Impact of interface sonification with touchless gesture command in a car

Ludovic Jaschinski[1], Sébastien Denjean[2], Jean-François Petiot[1], Frank Mars[1], & Vincent Roussarie[2]
[1]IRCCyN, UMR CNRS 6597, Ecole Centrale de Nantes, [2]PSA Peugeot Citroën France

Abstract

This experiment aims to study the impact of the sonification of a hand gesture controlled system on the driver behavior. The principle is to give an auditory feedback to the driver, in addition to a visual screen, in order to assist in-car devices interface manipulation. A touchless interface has been tested with a panel of 24 subjects on a driving simulator. Different tasks (pick up the phone, select an item in a list) involving the screen and the interface had to be performed by the user while driving. To study the contribution of the sound feedback on the drivers' behavior, two audio conditions were tested: with and without auditory feedback. The tasks were performed in lowly and highly demanding traffic conditions. Driving and gaze behavior as well as eye-tracking information were analyzed. Moreover, a questionnaire was used to obtain subjective measurements, such as ease of use and feeling of safety. The results show that the sonification helped drivers to feel safer and more focused on the road. This result was confirmed by gaze analysis, which shows that drivers look significantly less to the visual interface when a sound is present, leading to a safer use of the interface.

Introduction

The manipulation of in-vehicle information systems is a challenge in today's vehicle design. These systems are more and more interactive while their complexity is expanding, leading to *infotainement* systems that include many functions. When the user is engaged in a primary task (driving), the basic solution that involves a physical interaction with the device while looking at a visual display is an interaction way that can certainly be improved. In particular, auditory feedbacks may decrease the need for visual attention and free-hand gesture interaction eliminates the need for reaching the device. The sonification of information is an emerging discipline that exploits the capacity of sounds to convey information (Hermann et al., 2011) and many applications are proposed in particular in the car industry (Denjean et al., 2013). In this context, touchless interfaces with auditory feedback may be an interesting alternative to assist and even replace visual interfaces. (Kajastila & Lokki, 2013) showed that auditory interfaces can outpoint visual interfaces, in particular with free-hand (touchless) interaction. A study of multi-modal controls (visual, audio or visual+audio) of in-vehicle information

In D. de Waard, A. Toffetti, R. Wiczorek, A. Sonderegger, S. Röttger, P. Bouchner, T. Franke, S. Fairclough, M. Noordzij, and K. Brookhuis (Eds.) (2017). Proceedings of the Human Factors and Ergonomics Society Europe Chapter 2016 Annual Conference. ISSN 2333-4959 (online). Available from http://hfes-europe.org

systems is presented in (Jakus et al., 2015) with a driving simulator. The results showed that the visual and visual+audio interaction modes are the most fast and efficient, but no significant contribution of the audio to the visual-only mode on the driving performance and safety is noticed: the main contribution of audio is that it increased drivers' preferences. (Sodnik et al., 2008) presented a comparative study on the effectiveness and efficiency of audio-only or visual-only interfaces for the manipulation of a mobile phone while driving on a simulator. They demonstrated that audio-only interfaces were effective to control a mobile, particularly when spatialized auditory cues are used in the audio interface. A study with a real vehicle on the acceptability of a sonification for GPS navigation (Tardieu et al., 2015) showed that with sounds, the driving was estimated as safer by the drivers, even if it did not improve significantly the efficiency of the navigation.

In this context, our work aims to assess the contribution of sounds to a visual touchless interface. The objective is to study the impact of the interface sonification on the driver behavior, during the manipulation of an infotainement system in a car. An experiment was conducted on a driving simulator, with two main experimental conditions: manipulation of the interface with visual-only interaction, or with visual+audio interaction. Two categories of variables were observed: driving parameters, recorded by the simulator, but also the driver gaze behavior, recorded by means of an eyetracking system. After a presentation of the material and methods of the experiment, the results are presented and discussed.

Material and methods

Interface

A graphical interface with gestures command without contact (using a *Leap motion*[1] device) was implemented. Six functional categories were created to constitute the main menu of the infotainement system (phone, air conditioning, contacts, music, news and GPS). The main menu of the interface is given in figure 1.

