



# Sonification of in-vehicle interface reduces gaze movements under dual-task condition



Julien Tardieu <sup>a,\*</sup>, Nicolas Misdariis <sup>b</sup>, Sabine Langlois <sup>c</sup>, Pascal Gaillard <sup>d</sup>,  
Céline Lemerrier <sup>d</sup>

<sup>a</sup> MSHS-T USR3414, University of Toulouse and CNRS, Toulouse, France

<sup>b</sup> STMS Ircam-CNRS-UPMC, 1, place Igor Stravinsky, 75004 Paris, France

<sup>c</sup> Renault – Cognitive Ergonomics & HMI, 1 avenue du Golf, 78084 Guyancourt, France

<sup>d</sup> CLLE UMR5263, University of Toulouse and CNRS, Toulouse, France

## ARTICLE INFO

### Article history:

Received 21 February 2014

Accepted 18 February 2015

Available online 12 March 2015

### Keywords:

Sonification

Dual-task

Eye-tracking

## ABSTRACT

In-car infotainment systems (ICIS) often degrade driving performances since they divert the driver's gaze from the driving scene. Sonification of hierarchical menus (such as those found in most ICIS) is examined in this paper as one possible solution to reduce gaze movements towards the visual display. In a dual-task experiment in the laboratory, 46 participants were requested to prioritize a primary task (a continuous target detection task) and to simultaneously navigate in a realistic mock-up of an ICIS, either sonified or not. Results indicated that sonification significantly increased the time spent looking at the primary task, and significantly decreased the number and the duration of gaze saccades towards the ICIS. In other words, the sonified ICIS could be used nearly exclusively by ear. On the other hand, the reaction times in the primary task were increased in both silent and sonified conditions. This study suggests that sonification of secondary tasks while driving could improve the driver's visual attention of the driving scene.

© 2015 Elsevier Ltd and The Ergonomics Society. All rights reserved.

## 1. Introduction

According to recent surveys (<http://www.drive-safely.net/top-driving-distractions.html>), the use of external devices such as a telephone, a music player, or the radio is responsible for most driving distractions. Sending a text message or having a phone conversation while driving is undoubtedly unsafe and can have fatal consequences (Victoria et al., 2013). In this paper we focus on the use of in-car infotainment systems (ICIS), i.e., a device installed in the car to provide services such as navigation systems, music player or telephone. To minimize its impact on driver's visual attention, the ICIS is designed to have a visual display located in the middle of the dashboard, and a control device close to the gear stick. In the US, the National Highway Traffic Safety Administration (NHTSA) has recently published a series of guidelines for designing in-car devices with a limited impact on driver's visual attention (NHTSA, 2013). Nevertheless, the use of such technologies in cars is still responsible for frequent off-road glances: Sodhi et al. (2002)

found that tuning the radio while driving resulted in 42% of off-road glances (total task time of 21.1 s on average), Young et al. (2012) showed that using a portable music player not only resulted in increased off-road glances but also reduced a driver's ability to maintain a constant lane position. Kujala et al. (2013) found that most tasks performed on a touch screen while driving increased the number of rapid steering wheel movements. New solutions to further reduce even more the number and the duration of off-road glances are required.

Car driving is primarily a visual task (Sivak, 1996), whereas using an ICIS can rely on visual, auditory or both modalities. Visual-only ICIS has been largely addressed in the field of HMI ergonomics that provided design guidelines for improved information presentation (Singleton, 1971; Bastien and Scapin, 1992; Scapin and Bastien, 1997; Ziefle, 2010). Recently, Mitsopoulos-Rubens et al. (2011) demonstrated that list scrolling while driving significantly impaired driving performance (mean lane deviation and percentage of correct lane changes), but no gaze data were provided. Using an eye-tracker, Rydström et al. (2012) showed that both a touch screen and a rotary knob affected lateral control performance while doing alphanumeric input or list scrolling. Based on the occlusion

\* Corresponding author.

E-mail address: [julien.tardieu@univ-tlse2.fr](mailto:julien.tardieu@univ-tlse2.fr) (J. Tardieu).

technique, Baumann et al. (2004) indicated that a task performed with in-car navigation system must be portionable in 1–2 s chunks and the total task time is insufficient to evaluate the impact on driving performance. In other words, even if an ICIS has been optimized to be less visually demanding, it is still responsible for off-road glances and thus could impair the driving performances.

Sonification, as a way to display information using the only auditory modality, appears to be a strong candidate to reduce eye movements while driving. Introduced by Kramer (1992), sonification is based on at least three fundamental concepts: (1) *earcon*,<sup>1</sup> i.e., an abstract sound for which the sound/meaning relationship is arbitrary and must therefore be learned by the user (Blattner et al., 1989). Earcons are usually based on acoustic and musical characteristics: pitch, rhythm, duration, etc. (2) *auditory icon* (Gaver, 1986), i.e., a sound which establishes a direct link with the object or concept that it represents, by referring to an easily recognizable sound from our daily environment, and is accordingly almost directly comprehensible by the user. This approach is preferred in the context of driving since it reduces the time needed by the drivers to learn the sounds. (3) *spearcon* (speech-earcon), i.e., accelerated text to speech synthesis employed to facilitate fast scrolling in long menus (Walker et al., 2006, 2013).

