

COCKTAIL PARTY EFFECT & ATTENTION CAPTURE IN SEMI-AUTONOMOUS DRIVING

Crystal Tobias, Chen-Yung Su, Lutz Kolburg & Brian Lathrop
Volkswagen Group of America, Electronics Research Laboratory
Belmont, California, USA
Email: Brian.Lathrop@vw.com

Summary: A personal name has proven to be an effective stimulus to capture attention. The goal of this pilot study is to test if a personal name can be used as an effective audio warning for drivers of semi-autonomous vehicles. Participants drove a driving simulator in both manual and semi-autonomous driving conditions while doing a secondary task. An emergency situation was simulated, and participants were presented with a warning tone or his/her personal name. Reaction times for braking, steering, and eye disengagement were recorded. There was no significant main effect for cue type, a marginally significant interaction effect across driving condition and cue type, and a significant main effect for driving condition. These results suggest that engagement in a secondary task while driving semi-autonomously causes diverted driver attention to be at its highest. Importantly, however, the use of one's personal name shows promise in capturing attention back to the driving task and warrants deeper investigation for future research.

INTRODUCTION

Distraction, a form of inattention, causes great hazards for drivers. The problem lies in the driver losing focus on the driving task, and instead directing their attention elsewhere (NHTSA, 2009). Much effort has been dedicated to find solutions against distraction. These efforts are often aimed at developing better in-vehicle Human-Machine-Interface designs (Edworthy, 1996; Lee, 1996; Campbell et al., 2007), better warning strategies (Rhede et al., 2011, Fricke, 2009) or better timing for warning activations (Lee et al, 2002; Naujoks et al., 2012; Zarife et al., 2012). Since acoustic warnings have been established as an effective modality to communicate a high priority warning to the driver (Campbell et al., 2007), current topics of in-vehicle interface research largely focus on improving the auditory alert and interaction processes. For example, recent research has explored the combinations of the auditory modality with other modalities (Ho, 2005), cross-modality (Cao, 2010; Mahr, 2012) and the different parameters of auditory alerts (Kun et al., 2007), e.g. the loudness (Baldwin & May, 2011) or stimuli-response compatibility (Wang & Proctor, 2003; Barrow & Baldwin, 2009; Straughn et al., 2009).

There has also been interest in exploring the human interaction that takes place when a system employs self-automation for certain tasks. Within the field of automation research, researchers have desired to learn more about the ideal balance between system automation and human capabilities, such as reduced cognitive load and maintained situational awareness (Rouse, 1988; Sheridan & Parasuraman, 2005; Lee, 2006). In the case of semi-autonomous driving, the car can be expected to do the majority of the physical driving task, thus freeing up time for the driver. During this "free" time, the driver will likely engage in non-driving related tasks. Thus as the driver engages in these other tasks, s/he is no longer attending to the road as a normal driver

would, and becomes “out of the loop” to the driving scenario, namely losing situational awareness (Endsley, 1995). Not surprisingly, one of the latest goals towards improving future driver safety has revolved around this aspect of autonomous driving. There is a need to find a quick and efficient way to capture the driver's attention, especially when the driver is deeply engaged in another task.

This research explores the use of a personally significant acoustic semantic cue, more specifically, a personal name. By semantic, we refer to words of language that provide meaning. We hypothesize that a driver who hears his/her personal name during a state of distraction, will experience attention capture, divert from the distraction, and return his/her attention back to the driving scenario at a faster rate compared to hearing a standard warning tone. The idea of using personal names as a warning cue comes from the work by Cherry (1953) and his “cocktail party effect”. According to the cocktail party effect, it states that people have the ability to follow along in a conversation despite multiple conversations going on simultaneously, although this comes at the price of not knowing the content of the other conversations. Moray (1959) furthered Cherry’s work and discovered that a personal name could break through the attention barrier caused by the primary conversation. Therefore, we extend this body of research with this pilot study where we consider a personal name as an attention capturing cue in a semi-autonomous driving setting.

