Sound Decisions: How Synthetic Motor Sounds Improve Autonomous Vehicle-Pedestrian Interactions

Dylan Moore
djmoore3@alumni.stanford.edu
Center for Design Research,
Stanford University
Stanford, CA, USA

Rebecca Currano
bcurrano@stanford.edu
Center for Design Research,
Stanford University
Stanford, CA, USA

David Sirkin
sirkin@stanford.edu
Center for Design Research,
Stanford University
Stanford, CA, USA

ABSTRACT

Electric vehicles' (EVs) nearly silent operation has proved to be dangerous for bicyclists and pedestrians, who often use an internal combustion engine's sound as one of many signals to locate nearby vehicles and predict their behavior. Inspired by regulations currently being implemented that will require EVs and hybrid vehicles (HVs) to play synthetic sound, we used a Wizard-of-Oz AV setup to explore how adding synthetic engine sound to a hybrid autonomous vehicle (AV) will influence how pedestrians interact with the AV in a naturalistic field study. Pedestrians reported increased interaction quality and clarity of intent of the vehicle to yield compared to a baseline condition without any added sound. These findings suggest that synthetic engine sound will not only be effective at helping pedestrians to hear EVs, but also may help AV developers implicitly signal to pedestrians when the vehicle will yield.

CCS CONCEPTS

• **Human-centered computing → Interaction design theory, concepts and paradigms**; **Empirical studies in interaction design**; *Interface design prototyping*.

KEYWORDS

Pedestrian interaction, Autonomous vehicles, Sound design, Driverless cars, Ghostdriver, Wizard-of-Oz, External human-machine interfaces, Implicit interaction

ACM Reference Format:

Dylan Moore, Rebecca Currano, and David Sirkin. 2020. Sound Decisions: How Synthetic Motor Sounds Improve Autonomous Vehicle-Pedestrian Interactions. In *12th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '20), September 21–22, 2020, Virtual Event, DC, USA.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3409120.3410667

1 INTRODUCTION

Current human expectations of technology have been, to a surprising degree, inspired by film and television. Modern robot speech has been inspired by film characters such as R2D2 and Wall-E [39, 40, 54]. And the discomfort many people feel towards robots has likely been fueled by one too many dystopian Hollywood narratives [7].

Figure 1: We explored how EV motor sound influenced AV-pedestrian interactions in a naturalistic field study.

Films have also recently inspired another aspect of interaction with technology: synthetic vehicle motor sound. While no vehicle to date has reproduced the iconic Jetsons "putt putt" flying car sound, BMW has hired Hans Zimmer, who scored blockbuster films such as *Inception*, to design the sounds for its electric concept cars [25].

While the environmental benefits of electric vehicles (EVs) are numerous, their distinctly quiet motor sound compared with conventional internal combustion engines has proven problematic for vulnerable road users (VRUs), particularly visually impaired pedestrians. EVs are 10% more likely to be involved in pedestrian incidents and 51% more likely to be involved in bicycle accidents than combustion engine cars [5]. As a result, EVs and hybrid vehicles (HVs) in the US and the EU will be required to augment their motor sound [5, 15], which has shown to increase pedestrians' awareness [6, 17, 53]. Musicians are often enlisted to design these sounds. However, people's reactions to the real-world implementation of EV motor sound has not always been positive, as sometimes the sound is seen as too different or artful to be practical. We believe this problem could be minimized by involving users more integrally in the design process.

Analogous issues arise with the introduction of autonomous vehicles (AVs), where the absence of a driver removes potential signals used by VRU's to predict vehicle behavior [14]. Many researchers have explored alternative external human-machine interfaces (eHMI) including lights [13], displays [4, 9, 12, 19], auditory alerts [48], and projections [48]. However, such interfaces can be ignored [9], or over-relied on [22], and explicit signals such as

AutomotiveUI '20, September 21–22, 2020, Virtual Event, DC, USA

ACM ISBN 978-1-4503-8065-2/20/09…$15.00
https://doi.org/10.1145/3409120.3410667

recorded voices or alerts may be too jarring for pedestrians. Augmenting an AV's motor sound may provide an appropriate, subtle implicit eHMI signal that emphasizes the vehicle's natural slowing motion (already an implicit eHMI [28]) to VRUs without demanding attention.

Little research has been done to study how motor sound could (or will) function as an implicit eHMI for an AV in a naturalistic setting. As such, our contribution is twofold: (1) we provide methods to systematically evaluate candidate synthetic EV and HV motor sounds using online surveys (rather than simply relying on the designer's intuition), and (2) we show that synthetic motor sounds improve AV-pedestrian interactions in a naturalistic field study. This will help automotive designers easily gather user feedback on candidate motor sounds, and support people designing implicit eHMI signals such as an AV communicating its intent to yield.

2 RELATED WORK

The intentional design of vehicle motor sounds extends beyond popular media, and has been used to alter audible frequencies inside and outside a vehicle's cabin [44]. Harley Davidson even applied to trademark its V-Twin motor sound in 1994 [43], although the effort was dropped in 2000. But what about EVs, which on their own, generate very little motor sound?

2.1 Regulations for EV and HV motor sound

To ensure that synthetic motor sounds can be heard within the din of the urban environment, US regulations stipulate that they should include either two or four tones distributed across third-octave band frequencies from 315 Hz to 5 kHz, played while the vehicle moves at less than 30 km/h [34]. The tones must either be very loud when the vehicle is idle, or increase in volume as its speed increases. US regulations do not require that frequency to change dynamically with vehicle speed, although it was considered in depth [11], originally suggesting an increase of 1% per 1 km/h increase in speed. Existing rules require all cars of a particular make and model to sound the same, though there is ongoing discussion on allowing users to personalize their vehicle's sounds [21].

The EU's Acoustic Vehicle Alerting System (AVAS) regulations require EVs to play a 56 dB sound when traveling below 20 km/h or in reverse [3].

2.2 Design space for electric vehicle motor sound

Based on his experience testing electric vehicle sounds for BMW, Norman [35] offered four design principles for the design of EV motor sound:

- Indicate the presence of a vehicle
- Enable someone to locate the vehicle and determine its relative speed
- Minimize disruptive qualities, as the sound will be heard frequently
- Standardize sounds so people can identify EVs, while also allowing for car brand expression

Design firm usTwo mirrored Norman's approach in prototyping auditory alerts for vehicle exteriors, identifying three guiding principles in order of importance: (1) Safety of pedestrian, (2) Minimal noise pollution, (3) Brand expression [52]. In addition, they believed that the audio system could be context-aware, adapting to different locations, vehicle actions, road users, and environmental conditions in the surrounding area.