An ICIS is generally comprised of a small display presenting different menus and sub-menus organized in a large hierarchical structure. Sonification of hierarchical menus must address two issues: (1) sound representation of the hierarchical position of each item and (2) sound representation of the semantic content of each item. Earcon based sonifications are easily learned but not very well adapted to complex hierarchies (Brewster et al., 1995, 1998; Brewster, 1998). Leplâtre and Brewster (2000) and Leplâtre (2002) developed a new implementation of hierarchical earcons more closely linked to the hierarchical position of the items but too confusing in terms of semantic representation. On the other hand and as expected, auditory icons provide less arbitrary sonification as illustrated in Barrass (1998) and Conversy (1998). Auditory icons were also used in Gaver's Sonic Finder (Gaver, 1986), a computer auditory interface in which actions were mapped to everyday sound events (e.g., selecting a file mapped to the sound of an object being hit), added to another layer of information (e.g., file size with object size). In a previous study, we developed an original approach based on a combination of earcons and auditory icons (Langlois et al., 2010; Misdariis et al., 2011) and preference tests were performed in a driving simulator. The result was an improvement of the model achieved by adding synthesized speech at lower levels of the menu. A detailed presentation of this sonification is given in Section 2.5. The main limitation, the absence of gaze data, is addressed in the present paper.

Using a dual task paradigm, Jeon et al. (2009) measured the benefits of sonification on the performance to a visual primary task (a ball catching game) while navigating in a list of items. Results showed that sonification significantly improved both reaction times in the primary task and search times in the secondary task. However, no gaze data were collected to confirm that sonification reduced the visual demands of the secondary task. Sodnik et al. (2008) studied the navigation inside mobile phone menus (sonified or not) while driving in a simulator. The sonification was based on earcons mixed with synthesized speech and virtually placed in the space using a surround sound system. Results showed that sonification decreased the unsafe driving behaviors (e.g. unexpected deceleration) but did not improve menu navigation.

This review of related works reveals a lack of gaze data supporting the claim that sonification can efficiently reduce off-road glances while navigating in the menus of an ICIS. Our study addressed this issue with an experiment in which participants' gaze was measured while they performed two simultaneous tasks: (1) a visual primary task on a computer screen that simulates the sustained visual attention required when driving, and (2) a secondary navigation task in which participants navigated inside the menus of a realistic ICIS mock up (either sonified or not). The remainder of this article describes the experiment in Section 2, an analysis of the results in Section 3 and a discussion in Section 4.

## 2. Method

### 2.1. Participants

46 Participants (39 female, mean age 25.3, SD = 5.2) were recruited and paid for this study and gave consent prior to the experiment. Participants reported to have normal or corrected-to-normal vision and normal hearing. All the participants were right handed and had their driving license. Half of the participants performed the experiment with the silent interface (no sonification) and the other half with the sonified interface.

### 2.2. Apparatus

Fig. 1 presents the experimental apparatus. The participant was seated in front of the 17" monitor of a Tobii T120 eye-tracker that presented both the primary and the secondary tasks. All sounds were listened through a pair of Sennheiser HD380 Pro headphones. For optimal gaze tracking, participants were requested to keep their head positioned in front of the screen center at an approximate distance of 65 cm.

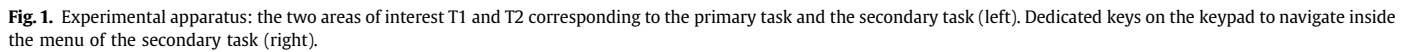
The two tasks were implemented on a desktop PC and displayed on a 34 cm × 27 cm monitor with a resolution of 1280 × 1024 pixels and a refresh rate of 60 Hz. According to the Renault engineering department, the dimensions of both tasks (see Fig. 1) proportionally reproduced what is usually found in a car:

- The angles between the driving scene and the visual display of an ICIS located in the middle of the dashboard of a small car (21° horizontal and a 10° vertical);
- The angular size of an ICIS with a 7" visual display;
- The angular size of the area in which drivers' gaze is mostly located (+6° horizontally and −6° vertically).

### 2.3. Primary task: target detection

Car driving is a very complex task in terms of visual attention, working memory, as well as physical activity. According to Waard (1996), most of the driving-related tasks are automated and the required attention is mainly based on visual resources. In addition, one common driving context is the situation that requires sustained attention to events occurring in the central vision area and for which immediate reaction is needed (e.g. reacting to other cars or traffic lights). This driving context has been employed in previous studies to measure the impact of using mobile phones while driving (Lamble et al., 1999; Strayer et al., 2003). As mentioned by Harvey et al. (2011), "A driver's ability to detect and respond to events and hazards in the driving environment can be used as a measure of the interference from secondary tasks". The primary task developed in this paper aimed at mimicking the sustained visual attention to sudden events that is required while driving.

<sup>1</sup> The name "earcon" comes from the term icon as an icon for the ear (Sumikawa, 1985).



The background color of both the monitor and the primary task window was set to a gray with a luminance of 160 cd/m<sup>2</sup>. This task was programmed using Processing,<sup>2</sup> also used to collect the gaze data<sup>3</sup> and the secondary task data via UDP.<sup>4</sup> The gaze data was sampled at a rate of 50 Hz.

The secondary task was a navigation inside a mock up of a real ICIS. The structure of the menu, as well as the visual elements was identical to the menus available in the ICIS (see [Figs. 3 and 4](#)).

- Primary task: reaction time to detect the visual target;
- Secondary task:
  - Time to reach the requested item by the participant;
  - Number of browsed items before the participant found the requested item;
- Gaze data:

<sup>5</sup> <http://cycling74.com>.

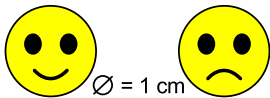


Fig. 2. Target (left) and distractor (right) used in the primary task. The diameter of both the distractor and the target was 1 cm.

