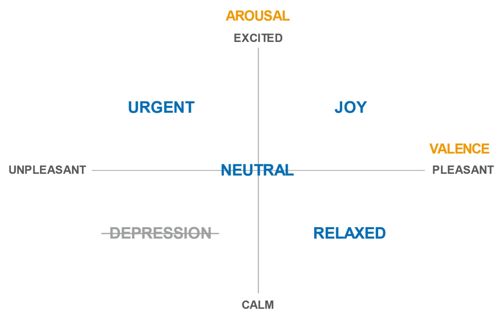

Figure 1: We used Russell's two-dimensional emotion model to design emotional voice assistants. Each assistant was designed to match the four dimensions and the origin (removing an inappropriate emotion).

emotional voice assistants to ensure comparability. All audio snippets used in the study were recorded by a voice actress.

We provide a brief description of each emotional voice assistant.

Joy. This emotion is expressed as a high level of excitement and pleasure. When talking to the assistant, she showed happy emotions, as if she had met something happy.

Relaxed. This emotion shows a high level of pleasure and a low level of excitement. The tone of voice is softer and slower when talking to this voice assistant.

Urgent. This emotion shows a high level of excitement and a low level of pleasure. The assistant speaks faster and with a higher tone of voice. And the speech is more brief and formal.

Neutral. This emotion is expressed as a moderate level of excitement and pleasure. When speaking to the assistant, her tone of voice was unemotional. The assistant simply follows orders and completes tasks.

Car Driving Simulation Study

We conducted a driving study (N = 26) in the lab to investigate the emotions of driving assistants by using a mid-fidelity driving simulator. The main question we answer with this experiment is whether the emotional voice assistant, based on different driving scenarios and the user's personality, has an advantage in comparison to a single emotional or personality in-car voice assistant. And whether this can play a positive role in the driving interaction experience.

This study used assistant emotion as a between-subject independent variable. Four emotional assistants (Joy, Relaxed, Urgent, Neutral) were tested in each participant's experiment. Additionally, we designed five different driving scenario tasks (Passenger assistance, Driving assistance, Navigation, In-car entertainment, Proactive assistance) as independent variables in the experiment for each participant (see Table 1).

As dependent variables we collected user personality traits (Big Five Inventory) (Carciofo et al., 2016) and subjective ratings on user experience (UEQ modules Attractiveness and Adaptability and Efficiency)(Laugwitz et al., 2008). Users were asked to experiment with five different driving scenario tasks in turn, with all four emotional voice assistants being tested in each

Table 1. Five driving scenarios and task contents.

Task	Access to amenities	Content
1	Passenger Assistance	Use the voice assistant to turn on the air conditioning and set it to 25 °C.
2	Driving Assistance	Use the voice assistant to call Xiao Ming.
3	Navigation	Use the voice assistant to navigate to the office.
4	In-car Entertainment	Use the voice assistant to play the music of Jay Chou's "Inari".
5	Proactive Assistance	Follow the voice assistant's instructions to complete the task of reversing into a parking space.

driving scenario task. Accordingly, users were required to perform a total of 20 task experiments (five scenario tasks, four emotional assistants per scenario task). At the end of the experiment, participants answered questions on the experienced emotional voice assistants in a semi-structured interview and offered personal insights into the future of in-car voice assistants.

Twenty-six participants aged 20–30 years who took part in the study (11 male, 15 female). All had previous experience of driving a car. None of them had expert knowledge on digital assistants, though most had used one in their everyday life (7 daily, 4 often, 12 rarely, 1 never). Overall, the majority of the participants were familiar with the voice controls as well as the in-car systems.

We conducted the experiment as a Wizard of Oz study, with the participants driving a mid-fidelity driving simulator that we had prepared, and the operator sitting in the back. The operator observed the display screen and the driver's behavior in order to trigger experiment use cases in appropriate situations.

The operator was equipped with a tablet that could run the speaker remotely, through which the audio files could be played on speaker on the right front of the user. A display was positioned directly in front of the user to show driving simulation scenes. A camera positioned in front of the user's left side, directed at the face, was used to record the user's behavior during the experiment. The driving simulator was used to simulate driving scenarios. The driving simulator has a force feedback steering wheel, an accelerator pedal and a brake pedal. Figure 2 shows the driving simulator setup.