HYPOTHESES

Based on the current research to date, we have the following hypotheses about this pilot experiment: 1) Under a high-load secondary task that occupies the audio channel, the personal name will capture the driver’s attention more readily than a standard warning tone. This will be demonstrated by faster reaction times for collision avoidance during a driving simulation. 2) Since the driver will be more engaged in the driving task during the manual driving mode, the overall reaction times to both acoustic warnings will be faster than the reaction times in the semi-autonomous driving mode. 3) An interaction effect will occur where the personal name will prove to be a better cue for attention capture during the semi-autonomous driving mode.

METHODS

General Methods

For the primary task, participants used a car simulator to drive on a simulated roadway. They were instructed to drive under a manual condition and a semi-autonomous condition. The latter would activate once the participants reached the freeway section of the simulated roadway, where the test would begin. Once activated, the system would drive the participant semi-autonomously through the simulation.

Upon reaching the freeway (in both driving conditions), participants began a secondary task. The secondary task consisted of the participant inputting a number sequence into a number pad as it was heard from a pair of speakers in the room. During the task, either a standard warning tone or the participant’s personal name would sound alerting the participant to a car that simultaneously appeared in the simulation before him/her. At this point, participants needed to react to avoid a collision. The recorded reaction times begin from the start of the acoustic cue and end with either the start of the driver’s response or the collision with the lead car. None of the participants were aware that either the acoustic cue or lead car would occur during the study.

Participants

Overall, 63 individuals from the San Francisco Bay Area in California participated in the study. The age bracket ranged from 22-35 years with a mean age of 27.98 years ($SD = 4.10$). The participants were experienced licensed drivers ($M = 11.19$ years; $SD = 4.26$ years) and most drove an average of 15,000 – 20,000 miles per year ($Mo = 19,000$ miles). All participants reported normal or corrected-to-normal vision and normal hearing ability. Due to technical problems and/or experienced simulator sickness, only data from 60 participants are included in the analysis, 26 female and 34 male. All were compensated \$120 for their time.

Apparatus

The study utilized the Drive Safety simulator software, run on three displays mounted in the front of a vehicle. Inside the simulator car, a number keypad replaced the traditional center console. In front of the simulator, speakers were stationed to the left and right and were used to play the audio of the simulation, secondary task, and the cues. Furthermore, participants wore a Dikablis eye-tracking system.

Study Design

This study uses a mixed study design. The driving modes are the within-subjects factor and the warning cues are the between-subjects factor. The independent variables, therefore, in this study are the two driving modes and the two cue types. For dependent variables, these are the reaction times (eye disengagement, steering, and braking) and the secondary task performance. A complete study session contains 20 experiment trials with 10 trials for each driving mode. Each driving mode has 5 trials with a cue. The second trial always held the first warning cue; however the order of the remaining warning cues was randomized with the rest of the trials.

The Secondary Task and Warning Cue

Participants performed the secondary task only on the freeway route. An indicator noted the start of the task via a “Ready” “Set” “Go” that successively appeared on the center screen. The task consisted of a sequence of numbers, distributed into 10 segments, spoken aloud through a text to speech (TTS) audio file. Each segment contained 8 digits and had a time span of 3.65 seconds. This was followed with a two second pause before the next segment began. An entire secondary task trial took around 54.5 seconds to complete. All digits ranged from 0 to 9, and sequences were produced randomly. All number sequences were prerecorded from a TTS software program using a female voice prior to the study.

For half of the participants, the individual’s proper first name was used (with the other half hearing the tone). All the participants in the name condition had a name with at least two syllables. The name cue was produced from a TTS software using a male voice at a speed of 221 words per minute. The tone cue was a Volkswagen series production tone and had a constant time span of 0.5 sec. This is consistent with the average time span of all the name cues. The volume for the name and tone cues was 78db. Both cues appeared with a time of collision rate of 2.5 seconds when the participant had his/her eyes fixed on the keypad.

Other

All driving modes, cues, secondary task sequences, start-points, and collision car appearances were randomized and counterbalanced. A post-questionnaire and the NASA – TLX questionnaire were filled out after each driving condition and a final-questionnaire was given at the end of the study to measure distraction, acceptance, and general subjective experience.