- Time spent in each area of the two areas of interest T1 and T2 (see Fig. 1), respectively the primary task window and the secondary task window;
- Number of gaze saccades, i.e., number of times the eyes moved to the secondary task window.

2.7. Procedure

2.7.1. Synopsis

This was a dual task experiment in which the participant was first asked to do the primary task alone (target detection, see

Table 1  
Sonification principles used to sonify the ICIS mock-up.

Level	Type	Navigation	Sonification
1	Icon based menu	Up/down Left/right	Non-speech sounds that are an extension of the concept of auditory icons: they are created from a mix of several sounds that are all semantically linked with the menu item.
2 + "Points of interests"	Lists	Up/down	Mix of the text-to-speech synthesis of the item and a shorter version of the auditory icon used for the parent menu: this new auditory icon both indicates the parent menu the item belongs to (with timbre similarity) and the deeper level of the item (with shorter duration). In addition, the list scrolling is represented by a pitch progression of the auditory icon. The same idea is used in the "Points of interests" sub-menu.
3+	Lists	Up/down	Text-to-speech mixed with a "click" sound that indicates scrolling.



Fig. 3. Four examples of the visual display used in the secondary task: Home screen (upper left), Music main menu (upper right), Navigation menu (lower left), New Destination Menu (lower right).

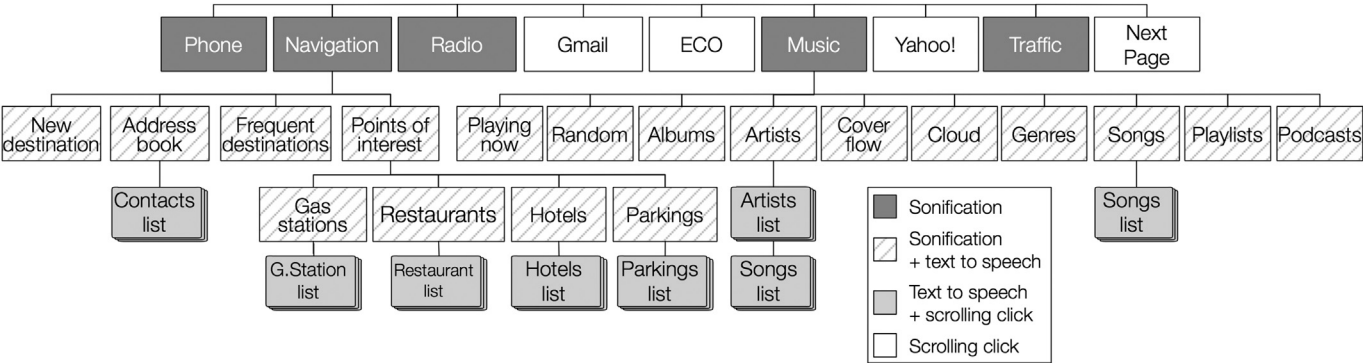


Fig. 4. Structure of the hierarchical menu and principles used for the sonification of each item. The menu is limited to Music and Navigation.



**Table 2**

Menu items that participants were asked to search in the menu of the ICIS mock up.

Item	Nb of browsed items	Length
1	9	Short
2	11	Short
3	12	Short
4	21	Medium
5	42	Medium
6	52	Medium
7	61	Long
8	94	Long
9	102	Long

Section 2.3), and after 1 min to find an item in the ICIS (secondary task, see Section 2.4) while doing the primary task simultaneously (see Fig. 5). The participant was instructed to give maximum priority to the primary task (complete instructions provided in Appendix). The experiment was made of three blocks of four trials: three items to search in the ICIS and one city name to type. The three blocks were presented in a balanced random order between participants.

**2.7.1.1. Menu items.** The items to be searched by the participants were chosen to get different path lengths inside the structure of the hierarchical menu, i.e., the minimum number of browsed items before arriving at the searched item. The nine chosen items were grouped in short, medium or long categories in accordance with their path length (see Table 2).

**2.7.1.2. City names to type.** In each block, participants were also asked to type a city name in the New Destination sub-menu. The 3 cities (Lyon, Genève, Paris) were different in path length: 22, 31 and 39 keystrokes, respectively.

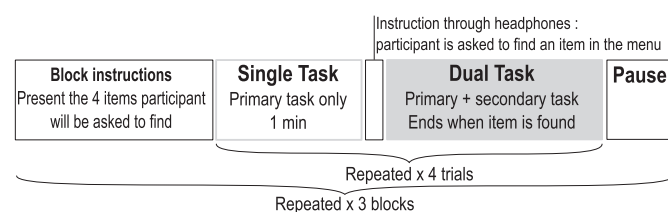
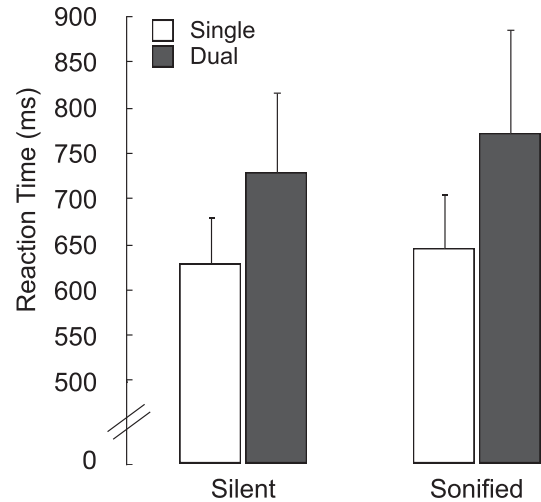
### 2.7.2. Training

The training phase was performed in two parts: (1) a complete presentation of the ICIS was given to the participant, followed by a few minutes of free use of the ICIS mock up (sonified or not, depending on the group), and (2) the participant ran one training block, with the same structure as an experimental block but with different items. This training block was repeated until the participant was comfortable with the dual task condition and could use the keyboard without looking at his/her fingers (controlled in real time with the eye-tracker). This training phase, added to the setting of the eye-tracker lasted between 15 and 20 min. The total duration of the experiment was approximately 1 h.