Figure 2. Main menu of the interface of the in-vehicle information system (French version)

[1] https://www.leapmotion.com

Different contactless gestures, necessary to control the interface, were programmed:

- *Selection* of a category in the main menu: the subject has to point with one finger (index finger) which category he/she wants to select (selection gesture)
- *Validation*: this is made by closing the fist (validation gesture)
- *Browsing* into a list of items (draggable carousel menu): this is done by sweeping the items on the right or left side (horizontal movement of the hand on the right or left) (sweeping gesture)
- *Back return* in the menu: this is done by rotating the hand the palm facing up (return gesture)

The general framework of the interface is given in figure 2. It allows a real time control of the screen and of the played audio samples, according to the gestures provided by the subject:

- The leap motion detects subject' gestures and sends the information to a general program (coded in python).
- The program updates the interface screen in real time and sends informations to the sound synthesizer (*Pure Data*[2] code).
- The syntheziser sends audio samples to the loudspeakers.

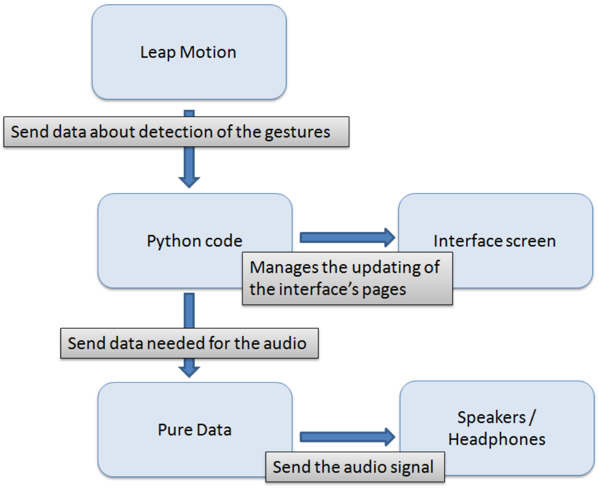

Figure 3. Framework of the interface for the realtime control of the screen and the audio samples

[2] https://puredata.info : Pure Data (aka Pd) is an open source visual programming language that processes and generates sounds based on a graphical interface.

Sonification

Different sounds were associated to the gestures of the user (sonification) in order to provide an audio feedback and help the handling of the interface. The sounds were designed from samples and sounds synthesized on the digital audio workstation *Reaper*[3], with MIDI notes. The following 11 sounds have been designed:

- 6 sounds corresponding to the 6 categories on the main menu, to specify the area pointed by the hand (selection gesture),
- a validation sound (validation gesture)
- a sweep / scrolling sound (sweeping gesture for browsing)
- a sweep / validation sound (sweeping gesture for validation)
- a back return sound (return gesture)
- a refusal sound (pop-up) (sweeping gesture for refusal)

All the sounds were very short (between 50 and 300ms) and they were designed to be easily recognized by the user.

Experiment

Material
The experience took place on the driving simulator of the IRCCyN laboratory, shown in figure 3.

Figure 4. Picture of the fixed-base simulator used for the driving tests

It is a fixed-base simulator, which consist of a compact size passenger car with actual instrument panel, clutch, brake and accelerator pedals, handbrake, ignition key, and an adjustable seat with seat belt. The visual environment was displayed on three 32-inch LCD monitors, each with a resolution of 1280□*720. One monitor was positioned in front of the driver, with two laterals inclined at 45 deg from the

[3] http://www.reaper.fm. Audio processing software

front one, viewed from a distance of about 1 m and covering 115 deg of visual angle. An additional screen was added to the simulator for the interface (figure 4).

Figure 5. Picture of the screen to display the interface and of the leapmotion to capture the gestures of the subject

The Leap Motion was placed in front of this additional screen and connected to the computer of the interface. Speakers were placed behind the simulator screens to play the sound samples. A Smarteye Pro 5 eye-tracking system, composed of 4 cameras placed under the three screens in front of the subject, was calibrated to measure the location and duration of gaze fixations (glances) during the tasks.

