Use the Right Sound for the Right Job: Verbal Commands and Auditory Icons for a Task-Management System Favor Different Information Processes in the Brain

Christiane Glatz

Max Planck Institute for
Biological Cybernetics
Tübingen, Germany
IMPRS for Cognitive &
Systems Neuroscience
Tübingen, Germany
christiane.glatz@tuebingen.mpg.de

Stas S. Krupenia

Scania CV AB
Soedertaelje, Sweden
stas.krupenia@scania.com

Heinrich H. Bülthoff

Max Planck Institute for
Biological Cybernetics
Tübingen, Germany
heinrich.buelthoff@tuebingen.mpg.de

Lewis L. Chuang

Max Planck Institute for
Biological Cybernetics
Tübingen, Germany
lewis.chuang@tuebingen.mpg.de

ABSTRACT

Design recommendations for notifications are typically based on user performance and subjective feedback. In comparison, there has been surprisingly little research on how designed notifications might be processed by the brain for the information they convey. The current study uses EEG/ERP methods to evaluate auditory notifications that were designed to cue long-distance truck drivers for task-management and driving conditions, particularly for automated driving scenarios. Two experiments separately evaluated naïve students and professional truck drivers for their behavioral and brain responses to auditory notifications, which were either auditory icons or verbal commands. Our EEG/ERP results suggest that verbal commands were more readily recognized by the brain as relevant targets, but that auditory icons were more likely to update contextual working memory. Both classes of notifications did not differ on behavioral measures. This suggests that auditory icons ought to be employed for communicating contextual information and verbal commands, for urgent requests.

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation:
User Interfaces – Input Devices and Strategies

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

CHI 2018 April 21–26, 2018, Montreal, QC, Canada

© 2018 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-5620-6/18/04.

DOI: <https://doi.org/10.1145/3173574.3174046>

Author Keywords

Notifications; Electroencephalography; Auditory Displays;
Autonomous Vehicles; In-Vehicle Interfaces

INTRODUCTION

Auditory notifications are used extensively by in-vehicle interfaces to inform the user of important events. Whilst parking, for example, decreasing beep intervals could communicate the nearing distance between a driver and an obstacle. This raises the question: How should auditory notifications be designed? Should they be verbal notifications that communicate instructions explicitly or should they be recognizable auditory icons that denote a critical scenario?

Previous researchers have generally agreed on the essential design guidelines for auditory notifications [46]. Auditory notifications need to be: (1) easily detectable [24, 45, 16], (2) readily discriminable against background noise [48, 39], (3) capture attention [29, 30], and, after all these requirements have been fulfilled (4) easily interpretable [45, 18, 70]. However, there are many ways in which auditory notifications can be designed to comply with these criteria. Preference between different designs is often determined by user studies that evaluate performance or subjective feedback. Unfortunately, such measures can often contradict from one study to another or fail to discriminate between different designs.

Verbal commands and auditory icons represent two general classes of auditory notifications that are commonly employed. They are favored over synthetic sounds [38, 13] because they are based on prior user experiences. Hence, they are easily learned in novel use settings for their intended meanings. Nonetheless, no clear consensus has been established for preferring either verbal commands or auditory icons.

Besides performance and subjective measurements, brain responses to notifications can also serve as a way to evaluate auditory notifications. Surprisingly, this approach is rarely employed (although, see [37]). In particular, electroencephalography (EEG) reveals how the brain processes information – namely, the extent to which notifications capture attention and are interpreted – regardless of how this eventually influences observable performance or subjective feedback. In other words, EEG provides a higher functional resolution of the processes that take place between stimulus presentation and the elicited response. Therefore, we can evaluate auditory notifications based on how user brains respond to them, and not merely on how users respond to them. The understanding that we can gain from the additional information provided by EEG can contribute towards the design of more appropriate and effective notifications that are easy to use.

The increasing reliance on automation to perform tasks is transforming the role of notifications. While notifications used to be prized for their effectiveness as a ‘call to action’, they are increasingly used to update and inform users on the current situation, or to assist users in supervising automation. This trend is especially prevalent in the context of automated vehicles, whereby auditory notifications have been specially designed to update drivers of automated trucks on prevailing road conditions and to remind them to supervise logistical tasks. Arguably, such notifications might not be readily evaluated by behavioral responses alone, but by how they are processed by the brain for conveyed information.

In the current work, we answer the following four research questions (RQ1-RQ4) using EEG analysis. RQ1: Are auditory icons or verbal commands more easily detected? RQ2: Are auditory icons or verbal commands more discriminable? RQ3: Do auditory icons or verbal commands capture more attention? RQ4: Do auditory icons or verbal commands result in more context-updating of what has to be done next?

To summarize, this paper makes the following contributions:

1. EEG demonstrates that both types of notifications are equally detectable and orient attention to the same extent or rather more in an applied context.
2. By using EEG, we are able to show that verbal commands are discriminated more easily than auditory icons, but that auditory icons are more likely to update contextual working memory.
3. Evaluating brain responses reveals that auditory notifications should follow a purpose-oriented design. That is, verbal commands seem more suitable for urgent requests while auditory icons should be used to communicate less pressing contextual information. Unlike behavioral measures, EEG gives insight in different stages of processing a notification. Notably, these insights are obtained in a passive and unobtrusive way and cannot be gained through behavioral results only.
4. The results of this study demonstrate that laboratory findings of auditory EEG studies can be extended to more realistic environments, such as driving simulators. Qualitative

comparisons suggest that results scale with experience and that the effect of notifications on brain responses increase with increasing relevance.

RELATED WORK

Challenges for Designing In-vehicle Auditory Displays

Sound design is a challenging task not only technically but even more so regarding human perception. Human’s auditory perception is influenced by a variety of factors such as emotions, memories, cognition, environment, previous experience, and the ability to understand speech [49]. In the design of auditory displays, sounds that communicate information to the user, factors like these are relied upon [43]. When not using words to communicate information but sounds, we talk about sonification [71]. In this regard, sound designers have to consider the important aspect of making sounds recognizable and identifiable for what they were designed. At the same time, auditory notifications need to convey the level of urgency without annoying the user [42]. Especially for in-vehicle notifications, auditory displays whose meaning is ambiguous can have a negative influence (i.e. higher perceived workload, slower response times) [72]. Hence auditory notifications should be evaluated, for example, based on their detectability, their position of presentation, their identifiability, and their conveyed meaning [69].

To optimize the design of auditory notifications, a two-stage pipeline has been proposed [39]. At an early first stage, designers consult with the user audience about the designed sound which feeds back into the design process. The second stage is an evaluation of the auditory display in a greater auditory context, simulating the use of the sound in its intended environment. While this is a step in the right direction for optimizing the design of effective auditory notifications, it relies on subjective user feedback, as well as performance measures, and does not reveal the actual influence auditory notifications have on the processing by the human brain.

Auditory Notifications

Auditory notifications can be categorized into two main groups, speech and non-speech. We differentiate non-speech sounds into auditory icons (representative sounds) and earcons (abstract synthesized sounds) [24, 32]. Unlike earcons, verbal commands and auditory icons are readily recognizable and do not require learning, which typically translates to faster responses [44].