## 3. Results

### 3.1. Reaction times in the primary task

The primary task data showed that participants made very few errors of detection in both silent and sonified conditions. No analysis of errors could be performed. The reaction times as a

**Fig. 5.** Experiment synopsis with three blocks of four trials.**Fig. 6.** Mean reaction time (ms) to the target detection in the primary task as a function of task and sound conditions (error bars represent  $\pm$ SEM).

function of the task condition (single versus dual task) are provided in Fig. 6. Data show an increase of the reaction time when the participant is in dual task condition compared to the single task condition. Moreover, this increase is similar in both sound conditions: 629–729 ms in silent condition, 646–772 ms in sonified condition. The data were analyzed using a mixed linear model<sup>6</sup> and demonstrated a significant effect of the dual versus single task ( $F(1,8290) = 1101.32$ ,  $p < 0.001$ ), but no significant effect of the sound condition ( $F(1,44) = 3.79$ ,  $p = 0.0578$ ). In addition, no interaction between sound condition and task was revealed in this analysis:  $F(1,8290) = 9.551$ ,  $p = 0.002$ .

### 3.2. Secondary task performance

#### 3.2.1. Time to target

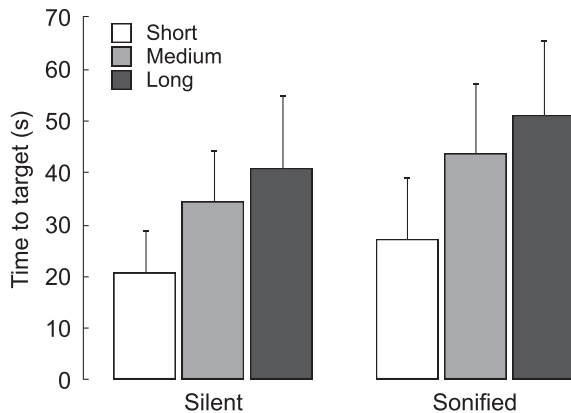
The time to target is the duration from the onset of the first navigation keystroke once the instructions have been played to the time when the participant successfully selected the requested item. As expected, Fig. 7 shows an increase of the time to target with item length in the silent condition. The same result is observed in the sonified condition. The data were analyzed using a mixed linear model and showed a significant effect of the item length,  $F(2,436) = 49.41$ ,  $p < 0.001$ . The sound condition is not significant  $F(1,38) = 3.70$  ( $p = 0.06$ ), and no interaction is apparent ( $F(2,436) = 0.96$ ,  $p = 0.38$ ).

#### 3.2.2. Path length

Path length is calculated as the number of items that have been browsed by the participant before he/she reached the requested item. It is comparative with the number of keystrokes used in previous works to measure the navigation efficiency (e.g., Jeon et al., 2009). Results are provided in Fig. 8. Just as seen with the results observed for the time to target, the path length increases with item length in both the silent and the sonified conditions. For example in silent condition, an average of 95 items were browsed to reach long items versus 14 in average for short items. The data were analyzed using a generalized linear model<sup>7</sup> since path length is not a continuous variable. The analysis revealed a significant effect of the item length,  $p < 0.001$ , as well as the sound condition,

<sup>6</sup> Using the lme function of nlme package in R version 3.1.1.

<sup>7</sup> Using the glm function of nlme package in R version 3.1.1.



**Fig. 7.** Mean time to target in seconds as a function of sound condition and item length ( $\pm$ SEM error bars).

$p < 0.001$ . The interaction between item length and sound condition is also significant,  $p < 0.001$ .

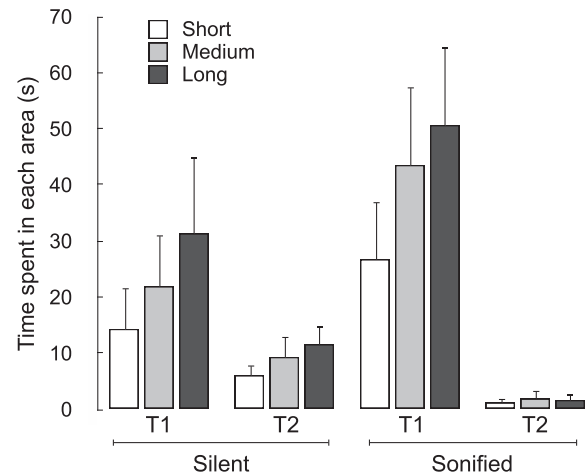
### 3.3. Gaze movements

Due to tracking problems that occurred with 6 participants, the analysis of gaze data was performed on 40 participants, 20 participants in each sound condition (silent and sonified).

#### 3.3.1. Areas of interest

The two areas of interest analyzed in this section correspond to the window of the primary task (T1) and the window the secondary task (T2). Fig. 9 shows the time spent in each area as a function of item length and sound condition. These results demonstrate that for short, medium and long items, the time spent in the primary task window (T1) increases with the sonification as the time spent in the secondary task window decreases. The proportion between the time spent in T1 and the time spent in T2 has been averaged across item length. In the silent condition, participants spent 28% of the time in T2 and 72% in T1, whereas in the silent condition they spent 4% of the time in T2 and 96% in T1.