Subject and experimental factors
24 subjects (16 men and 8 women – average age 24), students or researchers at Ecole Centrale de Nantes, participated to the tests. The subjects had to drive on a countryside road, on the same route for each scenario. The experimental factor was the sonification condition, with two levels: with sound (s) and without sounds (w-s). For each sonification condition, every course was carried out in two different conditions: with no visible danger on the road (free-flowing traffic) and with a disturbing traffic (a vehicle in front of the driver that drives slowly and brakes regularly). The later case should yield an increased mental workload (not measured). Four scenarios are thus envisaged:

- Scenario 1: no sound, free-flowing traffic
- Scenario 2: no sound, disturbing traffic
- Scenario 3: with sound, free-flowing traffic
- Scenario 4: with sound, disturbing traffic

Interface manipulation tasks
For each course of the scenario, the following four tasks had to be completed by the subject on the interface, in a straight line and in a curve:

- task 1: Acknowledge receipt of a message (pop-up), in a straight line

- task 2: Select a particular destination on the GPS, in a curve
- task 3: Select a particular destination on the GPS, in a straight line
- task 4: Acknowledge receipt of a phone call (pop-up), in a curve

Tasks 1 and 4 are simple: they require only one gesture from the subject (sweeping gesture). Tasks 3 and 4 are more complex: they require several different gestures (selection, browsing, validation, back return). The trigger of the tasks and the corresponding instructions were shown to the subject on the central screen of the simulator. The tasks had to be made always on the same location for every trial. The map of the course and the positioning of the tasks are presented in figure 5.

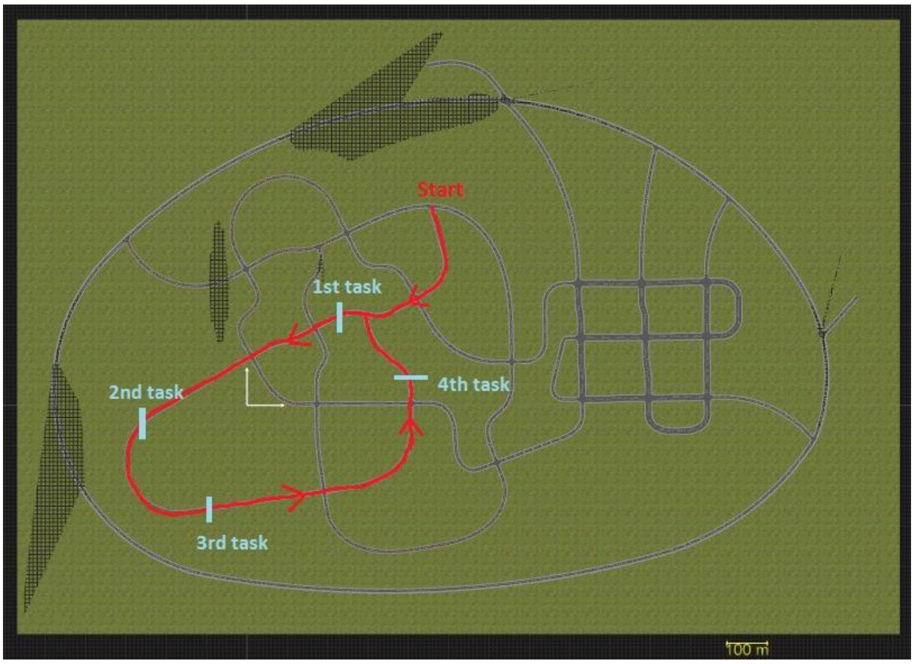

Figure 6. Map of the course and location of the four tasks

Test procedure
First, a training period was proposed to the participants so that they get acquainted with the *Leap Motion* and with the production of the gestures, with the manipulation of the interface, and with the simulator. No subject was familiar with the use of the *Leap motion*. For the interface, the subjects were trained without sound first, and then with sounds. Then, every subject made each of the four scenarios in a counterbalanced order.