Verbal Commands

We often rely on speech to communicate our intentions to each other. Once proficiency is acquired in a given language, verbal commands can be relied on to communicate complex messages that can be unambiguously interpreted [38, 13]. Thus, it is natural for humans to prefer speech notifications [44].

Nonetheless, verbal notifications face the risk of being masked by or confused with real conversations [50]. To circumvent this problem, contrivances could be introduced to make verbal notifications less human-like and more discriminable from real speech. This could be achieved by manipulating the pitch

or other spectral properties of verbal commands. Pilots were found to discriminate easily between natural speech and synthesized speech [63]. Thus, presenting notifications in synthesized speech could prevent verbal commands from being confused with real conversations in our environment.

Another shortcoming of speech is that it is harder to spatially localize than other sounds, presumably because of its smaller bandwidth [68]. However, this can easily be compensated by being interpreted unambiguously when it is used appropriately. Verbal commands can present unambiguous spatial information through their semantic context. For instance, presenting the word ‘front’ from a front speaker results in fast response times to potential head-on collisions [29]. However, if not used appropriately, speech can attract attention to the extent that it could interfere with other tasks such as driving [64].

Auditory Icons

Auditory icons are sounds that represent real world events. These are sounds with stereotypical associations with the object or event/action that created the sound. For example, the sound of a car horn could indicate a safety critical situation that requires immediate attention and action. Being familiar sounds, auditory icons are easily learned for their intended function.

An advantage of auditory icons is that they are not easily masked by background speech [38, 66, 30]. For instance, auditory icons are unlikely to be confused with a radio jockey’s monologues. Nonetheless, auditory icons, like skidding tires or the car horn, can still be confused with real environmental occurrences. Also, auditory icons have been shown to be more likely in generating false alarms than abstract notifications [26]. This is most likely due to the fact that humans might have overlearned certain cues (e.g., car horn). Given that background experiences are likely to be different across different users, auditory icons might be challenging to calibrate for their conveyed urgency.

Auditory icons are susceptible to misinterpretation because a single sound can represent more than one meaning [24, 45]. Depending on previous experience and the use-context, auditory icons can be recognized as an object (that generates the sound) or as the action that generated the sound [45, 23, 44]. For example, the sound of screeching tires can be interpreted either as a proximal collision vehicle or as a command for braking. In complex operations, auditory icons might not be the appropriate notification [25, 28]. However, [3] used skidding tires and a car horn honk in highly safety critical situations, namely to signal impending collisions. The successful use of auditory icons in this case might be due to the straightforward association of meaning. In establishing guidelines for designing auditory icons, [45] suggests that auditory icons usability is highly affected by their identifiability. Nonetheless, the recognition accuracy for auditory icons, but not response times, can be significantly improved if users are aware of the icons’ design [40]. This suggests that auditory icons should use mappings that do not have multiple interpretations and can be easily associated with the events they are representing.

Verbal Commands versus Auditory Icons

Auditory notifications can be evaluated on two different performance measures. First, effective notifications are believed to elicit faster reaction times. Second, effective notifications are more accurately detected from background noise and discriminated from other notifications. Previous research has compared the effectiveness of verbal commands and auditory icons across different scenarios and has generally found mixed support for either sound type.

To begin, while some studies have found faster responses for verbal notification [40, 29, 21], others found faster responses for auditory icons [24, 60]. Other studies have found a response time preference for neither auditory icons nor verbal notifications [44, 70]. Overall, this would suggest that the processing for both sounds is not different. Some reasons for these mixed findings could be the context in which they were presented as well as the type of task participants were asked to do. Some studies, for example, took place in a simulated driving context and asked participants to avoid collisions by breaking while others required participants to match a presented sound to a description of the object/action generating this sound on a desktop PC.

Similarly, measures for response accuracy also provide mixed support for either verbal commands or auditory icons. [9] found that people were more accurate at matching auditory icons to visual context than words. This could be due to a trade-off between response times and discrimination sensitivity since reaction times were faster for nouns than auditory icons in this study. In contrast, [61, 51] found higher accuracy when cuing with verbal commands than with auditory icons. [38], on the other hand, found no difference in accuracy for verbal commands compared to auditory icons. These different findings could be due to the different experimental tasks participants were asked to complete. On the other hand it could also depend on the type of auditory notifications employed. If time was not stressed, participants could respond when they had fully evaluated the auditory stimulus, even if the auditory icon was not readily interpretable at first.

Given that there will always be background noise, including meaningful sounds; it is worthwhile to use prominent and highly discriminable auditory notifications. Previous research has shown that distractor sounds have a larger negative impact on masking auditory icons than verbal commands [60]. That is, distractors were more likely to interrupt the processing of auditory icons than verbal commands. This suggests that non-verbal processing is more likely to be affected by increases in processing workload, often present in stressful and urgent situations.

Event-Related EEG Potentials

Brain responses can be used to evaluate auditory notifications, especially in terms of how they are processed for information. One prominent EEG measure is the event-related potential (ERP) that represents brain activity that proceeds from the presentation of a given stimuli. An ERP is an average waveform of negative and positive voltage deflections, which can be functionally related to different stages of information processing of the presented stimuli [41]. Auditory stimuli characteristically

elicit a series of evoked potentials, namely N1, P2, P3a, and P3b components which are associated with how the presented sound is processed [8, 54, 35]. Thus, amplitudes of these potentials can provide insight into how the brain processes a given sound.

N1

The N1 is the first negative deflection of the ERP waveform, e.g. [41, 54, 74]. It reflects early involuntary sensory processing and is highly sensitive to a sound's physical properties (e.g., pitch, intensity) [41, 65]. It is a reliable indicator to the perceptual detection of a sound's presentation [73, 34, 54].

P2

The P2 is the second positive deflection, prominent at frontal and central electrode sites. It is often reported as part of the N1-P2 complex, also called 'vertex-potential' [8, 54]. In the current context, we interpret the P2 component as a measure for discriminability. It is elicited by attended and not attended stimuli [8, 55]. This means that it is evoked involuntarily by both, target and non-target stimuli of an oddball paradigm. The difference between the P2 of target and non-target stimuli is that the P2 amplitude is larger for stimuli containing target features [41]. This fits the theory of P2 reflecting object discrimination [47, 22]. P2 amplitudes are larger after one learns to discriminate a stimulus from other target stimulus [2]. In this respect, a larger P2 amplitude reflects target identification and only subsequently a P3 is elicited. Once this classification has taken place, the resolved information can be transmitted to higher cortical areas to be evaluated further [54]. If the stimulus being processed does not contain target features, no P3 is elicited and, hence, further cognitive evaluation of the stimulus is stopped.

P3a

The P3a refers to the third positive peak that is observed at frontal areas. It is evoked by unexpected stimuli regardless and decreases in amplitude to a surprising stimulus if it is presented repeatedly [59]. It is sometimes referred to as the novelty P3 and is believed to reflect an automatic orienting response to interesting information [37, 19, 57].

P3b

The P3b refers to the third positive peak that is observed at centro-parietal regions. It is sensitive to the presentation of task-relevant stimuli, especially those that occur infrequently [15, 56]. In the current context, P3b is treated as a measure for context-updating, a process that underlies how we update our situational understanding when unexpected events occur [14, 20, 57]. When an interesting stimulus is recognized, which is different from the standard background, our brains update our mental representations of the environment. This updating process is reflected by larger P3b amplitudes for target stimuli than standard stimuli. Furthermore, P3b amplitudes are also treated as an index for working memory load, whereby larger P3bs are associated with less mental effort [6].