RESULTS

The Warning Cue

First of all it has to be clarified, that the participants perceived both cue types as a warning. The subjective results from the post-questionnaire show that 85% of the participants perceived their name as a warning, and 86% perceived the tone as a warning when driving manually. However, under the semi-autonomous mode, only 77% of the participants recognized the tone as a warning cue, but 89% said that the name cue sounded like a warning. A chi-square test was performed to examine the relation of cue type and the perception, independently from the driving mode. The relationship between these variables are non-significant ($\chi^2(2, N = 112) = .895, p = .597$). Both cue types had the same probability to be perceived as a warning.

These findings align with other participant responses. Approximately 81% of the participants agreed or strongly agreed to the statement, “I perceived the cue to be urgent” when the name was used as a warning signal. Statistical analysis revealed that there was no perceived difference when the tone was used as a warning signal instead ($t(61) = .941, p = .350$). All the other participants, who did not perceive the cues as a warning, perceived it at least as an indicator. However, all of the participants took action when they perceived the cue.

Work Load and Secondary Task

We assessed the cognitive workload of the secondary task using the NASA-TLX questionnaire. The data suggest that the secondary task caused a high cognitive workload in both driving modes. A MANOVA with repeated measure (driving modes as a within-participant factor) revealed a statistically significant difference over all dimensions of the questionnaire ($F(6,51) = 3.16, p = .010$). Following univariate tests, results show that the mental demand due to the secondary task was perceived higher when driving manually ($F(1,56) = 4.28, p = .043$).

In addition, participant performance in the secondary task underwent analysis. The Levenshtein distance was used as a measure of accuracy in the secondary task (Levenshtein, 1966). Using this method, every deviation from the given order of numbers performed by the participant was penalized by one distance point. Thus, a greater distance represents less accuracy in the secondary task performance, and therefore can be interpreted as an effect of higher workload and less engagement in the secondary task. To compare the distances, we conducted a one-way ANOVA with repeated measure, with the driving mode as a within-participant factor. Results show a marginally significant difference between the modes ($F(1,56) = 3.48, p = .067$). Participants completing the secondary task in the manual driving mode were less accurate compared to the semi-autonomous mode ($M = 9.01$ and 7.42 ; $SD = 9.72$ and 6.85 respectively). Thus, this conveys a higher level of engagement and better performance by the participant during semi-autonomous driving.

Reaction Times

We conducted a 2 x 2 mixed ANOVA with cue type as the between-participant factor and driving mode as the within-participant factor. For the first analysis, we looked at the fastest reaction times among all possible reaction types (eye disengagement, steering, braking) for all ten warning events per participant, with special interest in the interaction between the cues and driving modes. The fastest reaction time was used since we wanted to capture the first instance of driver response for collision avoidance, and not all participants reacted in the same manner (steering, braking, etc.).

Overall, no significant main effects were found for either the cue type ($F(1,54) = .95, p = .335$) or mode type ($F(1,54) = 2.87, p = .09$). In regards to an interaction effect, results were marginally significant ($F(1,54) = 3.87, p = .054$). Reaction time did not differ much under manual mode between cue type, however participants under the semi-autonomous mode who had the name cue performed faster than those with the tone cue ($M = 554\text{ms}, M = 679\text{ms}; SD = 183\text{ms}, SD = 285\text{ms}$ for name and tone cues respectively). Therefore, the interaction effect suggests that the differences in reaction times are most disparate in the semi-autonomous mode, in favor of the name cue.

Reaction Times – First Warning Event

In the next step we used the fastest reaction times from only the first warning event of each participant exclusively. The reason for doing so, was the expected and observed learning curve (Fig.3). Participants performed faster after the first warning trial. Therefore, to get a grasp of the initial reaction times of the cues, data from the first warning trial were analyzed separately.

We conducted a 2 x 2 mixed ANOVA, with the cue type and driving modes as fixed factors. No statistically significant main effect for cue type could be found ($F(1,57) = 2.36, p = .130$). However, there was a statistically significant main effect for mode type ($F(1,57) = 15.21, p < .001$). As expected, the participants reacted slower when driving in semi-autonomous mode for both cue types ($M = 525\text{ms}, M = 1022\text{ms}; SE = 87\text{ms}, SE = 93\text{ms}$ for manual and semi-autonomous modes respectively). The interaction between cue type and mode type was not significant ($F(1,57) = 2.36, p = .134$).