The data were analyzed using a mixed linear model and showed significant effect of the sound condition on the time spent in T1 ( $F(1,38) = 12.77$ ,  $p < 0.001$ ) and the time spent in T2 ( $F(1,38) = 78.75$ ,  $p < 0.001$ ). The item length had also a significant effect on the time spent in T1 ( $(2,910) = 39.51$ ,  $p < 0.001$ ) and T2



**Fig. 9.** Mean time spent by participants' gaze in each area of interest as a function of sound condition and item length ( $\pm$ SEM error bars).

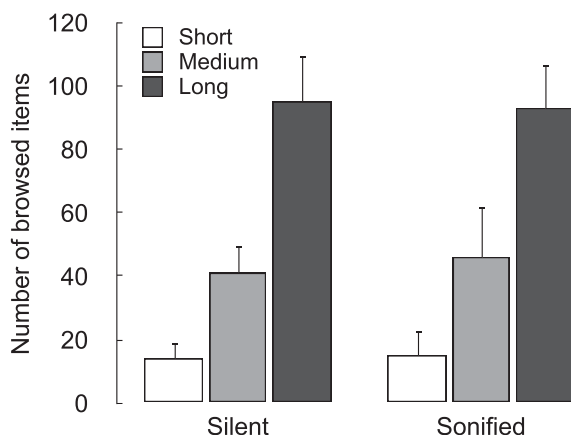
( $F(2,910) = 14.35$ ,  $p < 0.001$ ). No interaction between sound condition and item length was revealed in the analysis.

#### 3.3.2. Gaze saccades

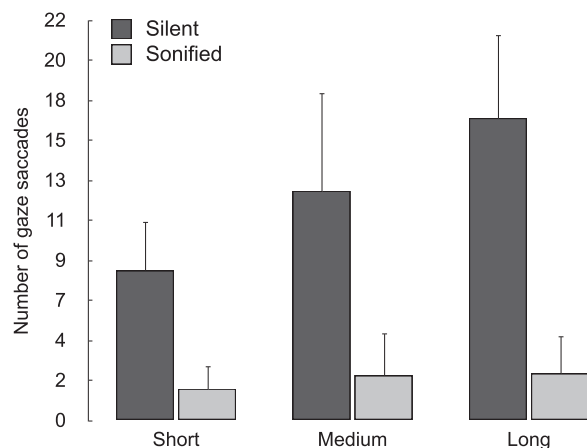
The number of gaze saccades is calculated as the number of times participant's gaze moved to the secondary task window during the dual task condition. Fig. 10 shows the number of gaze saccades as a function of the item length and the sound condition. As expected from the previous result, the number of gaze saccades decreases with sonification for all item lengths. As for the path length, the saccades data were analyzed using a generalized linear model. The analysis showed a significant effect of sound condition ( $p < 0.001$ ), as well as item length ( $p < 0.001$ ), but no interaction ( $p = 0.007$ ).

## 4. Discussion

This paper presents a laboratory study that examined the gaze movements of 40 participants requested to perform navigation tasks in a hierarchical menu (secondary task), while prioritizing their attention on a visual detection task (primary task). This primary task was designed to produce a sustained visual attention to sudden visual events, which is one component of the visual



**Fig. 8.** Mean path length as the number of browsed item before reaching the requested item ( $\pm$ SEM error bars).



**Fig. 10.** Mean number of gaze saccades, i.e., number of times the eyes moved to the secondary task window ( $\pm$ SEM error bars).

attention required when driving. The secondary task was a realistic mock-up of an ICIS available in Renault cars. Two conditions were compared: the ICIS was either silent (with only a visual interface) or sonified (with a visual and auditory interface) according to principles developed in a previous study (Langlois et al., 2010; Misdariis et al., 2011).

The gaze data revealed a very significant impact of the sonification on the time spent in either the primary task or the secondary task. Indeed, in dual task condition, the proportion of time spent looking at the primary task increased from 72% to 96% between the silent and the sonified conditions. Moreover, the number of gaze saccades was also significantly reduced: for example with long items, the analysis revealed 16.6 saccades on average without sonification versus 2.6 saccades on average with sonification. According to recommendations made by Zwahlen et al. (1988), there must be less than four saccades of less than 1.2 s to ensure a sufficient level of safety. Our findings show that the ICIS does not satisfy this requirement when the interface is not sonified, whereas it is fully acceptable when the interface is sonified. In the silent condition, our results agree with Sodhi et al. (2002) in which participants took 21.1 s in average to perform a radio tuning task while driving, with 9 s of off road glance: in our data, medium items were reached in 21.9 s with 9.3 s of off road glance. In other words, we showed that sonification helped participants to keep their eyes focused on the primary task while doing the menu navigation almost exclusively by ear. This result supports the findings of previous works on sonification in dual task situations (Jeon et al., 2009; Sodnik et al., 2008) in which participants were able to perform both tasks efficiently, but nothing was reported on participants' visual behavior. With the support of gaze data, our study demonstrated that sonification helped participants to keep their eyes on the primary task most of the time. Sonification also made the ICIS fully acceptable regarding safety recommendations. Further dual task experiments are required to adapt these results in a more ecologically valid driving environment such as driving simulator or real driving.