To date, little research has been conducted to investigate how verbal commands might be processed differently by the brain, compared to auditory icons. Behavioral studies have mixed

results given that they tend to differ depending on the experimental task and context. The notifications that are used here were designed for an auditory display of a highly automated truck environment [36], for which verbal commands have been claimed to deliver faster responses than auditory icons [21]. However, the underlying reason for this has been unclear. The current study was designed to look specifically at how the brain might process these auditory notifications differently, depending on whether they were verbal commands or equivalent auditory icons. For this purpose, we measured the EEG activity of naïve participants (Experiment 1) and professional truck drivers (Experiment 2). The current EEG dataset has been previously analyzed for differences between the two participant groups and have shown that both groups respond to these notifications similarly as a whole [7]. While professional truck drivers responded slower in general, this was not due to fundamental differences in brain responses to the auditory notifications. In fact, the current analyses show similar EEG/ERP waveforms between the two participant groups. In contrast to previous work, this current work focuses specifically on how verbal commands and auditory icons are processed differently in the brain. Although both types of auditory notifications produce similar brain responses, significant differences exist in specific ERP components, which suggests that they should be employed for different purposes.

STUDY METHODS

This study compared auditory icons to verbal commands in a controlled experimental laboratory environment. It was a within-subject design that used separable EEG/ERP components to evaluate these notifications for how well they were detected (N1), discriminated for against other sounds (P2), captured attention (P3a), and updated contextual working memory (P3b). The whole experiment lasted 2.5 hours, which included training, preparation time, and debriefing. The experimental procedure was approved by the Ethics Council at the University Hospital Tübingen.

Two experiments comprised this study. Experiment 1 was performed on university students ($N=15$; mean age= 26.1 ± 4.0 years; 9 males) and a follow-up experiment 2, on professional truck drivers (mean age= 41.4 ± 12.1 years; 13 males). Experiment 1 was conducted in a psychophysical laboratory setting and experiment 2, in a high perceptual fidelity fixed-based truck simulator.

Participants

All participants reported no known hearing deficits, normal (or corrected-to-normal) vision, and no history of neurological problems. They provided signed consent to written instructions, and were remunerated for their voluntary participation.

Stimuli and apparatus

Auditory notifications

The auditory notifications (duration: 500 ms) were adapted from target sounds that were originally designed for the in-vehicle interface of an autonomous truck cabin [21, 36]. There were 12 notifications in total, 6 verbal commands and 6 auditory icons that were complements of each other. They were designed to remind truck drivers to perform certain tasks and

Figure 1. Experiment 1's ERP responses (left) with scalp topography plots (right) of statistically significant differences across time and electrodes respectively. Left: ERP waveforms are averaged across the frontal (pink) and parietal (green) electrodes and deflections are labeled for N1, P2, P3a, and P3b. The shaded areas between the two waveforms indicate time-regions that are significantly different. Right: The scalp topographies show the EEG activity to verbal commands and auditory icons at time-ranges A and B. Electrodes that are significantly different are represented by white dots.

of driving conditions at the appropriate times. These verbal commands (*auditory icons*) were “system” (*synthetic tone*), “convoy” (*train whistle*), “driver” (*human whistle*), “weather” (*raindrop*), “road” (*ground rumbling*), and “traffic” (*car horn*). Verbal commands were in Swedish, which was the mother tongue of the professional drivers in Experiment 2 but not the student volunteers of Experiment 1. However, the student volunteers were extensively briefed on the auditory stimuli and practiced discriminating them until they were 80% accurate, prior to testing.

Ninety distractor sounds were created. Each distractor was a simultaneous presentation of 2 verbal commands and 2 auditory icons, played in reverse and their loudness adjusted to be comparable to the notification targets.

Experiment 1: Psychophysics laboratory

The experimental laboratory was a dark room which was insulated against external sounds. The visualization was presented on a desktop screen (ViewPixx Screen, 60.5 x 36.3 resolution; 120 Hz) at a fixed distance of 45 cm from the participant who was in a chin-rest. The experiment was controlled with customized software (MATLAB 8.2.0.701, R2013b) and Psychophysics Toolbox 3.0.12 [5, 52, 33]. An ASIO 2.0 compatible sound card was used to control sound presentation (SoundBlaster ZxR; Creative Labs). The auditory stimuli were presented via stereo speakers, one paced on the left the other on the right side of the desktop display.

Experiment 2: High fidelity truck simulator

Professional truck drivers sat in a driving simulator that consisted of a realistic truck cabin. This contained a steering

wheel, dashboard with instruments, and a pneumatic seat. The visualization consisted of an automated drive on the highway from Linköping to Norrköping with minimal traffic. The highway had two lanes, one for each antagonistic traffic direction. A three wall display (approx. 150 deg field-of-view) presented the frontal visualization (450 cm distance to head). Two vertically-aligned displays were attached to the outside of the cabin, to simulate side mirrors for displaying the rear traffic scene. Using OpenDrive8 road network files (xodr) and an additional file describing the landscape (xml), a customized graphical engine (i.e., VISIR) rendered the presented visualization. Buttons located on the left and right on the steering wheel collected the participants' behavioral responses. An ASIO 2.0 compatible sound card (RME HDSP 9632; RME Intelligent Audio Solutions) controlled the presentation of auditory notifications. A 5.1 surround sound system, installed in the truck cabin, was used to display the sounds.

Task and Procedure

During testing all participants observed an automated driving scene that was visually presented. They had to respond to auditory notifications whenever one was presented, with a button press using either their left or right index fingers. Six notifications (i.e., 3 complementary pairs of verbal commands and auditory icons) were pre-assigned to a left index-finger press and the remaining six, to a right index-finger press. In other words, each button corresponded to three events, which were represented by a verbal command as well as an auditory icon. Button press assignment was randomized across participants. Prior to testing, all participants practiced until they were able to achieve accuracy levels of at least 80% in this task.

Approximately 980 sounds were presented throughout the experiment testing. All sound presentations were separated by a time-interval randomly selected from a uniform range of 2300–2700 ms. Twenty percent of these were target notifications and eighty percent were distractors. No button presses were necessary when distractors were presented. Target notifications were evenly divided for verbal commands and auditory icons. Participants had 2000 ms to respond after each auditory notification was presented.

EEG recording

The EEG was recorded using 59 active electrodes mounted to the scalp using an elastic cap according to the international 10–20 system (ActiCap System, Brain Products GmbH, Munich, Germany). Four additional electrodes were used to record the vertical and horizontal electrooculogram from the right and left canthi as well as above and below the left eye. All signals were recorded with the online reference FCz and AFz as the ground. EEG signals were digitized with a sampling rate of 1000 Hz. Electrode gel was applied to each electrode to ensure an impedance below 20 k Ω . A parallel port connection between recording PC and experimental PC synchronized the EEG recording with the experimental events, such as the sound onset and button press.