The experiment took place in an indoor lab. Driving scenarios are urban roads in a driving simulation game (City Car Driving). The first four driving scenarios (Passenger assistance, Driving assistance, Navigation, In-car entertainment) are free riding in the urban area and the last one (Proactive assistance) is in an urban car park. Firstly the operator will demonstrate the operation of the simulator to ensure that the user knows how to drive using the driving simulator. Before the experiment users were allowed to enter the driving simulation game to familiarize themselves with the vehicle and the roads for 5 minutes. There are pedestrians and other vehicles in the urban

Figure 2: The driving simulator setup. Participants performed the driving simulation in front of a table.

road scene to simulate realistic urban road conditions. Additionally there are virtual traffic lights, buildings, parks and other facilities that simulate realistic scenarios, and the user is required to obey the traffic rules within the scenario when driving.

At the lab site, we introduced participants to the concept of intelligent voice assistants and to the procedure. They answered the general questions on demographics and Big Five Inventory questionnaire and signed a declaration of consent. After 5 minutes of familiarization with the driving simulation game, the experiment began. Each participant experienced 20 rides: five driving scenario tasks (Passenger assistance, Driving assistance, Navigation, In-car entertainment, Proactive assistance) and four emotional voice assistants (Joy, Relaxed, Urgent, Neutral) per driving scenario task. The order of experienced emotional assistants was alternated between participants to prevent sequence effects.

The ride experiences are divided into five groups according to the driving scenario tasks: Passenger assistance, Driving assistance, Navigation, In-car entertainment, as well as Proactive assistance. In each ride experience group, participants had to interact with each of the four emotional assistants to complete the task. After each ride, participants rated the questionnaire on subjective user experience.

RESULTS

We analysed user experience, gender influence, user personality influence and user feedback of the emotional voice assistant.

In terms of external environmental influences, we compared five driving scenario tasks to analyse user preferences and experiences of speech emotions in different driving scenario tasks. From the user aspect, we compared the differences in male and female preferences for speech emotions. Besides this, we also collected user personality traits (Big Five Inventory), which were used to study the differences in their preferences for voice emotions. All values reported in Table 2 were collected on a scale from −3 (low) to +3 (high).

User Experience

Participants provided subjective UX evaluations via a questionnaire incorporating the modules Attractiveness, Adaptability and Efficiency of the UEQ.

The average of the three combined ratings of attractiveness, adaptability and efficiency is what we call user satisfaction. In direct comparison, the Joy emotional assistant is rated the highest in user satisfaction during the passenger assistance scenario task ($p < 0.05$). However, in the scenario of in-car entertainment, Urgent is assessed to have the lowest level of user satisfaction ($p < 0.01$).

Gender Influence

We evaluated the effect of gender on users' voice preferences in different driving scenario tasks. In this case, gender has an impact on driving style and emotion (Gwyther and Holland, 2012). Figure 3 shows that in the scenario

Table 2. User ratings for experienced emotional assistant Joy, Relaxed, Urgent, and Neutral in five driving scenarios.

		Joy		Relaxed		Urgent		Neutral	
		M	SD	M	SD	M	SD	M	SD
1. Passenger assistance	Attractiveness	1.65	1.54	0.73	1.74	0.19	1.57	−0.19	1.59
	Adaptability	1.69	1.20	1.27	1.06	1.81	0.88	0.92	1.14
	Efficiency	2.12	0.75	2.04	0.90	2.35	0.78	1.92	0.96
	Avg.	1.82		1.35		1.45		0.88	
2. Driving assistance	Attractiveness	0.73	1.93	0.54	1.55	0.38	1.50	1.00	1.47
	Adaptability	1.00	1.47	1.19	0.83	1.31	1.32	1.88	0.89
	Efficiency	1.92	1.24	1.69	0.91	2.12	0.93	2.27	0.65
	Avg.	1.22		1.14		1.27		1.72	
3. Navigation	Attractiveness	0.58	1.90	0.23	1.60	0.69	1.10	1.42	1.01
	Adaptability	0.69	1.61	1.08	1.07	1.65	1.17	1.73	1.09
	Efficiency	1.88	0.93	1.92	0.87	2.00	0.88	2.04	0.85
	Avg.	1.05		1.08		1.45		1.73	
4. In-car entertainment	Attractiveness	1.42	1.60	1.04	1.56	0.08	1.44	1.12	1.19
	Adaptability	1.65	1.27	1.65	0.92	0.31	1.64	1.46	1.25
	Efficiency	2.19	0.73	2.00	1.07	1.96	1.09	2.15	0.77
	Avg.	1.76		1.56		0.78		1.58	
5. Proactive assistance	Attractiveness	0.85	1.77	0.42	1.76	0.35	1.66	0.50	1.62
	Adaptability	0.81	1.49	0.54	1.71	1.35	1.24	1.35	1.41
	Efficiency	1.38	1.36	1.08	1.47	1.85	0.99	1.62	1.21
	Avg.	1.01		0.68		1.18		1.15	

tasks of Passenger assistance, Navigation and Proactive assistance, gender played a role in users' preference for voice emotion. All values reported in Figure 3 were collected on a scale of −3 (low) to +3 (high).