On the other hand, both time to target and number of browsed items in the secondary task were not significantly improved by the sonification (see Figs. 7 and 8). This is coherent with our previous study in which the navigation duration was increased with the sonification (Langlois et al., 2010; Misdariis et al., 2011). This is also consistent with findings of Lepître and Brewster (2000) in which navigation performances in sonified mobile phone menus were not significantly different from the navigation without sonification. However, this contrasts the benefit of sonification for menu navigation revealed in many works as noted in the introduction. This can be explained by the fact that most of these works did not use a dual-task paradigm, so our results suggest that sonification benefits depend on the presence of another task performed simultaneously. In Jeon et al. (2009) though, sonification improved participants' performance to menu navigation while doing a primary task. The work of Lee and Spence (2008) also revealed sonification benefits when using a phone while driving. However, the instructions given to the participant prior to the experiment were different from the ones we gave in the present work: in both Jeon et al. (2009) and Lee and Spence (2008), the participants were instructed to perform both tasks as fast as possible, whereas we explicitly asked the participants to perform the secondary task while remaining as fast as possible to the primary task. This suggests that a dual task experiment in a driving simulator should consider what task must be prioritized: the driving or the secondary task.

As regards the reaction times to the primary task, results indicated that in single task condition participants took an average of 629 ms to detect the target (see Fig. 6), versus an average of 729 ms in dual task condition when the interface was not sonified. The

629 ms baseline reaction time is consistent with a previous study by Thorpe et al. (1996) who observed mean reaction times between 385 ms and 567 ms in a go/no-go detection task. The slower reaction times we observed could be a result of the random positions of the targets and distractors. In addition, the increased reaction time induced by the secondary task is also similar to previous findings on drivers' attention (Lamble et al., 1999; Strayer et al., 2003; Jeon et al., 2009). According to gaze data (Figs. 9 and 10), in silent condition this increase is due to the time spent looking at the ICIS. However, in the sonified condition the reaction times were also increased (from 646 ms to 772 ms on average, see Fig. 6), but the participants used the ICIS almost exclusively by ear as evidenced by the gaze data. In other words, even if participants were staring at the primary task, reaction times in the primary task were still impaired due to the secondary task that was almost exclusively auditory. This finding is discussed below.

On one hand, this finding contrasts with two of the aforementioned research (namely Jeon et al., 2009; Sodnik et al., 2008) in which the performance to the primary task was improved when the secondary task was sonified. However, our study differs from those in several crucial ways. In Sodnik et al. (2008), the driving performance was measured as the number of unsafe driving behaviors. This differs from our study in which performance was measured with reaction times to sudden events. In Jeon et al. (2009), the primary task (a ball-catching game) did not imply sustained attention as in our primary task since it was possible to anticipate the act of catching. Furthermore, they used a secondary task (finding an item in a long list) that was less complex and less realistic than the interface navigation task employed in this research. As mentioned previously, these findings indicate the importance of future work in a more realistic driving context to more accurately evaluate the benefits of a sonified secondary task depending on the complexity of the driving situation and the secondary task.

On the other hand, this finding is consistent with previous studies in which a focused attention on a phone conversation impaired drivers' performances (Strayer et al., 2003; Strayer and Johnston, 2001). In Strayer et al. (2003), gaze data illustrated that phone conversation reduced drivers' attention to visual signs even if they were visually fixated. This finding is likely related with the inattention blindness (Neisser and Becklen, 1975; Mack and Rock, 1998; Simons and Chabris, 1999; Most et al., 2005), referring to failures to detect obvious elements of the visual scene due to attentional engagement in another visual task. This phenomenon also exists in the auditory modality as revealed by Dalton and Fraenkel (2012). Our findings suggest that this phenomenon could be extended to a cross modal effect, i.e., a failure to detect unexpected visual stimuli when engaged in an auditory task, even though our study showed a decreased speed in detecting rather than a failure. This could be further investigated by increasing the difficulty in the primary task to observe more detection failures. Future research could also include a real time measurement of workload (using pupil dilatation for example) that could reveal significant differences as suggested in Jeon et al. (2009) as well as Langlois et al. (2010) and Misdariis et al. (2011) in which the driving comfort perceived by the participants in a driving simulator was increased with sonification.

## 5. Conclusion

This work examined the use of sonification as a way to reduce gaze movements towards the visual display of an in-car infotainment system (ICIS) in order to ensure a greater level of safety. A laboratory experiment was performed in which participants were requested to prioritize a primary task while performing a secondary task simultaneously. The primary task was a visual target

detection task on the computer involving a sustained visual attention to sudden events, and the secondary task was a mock up of an ICIS which was either sonified or not. The sonification used a combination of earcons, auditory icons and text to speech. Eye-tracking measurements demonstrated that participants used the sonified interface almost exclusively by ear, while staring at the prioritized primary task (96% of the time on average). Comparatively, when the interface was not sonified, participants spent 72% of the time on average looking at the primary task. On the other hand, reaction times in the primary task as well as menu navigation in the secondary task were not significantly impacted by the sonification. This study has implications in the design of auditory interfaces for dual-task contexts, and suggests the requirements of further investigations to evaluate how these results are contrasted in more realistic driving contexts such as a driving simulator or real driving.

## Acknowledgments

This study was part of the “Findability” research project supported by Renault at Cognitive Ergonomics & HMI department. The authors wish to thank Matt Coler (INCAS<sup>3</sup>) for his editorial assistance, as well as two anonymous reviewers for their constructive comments on earlier versions of this article.

## Appendices

### Instructions

The quoted text below is an example of the instructions given to the participants before each block, it was presented visually on the screen, the participant had to hit a key to start the block. The text was presented in French and is translated to English by the author.

“Your task is to keep your attention on the primary task and to hit the space bar as soon as you see the [picture of the smiley (Fig. 2 left)]. Try to be as fast as you can.