Automotive manufacturers have approached EV sound design by partnering either with musicians, as BMW did with Hans Zimmer [25], or with sound design firms. Nissan worked with the firm Man Made Music to produce a sound concept called "Canto", which is a stylized add-9 chord [32] . Jaguar turned to Atlanta-based musician Richard Devine [1], to create an "elegant, classy, and modern" sound. Jaguar noted that "Initial attempts to create a noise inspired by the sound of sci-fi spacecraft had to be shelved after pedestrians reacted by looking up to the sky, rather than at the road" [2].

Gathering user feedback when designing sound is critical, as sonic experiences are highly subjective. Transport for London retained a new design firm after reports of the prototype sounds for its electric buses being too futuristic [37]. usTwo [52] conducted a virtual reality experiment of different auditory samples, finding that sudden changes in sound communicated risk, and that melodic sound was distracting. Petiot et al. [36] found that users preferred an EV to sound like a conventional car, or to make no sound at all. Swart [47] found that variation in user evaluation of EV motor sounds was largely explained by the terms "Power", "Comfort", and "Futuristic".

2.3 Motor sound design in other contexts

Motor sounds are a common example of *consequential sound*, or sound made as a result of a system's movement or operation. This is distinct from *intentional sound* such as a beep or phone notification, which are intentional noises to communicate something about the system [23]. Consequential sound has been thoroughly studied in product sound design, demonstrating that experiences such as vacuuming [26], eating a potato chip [46], boiling water in a kettle [16], and popping a soda can [45] are all influenced by the sounds that products make.

The human-robot interaction community has explored how servomotor sound characteristics can influence how people perceive robots [18, 29, 31, 49]. Others have shown that low frequency *infrasound* can emphasize a robot's expression [50], and how the spacial distribution of sound can influence interaction [41]. Moreover, consequential sound is an importance source of robot localization [8]. These researchers have used a combination of online [29, 31, 49] and in-person [8, 18] study methods to tease out small differences in sound that can naturally be adapted to the EV sound design space.

Automotive manufacturers are no stranger to manipulating sound. For instance, BMW has played a vehicle's motor sound through its interior speakers to enhance its perceived performance [20]. This relates to the concept of *blended sonification* [51], where sound changes based on the desired interaction. For instance, a door knock may sound differently depending on whether or not the room being visited is occupied.

Figure 2: The 14 sounds chosen to explore in Survey 1. Each sound is shown with its original key and transposed (if applicable) to ensure that each bass note was F2 at the start of the clip, shifting down one octave to F1 at the end.

3 DESIGNING SYNTHETIC MOTOR SOUND

To choose a sound for our on-road study, we conducted a two-part online study. We collected input on the acoustic characteristics and appropriateness of 14 sounds comprised of different chords. While these sounds are neither a comprehensive nor necessarily optimal set for all conditions, these pilot studies prototyped ways that EV designers may explore sound, and methods to gather user feedback that informed our final choice of sound for the on-road study.

3.1 Candidate sound generation

Given the immense design space for sound, we could not explore all dimensions, and needed to narrow our focus. Existing EV motor sounds are typically comprised of multiple notes (chords), and each individual note can be defined by three primary dimensions: frequency, intensity, and timbre (or tone color), which are known to interact with one another [33]. Government regulations provide guidance on the frequency ranges and intensity of sounds, so we chose to focus our exploration on other dimensions that could vary across Evs. Timbre is a very complex and subtle dimension, so we chose the most direct approach and chose a timbre resembling an electric motor.

This left us to explore harmonic content and voicing of the motor sounds' chords. We benchmarked a number of different sounds used in existing EVs, such as the Nissan Canto sound [32], chords and voicings from classical music, and train horns. All sounds were tonal and arrhythmic, and chords with adjacent lower notes often created a slight pulsing, reminiscent of an engine. The set included major, minor, augmented, diminished, and open chord types. Each sound is written in musical notation in Figure 2, and the sound clips are also available as a video figure[1].

The sounds were recorded in Apple GarageBand using the SynthOne iPad app, built on the open source AudioKit platform[2] that we used later to generate sound for the on-road study. The following elements were held constant across all sounds:

- Timbre: Tone generators were set as square waves with a low-pass cutoff filter around 690 Hz, to mimic the sound of an electric motor (determined through significant trial and error of all synthesizer parameters)
- Frequency shift: A linear 1 octave shift over the 5 second clip of the sound—prior work has shown that a frequency shift emphasizes when a robot slows down on a sidewalk [30]
- Base note: At the end of the frequency shift, the lowest frequency in each candidate sound was F1 (43.65 Hz)

All sounds were recorded at the same volume. The one octave pitch shift roughly matched the frequency shift and timing of the vehicle slowing and stopping in the on-road study.

3.2 Survey 1 Setup

Our first objective was to evaluate appropriateness of the sounds to our application, and to understand underlying descriptors associated with more appropriate sounds.

[1]https://vimeo.com/439646142
[2]https://audiokit.io

3.2.1 Procedure. Participants first read instructions, listened to a test sound, and listened to a recording with an audio passcode to ensure that they had working audio. We asked participants to keep their audio level fixed while completing the survey. Each subsequent page played a candidate motor sound, and asked participants to imagine the sound coming from the car shown in an image (Figure 1) and rate how well the sound fit the car. Then, following the Check-All-That-Apply (*CATA*) protocol, participants selected words they associated with the given sound from a list of bipolar pairs: annoying/pleasant, rough/smooth, alarming/relaxing, weak/strong, happy/sad, intimidating/comforting, expensive/inexpensive, simple/complex. Participants evaluated all 14 candidate sounds. The last page collected general demographic questions.

3.2.2 Measures. The primary measures for Survey 1 were the following:

- *Appropriateness*: How well the particular sound fit the desired application (on a scale from 1 to 9).
- *CATA frequency*: How often a particular descriptive word was associated with a particular sound.

3.2.3 Participants. $N = 50$ members of Prolific, a crowdsourcing research platform similar to Amazon Mechanical Turk, participated in Survey 1. Participants were 68% male and 32% female, with a mean age of 31 years old. They were geographically spread between the United States (26%) and many other countries around the world. Participants received $1.10 for the estimated 10 minute survey (actual average completion time was 8.5 minutes). Nine participants were excluded for not using headphones as the instructions requested.