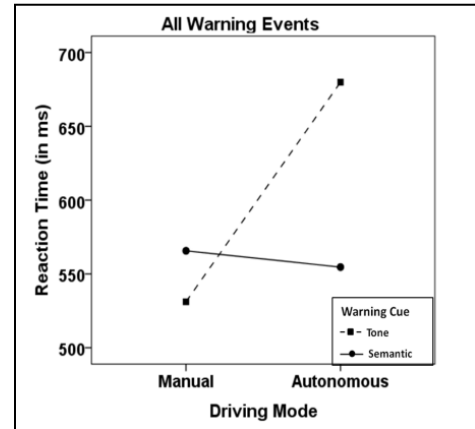


Figure 1. Fastest reaction time for all participants & all warning events

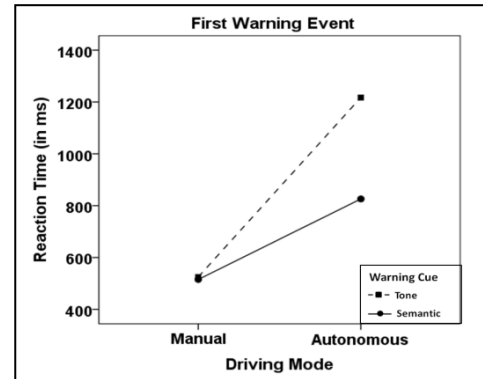


Figure 2. Fastest reaction time for first warning event

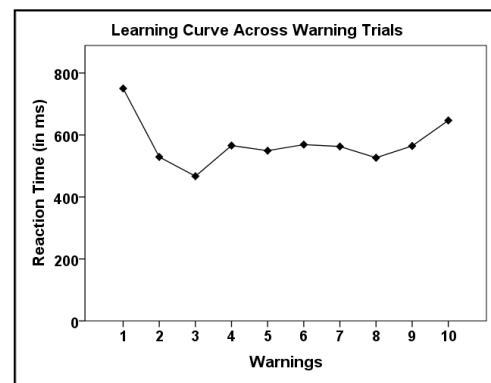


Figure 3. Trend of the reaction times overall

Subjective Ratings

Results show a high acceptance for the personal name cue. For instance, 78% of the participants exposed to the name cue agreed or strongly agreed to the statement, “I liked the use of the cue in the study”. Statistical analysis revealed high acceptance of the name as a warning cue, comparable to that of the tone ($F(1,61) = .015, p = .904$).

Participants also received general questions asking for their opinion on the name cue. Only 19% of the participants agreed or strongly agreed to the statement, “I don’t like systems using my name”. The responses to the statement “I prefer name for non-urgent cues” are mixed. About 29% of the participants agreed or strongly agreed, but 28% of the participants also disagreed or strongly disagreed. When participants were asked if they would like to have their own car use the personal name cue as a warning system 77% agreed or strongly agreed.

DISCUSSION

This study sought to investigate the effectiveness of a personal name in capturing attention against that of a warning tone, with special interest in a semi-autonomous driving situation. Our results indicate that there is a delay in overall reaction times during the semi-autonomous driving condition, but this is only significant during the first warning trials. The delay comes to no surprise, as we hypothesized that people would need more time to process and understand their surroundings when resuming the driving task from a state of deep distraction. At least under manual mode, participants need to occasionally look at the road in order to drive straight and maintain their position in the middle lane. Thus in manual mode, they can potentially gain a better understanding of their overall surroundings with those brief glances. We suspect that this reaction time delay proved significant only for the first warning trials because participants were unaware of the upcoming obstacle that would appear and allowed themselves to be absorbed in the secondary task. After the first warning sounded, participants likely expected the cue and obstacle to appear again. When considering interaction effects, the name cue does have a marginal advantage over the tone cue, especially in the semi-autonomous mode.