EEG data processing and analysis

To analyze the EEG data, MATLAB (8.2.0.701, R2013b) and EEGLAB v.14.0.0 (<https://sccn.ucsd.edu/eeglab/>), an open source software to analyze electrophysiological data, was used [10]. Before analyzing the ERP to the auditory stimuli, the data was preprocessed for every subject according to the following steps. To reduce the computational costs, the data recorded at 1000Hz was downsampled to 250Hz. To remove any slow drifts, a high-pass filter (cut-off = 0.5 Hz) was applied subsequently to the data. Using CleanLine, a plugin in EEGLAB, 50 Hz electrical noise picked up from the environment when recording electrical brain activity was removed. Then, bad channels, e.g. channels with flat lines, are removed using artifact subspace reconstruction. Following these cleaning steps, the data is re-referenced offline to the common average reference and then submitted to the Adaptive Mixture ICA (AMICA, [11]). This algorithm decomposes the electrical activity recorded at sensor level (electrodes) into source-resolved activity, also called independent components (ICs). These ICs were subject to equivalent dipole estimation. A MNI Boundary Element Method head model was used to fit an equivalent dipole to the ICs [53]. IC dipoles with location outside the brain, as well as, ICs with a residual variance larger than 15% were excluded. Next, the ICs of all participants were grouped into 30 clusters using k-means based on their power spectrum. These clusters were then inspected for non-cortical electrical activity such as eye-related activity, muscle-related activity, line noise, and unresolved components. Clusters, containing such non-cortical activity, were determined based on their power spectrum, their scalp topography, and their dipole location in a volumetric brain model. These non-cortical activity clusters, present across the group of participants, were removed from the EEG data (for examples, see [7]). Finally, this EEG data for cortical activity was backprojected to the

sensor level, and analyzed for potential differences between verbal commands and auditory icons.

ERPs were computed for each participant, and for every electrode, by extracting an epoch of EEG activity around the notification presentation. The presentation onset of the notifications was the trigger event for an epoch that consisted of 500 ms of baseline activity pre-trigger and 1000 ms of brain response post-trigger. All epochs that belonged to either verbal commands or auditory icons were mean-averaged for each electrode. We further grouped the frontal and parietal electrodes into two separate groups for visualization (see Figures 1 and 2, right). These group averaged waveforms depict distinct ERP components (i.e., N1, P2, P3a, and P3b) that serve as established neural correlates for perceptual and cognitive mechanisms. With regards to auditory information processing, they relate to detection (N1), discrimination (P2), attentional capture (P3a), and context-updating (P3b).

We performed mass-univariate analysis (MUA) to statistically evaluate EEG differences between verbal commands and auditory icons [27]. Simply put, this method compares the two conditions at every electrode and time-point and performs a t-test for significant differences. A false discovery rate procedure (i.e. FDR-BH; [4]) was applied to control for multiple comparisons.

RESULTS

Behavioral performance

Behavior performance was evaluated in terms of discrimination sensitivity and correct response times. All scores were submitted to a within-subjects t-test and the Bayes Factor (BF_{01}) was calculated for the likelihood of the null-hypothesis relative to the alternative-hypothesis. The behavioral data of one participant from Experiment 2 had to be excluded due to missing button presses.

Discrimination sensitivity (d') was computed for each participant as the difference between the z-score of correct recognition and false recognition. In Experiment 1, d' scores were not significantly different for verbal icons and auditory icons ($t_{14} = 0.17, p = 0.87, \text{Cohen's } d = 0.04$), respectively 3.82 ± 1.15 and 3.78 ± 1.10 . The null-hypothesis was favored by a Bayes Factor of 3.7. In Experiment 2, d' scores were not significantly different for verbal icons and auditory icons ($t_{13} = 1.53, p = 0.15, \text{Cohen's } d = 0.41$), respectively 2.45 ± 1.44 and 2.17 ± 1.12 . The null-hypothesis was marginally favored by a Bayes Factor of 1.4.

Response times were calculated for all correct responses that occurred within 2500 ms of notification onset, for each participant. In Experiment 1, participants were not significantly faster in responding to verbal commands compared to auditory icons ($t_{14} = 0.68, p = 0.51, \text{Cohen's } d = 0.18$), respectively 1070 ± 131 and 1094 ± 182 ms. The null-hypothesis was favored by a Bayes Factor of 3.1. In Experiment 2, participants were not significantly faster in responding to verbal commands compared to auditory icons ($t_{13} = 0.46, p = 0.65, \text{Cohen's } d = 0.12$), respectively 1236 ± 101 and 1251 ± 156 ms. The null-hypothesis was favored by a Bayes Factor of 3.3.

Figure 2. Experiment 2's ERP responses (left) with scalp topography plots (right) of statistically significant differences across time and electrodes respectively. Left: ERP waveforms are averaged across the frontal (pink) and parietal (green) electrodes and deflections are labeled for N1, P2, P3a, and P3b. The shaded areas between the two waveforms indicate time-regions that are significantly different. Right: The scalp topographies show the EEG activity to verbal commands and auditory icons at time-ranges A and B. Electrodes that are significantly different are represented by white dots.

To summarize, the current behavioral results do not argue in favor of either verbal commands or auditory icons.

EEG/ERP responses

The EEG/ERP activity elicited by verbal commands and auditory icons were similar in general morphology, latency, and scalp distribution in the anterior-posterior dimension for both Experiments 1 and 2 (Figs. 1 and 2). Statistically significant differences were revealed in the EEG/ERP activity generated by auditory icons and verbal commands in the frontal as well as parietal electrodes¹. The frontal group of electrodes are: *F5, F3, F1, Fz, F2, F4, F6, FC5, FC3, FC1, FC2, FC4, FC6*. The parietal group of electrodes are: *P5, P3, P1, Pz, P2, P4, P6, CP5, CP3, CP1, CPz, CP2, CP4, CP6*.

In Experiment 1, student participants showed significant differences in their EEG/ERP responses to these notifications, even though the verbal commands were not presented in their native language nor did they have contextual meaning. Specifically, the amplitude of the P2 component (236–304 ms; frontal electrodes) was significantly larger for verbal commands than for auditory icons. This suggests that verbal commands were more discriminable than auditory icons from presented sounds. The P3b amplitude (512–640 ms; parietal electrodes) was significantly larger for auditory icons than for verbal commands. This suggests that auditory icons induced more context updating than verbal commands did.

In Experiment 2, the professional truck drivers generated similar results as the naïve participants of Experiment 1. Similarly,

¹Electrode labels are in italics to distinguish them from ERP component labels.

verbal commands generated larger P2 component deflections (212–352 ms; frontal electrodes) than auditory icons, and auditory icons generated larger P3b deflections (412–624 ms) than verbal commands. However, EEG/ERP activity in the frontal electrodes revealed that auditory icons generated larger N1 deflections (160–212 ms) than verbal commands, and that verbal commands generated larger P3a deflections (352–468 ms) than auditory icons. Differences in N1 deflections suggest that auditory icons were more likely to be detected against the general auditory background. Differences in P3a deflections suggest that verbal commands were more likely than auditory icons to capture observer attention.

DISCUSSION

The design of auditory displays faces the challenges of incorporating human perception to make notifications effective for their designed purpose (see Sect. 1 and 2). Behavioral performance measures are limited in discriminating between notifications for the various purposes that they might be designed for. Some notifications might be designed for the purposes of being highly detectable while others might be designed to communicate a given context or scenario. Performance measures, e.g., response times or accuracy, do not discriminate for how the brain processes notifications for information.