3.3 Survey 1 Results

3.3.1 Little variance in appropriateness. The mean appropriateness of all sounds hovered around the midpoint of the scale ($M = 4.4$ out of 9, $SD = 2.1$). However, individual scores ranged from 1 to 9 for each sound, suggesting substantial variation in individual preference. There were no statistically significant differences in the mean appropriateness of the sound to the car.

One trend was that of the seven major tonality chords, five were rated in the bottom half of the appropriateness measure. Conversely, all four of the sounds having open tonality were rated in the top half of the scale. While the lack of statistical significance implies that we cannot lean on these differences too much, the general trend suggested that chords with less distinct tonalities may fare better as EV motor sounds.

3.3.2 Correspondence analysis of CATA data. Following methods from online research evaluating motor sounds [29], we used correspondence analysis of the CATA data to identify general themes and descriptors that explain variance between candidate sounds (Figure 3). In a correspondence analysis, descriptors are more likely to be associated with the sound they are visually closest to. Looking at the layout of these descriptors provides insight into words that explained more variance across the sample. First, the horizontal dimension (explaining 31.7% of variance) was spanned by rough/strong and happy/comforting. The vertical dimension (explaining 18.5% of variance) was spanned by pleasant/alarming.

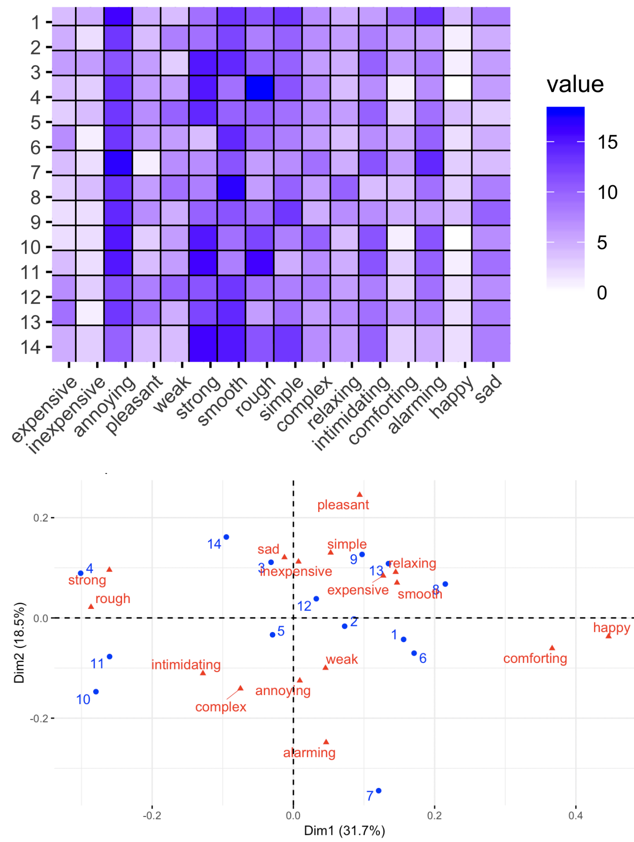

Figure 3: (Top) The contingency table of the CATA data show how often a descriptive word was associated with a given sound. Some descriptors, like "happy" were seldom selected. (Bottom) The Correspondence Analysis of the CATA data reveals some preliminary trends in the data, such as distinctions between alarming and pleasant sounds, and helps to visually identify sounds with desirable characteristics.

3.4 Survey 2 Setup

Several participants from Survey 1 mentioned that it was difficult to hear differences between sounds played one at a time. So for our next iteration, we chose to follow a pairwise comparison method previously used to study motor sound in an online setting [31].

To prevent survey fatigue, we downselected from 14 to 7 sounds, choosing the 7 sounds rated as most appropriate in Survey 1 (Sounds 3, 4, 5, 8, 12, 13, 14), which also were on the top half of the correspondence analysis (Figure 3), and were more often associated words with positive characteristics such as pleasant and smooth. We avoided sounds associated with undesirable characteristics of intimidating, complex, annoying, weak, and alarming.

3.4.1 Procedure. Participants received the introduction and context on EV sound as in Survey 1, and confirmed their working sound with an audio passcode. On each subsequent page, participants saw Figure 1 above pairs of sounds (labeled Sound A and B) and were asked to rate which sound they preferred to hear on an autonomous electric vehicle, using the following scale: "Definitely A", "Probably A", "No preference/Can't tell", "Probably B", "Definitely B". Participants saw all 21 possible pairs of sounds in a randomized order.

3.4.2 Measure. *Preference* for the sound, equal to the total number of points that a sound received, served as the primary metric in Survey 2. Sounds received points based on each response, with Definitely A/B receiving 2 points, Probably A/B receiving 1 point, and No preference/Can't tell receiving 0 points. Due to slightly incomplete data from participants skipping questions, scores were normalized to account for the number of times participants answered the question.

3.4.3 Participants. Participants were $N = 50$ members of Prolific who were 70% male and 30% female, with a mean age of 31 years. They were geographically spread between the United States (24%) and many other countries around the world. Participants received $1.11 for the estimated 10 minute survey (actual mean completion time was 8.5 minutes). Ten participants, who did not use headphones as the instructions requested, were excluded from the data.

3.5 Survey 2 Results

The sound with the most preference points after all comparisons (shown in Figure 4) was Sound 14: an open fifth chord (F-C-F)—it was not major, minor, augmented, diminished, or a cluster of notes. Sound 14 was not particularly distinctive in correspondence analysis. However, it was rated as the most appropriate sound in Survey 1.

While Sound 14 received 41% more points than the next highest rated sound, other samples also received a substantial number of points, thus this preference was not unanimous. Interestingly, the top three sounds were all relatively simple open chords, followed by two major chords, followed by more dissonant augmented and diminished chords.

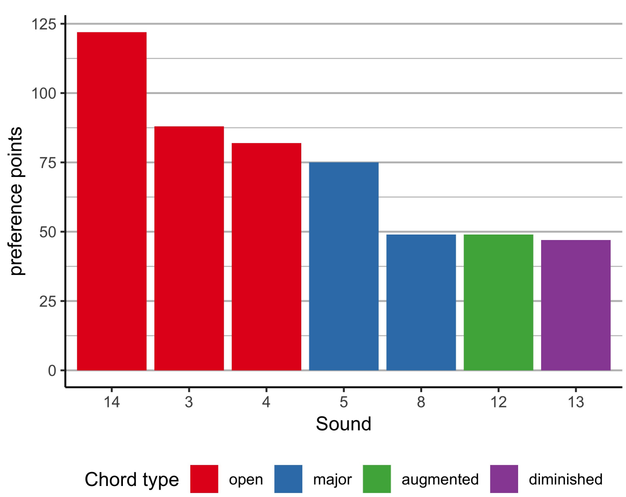

Figure 4: After tallying the pairwise ranking preference points, Sound 14 was the most preferred sound for our context.