Dependant variables (DV)
The following dependent variables were considered, for each subject, each task and each scenario, between the trigger of the task and its completion by the subject:
Driving variables (9 variables):

- Completion time of the task,

- Average and standard deviation (SD) of the lateral deviation of the vehicle,
- Number of steering wheel reversal,
- Average and standard deviation (SD) of the speed of the vehicle,
- Average and standard deviation of longitudinal acceleration,
- Inter-vehicular time.

Eye tracking variables (5 variables):

- Glance duration on the interface screen and number of glances,
- Glance rate (corresponding to the glance duration divided by the time for performing the task) on the interface screen,
- Maximum time of glance duration on the interface screen (without looking back on the simulator screens),
- Number of eye movements from the simulator screens to the interface screen.

Questionnaire

A semi-directive questionnaire was proposed to the participants after the completion of each scenario. Different questions were formulated to assess:

- the global impression: ease of use, feeling of safety, feeling of not looking at the road,
- the impression on the sound design: is it disturbing, is it appreciated?
- two open questions, one on the global interface and one on the sound design

Data analysis

Descriptive statistics on the DV showed that few spurious data were present in the recordings. To reduce the effect of these outliers on the DV, a winsorisation was applied on each DV (extreme values above the 95^{th} percentile and below the 5^{th} percentile were setted to the 95^{th} and 5^{th} percentile respectively) (Wilcox, 2014). Each dependent variable (DV) was next analysed with ANOVA (Næs et al., 2010). Two ANOVA were carried out:

- one-way ANOVA with the factor “traffic condition” (two levels: free-flowing traffic; disturbing traffic)
- two-ways ANOVA with interaction with the factors “subject” (24 levels) and “sonification condition” (two levels: with sounds (s) and without sounds (w-s))

The significance of the effects was analysed with the Fisher test (type III sum of square) and the associated p-value (significance level: $p<.05$)

Results and discussion

Effect of the factor “traffic conditions” (one way ANOVA)

The results show that, for all the DVs considered, there is no significant effect of the “traffic conditions” (one way ANOVA, $p>.05$). The subjects were very focused on the road even when the traffic condition was not difficult because they wanted to respect as well as possible the driving instructions. For this reason, this factor is ignored in the following analyses.

Effects of the factors "subject" and "sonification conditions" (two-ways ANOVA)

The results show that the factor "subject" was significant (p<.05) for most of the DVs (detailed results are not reported here for concision). This is a sign of significant inter-individual differences in the driving performances and in the management of the interface. This result was expected and does not need particular comments. The interactions subject*sonification was almost never significant.

Table 1 presents the results of the two-way ANOVAs for the factor "sonification condition", for the two categories of DVs (Driving and Eye tracking). When significant, the sign of the effect (difference between the level with sound (s) and without sound (w-s)) is mentioned with the relation w-s>s or w-s<s).

*Table 1. Results of ANOVAs: significance of the "sonification condition" for the different dependent variables (without sound: w-s – with sound: s). F-test: * p<.05 - **p<.01 – n.s. : not significant (p>.05)*

	Dependent variable DV	*Task 1*	*Task 2*	*Task 3*	*Task 4*
Driving variables	Completion time	n.s.	n.s.	n.s.	n.s.
	Average lateral deviation	n.s.	n.s.	n.s.	n.s.
	SD of lateral deviation	** w-s > s	n.s.	n.s.	n.s.
	Number of steering wheel reversals	n.s.	n.s.	n.s.	n.s.
	Average speed	n.s.	n.s.	n.s.	n.s.
	SD of speed	n.s.	n.s.	n.s.	n.s.
	Average longitudinal acceleration	** w-s < s	n.s.	n.s.	n.s.
	SD of longitudinal acceleration	n.s.	n.s.	** w-s > s	** w-s > s
	Inter-vehicular time	n.s.	n.s.	n.s.	** w-s < s
Eye tracking variables	Glance duration on the interface screen	n.s.	** w-s > s	** w-s > s	n.s.
	Number of glances	n.s.	n.s.	n.s.	n.s.
	Glance rate	n.s.	** w-s > s	** w-s > s	** w-s < s
	Maximum time of glance duration	n.s.	** w-s > s	** w-s > s	n.s.
	Number of road-interface eye movements	n.s.	** w-s > s	** w-s > s	n.s.