To summarize, we evaluated verbal commands and auditory icons that were especially designed for in-vehicle information displays related to aspects of task management and context-updating in highly automated trucks [21, 36]. We were motivated to do so according to the guidelines for auditory displays [46], which states that auditory displays ought to be highly detectable in terms of their physical properties (N1), support

learned discrimination from other target notifications (P2), have the potential for capturing attention (P3a), and communicate for its intended purpose (e.g., updating contextual working memory; P3b). Experiment 1 tested naïve students, as is often the case during design and prototyping phases, and Experiment 2 tested professional truck drivers, as is often the case when performing an evaluation and validation. Both types of notifications were effective and did not significantly differ in terms of response times or accuracy. However, verbal commands are more easily discriminable from other target notifications at an early perceptual stage (i.e., larger P2 component), and auditory icons are more likely to update contextual working memory (i.e., larger P3b). This is consistently true regardless of testing environment or participant groups. In fact, discriminable and significant trends in EEG/ERP waveforms are amplified in professional truck drivers. Professional truck drivers also show a neural preference for verbal commands with regards to their detection (N1) and attentional capture properties (P3a), which is presumably driven by a familiarity with the given language.

Therefore, we advocate the use of both verbal commands and auditory icons. However, they should be employed according to the job that they are intended for. We suggest that verbal commands should be employed in critical situations that require immediate action while auditory icons seem more appropriate to notify the user of non-urgent environmental updates. In the context of highly automated vehicles, verbal commands should be used for time-critical situations that cannot afford ambiguity, such as *'low fuel'*, while auditory icons might be better employed in indicating that driving conditions are changing, such as the sound of a thunderstorm for inclement weather.

Justification for the Current Interpretation

Our recommendations for using verbal commands and auditory icons for different purposes are based on the following reasons, based on the EEG/ERP responses to these auditory notifications. We do not address the performance results given that they are shown to be equally effective in terms of discriminability and correct response times.

To begin, we do not emphasize the detectability of either notification types (in terms of their physical properties). This is because the timing of N1s were identical in both Experiments 1 and 2. Interestingly, the amplitude of N1 was larger for auditory icons in Experiment 2. We believe that this reflects the larger variability of the spectral properties of auditory icons relative to verbal commands, which renders it more detectable against a richer (i.e., noisier) background. While this could be treated in favor of auditory icons, we believe that the consistently larger amplitudes in P2 components for learned notification discriminability across both experiments compensates for this minor advantage.

Next, the frontal P2 component is larger for verbal commands than for auditory icons. P2 is believed to reflect learned object discrimination [8, 47, 22]. In this regard, P2's amplitude indicates the efficiency in recognizing the associated notification as a discriminable target, relative to other target notifications. Trained discriminability is known to have an effect on P2

amplitude. For example, musicians who are trained to discriminate sounds for pitch and timbre generate larger P2s than non-musicians, especially for musical sounds [62]. In a similar fashion, most of us are highly trained to discriminate between different verbal sounds for their intended meanings and associations. This reflects the natural advantage rendered by the use of verbal commands over auditory icons. The current findings suggest that even if certain auditory icons are determined by sound designers as being highly discriminable and recognizable as targets, they should be matched to the standards of verbal commands, which is quantitatively measurable in terms of P2 component amplitude.

Thirdly, we believe that verbal commands capture attention more readily than auditory icons. P3a amplitudes are indicative of an involuntary orienting response to surprising and novel events [57]. While the larger P3a amplitude for verbal commands was not significant in Experiment 1 (Figure 1), it was for the participants of Experiment 2. We believe that this was because the professional truck drivers understood the verbal commands and their operational implications more readily, which increased the potential of verbal commands's in capturing attention (Figure 2).

Last, but not least, the P3b component reflects the updating of one's mental representation of relevant information [14]. P3b amplitudes are larger when a task-relevant event occurs that is different from one's expectation. For this reason, it is believed to underlie context-updating [57]. Related to this, P3b amplitudes have also been used to evaluate for working memory load or mental workload [6]. Larger P3b are associated with low mental workload and smaller P3b, with high mental workload. In the current context, this would suggest that auditory icons are more memorable and result in stronger context-updating than verbal commands, in a way that requires significantly less mental effort.

Auditory displays are designed to capture attention and to clearly communicate events. On the one hand, notifications that are readily recognized as task-relevant targets and capture attention are necessary for urgent events. On the other hand, notifications to indicate changing circumstances are also required to assist in updating a user's situational awareness. The current EEG/ERP results indicate that verbal commands are more discriminable and better at capturing attention than auditory icons. This suggests that verbal commands should be used in critical situations where quick action is required. Previous research based on behavioral results are in agreement with our conclusion. For example, verbal notifications are claimed to be especially effective in stressful situations because speech is processed automatically [24]. The current EEG/ERP results also suggest that auditory icons result in less effortful context-updating than verbal commands. Hence, auditory icons appear to be more suitable in communicating environmental circumstances that are less urgent. [17, 32] have suggested using auditory icons as notifications that inform and advise about background events. Their recommendation based on behavioral results agree with our current findings. It is worth noting that auditory icons are also believed to produce higher compliance levels than verbal commands, if they are

understood in the first place [16]. In conclusion, our results advocate that different types of notification should be designed in accordance to their intended purpose. This agrees with previous findings that, until now, have been mixed due to the imprecision of behavioral results in discriminating for how these notifications might be processed by the brain.

Limitations of the Current Study

One might argue that we only observe the changes in the ERPs because the neural origins of auditory icons and verbal commands are possibly oriented differently. This is a known limitation of ERP analysis, namely that it has limited spatial resolution for localizing brain regions that give rise to detected activity. However, this line of argument is unlikely. Previous work that relied on neuroimaging with better spatial resolution (i.e., fMRI) have demonstrated that words and object sounds involve the same neural region for processing information content [12]. In addition, verbal commands and auditory icons produce equivalent scalp topographies [9], which we have also observed in the current experimental paradigm. Therefore, it is unlikely that we observed the current results because verbal commands and auditory icons were not processed by identical neural regions.

The differences between the participant groups of Experiments 1 and 2 were intended to reflect the different stages of notification development, namely design and prototyping (Experiment 1) and validation (Experiment 2). Nonetheless, some differences between the participant groups might be of concern. In particular, the student volunteers in Experiment 1 were not proficient in the verbal commands. In spite of this, we note that their brain responses to verbal commands and auditory icons replicated in Experiment 2, which employed truck drivers. One reason is that the native language of Experiment 1's participants was highly similar to Experiment 2's participants. Another reason is that the current task focused on the participants' ability to recognize notifications and to respond appropriately, regardless of the extent to which they understood the implications of the notifications. Professional truck drivers who understood the language and the implications of the notifications showed stronger neural discriminations between verbal commands and auditory icons. Nonetheless, we can argue that the trends observed in Experiment 1 are likely to be generalizable trends while those in Experiment 2 are cultural and profession specific. The current study does not directly compare the two participant groups for their discrimination ability of the given sounds. Therefore, we do not make any inferences concerning either group's proficiency in discriminating the notifications from one another. The focus of this work is in evaluating how verbal commands and auditory icons are responded to at the level of information processing (i.e., brain responses).