After a certain amount of time, a spoken message will indicate the menu item you'll have to find. Go to the requested item while remaining as good as possible in the primary task.

Below are the four items that you'll have to find, in random order (click an item to listen to the corresponding spoken message):

- Go and find Aurélie Savary in the Address Book
  - Go and find the restaurant Georges
  - Go and find the Artist Vanessa Paradis and then the song Divine Idylle
  - Type in the city LYON in the menu New Destination
- Hit any key to start.”

## References

- Barrass, S., 1998. Auditory Information Design (PhD thesis). The Australian National University.
- Bastien, J.M.C., Scapin, D.L., 1992. A validation of ergonomic criteria for the evaluation of human–computer interfaces. *Int. J. Hum. Comput. Interact.* <http://dx.doi.org/10.1080/10447319209526035>.
- Baumann, M., Keinath, A., Krems, J.F., Bengler, K., 2004. Evaluation of in-vehicle HMI using occlusion techniques: experimental results and practical implications. *Appl. Ergon.* 35 (3), 197–205. <http://dx.doi.org/10.1016/j.apergo.2003.11.011>.
- Blattner, M., Sumikawa, D., Greenberg, R., 1989. Earcons and icons: their structure and common design principles. *Hum. Comput. Interact.* [http://dx.doi.org/10.1207/s15327051hci0401\\_1](http://dx.doi.org/10.1207/s15327051hci0401_1).
- Brewster, S.A., 1998. Using nonspeech sounds to provide navigation cues. *ACM Trans. Comput.-Hum. Interact.* 5 (3), 224–259. <http://dx.doi.org/10.1145/292834.292839>.
- Brewster, S.A., Wright, P.C., Edwards, 1995. Experimentally derived guidelines for the creation of earcons. In: *Adjunct Proceedings of HCI'95*, Huddersfield, UK.
- Brewster, S.A., Capriotti, A., Hall, C.V., 1998. Using compound earcons to represent hierarchies. *HCI Lett.* 1 (1), 6–8.
- Conversy, S., 1998. Ad-hoc Synthesis of Auditory Icons. *Proceedings of the 5th International Conference on Auditory Display*, University of Glasgow, U.K.
- Dalton, P., Fraenkel, N., 2012. Gorillas we have missed: sustained inattention deafness for dynamic events. *Cognition* 124 (3), 367–372. <http://dx.doi.org/10.1016/j.cognition.2012.05.012>.
- Gaver, W., 1986. Auditory icons: using sound in computer interfaces. *Hum. Comput. Interact.* [http://dx.doi.org/10.1207/s15327051hci0202\\_3](http://dx.doi.org/10.1207/s15327051hci0202_3).
- Harvey, C., Stanton, N.A., Pickering, C.A., McDonald, M., Zheng, P., 2011. A usability evaluation toolkit for in-vehicle information systems (IVISs). *Appl. Ergon.* 42 (4), 563–574. <http://dx.doi.org/10.1016/j.apergo.2010.09.013>.
- Jeon, M., Davison, B.K., Nees, M.A., Wilson, J., Walker, B.N., Street, C., 2009. Enhanced auditory menu cues improve dual task performance and are preferred with in-vehicle technologies. In: Schmidt, A., Dey, A., Seder, T., Juhlin, O., Kern, D. (Eds.), *Billboard*. ACM, pp. 91–98. <http://dx.doi.org/10.1145/1620509.1620528>.
- Kramer, G., 1992. In: Kramer, G. (Ed.), *Some Organizing Principles for Representing Data with Sound*. Addison Wesley Longman.
- Kujala, T., Silvennoinen, J., Lasch, A., 2013. Visual-manual in-car tasks decomposed: text entry and kinetic scrolling as the main sources of visual distraction. In: *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*. ACM, New York, NY, USA, pp. 82–89. <http://dx.doi.org/10.1145/2516540.2516562>.
- Lamble, D., Kauranen, T., Laakso, M., Summala, H., 1999. Cognitive load and detection thresholds in car following situations: safety implications for using mobile (cellular) telephones while driving. *Accid. Anal. Prev.* 31 (6), 617–623.
- Langlois, S., Loiseau, S., Tardieu, J., Cera, A., Misdariis, N., 2010. Evaluation de la sonification d'un système multimédia automobile. In: *Conference Internationale Francophone Sur l'Interaction Homme-Machine*. ACM, Luxembourg, Luxembourg, pp. 173–180. <http://dx.doi.org/10.1145/1941007.1941039>.
- Lee, J.-H., Spence, C., 2008. Assessing the benefits of multimodal feedback on dual-task performance under demanding conditions. In: *Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction*, vol. 1. British Computer Society, Swinton, UK, pp. 185–192.
- Leplatre, Grégory, 2002. The Design and Evaluation of Non-Speech Sounds to Support Navigation in Restricted Display Devices (PhD thesis).
- Leplatre, G., Brewster, S., 2000. Designing Non-speech Sounds to Support Navigation in Mobile Phone Menus. *Proceedings of the 6th International Conference on Auditory Display*, Atlanta, GA, USA.
- Mack, A., Rock, I., 1998. *Inattention Blindness*. The MIT Press, Cambridge, MA, US, p. xiv, 273.
- Misdariis, N., Tardieu, J., Langlois, S., Loiseau, S., 2011. Menu sonification in an automotive media center. In: *Human-Computer Interactions in Transport*. John Wiley & Sons, Inc., pp. 233–281. <http://dx.doi.org/10.1002/9781118601907.ch8>.
- Mitsopoulos-Rubens, E., Trotter, M.J., Lenné, M.G., 2011. Effects on driving performance of interacting with an in-vehicle music player: a comparison of three interface layout concepts for information presentation. *Appl. Ergon.* 42 (4), 583–591. <http://dx.doi.org/10.1016/j.apergo.2010.08.017>.
- Most, S.B., Scholl, B.J., Clifford, E.R., Simons, D.J., 2005. What you see is what you set: sustained inattention blindness and the capture of awareness. *Psychol. Rev.* 112 (1), 217–242. <http://dx.doi.org/10.1037/0033-295X.112.1.217>.
- National Highway Traffic Safety Administration, 2013. *Visual-Manual NHTSA Driver Distraction Guidelines for in-Vehicle Electronic Devices* (NHTSA-2010-0053).
- Neisser, U., Becklen, R., 1975. Selective looking: attending to visually specified events. *Cogn. Psychol.* 7 (4), 480–494. [http://dx.doi.org/10.1016/0010-0285\(75\)90019-5](http://dx.doi.org/10.1016/0010-0285(75)90019-5).
- Rydström, A., Broström, R., Bengtsson, P., 2012. A comparison of two contemporary types of in-car multifunctional interfaces. *Appl. Ergon.* 43 (3), 507–514. <http://dx.doi.org/10.1016/j.apergo.2011.08.004>.
- Scapin, D.L., Bastien, J.M.C., 1997. Ergonomic criteria for evaluating the ergonomic quality of interactive systems. *Behav. Inf. Technol.* <http://dx.doi.org/10.1080/014492997119806>.
- Simons, D.J., Chabris, C.F., 1999. Gorillas in our midst: sustained inattention blindness for dynamic events. *Perception* 28, 1059–1074.
- Singleton, W.T., 1971. The ergonomics of information presentation. *Appl. Ergon.* 2 (4), 213–220. [http://dx.doi.org/10.1016/0003-6870\(71\)90108-6](http://dx.doi.org/10.1016/0003-6870(71)90108-6).
- Sivak, M., 1996. The information that drivers use: is it indeed 90% visual? *Perception* 25 (9), 1081–1089. <http://dx.doi.org/10.1068/p251081>.
- Sodhi, M., Reimer, B., Llamazares, I., 2002. Glance analysis of driver eye movements to evaluate distraction. *Behav. Res. Methods Instrum. Comput. J. Psychon. Soc. Inc.* 34 (4), 529–538.
- Sodnik, J., Dicke, C., Tomazić, S., Billingham, M., 2008. A user study of auditory versus visual interfaces for use while driving. *Int. J. Hum. Comput. Stud.* 66 (5), 318–332. <http://dx.doi.org/10.1016/j.ijhcs.2007.11.001>.
- Strayer, David L., Drews, F.A., Johnson, W.A., 2003. Cell phone-induced failures of visual attention during simulated driving. *J. Exp. Psychol. Appl.* 9 (1), 23–32. <http://dx.doi.org/10.1037/1076-898X.9.1.23>.
- Strayer, D.L., Johnson, W.A., 2001. Driven to distraction: dual-task studies of simulated driving and conversing on a cellular telephone. *Psychol. Sci. J. Am. Psychol. Soc./APS* 12 (6), 462–466.