While 50% of participants in Survey 2 reported being at least beginner musicians, both non-musicians' and musicians' ranked order preferences for the sound were the same overall.

3.6 Pilot Survey Summary

After generating 14 candidate sounds with a synthesizer and collecting user feedback in two rounds of online surveys, we arrived at a single motor sound to use in the on-road study: an open fifth chord with the notes F-C-F. While by no means a universally optimal sound, this two part evaluation process gave us confidence that we were choosing a sound that was appropriate to the context.

4 EVALUATING EV MOTOR SOUND ON THE ROAD WITH GHOSTDRIVER

With an appropriate synthetic motor sound prototyped, we could then implement the motor sound in an on-road study to evaluate how pedestrians reacted to the sound on an AV. To do so, we conducted a naturalistic field study using a Wizard-of-Oz AV setup. We added a Bluetooth speaker underneath the hood of the vehicle to play the synthetic motor sound as it drove around a fixed course, and then observed and interviewed pedestrians who interacted with the vehicle.

4.1 Simulated AV (Ghostdriver) Setup

Following the Ghostdriver protocol [10, 24, 28, 42], we equipped a hybrid vehicle (a 2014 Toyota Prius, shown in Figure 5) with cameras, fake LiDAR sensors, and decals reading "Stanford Driverless Vehicle". These props drew attention to the vehicle and increased the likelihood that pedestrians would perceive and interact with the vehicle as *driverless*, but did not hinder the goal of studying how an AV could communicate through EV sound.

As an HV, the Prius uses an electric motor at low speeds, so the combustion engine consistently turned off as the vehicle approached the crosswalk of interest. Thus, it served as a suitable platform to test the impact of synthetic motor sound on pedestrians' impressions of AVs.

A custom costume built out of a car seat cover (Figure 5) hid the driver so that from the outside, the car appeared to be driverless. Eight GoPro cameras captured each interaction from different angles: five captured the car's perspective and three captured the crosswalk of interest from the front and two sides.

Figure 5: (Left) The test vehicle (a Toyota Prius) mocked up with decals, fake LiDAR, and cameras. (Right) The Ghostdriver costume hid the driver from pedestrians.

4.2 Live motor sound implementation

To increase believability of the sound, we implemented a system to modulate the sound's frequency and volume in response to the vehicle's speed.

To do so, we connected an OBDLink scanner to the Prius' OBD2 port, which transmitted the car's speed via Wifi to an iPhone XS. This phone ran a custom iOS app built on the AudioKit platform that generated the F-C-F sound, and varied the frequency 3% in pitch per 1 km/hr of speed change.

The sound played out of a JBL Charge 2 Bluetooth Speaker taped underneath the hood of the Prius. The metal hood resonated making the sound's exact source difficult to localize, mimicking realistic motor sound.

After pilot testing on our course, we found that the low frequency range we used in the online studies was relatively difficult to distinguish from the surrounding road noise, so we increased the sound's fundamental frequency. This final sound had a base note of F4 (349.23 Hz) when the vehicle's speed was 0 km/h. The highest base note frequency was approximately 820 Hz at a speed of 45 km/h.

4.3 Study Protocol

We drove the Ghostdriver car in a loop between two roundabouts on a university campus, repeatedly passing through a crosswalk that did not have a stop sign to moderate traffic—the crosswalk shown in Figure 1. Researchers stood at each end of the crosswalk to conduct interviews with pedestrians who interacted with the Ghostdriver car.

If a pedestrian interacted with our vehicle and agreed to be interviewed, they were asked to describe their experience and answer both qualitative questions (about the vehicle's behavior and perceived autonomy) and quantitative questions (Likert bipolar word pairs evaluating the interaction on numerous dimensions on a scale from 1 to 5). If participants had not mentioned the vehicle's sound by the end of the interview, interviewers asked if they noticed anything about the vehicle's sound. The interview protocol is included in the ACM Digital Library.

4.4 Video analysis

We analyzed video of each interaction from both the car's perspective and the crosswalk perspective. The following variables helped to characterize the interactions:

- *Stopping distance*: The distance between the front bumper and the edge of the crosswalk at the point in time when the car reached its lowest speed[3]. To estimate this distance, we created a visual template by measuring the distance to landmarks in the scene from the camera's perspective on the side of the crosswalk.
- *Hesitation behavior*: Whether or not the pedestrian slowed down or stopped before crossing the street.
- *Arrival order*: Whether the car or pedestrian arrived first to the crosswalk. If it was too difficult to tell which interactant arrived first, we coded it as "same", however this was combined with the pedestrian group for statistical analysis.

[3]For example, when the car stopped, or when it began to crawl at a slow speed towards the crosswalk.

In addition to examining how pedestrians reacted to the Ghostdriver car, we also examined how they reacted to conventional cars passing through the same intersection. These cars were not driven by researchers, they just passed through the intersection and interacted with pedestrians while we recorded video of our experiment.

Video coding was completed in pairs by the authors and other researchers in our lab, and the scheme was modified and iterated on as we sifted through the video data set.

4.5 Measures of interaction

We conducted a principal component analysis of the survey items to group aligned responses into *interaction quality* and *intent*:

- *Interaction quality*: how comfortable and safe the pedestrian felt during the interaction (α = .77), from the following interview items:
  - Interaction was uncomfortable/comfortable
  - Interaction was unpleasant/pleasant
  - Interaction was unsafe/safe
- *Intent*: how well the car communicated that it would stop and patiently wait (α = .74), from the following interview questions:
  - Vehicle was aware of me/not aware of me
  - Vehicle wanted me to stop walking/keep walking
  - Vehicle wanted me to take my time/hurry up
  - I knew the vehicle would stop for me: Definitely yes/no
  - The vehicle was trustworthy/untrustworthy

We also considered qualitative feedback from interviewees on how they felt during the interaction, whether they perceived the vehicle as autonomous or driven by a human, and evaluated whether they hesitated before crossing in the video analysis.