CONCLUSION AND FUTURE WORK

Taken together, the current work contributes by showing that auditory notifications can be evaluated and functionally discriminated for how they are processed by the brain for information. This has implications for the operational context as well as design. Choices for which notifications to use for which purpose can be based not only in terms of response

times and discrimination accuracy, which is not necessarily the operational objective, but in terms of how the notifications are: (1) detected against the auditory scene, (2) discriminated against other notification targets, (3) likely to capture attention, and (4) capable of updating contextual working memory.

To date, most studies have questioned whether verbal commands or auditory icons serve better as notifications, namely in terms of how well they elicit a speeded and accurate response. The current findings suggest that this question, while well-intentioned, is misplaced. Our results demonstrate that verbal commands and auditory icons have different qualities. While verbal commands are better discriminated against other notifications, auditory icons can update contextual working memory with less effort. Practically speaking, this suggests that verbal icons are ideally used for time-critical information where there is no leeway for ambiguity, e.g., collision warnings. Meanwhile, auditory icons are likely to be more effective in communicating contextual information, such as entry into a poorly maintained road section or changing weather conditions. In other words, verbal commands and auditory icons should be used as complementary (and not competing) notifications.

Previous research has recommended using auditory icons to notify users of environmental events [17, 32]. More specifically, auditory icons have been suggested to enhance situational awareness [1, 31]. For example, a walking sound can more effectively indicate a nearing pedestrian. In addition, auditory icons might be favored because it is believed that they can be processed in parallel to other auditory events [1]. These findings so far converge with our current results and interpretation. Nonetheless, there are works that do not. For example, contrary to our current believe, that verbal commands capture attention, some work have shown that certain auditory icons (i.e., car horn) result in significantly faster response times (e.g., [24]). We might account for this by the fact that some auditory icons are overlearned to indicate danger. It should be noted that verbal processing is known to differ for different word classes (i.e., verbs, nouns) [58, 67]. The current study only uses nouns for verbal commands and, thus, future studies should verify whether verbal commands attract attention preferentially for all word classes, relative to auditory icons. In this work, we present EEG/ERP evidence that discriminates for how auditory icons and verbal commands are processed by the brain. Nonetheless, we do not doubt that nuances in how auditory notifications are engineered could ultimately render an auditory icon attention-grabbing and/or a verbal command more suited for communicating context. Our current results contribute by providing a starting point for understanding what type of sounds ought to be employed for which purposes, bearing in mind the brain's likely response to them.

The participants in Experiment 1 possessed neither a language proficiency for the verbal commands nor an expert understanding of the operational tasks that the notifications indicated. Therefore, the EEG/ERP differences (i.e., P2, P3b) found between verbal commands and auditory icons can be considered as general differences between the two notification

classes. In contrast, Experiment 2 was performed on professional truck drivers in a highly realistic test environment. A comparison between the two experiments reveals that these differences in brain responses scale with realism and user proficiency. Thus, the current approach of evaluating notification designs on the basis of brain responses is robust, even when behavioral responses do not differ. Notifications that are first designed in sterile lab environments could also be evaluated for the EEG/ERP responses that they elicit. This would narrow down the candidates for deployment and validation in high fidelity simulation environments or field-testing. Besides this, EEG/ERP methods could also be used to discriminate between different instantiation of the same target notification. One example would be to determine the preferability of semantically comparable verbal commands, such as *tank* or *fuel*.

To conclude, the current work suggests that verbal commands and auditory icons serve different purposes, at least from the standpoint of how they are processed by the brain. Thus, evaluations that directly compare them in terms of performance measures might not be appropriate. This might also explain the mixed evidence from previous studies in support of either auditory notifications. The growing accessibility of brain recording methods (i.e., EEG) mean that the current approach can be used to support finer functional discriminations for notifications and can be effectively deployed, even in challenging deployment scenarios such as high fidelity truck simulators.

ACKNOWLEDGMENTS

We thank the reviewers for their valuable feedback, which was helpful in revising the paper. We also thank K-Marie Lahmer and Rickard Leandertz for their assistance in data collection, and BrainProducts GmbH (Munich, Germany) for loaning us the necessary equipment for this study. This work was supported by the German Research Foundation through SFB/Transregio 161 projects, as well as by Scania CV AB, Sweden.

REFERENCES

1. Matt Adcock and Stephen Barrass. 2004. Cultivating Design Patterns for Auditory Displays. In *Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display*. 4–7.
2. Claude Alain. 2007. Breaking the wave: effects of attention and learning on concurrent sound perception. *Hearing research* 229, 1 (2007), 225–236.
3. Steven M Belz, Gary S Robinson, and John G Casali. 1999. A new class of auditory warning signals for complex systems: Auditory icons. *Human Factors* 41, 4 (1999), 608–618. DOI: <http://dx.doi.org/10.1518/001872099779656734>
4. Yoav Benjamini and Yosef Hochberg. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)* (1995), 289–300.
5. David H Brainard. 1997. The Psychophysics Toolbox. *Spatial vision* 10, 4 (1997), 433–436. DOI: <http://dx.doi.org/10.1163/156856897X00357>
6. Anne-Marie Brouwer, Maarten A Hogervorst, Jan BF Van Erp, Tobias Heffelaar, Patrick H Zimmerman, and Robert Oostenveld. 2012. Estimating workload using EEG spectral power and ERPs in the n-back task. *Journal of neural engineering* 9, 4 (2012), 045008.
7. Lewis Chuang, Christiane Glatz, and Stas Krupenia. 2017. Using EEG to Understand why Behavior to Auditory In-vehicle Notifications Differs Across Test Environments. In *Proceedings of the 9th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (Automotive'UI 17)*. ACM, New York, NY, USA.
8. Kate E Crowley and Ian M Colrain. 2004. A review of the evidence for P2 being an independent component process: Age, sleep and modality. *Clinical Neurophysiology* 115, 4 (2004), 732–744. DOI: <http://dx.doi.org/10.1016/j.clinph.2003.11.021>
9. A Cummings, Rita Čeponiene, Alain Koyama, Ayşe Pinar Saygin, Jeanne Townsend, and Frederic Dick. 2006. Auditory semantic networks for words and natural sounds. *Brain Research* 1115, 1 (2006), 92–107. DOI: <http://dx.doi.org/10.1016/j.brainres.2006.07.050>
10. Arnaud Delorme and Scott Makeig. 2004. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods* 134, 1 (2004), 9–21.
11. Arnaud Delorme, Jason Palmer, Julie Onton, Robert Oostenveld, and Scott Makeig. 2012. Independent EEG sources are dipolar. *PloS one* 7, 2 (2012), e30135.
12. Frederic Dick, Ayşe Pinar Saygin, Gaspare Galati, Sabrina Pitzalis, Simone Bentrovato, Simona D'Amico, Stephen Wilson, Elizabeth Bates, and Luigi Pizzamiglio. 2007. What is Involved and What is Necessary for Complex Linguistic and Nonlinguistic Auditory Processing: Evidence from Functional Magnetic Resonance Imaging and Lesion Data. *Journal of Cognitive Neuroscience* 19, 5 (2007), 799–816. DOI: <http://dx.doi.org/10.1162/jocn.2007.19.5.799>
13. Tilman Dingler, Jeffrey Lindsay, and Bruce N Walker. 2008. Learnability of Sound Cues for Environmental Features: Auditory Icons, Earcons, Spearcons, and Speech. *14th International Conference on Auditory Display* (2008), 1–6.
14. Emanuel Donchin and Michael G H Coles. 1988. Is the P300 component a manifestation of context updating? *Behavioral and brain sciences* 11, 3 (1988), 357–374.
15. Connie C Duncan-Johnson and Emanuel Donchin. 1977. Effects of a priori and sequential probability of stimuli on event-related potential. In *Psychophysiology*, Vol. 14. Cambridge Univ Press 40 West 20th Street, New York, NY 10011-4211, 95–95.