- Suied, C., Susini, P., McAdams, S., 2008. Evaluating warning sound urgency with reaction times. *J. Exp. Psychol. Appl.* 14 (3), 201–212. <http://dx.doi.org/10.1037/1076-898X.14.3.201>.
- Sumikawa, D.A., 1985. Guidelines for the Integration of Audio Cues into Computer User Interfaces (Technical report no. UCRL 53656). Lawrence Livermore National Laboratory.
- Thorpe, S., Fize, D., Marlot, C., 1996. Speed of processing in the human visual system. *Nature* 381 (6582), 520–522. <http://dx.doi.org/10.1038/381520a0>.
- Victoria, K. Lee, Champagne, C.R., Francescutti, L.H., 2013. Fatal distraction: cell phone use while driving. *Can. Fam. Physician* 59 (7), 723–725.
- Waard, D. De, 1996. The Measurement of Drivers' Mental Workload (Unpublished doctoral dissertation). University of Groningen, Traffic Research Centre, Haren, The Netherlands.
- Walker, B.N., Nance, A., Lindsay, J., 2006. Spearcons: Speech-based Earcons Improve Navigation Performance in Auditory Menus. Proceedings of the 12th International Conference on Auditory Display, pp. 63–68. London, UK.
- 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. *Hum. Factors* 55 (1), 157–182.
- Wierwille, W.W., 1993. An initial model of visual sampling of in-car displays and controls. In: Gale, A.G., Brown, I.D., Haslegrave, C.M., Kruysse, H.W., Taylor, S.P. (Eds.), *Vision in Vehicles IV*. Elsevier Science, Amsterdam, NL, pp. 271–279.
- Young, K.L., Mitsopoulos-Rubens, E., Rudin-Brown, C.M., Lenné, M.G., 2012. The effects of using a portable music player on simulated driving performance and task-sharing strategies. *Appl. Ergon.* 43 (4), 738–746. <http://dx.doi.org/10.1016/j.apergo.2011.11.007>.
- Ziefle, M., 2010. Information presentation in small screen devices: the trade-off between visual density and menu foresight. *Appl. Ergon.* 41 (6), 719–730. <http://dx.doi.org/10.1016/j.apergo.2010.03.001>.
- Zwahlen, H.T., Adams, C.C., DeBals, D.P., 1988. Safety aspects of CRT touch panel controls in automobiles. In: *Vision in Vehicles II. Second International Conference on Vision in Vehicles*.