4.6 Participants

We interacted with $N = 84$ pedestrians while playing the synthetic motor sound, gathering a total of $N = 43$ participant interviews. Participants ranged in age from 18 to 63 years old ($M = 26$). Participants were from the US (74%), Asia (9%), Europe (7%), or did not say (9%).

We compared these data to a baseline condition of an earlier iteration of the study which used an identical protocol, but did not incorporate synthetic motor sound [28]. In the baseline condition, we interacted with $N = 72$ pedestrians without any added motor sound, and conducted $N = 41$ interviews. We also compared video of pedestrians' reactions to the Ghostdriver car in this study to video of $N = 155$ interactions with conventional vehicles.

4.7 Results

In all conditions, 100% of pedestrians crossed in front of the Ghostdriver vehicle. To dive deeper into the impact of synthetic motor sound compared to the baseline condition, we analyzed data using linear models for continuous dependent variables and generalized linear models for binomial dependent variables. We used R Studio 1.1.419 and R version 3.5.1, [38] using the base `stats` package functions `lm()` and `glm()`.

4.7.1 Interaction quality higher than baseline. We fit a linear model to predict interaction quality from the factors condition (Ghostdriver baseline or Ghostdriver with sound), car stopping distance, and perceived vehicle operation (driven by a human, autonomously, or N/A). Given only 5 data points in the N/A category, we excluded those participants from this statistical analysis. Adding additional factors did not explain sufficient additional variance (after comparing models with ANOVA).

As shown in Figure 6, the sound condition saw an increased interaction quality, $b = 0.39,\ t = 2.49,\ p = .011$, while controlling for variance in car stopping distances, which were on average slightly farther away in the sound condition than the baseline condition. There was no significant difference between those who saw the car as controlled autonomously or controlled by a human before crossing, $b = 0.14,\ t = -0.84,\ p = .402$, nor did car stopping distance have a significant effect, $b = 0.01,\ t = 0.89,\ p = .378$.

4.7.2 Intent clearer than baseline. We fit a linear model to predict intent from the following factors: condition, car stopping distance, and perceived operator. We excluded those who did not report whether the car was controlled autonomously or by a human due to a small sample size in that group.

Also shown in Figure 6, mean ratings of intent were higher in the sound condition, $b = 0.49,\ t = 3.39,\ p = .001$. The car's stopping distance from the crosswalk was a significant factor, with intent being clearer the closer the car stopped to the crosswalk, $b = 0.04,\ t = 3.15,\ p = .002$. There were no significant differences between groups who saw the car as controlled autonomously or by a human, $b = -0.13,\ t = -0.94,\ p = .349$.

4.7.3 No significant difference in hesitation between all conditions. We fit a generalized linear model to predict likelihood of hesitation from the following factors: condition (Ghostdriver baseline, Ghostdriver sound, *and conventional cars*), and arrival order when the vehicle came to a full stop.

As shown in Figure 7, pedestrians were significantly less likely to hesitate before crossing when the car stopped before they entered the crosswalk, $b = -3.01,\ t = -4.84,\ p < .0001$. As fewer than 3% of people hesitated before crossing in front of any car

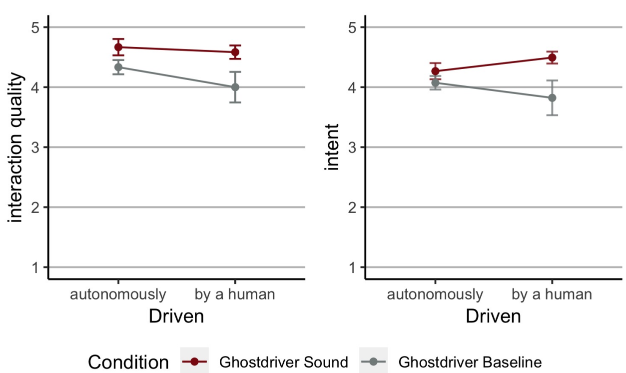

Figure 6: Interaction quality (left) and intent (right) were higher in the sound condition compared to the baseline condition. Error bars represent +/- 1 standard error.

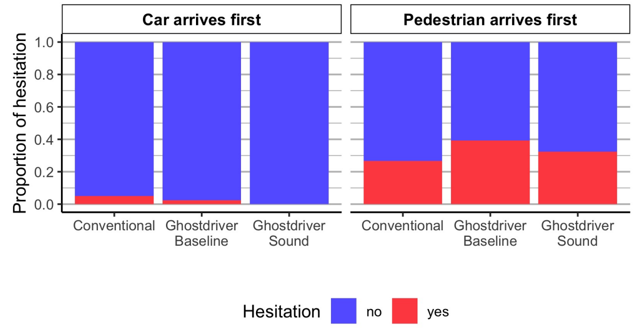

Figure 7: Pedestrians were significantly more likely to hesitate before crossing if they arrived at the intersection before the approaching car stopped. Pedestrians were slightly more likely to hesitate before crossing in front of the Ghostdriver car than conventional cars, though this was somewhat ameliorated by adding the electric vehicle motor sound.

(Conventional or Ghostdriver) when the car arrived first, we can zoom in at the subset of instances where pedestrians arrived at the intersection first. Pedestrians hesitated before 27% of conventional cars compared to 32% in the Ghostdriver sound condition, however this difference is not significant, $b = -0.03,\ z = -0.06,\ p = .956$. When comparing these to pedestrians hesitating before crossing in front of the baseline Ghostdriver car 39% of the time, this difference is also not significant, $b = 0.36,\ z = 0.73,\ p = .469$.

4.7.4 Qualitative commentary. No participants indicated the car's motor sound as a reason why they crossed in front of it. However, one mentioned the sound and spoke positively about it without prompting. This participant said, "The first thing I noticed was the noise, some sort of projected sound...I think it's a good idea to get people used to [the sound]". This participant also expressed very positive views of autonomy, noting that "Autonomous vehicles don't text and drive, they don't drink and drive. I would probably feel safer stepping in front of one of them than the other people who sometimes drive around campus."

After prompting, 27% of interviewees commented that the sound was different than that of a conventional car. Of those who noticed the sound, some described it as fitting in with the car, "I definitely heard a sound. I'd say it was on the quieter end, like a normal engine for the most part, for a modern car, for something that's newer." While not all pedestrians explicitly noticed the sound, it is still possible that it influenced people on a subconscious level.

5 DISCUSSION

We tested adding synthetic motor sound as an implicit signal for pedestrians to know when to cross in front of an AV in a naturalistic field study, and found that the sound seemed to have a positive effect on interactions. We did not exhaustively study every possible motor sound, but we believe that our experience with the sound generation and evaluation process can inform future sound design and AV eHMI design.