The sonification barely influenced the lateral control of the vehicle. The only exception was a significant reduction of the SD of lateral position with auditory feedback, but only for task 1. On the other hand, longitudinal control was influenced by sound in different ways depending on the task: increased average acceleration in task 1, decreased SD of acceleration in task 3, decreased SD of acceleration and increased inter-vehicular time in task 4. Although this pattern of result is not entirely consistant, it suggests a moderate facilitation of vehicular control with sonification.

Table 1 shows that the presence of sounds has a significant impact on the eye-tracking variables, mainly for the longer tasks (task 2 and 3: manipulation of GPS in a curve and in a straight line). In these cases, the presence of sounds significantly decreased the glance duration, the glance rate, the maximum time of glance duration and the number of eye movements toward the interface. The subjects spent less time looking at the interface with sounds than without, increasing driving safety. This effect was not observed for the shorter tasks (tasks 1 and 4 – reaction to a pop-up message). For task 4 (pop-up in a curve), the glance rate on the interface was even higher with sounds than without. For these two short tasks, to study the reactions of the driver after the task completion (not considered in the previous DVs), we looked at the number of visual controls of the driver on the interface after the task completion (controls to verify that the pop-up actually disappeared). The proportion of visual controls for each task and each condition is given in table 2.

Table 2. Proportion of visual controls on the interface after completion of tasks 1 and 4 in the two sonification conditions (without sound: w-s – with sound: s). Significance test of the difference (unilateral test of two proportions z-test - p<.05 - **p<.01 – n.s. : not significant (p>.05))*

	Task 1		*Task 4*	
	w-s	s	w-s	s
Proportion of visual control after the completion of the task	21/96	15/96	17/96	5/96
p-value (unilateral z-test of two proportions)	n.s. (p>.17)		** (p<.01)	

The results show that for task 4, the presence of sounds significantly decreased the proportion of visual controls after the task completion, and hence increases the driving safety.

Questionnaire

The questionnaire reports the subjects' subjective assessments, stated just after the experiment. The analysis of the response can be used to give indications to better understand the causes of the significant effects of the sonification condition on the DVs, based on the feelings of the drivers. For the questions about feeling of safety and feeling of not looking at the road, the responses rates are given in figure 6 and 7.

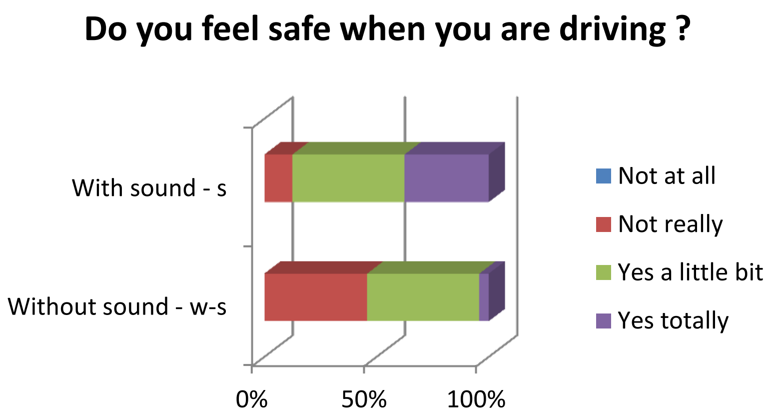

Figure 6. Proportion or responses to the question about safety

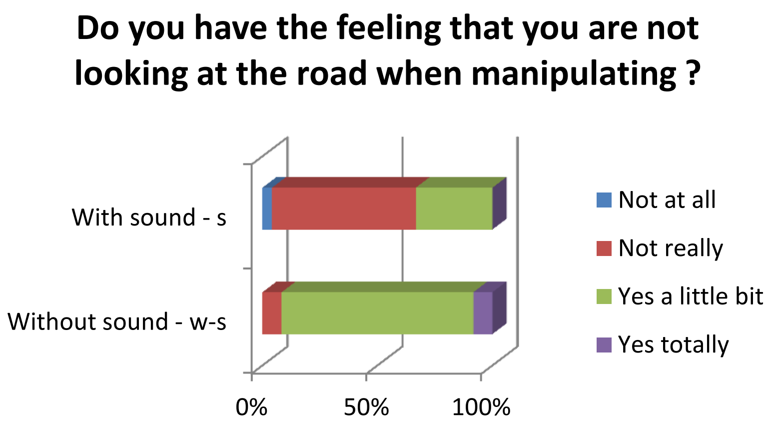

Figure 7. Proportion or responses to the question about "looking at the road"