In the scenario task of passenger assistance, the satisfaction rating for Relaxed was higher by female users than by males (Relaxed ranked second very close to first for females while it was very close to lowest for males). In the scenario task of Navigation, females rated Relaxed higher than males (Relaxed was rated second by females, but worst by males). For the Proactive Assistance scenario task, Joy was rated highest by females but lowest by males in terms of user satisfaction.

In the tasks involving driving assistance and in-car entertainment, the user satisfaction rating tendencies of men and women for emotional voice assistants were relatively consistent.

User Personality

After all scenario tasks, participants evaluated the experienced emotional voice assistants using a semantic differential scale with 4 dimensions. Figure 4 shows that all the emotional voice assistants were perceived as we designed.

Figure 5 shows the mean personality trait scores of participants, grouped by their preferred emotional voice assistant. We obtained the participants' preference for the voice assistants based on their user experience scores of the four voice assistants. The results were as follows: of the 26 participants, 13 preferred Joy, 4 preferred Urgent and 7 preferred Neutral, leaving 2 experiencers who preferred Relaxed. As such a small sample is considered to have

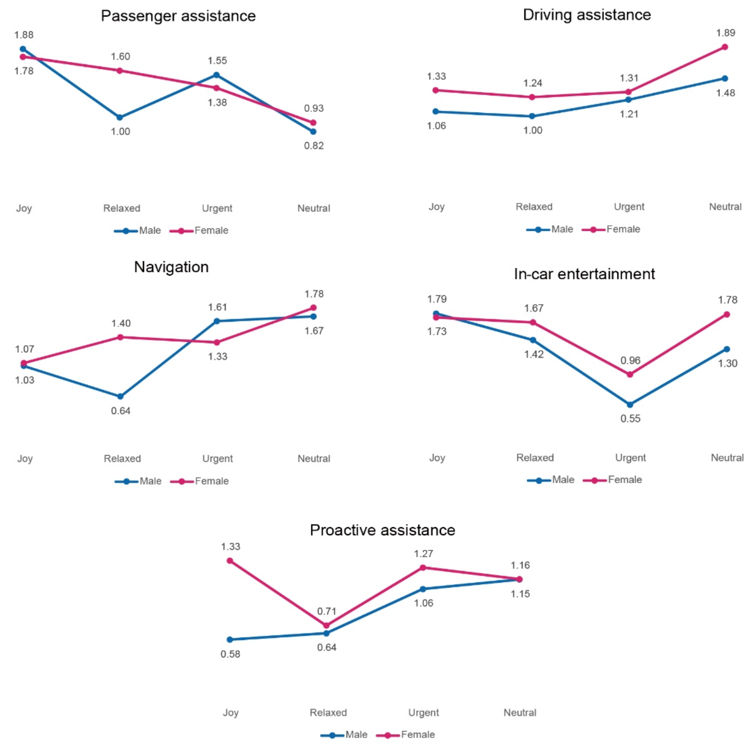

Figure 3: Male and female ratings of voice assistants in five driving tasks.

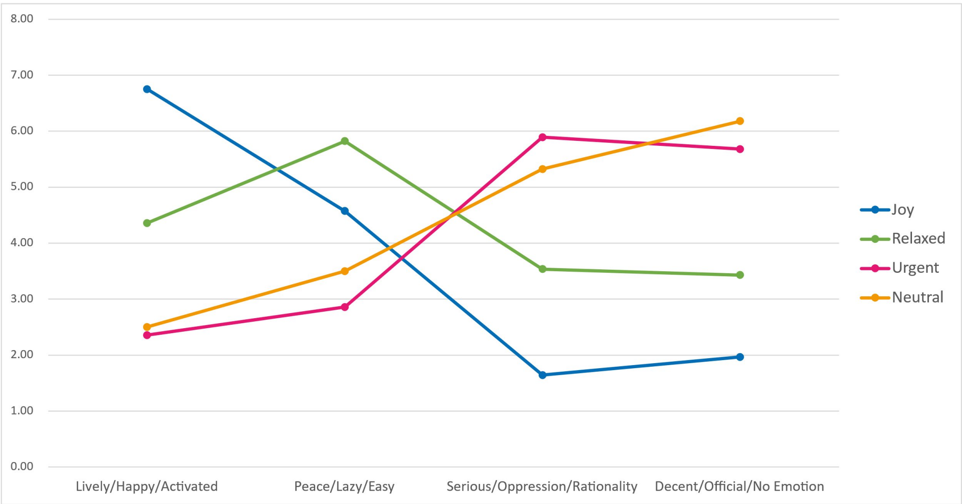

Figure 4: The semantic differences of perceived assistant emotion scale, which shows that participants' perceptions were consistent with our design.

a large error, we have omitted the results for this emotion in the comparison that follows.