5.1 Motor sound and AV-pedestrian interactions

Adding motor sound seemed to augment the pre-existing implicit eHMI of the car's slowing motion, more clearly communicating to pedestrians that the vehicle would yield. Both quality of interaction and clarity of intent showed slightly higher values in the sound condition compared to the baseline, even though only 27% of participants expressed that they noticed the sound in the interview.

The fact that few people mentioned the sound without prompting suggests that the sound was an appropriate fit for the car, and did not stand out as strikingly different, unexpected, or out of its natural environment. If the sound were more novel (such as the Jetsons' car), more people would probably have noticed the sound, but it would have been a less realistic implementation. We sought to create a sound that reflected the impending government regulations, and was also inspired by existing concepts that will likely enter the market in the near future. In doing so, we have added more data suggesting that synthetic motor sound on an EV can improve interactions with pedestrians. Building on this familiar interaction pattern with AVs offers an avenue to mitigate potential challenges that pedestrians may face due to the loss of signals such as eye contact from drivers.

That is not to say that adding motor sound assuaged all potential problems. With at least one person expressing surprise that the Ghostdriver car stopped, more will need to be done to assure pedestrians that AVs are safe.

5.2 Process guidelines for sound generation

The combination of exploratory survey methods in Survey 1 and pairwise comparison in Survey 2 enabled us to quickly reduce a set of candidate sounds to one that users reported was the most appropriate. This strategy could be applied to a larger set of sounds, and a larger population, with automotive manufacturers targeting online surveys to their desired demographics. The pairwise comparisons between sounds are key to teasing out small differences between sounds, which would otherwise be difficult for people to discern, particularly in online settings [29].

On-road testing of sound is also critical, as ambient noise in the environment may mask certain frequencies, and the experience of encountering the vehicle and hearing sound in person will differ from viewing a photo and hearing the sound through headphones. This was evidenced in our study, by the need to change sound frequency after initial on-road trials. Therefore, we advise simultaneously, and iteratively testing sounds in both on-road and online contexts. Increasing the frequency of the sound changes how people hear it, and we cannot say with confidence that our on-road sound would be the most preferred if we re-ran the online study at higher frequencies.

5.3 Design guidelines for EV sound characteristics

User testing of sound is critical, as every application and context will be different. As Jaguar and Transport for London reported, users may have very different perceptions of appropriate sound than their designers. Users may react to sounds being too futuristic or complex. We observed this as well—results from our studies suggest that open chords are more likely to be preferred by users than major, minor, augmented, or diminished chords.

We expected participants to prefer major chords, as they are associated with happy and joyful music in the classical tradition. However, this was not the case. Perhaps melodic motor sounds are too distracting or notable, or fatiguing to listen to for long periods. How this will play out in industry remains to be seen: Nissan's Canto sound was an open ninth chord, while Transport for London is testing a major chord.

Our pilot testing also clearly revealed that shifting the pitch and volume of the sound made it easier to hear above the din of a busy street.

5.4 Future Work and Limitations

Given the nearly infinite design space for sound, we could not exhaustively evaluate all types of sound, but focused instead on chord type and voicing as two of many interesting dimensions. We hope that future work will explore additional sounds and systematically evaluate additional dimensions.

Our on-road sound had a higher pitch than the sounds piloted in our surveys, though they had the same tonality. It is possible that if we had conducted the surveys with higher pitched sounds, the results might be different. However, regardless of the approach for designing the sound, our results suggest that the sound implemented in the on-road setting was effective in improving interactions with pedestrians, which was the primary goal of the study.

In any naturalistic field study, there are many factors that could influence an interaction. We have explored several potential confounds such as weather, time of year, and day of week, and found no significant influence on results. However, the demonstrated effect may still have been caused by external factors.

While regulations were motivated primarily by EVs, they also pertain to HVs. The Toyota Prius, used in both the sound and baseline conditions of this study, is significantly quieter (in electric mode) than conventional vehicles, but in conventional mode is much louder than an EV. The effect of the added sound would likely be more pronounced when applied to a fully electric vehicle.

6 CONCLUSION

We prototyped an online method to evaluate EV motor sound candidates, showing that users likely prefer an open chord. We then implemented that sound in an on-road naturalistic field study demonstrating that the benefits of adding synthetic motor sound to EVs also translate to AVs-pedestrian interactions.

Whereas art and media can inspire EV motor sound, designers should also gather user input before a sound is fully implemented. Individuals vary significantly in their preferences for, and sensitivity to, sound, and it can take many iterations to design a motor sound that appeals to most. The online elicitation and data analysis methods used in this paper effectively evaluated a set of candidate motor sounds, reducing the set to an optimal choice for our application.

As a form of implicit eHMI for AVs, motor sound can augment the visual cue of an approaching AV, emphasizing to pedestrians when it will slow and yield. Such a signal would be less likely than an explicit signal to draw attention to an AV behaving differently than other cars, thereby reducing the likelihood of the AV receiving antagonistic behavior from other road users [27]. Critical to this aim is designing a sound that is audible above the hum of a busy street, but not so distracting as to draw attention to a stopped vehicle. Thus, successful EV and AV sound design will benefit from a healthy combination of both science and art.

ACKNOWLEDGMENTS

This study was supported by Robert Bosch, LLC and Daimler AG. We thank the many students and researchers at the Center for Design Research who helped to interview participants. This study was conducted under Stanford IRB protocol #32896.

REFERENCES

[1] 2018. The Future Sound of Cars with Richard Devine: How Jaguar's I-PACE Is Redefining the Engine Sound.

[2] 2018. Sound of Janguar I-Pace Protects Road Users. https://media.jaguar.com/news/2018/10/sound-jaguar-i-pace-protects-road-users.

[3] 2019. *Electric and Hybrid Cars: New Rules on Noise Emitting to Protect Vulnerable Road Users | Internal Market, Industry, Entrepreneurship and SMEs*. Technical Report. European Commission.

[4] Claudia Ackermann, Matthias Beggiato, Sarah Schubert, and Josef F. Krems. 2019. An Experimental Study to Investigate Design and Assessment Criteria: What Is Important for Communication between Pedestrians and Automated Vehicles? *Applied Ergonomics* 75 (2019), 272–282. https://doi.org/10.1016/j.apergo.2018.11.002

[5] National Highway Traffic Safety Administration. 2016. Federal Motor Vehicle Safety Standards; Minimum Sound Requirements for Hybrid and Electric Vehicles. *Federal Register* 80, 240 (Dec. 2016).