Generally, participants were open-minded to the personal name cue, as shown through the 77% who agreed that they would be interested in owning a vehicle warning system that uses their personal name. The best implementation of this name cue would need to be researched, as participants’ feedback on its ideal use was split between urgent and non-urgent purposes. Furthermore, when asked about the perception of the cue as a warning, participants were less likely to choose the tone as a warning under semi-autonomous mode.

As a follow up to this study, a future study design could allow the participants to gain exposure to both cues in a within-subjects design. By doing so, participants would be able to directly compare the cues. Future iterations of this study could also test the use of a personal name against other semantic cues that give explicit instructions to drivers, such as “attention” or “look up”. A combination of the personal name and other cues may improve overall driver safety. For instance, the name could be paired with directional cues, such as “Sarah! Steer Right!” or with directional haptics. This study served as a pilot study for exploration into the personal name cue and to learn of its potential value. More research would be needed to explore the conceptualization and design principle of an acoustical semantic cue in driver assistance systems.

REFERENCES

- Baldwin, C. L., & May, J. F. (2011). Loudness interacts with semantics in auditory warnings to impact rear-end collisions. *Transportation Research, Part F 14*, 36–42.
- Barrow, J. H., & Baldwin, C. L. (2009). Verbal-Spatial Cue Conflict: Implications For The Design of Collision-Avoidance Warning Systems. *Proceedings of International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 5, 405–411.
- Campbell, J. L., Richard, C. M., Brown, J. L., & McCallum, M. (2007) *Crash Warning System Interfaces: Human Factors Insights and Lessons Learned*. Washington, DC: NHTSA.
- Cao, Y., Mahr, A., Castronovo, S., Theune, M., Stahl, C., & Müller, C. (2010). Local Danger Warnings for Drivers: The Effect of Modality and Level of Assistance on Driver Reaction. *Proceedings of International Conference on Intelligent User Interfaces*, 239–248.
- Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, With One and With Two Ears. *The Journal of the Acoustical Society of America*, 25(5), 975–979.
- Edworthy, J., & Adams, A. (1996). *Warning Design: A Research Prospective*. London: Taylor & Francis.
- Fricke, N. (2009). *Gestaltung zeit- und sicherheitskritischer Warnungen im Fahrzeug* (Dissertation). Technische Universität Berlin, Berlin.
- Ho, C., Spence, C., & Tan, H. Z. (2005). Warning Signals Go Multisensory. *Proceeding of International Conference on Human-Computer Interaction*, 11.
- Kun, A., Peak, T., & Zeljko, M. (2007). The Effect of Speech Interface Accuracy on Driving Performance. *Proceedings of Interspeech*.
- Lee, J. D., Gore, B. F., & Campbell, J. L. (1999). Display Alternatives For In-Vehicle Warning and Sign Information: Message Style, Location, and Modality. *Transportation Human Factors*, 1(4), 347–375.
- NHTSA (2009). *Distracted Driver*. Available under <http://www.distraction.gov/index.html>.
- Mahr, A. (2012). Crossmodal Speech Priming of Visual Objects: A New Approach For Urgent Information Transfer From Car to Driver. *it - Information Technology*, 54(4).
- Moray, N. (1959). Attention in Dichotic Listening: Affective Cues and the Influence of Instructions. *Journal of Experimental Psychology*, 56–60.
- Lee, J.D., McGehee, D.V., Brown, T.L., and Reyes, M.L. (2002). Collision Warning Timing, Driver Distraction, and Driver Response to Imminent Rear-End Collisions in a High-Fidelity Driving Simulator. *Human Factors*, 44 (2).
- Levenshtein, V.I. (1966). Binary Codes Capable of Correcting Deletions, Insertions, and Reversals. *Soviet Physics-Doklady*, 10(8).
- Straughn, M. S., Gray, R., & Tan, H. Z. (2009). To Go or Not To Go: Stimulus-Response Compatibility For Tactile and Auditory Pedestrian Collision Warnings. *IEEE Transactions on Haptics*, 2(2), 111-117.
- Wang, D.-Y. D., & Proctor, R. W. (2003). Stimulus-Response Compatibility Effects For Warning Signals and Steering Response. *Proceedings of International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design*, 2, 226–230.