The results show that the proportions of responses between the two conditions are significantly differents for the two questions (multinomial goodness of fit test, $p<.01$). With the presence of sounds, the subjects judged the interface as easier to use and had the feeling of looking more at the road.

Unsurprisingly, 100% of participants preferred to use the interface with sound than without sound. This supports the idea that the sound is a useful parameter for the manipulation of the interface. Finally, concerning the quality of sound design, the participants considered that the sounds were not troublesome and 91% appreciated the sound.

Discussion

Auditory feedback for touchless interaction marginally influenced driving behaviour, mainly in the longitudinal control of the vehicle. One reason may be that the subjects were very much engaged in the driving task during the whole experiment, with or without sounds. Inter-individual differences between subjects were large, possibly due to differences in driving style and experience. The influence of the sound in that context was relatively weak. Furthermore, for the short popups tasks (tasks 1 and 4), it was difficult to observe possible changes in behaviour in the 1 or 2 seconds the tasks required to be completed. In particular, before the validation, the sound did not bring any information. It was only useful for the confirmation of the task completion.

On the other hand, the interest of auditory feedback for the manipulation of the touchless interface was clearly evidenced by the gaze analysis. Indeed, the sound influenced gaze behaviour in a very significant way. With the sound, the participants looked much less at the interface when navigating in the menus or to verify that a call/message was refused or accepted. Furthermore, the lower number of eye movements between the interface and the road and the shorter glance duration in the “sound” condition shows that the users are feeling much less confident in their navigation. They could consequently pay more attention to the road and the driving task.

The results of the questionnaire confirm this analysis: with sounds, users are feeling safer, thanks to a sound design that meets the requirements (easily interpretable by user, bringing information needed to the navigation and relieving the visual workload).

Conclusion

The study showed that the presence of auditory feedbacks for the manipulation of a touchless interface of an infotainment system in a car significantly decreased the eyes-on-the-interface time. The sonification provides useful information on the manipulation of the interface, information that can only be obtained through vision in the absence of sound. It may increase safety because the driver can be more attentive to the road. The interface seems also easier to use with sounds, and the sound allows a more user-friendly experience.

References

Denjean, S., Roussarie, V., Ystad, S., & Kronland-Martinet, R. (2013). An innovative method for the sonification of quiet cars. *The Journal of the Acoustical Society of America*, *134*, 3979.

Hermann T., Hunt A., & Neuhoff J.G. (2011). *The sonification Handbook*. Logos Publishing House, Berlin, Germany. ISBN 978-3-8325-2819-5.

Jakus, G., Dicke, C., & Sodnik, J. (2015). A user study of auditory, head-up and multi-modal displays in vehicles. *Applied Ergonomics*, *46*, 184-192.

Kajastila, R., & Lokki, T. (2013). Eyes-free interaction with free-hand gestures and auditory menus. *International Journal of Human-Computer Studies*, *71*, 627-640.

Tardieu, J., Misdariis, N., Langlois, S., Gaillard, P., & Lemercier, C. (2015). Sonification of in-vehicle interface reduces gaze movements under dual-task condition. *Applied Ergonomics*, *50*, 41-49.

Næs, T., Brockhoff, P. B., & Tomic, O. (2010). *Statistics for Sensory and Consumer Science*. John Wiley & Sons Ltd. ISBN: 978-0-470-51821-2.

Sodnik, J., Dicke, C., Tomazic, S., & Billinghurst, M. (2008). A user study of auditory versus visual interfaces for use while driving. *International Journal of Human-Computer Studies*, *66*, 318-332.

Wilcox, R.R. (2014). *Winsorized Robust Measures*. Wiley StatsRef: Statistics Reference Online.