16. Judy Edworthy. 1994. The design and implementation of non-verbal auditory warnings. *Applied Ergonomics* 25, 4 (1994), 202–210. DOI: [http://dx.doi.org/10.1016/0003-6870\(94\)90001-9](http://dx.doi.org/10.1016/0003-6870(94)90001-9)
17. Judy Edworthy and Rachael Hards. 1999. Learning auditory warnings: The effects of sound type, verbal labelling and imagery on the identification of alarm sounds. *International Journal of Industrial Ergonomics* 24, 6 (1999), 603–618.
18. Judy Edworthy and Elizabeth Hellier. 2006. Alarms and human behaviour: Implications for medical alarms. *British Journal of Anaesthesia* 97, 1 (2006), 12–17. DOI: <http://dx.doi.org/10.1093/bja/ae1114>
19. Carles Escera, Kimmo Alho, Erich Schröger, and István Winkler Winkler. 2000. Involuntary attention and distractibility as evaluated with event-related brain potentials. *Audiology and Neurotology* 5, 3-4 (2000), 151–166.
20. Monica Fabiani, Demetrios Karis, and Emanuel Donchin. 1986. P300 and recall in an incidental memory paradigm. *Psychophysiology* 23, 3 (1986), 298–308.
21. Johan Fagerlönn, Stefan Lindberg, and Anna Sirkka. 2015. Combined Auditory Warnings For Driving-Related Information. In *Proceedings of the Audio Mostly 2015 on Interaction With Sound (AM '15)*. ACM, New York, NY, USA, Article 11, 5 pages. DOI: <http://dx.doi.org/10.1145/2814895.2814924>
22. Luis García-Larrea, Anne Claire Lukaszewicz, and François Mauguière. 1992. Revisiting the oddball paradigm. Non-target vs neutral stimuli and the evaluation of ERP attentional effects. *Neuropsychologia* 30, 8 (1992), 723–741. DOI: [http://dx.doi.org/10.1016/0028-3932\(92\)90042-K](http://dx.doi.org/10.1016/0028-3932(92)90042-K)
23. William Gaver. 1989. The SonicFinder: An Interface That Uses Auditory Icons. *Human-Computer Interaction* 4, 1 (1989), 67–94. DOI: http://dx.doi.org/10.1207/s15327051hci0401_3
24. Robert Graham. 1999. Use of auditory icons as emergency warnings: evaluation within a vehicle collision avoidance application. *Ergonomics* 42, 9 (Sept. 1999), 1233–48. DOI: <http://dx.doi.org/10.1080/001401399185108>
25. Robert Graham, S J Hirst, and C Carter. 1995. Auditory icons for collision-avoidance warnings. In *Intelligent Transportation: Serving the User Through Deployment. Proceedings of the 1995 Annual Meeting of ITS America*.
26. Rob Gray. 2011. Looming Auditory Collision Warnings for Driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 53, 1 (2011), 63–74. DOI: <http://dx.doi.org/10.1177/0018720810397833>
27. David M Groppe, Thomas P Urbach, and Marta Kutas. 2011. Mass univariate analysis of event-related brain potentials/fields I: a critical tutorial review. *Psychophysiology* 48, 12 (Dec. 2011), 1711–25. DOI: <http://dx.doi.org/10.1111/j.1469-8986.2011.01273.x>
28. Ellen Haas and Jeffrey Schmidt. 1995. Auditory icons as warning and advisory signals in the US Army Battlefield Combat Identification System (BCIS). In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, Vol. 39. SAGE Publications Sage CA: Los Angeles, CA, 999–1003.
29. Cristy Ho and Charles Spence. 2005. Assessing the effectiveness of various auditory cues in capturing a driver's visual attention. *Journal of Experimental Psychology: Applied* 11, 3 (2005), 157–74. DOI: <http://dx.doi.org/10.1037/1076-898X.11.3.157>
30. Cristy Ho and Charles Spence. 2006. Verbal interface design: Do verbal directional cues automatically orient visual spatial attention? *Computers in Human Behavior* 22, 4 (2006), 733–748. DOI: <http://dx.doi.org/10.1016/j.chb.2005.12.008>
31. Mandana L N Kazem, Janet M Noyes, and Nicholas J Lieven. 2003. Design Considerations for a Background Auditory Display to Aid Pilot Situation Awareness. In *Proceedings of the 2003 International Conference on Auditory Display*. 6–9.
32. Peter Keller and Catherine Stevens. 2004. Meaning From Environmental Sounds: Types of Signal-Referent Relations and Their Effect on Recognizing Auditory Icons. *Journal of Experimental Psychology: Applied* 10, 1 (2004), 3–12. DOI: <http://dx.doi.org/10.1037/1076-898X.10.1.3>
33. Mario Kleiner, David Brainard, Denis Pelli, Allen Ingling, Richard Murray, and Christopher Broussard. 2007. What 's new in Psychtoolbox-3 ? *Perception* 36, 14 (2007), 1.
34. Sonja A Kotz. 2013. Electrophysiological Indices of Speech Processing. In *Encyclopedia of Computational Neuroscience*. 1–5. DOI: <http://dx.doi.org/10.1007/978-1-4614-7320-6>
35. Nina Kraus and Trent Nicol. 2008. Auditory evoked potentials. In *Encyclopedia of Neuroscience*. Springer, 214–218.
36. Stas Krupenia, Anna Selmarker, Johan Fagerlönn, Katarina Delsing, Anders Jansson, Bengt Sandblad, and Camilla Grane. 2014. The Methods for Designing Future Autonomous Systems' (MODAS) project: Developing the cab for a highly autonomous truck. In *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics (AHFE2014)* (Krakow, Poland). 19–23.
37. Yi-Chieh Lee, Wen-Chieh Lin, Jung-Tai King, Li-Wei Ko, Yu-Ting Huang, and Fu-Yin Cherng. 2014. An EEG-based approach for evaluating audio notifications under ambient sounds. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14* (2014), 3817–3826. DOI: <http://dx.doi.org/10.1145/2556288.2557076>