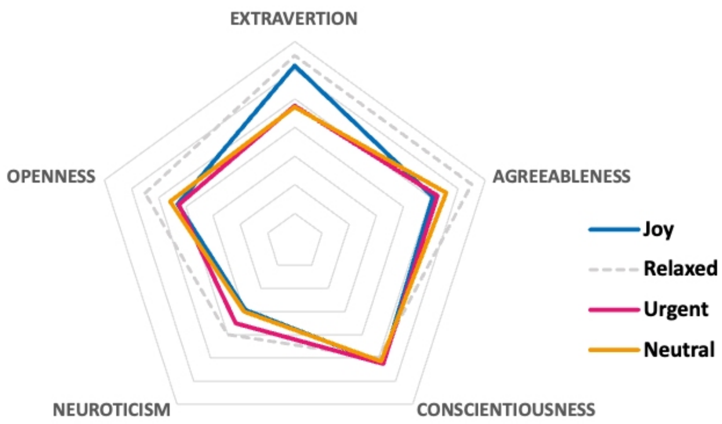

Figure 5: Mean big five scores for participants, split by their preferred emotional voice. Users who chose the joy were significantly more extraverted than all other groups.

In pairwise comparisons, participants who preferred Joy scored significantly higher in extraversion than all other groups ($p < 0.05$). The results show that more extroverted people may prefer lively, happy and activated assistants. In terms of agreeableness, conscientiousness, neuroticism and openness, there were no significant differences between the participants who preferred different voice assistants.

CONCLUSION

This paper explored the impact of different emotional voice assistants on user satisfaction (attractiveness, adaptability, efficiency) in five driving scenarios and tasks. The results show that adapting the right emotional voice to different driving situations and user personalities has a positive impact on increasing user satisfaction. Our research has found that it is necessary to adapt voice emotion based on context-aware, e.g., voice should be serious and brief in situations related to driving performance and driving safety, while in situations not related to driving performance (e. g. in-car entertainment) the voice assistant can be more lively and chatty. And the diversity of users also leads to differences in voice preferences, with gender and extraversion having a significant impact on voice preference.

Follow-up work should study more emotional speech samples and driving scenarios to obtain a model for matching emotional speech to driving contexts. And work needs to be attempted in different linguistic and cultural contexts in order to generalize the findings to users with different cultural backgrounds.

REFERENCES

Braun, M., Mainz, A., Chadowitz, R., Pfleging, B., Alt, F., 2019. At Your Service: Designing Voice Assistant Personalities to Improve Automotive User Interfaces, in: Proceedings of the 2019 CHI Conference on Human Factors in Computing

Systems. Presented at the CHI '19: CHI Conference on Human Factors in Computing Systems, ACM, Glasgow Scotland Uk, pp. 1–11. https://doi.org/10.1145/3290605.3300270

Carciofo R., Yang J., Song N., Du F., Zhang K., 2016. Psychometric Evaluation of Chinese-Language 44-Item and 10-Item Big Five Personality Inventories, Including Correlations with Chronotype, Mindfulness and Mind Wandering. PLOS ONE 11, e0149963. https://doi.org/10.1371/journal.pone.0149963

Gwyther, H., Holland, C., 2012. The effect of age, gender and attitudes on self-regulation in driving. Accident Analysis & Prevention 45, 19–28. https://doi.org/10.1016/j.aap.2011.11.022

Jonsson, M., Dahlbäck, N., 2009. Impact of voice variation in speech-based in-vehicle systems on attitude and driving behaviour, in: Human Factors and Ergonomics Society Europe Chapter (HFES).

Laugwitz, B., Held, T., Schrepp, M., 2008. Construction and Evaluation of a User Experience Questionnaire, in: Holzinger, A. (Ed.), HCI and Usability for Education and Work. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 63–76.

Posner, J., Russell, J. A., Peterson, B. S., 2005. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and psychopathology 17, 715–734. https://doi.org/10.1017/S0954579405050340

Strohmann, T., Siemon, D., Robra-Bissantz, S., 2019. Designing virtual in-vehicle assistants: Design guidelines for creating a convincing user experience. AIS Transactions on Human-Computer Interaction 11, 54–78. https://doi.org/10.17705/1thci.00113