[6] Ercan Altinsoy. 2013. The Detectability of Conventional, Hybrid and Electric Vehicle Sounds by Sighted, Visually Impaired and Blind Pedestrians. In *Inter Noise*. Innsbruck, Austria.

[7] Christoph Bartneck. 2004. From Fiction to Science – A Cultural Reflection of Social Robots. In *Proceedings of the CHI2004 Workshop on Shaping Human-Robot Interaction*.

[8] E. Cha, N.T. Fitter, Y. Kim, T. Fong, and M.J. Matari. 2018. Effects of Robot Sound on Auditory Localization in Human-Robot Collaboration. In *ACM/IEEE International Conference on Human-Robot Interaction*. 434–442. https://doi.org/10.1145/3171221.3171285

[9] Michael Clamann. 2017. Evaluation of Vehicle-to-Pedestrian Communication Displays for Autonomous Vehicles. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 57, 3 (2017), 407–434.

[10] Rebecca Currano, So Yeon Park, Lawrence Domingo, Jesus Garcia-Mancilla, Pedro Santana-Mancilla, Victor Gonzalez, and Wendy Ju. 2018. Vamos! Observations of Pedestrian Interactions with Driverless Cars in Mexico. In *10th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '18),*. ACM, Toronto, Canada. https://doi.org/10.1145/3239060.3241680

[11] Gayle Dalrymple. 2013. *Minimum Sound Requirements for Hybrid and Electric Vehicles*. Technical Report. US Department of Transportation - National Highway Traffic Safety Administration. 114 pages.

[12] Koen de Clercq, Andre Dietrich, Juan Pablo Núñez Velasco, Joost de Winter, and Riender Happee. 2019. External Human-Machine Interfaces on Automated Vehicles: Effects on Pedestrian Crossing Decisions. *Human Factors* 61, 8 (Dec. 2019), 1353–1370. https://doi.org/10.1177/0018720819836343

[13] Debargha Dey, Azra Habibovic, Bastian Pfleging, Marieke Martens, and Jacques Terken. 2020. Color and Animation Preferences for a Light Band eHMI in Interactions Between Automated Vehicles and Pedestrians. (2020), 13.

[14] Debargha Dey and Jacques Terken. 2017. Pedestrian Interaction with Vehicles: Roles of Explicit and Implicit Communication. In *AutoUI Adjunct Proceedings*. 109–113. https://doi.org/10.1145/3122986.3123009

[15] Rachel England. 2018. European EVs Must Be Fitted with Sound Emitters by 2021. https://www.engadget.com/2018/05/08/european-evs-fitted-with-sound-emitters-by-2021/.

[16] Anna Fenko, Hendrik N J Schifferstein, and Paul Hekkert. 2011. Noisy Products: Does Appearance Matter? *International Journal of Design* 5, 3 (2011), 77–87.

[17] Sylvain Fleury, Éric Jamet, Vincent Roussarie, Laure Bosc, and Jean Christophe Chamard. 2016. Effect of Additional Warning Sounds on Pedestrians' Detection of Electric Vehicles: An Ecological Approach. *Accident Analysis and Prevention* 97 (2016), 176–185. https://doi.org/10.1016/j.aap.2016.09.002

[18] Emma Frid, Roberto Bresin, and Simon Alexanderson. 2018. Perception of Mechanical Sounds Inherent to Expressive Gestures of a NAO Robot - Implications for Movement Sonification of Humanoids. In *15th Sound and Music Computing Conference*. Limassol, Cyprus.

[19] Azra Habibovic, Victor Malmsten Lundgren, Jonas Andersson, Maria Klingegård, Tobias Lagström, Anna Sirkka, Johan Fagerlönn, Claes Edgren, Rikard Fredriksson, Stas Krupenia, Dennis Saluäär, and Pontus Larsson. 2018. Communicating Intent of Automated Vehicles to Pedestrians. *Frontiers in Psychology* 9 (Aug. 2018). https://doi.org/10.3389/fpsyg.2018.01336

[20] Drew Hartwell. 2015. America's Best-Selling Cars and Trucks Are Built on Lies: The Rise of Fake Engine Noise. *The Washington Post* (2015).

[21] Andrew J. Hawkins. 2019. Electric Car Owners Could Choose Which Fake Sounds Their Cars Make under New Proposal. https://www.theverge.com/2019/9/16/20869035/electric-car-ev-fake-noise-nhtsa.

[22] Anees Ahamed Kaleefathullah, Natasha Merat, Yee Mun Lee, Yke Bauke Eisma, Ruth Madigan, Jorge Garcia, and Joost de Winter. [n.d.]. External Human-Machine Interfaces Can Fail! An Examination of Trust Development and Misuse in a CAVE-Based Pedestrian Simulation Environment. ([n. d.]), 26.

[23] Lau Langeveld, René van Egmond, Reinier Jansen, and Elif Özcan. 2013. Product Sound Design: Intentional and Consequential Sounds. In *Advances in Industrial Design Engineering*. InTech, 47–73. https://doi.org/10.5772/3415

[24] Jamy Li, Rebecca Currano, David Sirkin, David Goedicke, Hamish Tennent, Aaron Levine, Vanessa Evers, and Wendy Ju. 2020. On-Road and Online Studies to Investigate Beliefs and Behaviors of Netherlands and US Pedestrians Encountering Hidden-Driver Vehicles. In *15th Annual ACM/IEEE International Conference on Human-Robot Interaction*. Cambridge, England.

[25] Andrew Liptak. 2019. Hans Zimmer Designed the Sound for BMW's Futuristic Concept Car. https://www.theverge.com/2019/6/29/19914287/bmw-hans-zimmer-design-bmw-vision-m-next-sound-profile-blade-runner-cyberpunk-3d-print.

[26] Richard H. Lyon. 2003. Product Sound Quality: From Perception to Design. *Sound and Vibration* 37, 3 (2003), 18–23. https://doi.org/10.1121/1.4743110

[27] Dylan Moore, Rebecca Currano, Michael Shanks, and David Sirkin. 2020. Defense Against the Dark Cars: Design Principles for Griefing of Autonomous Vehicles. In *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction - HRI '20*. Cambridge, UK. https://doi.org/10.1145/3319502.3374796

[28] Dylan Moore, Rebecca Currano, G. Ella Strack, and David Sirkin. 2019. The Case for Implicit External Human-Machine Interfaces for Autonomous Vehicles. In *AutoUI 2019*. Utrecht, Netherlands. https://doi.org/10.1145/3342197.3345320

[29] Dylan Moore, Tobias Dahl, Paula Varela, Wendy Ju, Tormod Næs, and Ingunn Berget. 2019. Unintended Consonances: Methods to Understand Robot Motor Sound Perception. In *CHI*. Glasgow, UK, 1–12.