38. Ying K Leung, Sean Smith, Simon Parker, and Russell Martin. 1997. Learning and retention of auditory warnings. *Proceedings of the Third International Conference on Auditory Display* (1997).
39. Mats Liljedahl and Johan Fagerlönn. 2010. Methods for sound design: a review and implications for research and practice. In *Proceedings of the 5th Audio Mostly Conference: A Conference on Interaction with Sound*. ACM, 2.
40. Paul A Lucas. 1994. An evaluation of the communicative ability of auditory icons and earcons. Georgia Institute of Technology.
41. Steven J Luck. 2005. An Introduction to the Event-Related Potential Technique (Cognitive Neuroscience). (2005).
42. Dawn C Marshall, John D Lee, and P Albert Austria. 2007. Alerts for in-vehicle information systems: Annoyance, urgency, and appropriateness. *Human factors* 49, 1 (2007), 145–157.
43. David K McGookin and Stephen A Brewster. 2004. Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition. *ACM Transactions on Applied Perception (TAP)* 1, 2 (2004), 130–155.
44. Denis McKeown. 2005. Candidates for within-vehicle auditory displays. Georgia Institute of Technology.
45. Elizabeth D Mynatt. 1994. Designing with auditory icons: how well do we identify auditory cues?. In *Conference companion on Human factors in computing systems*. ACM, 269–270.
46. Michael A Nees and Bruce N Walker. 2011. Auditory displays for in-vehicle technologies. *Reviews of human factors and ergonomics* 7, 1 (2011), 58–99.
47. Gerald Novak, Walter Ritter, and Herbert G Vaughan. 1992. Mismatch detection and the latency of temporal judgments. *Psychophysiology* 29, 4 (1992), 398–411.
48. Arne Nykänen. 2008. *Methods for product sound design*. Ph.D. Dissertation. Luleå tekniska universitet.
49. Casey O’Callaghan. 2009. Auditory perception. (2009).
50. Eunmi L Oh and Robert A Lutfi. 1999. Informational masking by everyday sounds. *The Journal of the Acoustical Society of America* 106, 6 (1999), 3521–3528. DOI : <http://dx.doi.org/10.1121/1.428205>
51. Guido Orgs, Kathrin Lange, Jan Henryk Dombrowski, and Martin Heil. 2006. Conceptual priming for environmental sounds and words: An ERP study. *Brain and Cognition* 62, 3 (2006), 267–272. DOI : <http://dx.doi.org/10.1016/j.bandc.2006.05.003>
52. Denis G Pelli. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. (1997). DOI : <http://dx.doi.org/10.1163/156856897X00366>
53. Caterina Piazza, Makoto Miyakoshi, Zyenab Akalin-Acar, Chiara Cantiani, Gianluigi Reni, Anna Maria Bianchi, and Scott Makeig. 2016. An Automated Function for Identifying EEG Independent Components Representing Bilateral Source Activity. *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016, IFMBE Proceedings* 57 (2016), 105–109. DOI : <http://dx.doi.org/10.1007/978-3-319-32703-7>
54. Terence W Picton. 2010. *Human auditory evoked potentials*. Plural Publishing.
55. Terence W Picton. 2014. Auditory event-related potentials. *Encyclopedia of Computational Neuroscience* (2014), 1–6.
56. Terence W Picton and Steven A Hillyard. 1974. Human auditory evoked potentials. II: effects of attention. *Electroencephalography and clinical Neurophysiology* 36 (1974), 191–199. DOI : [http://dx.doi.org/10.1016/0013-4694\(74\)90156-4](http://dx.doi.org/10.1016/0013-4694(74)90156-4)
57. John Polich. 2007. Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology* 118 (2007), 2128–2148. DOI : <http://dx.doi.org/10.1016/j.clinph.2007.04.019>
58. Friedemann Pulvermüller. 1999. Words in the brain’s language. *Behavioral and Brain Sciences* 22, 1999 (1999), 253–336. DOI : <http://dx.doi.org/10.1017/S0140525X9900182X>
59. Brock R Riggins and John Polich. 2002. Habituation of P3a and P3b from visual stimuli. *The International Journal of Creativity & Problem Solving* 12, 1 (2002), 71–81.
60. Ayşe Pinar Saygin, Frederic Dick, and Elizabeth Bates. 2005. An on-line task for contrasting auditory processing in the verbal and nonverbal domains and norms for younger and older adults. *Behavior Research Methods* 37, 1 (2005), 99–110.
61. Ayşe Pinar Saygin, Frederic Dick, Stephen W Wilson, Nina F Dronkers, and Elizabeth Bates. 2003. Neural resources for processing language and environmental sounds: Evidence from aphasia. *Brain* 126, 4 (2003), 928–945. DOI : <http://dx.doi.org/10.1093/brain/awg082>
62. Antoine Shahin, Daniel J Bosnyak, Laurel J Trainor, and Larry E Roberts. 2003. Enhancement of neuroplastic P2 and N1c auditory evoked potentials in musicians. *Journal of Neuroscience* 23, 13 (2003), 5545–5552.
63. Carol A Simpson and Kristine Marchionda-Frost. 1984. Synthesized speech rate and pitch effects on intelligibility of warning messages for pilots. *Human factors* 26, 5 (1984), 509–517. DOI : <http://dx.doi.org/10.1177/001872088402600503>
64. Charles Spence and Liliana Read. 2003. Speech shadowing while driving: On the difficulty of splitting attention between eye and ear. *Psychological science* 14, 3 (2003), 251–256.

65. Mitchell Steinschneider and Michelle Dunn. 2002. Electrophysiology in developmental neuropsychology. *Handbook of neuropsychology* 8, 1 (2002), 91–146.
66. David L Strayer and William A Johnston. 2001. Driven to Distraction: Dual-Task Studies of Simulated Driving and Conversing on a Cellular Telephone. *Psychological Science* 12, 6 (2001), 462–466. DOI: <http://dx.doi.org/10.1111/1467-9280.00386>
67. Anna Szekely, Simonetta D’Amico, Antonella Devescovi, Kara Federmeier, Dan Herron, Gowri Iyer, Thomas Jacobsen, L Arévalo Anal’a, Andras Vargha, and Elizabeth Bates. 2005. Timed action and object naming. *Cortex* 41, 1 (2005), 7–25.
68. Tuyen V Tran, Tomasz Letowski, and Kim S Abouchacra. 2000. Evaluation of acoustic beacon characteristics for navigation tasks. *Ergonomics* 43, 6 (2000), 807–827.
69. Kai Tuuri, Manne-Sakari Mustonen, and Antti Pirhonen. 2007. Same sound–different meanings: A novel scheme for modes of listening. *Proceedings of Audio Mostly* (2007), 13–18.
70. Pernilla Ulfvengren. 2003. *Design of natural warning sounds in human-machine systems*. Ph.D. Dissertation. KTH.
71. Bruce N Walker and Michael A Nees. 2011. Theory of sonification. *The sonification handbook* (2011), 9–39.
72. Emily E Wiese and John D Lee. 2007. Attention grounding: a new approach to in-vehicle information system implementation. *Theoretical Issues in Ergonomics Science* 8, 3 (2007), 255–276.
73. Istvan Winkler, Susan L Denham, and Carles Escera. 2013. Auditory Event-related Potentials. In *Encyclopedia of Computational Neuroscience*. 1–29. DOI: <http://dx.doi.org/10.1007/978-1-4614-7320-6>
74. David L Woods. 1995. The component structure of the N1 wave of the human auditory evoked potential. *Electroencephalography and Clinical Neurophysiology-Supplements Only* 44 (1995), 102–109.