[30] Dylan Moore and Wendy Ju. 2018. Sound as Implicit Influence on Human-Robot Interactions. In *Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction - HRI '18*. ACM Press, Chicago, IL, USA, 311–312. https://doi.org/10.1145/3173386.3176918

[31] Dylan Moore, Hamish Tennent, Nikolas Martelaro, and Wendy Ju. 2017. Making Noise Intentional: A Study of Servo Sound Perception. In *Human Robot Interaction*. Vienna, Austria.

[32] Man Made Music. 2018. How Silent Electric Vehicles Get Their Sound. https://www.fastcompany.com/video/how-silent-electric-vehicles-get-their-sound/PYPTnTl4.

[33] John G Neuhoff, Gregory Kramer, and Joseph Wayand. 2000. Sonification and the Interaction of Perceptual Dimensions: Can the Data Get Lost in the Map?. In *Proc. Int. Conf. on Auditory Display*. Atlanta, GA, USA, 6.

[34] NHTSA. 2016. NHTSA Sets 'Quiet Car' Safety Standard to Protect Pedestrians. https://www.nhtsa.gov/press-releases/nhtsa-sets-quiet-car-safety-standard-protect-pedestrians.

[35] Don Norman. 2014. What Noise Does the Electric Car Make? *MIT Technology Review* (Feb. 2014), 8.

[36] Jean-François Petiot, Bjørn G. Kristensen, and Anja M. Maier. 2013. How Should an Electric Vehicle Sound? User and Expert Perception. In *Volume 5: 25th International Conference on Design Theory and Methodology; ASME 2013 Power Transmission and Gearing Conference*. American Society of Mechanical Engineers, Portland, Oregon, USA, V005T06A028. https://doi.org/10.1115/DETC2013-12535

[37] Jon Porter. 2019. London's Electric Buses Are Getting Fake Noise, and It's Positively Psychedelic. https://www.theverge.com/tldr/2019/12/20/21031524/london-electric-buses-artificial-fake-noise-safety-sound.

[38] R Development Core Team. 2016. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing Vienna Austria* 0 (2016), {ISBN} 3–900051–07–0. https://doi.org/10.1038/sj.hdy.6800737

[39] R Read and T Belpaeme. 2012. How to Use Non-Linguistic Utterances to Convey Emotion in Child-Robot Interaction. *Human-Robot Interaction (HRI), 2012 7th ACM/IEEE International Conference on* (2012), 219–220. https://doi.org/10.1145/2157689.2157764

[40] Robin Read and Tony Belpaeme. 2016. People Interpret Robotic Non-Linguistic Utterances Categorically. *International Journal of Social Robotics* 8, 1 (2016), 31–50. https://doi.org/10.1007/s12369-015-0304-0

[41] Frederic Anthony Robinson, Oliver Bown, and Mari Velonaki. 2020. Implicit Communication through Distributed Sound Design: Exploring a New Modality in Human-Robot Interaction. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. ACM, Cambridge United Kingdom, 597–599. https://doi.org/10.1145/3371382.3377431

[42] Dirk Rothenbücher, Jamy Li, David Sirkin, Brian Mok, and Wendy Ju. 2016. Ghost Driver: A Field Study Investigating the Interaction between Pedestrians and Driverless Vehicles. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium On*. 795–802. https://doi.org/10.1109/ROMAN.2016.7745210

[43] Michael B Sapherstein. 1998. The Trademark Registrability of the Harley-Davidson Roar: A Multimedia Analysis. In *Boston College Intellectual Property & Technology Forum*. 8.

[44] Joachim Scheuren, Rolf Schirmacher, and Josef Hobelsberger. 2012. Active Design of Automotive Engine Sound. In *Inter Noise*. 6.

[45] Charles Spence and Qian Wang. 2015. Sensory Expectations Elicited by the Sounds of Opening the Packaging and Pouring a Beverage. *Flavour* 4, 1 (2015), 35. https://doi.org/10.1186/s13411-015-0044-y

[46] Charles Spence and Massimiliano Zampini. 2006. Auditory Contributions to Multisensory Product Perception. *Acta Acustica united with Acustica* 92, 6 (2006), 1009–1025.

[47] Daniël Johannes Swart. 2018. *The Psychoacoustics of Electric Vehicle Signature Sound.* Ph.D. Dissertation. Stellenbosch University.

[48] M Sweeney, T Pilarski, W Ross, and C Liu. 2017. Light Output System for a Self-Driving Vehicle. US Patent Office Application No. US20180072218A1.

[49] Hamish Tennent, Dylan Moore, Malte Jung, and Wendy Ju. 2017. Good Vibrations: How Consequential Sounds Affect Perception of Robotic Arms. In *RO-MAN 2017 - 26th IEEE International Symposium on Robot and Human Interactive Communication*, Vol. 2017-Janua. Lisbon, Portugal, 928–935. https://doi.org/10.1109/ROMAN.2017.8172414

[50] Raquel Thiessen, Daniel J Rea, Diljot S Garcha, Cheng Cheng, and James E Young. 2019. Infrasound for HRI: A Robot Using Low-Frequency Vibrations to Impact How People Perceive Its Actions. In *HRI*. 8.

[51] René Tünnermann, Jan Hammerschmidt, and Thomas Hermann. 2013. Blended Sonification: Sonification for Causal Information Interaction. In *19th International Conference on Auditory Display*. Lodz, Poland.

[52] usTwo. 2017. A Glance At The Future Of External Vehicular Sound. https://www.ustwo.com/blog/future-of-external-vehicular-sound.

[53] Nozomiko Yasui. 2019. Subjective Evaluations of Detectability of Alert Sound for Electric and Hybrid Electric Vehicle under Actual Environment. In *Inter Noise*. Madrid, Spain.

[54] Selma Yilmazyildiz, Robin Read, Tony Belpeame, and Werner Verhelst. 2016. Review of Semantic-Free Utterances in Social Human–Robot Interaction. *International Journal of Human-Computer Interaction* 32, February (2016), 63–85. https://doi.org/10.1080/10447318.2015.1093856