Auditory localization of multiple stationary electric vehicles

Citation for the original published paper (version of record):
Müller, L., Forssén, J., Kropp, W. (2025). Auditory localization of multiple stationary electric vehicles. Journal of the Acoustical Society of America, 157(3): 2029-2041.
http://dx.doi.org/10.1121/10.0036248

(article starts on next page)

The Journal of the Acoustical Society of America

MARCH 24 2025

Auditory localization of multiple stationary electric vehicles

Leon Müller ; Jens Forssén ; Wolfgang Kropp

J. Acoust. Soc. Am. 157, 2029–2041 (2025)
https://doi.org/10.1121/10.0036248

JASA ARTICLE

Auditory localization of multiple stationary electric vehicles

Leon Müller,[a)] Jens Forssén, and Wolfgang Kropp
Division of Applied Acoustics, Chalmers University of Technology, Gothenburg, 41296, Sweden

ABSTRACT:
Current regulations require electric vehicles to be equipped with acoustic vehicle alerting systems (AVAS), radiating artificial warning sounds at low driving speeds. The requirements for these sounds are based on human subject studies, primarily estimating detection time for single vehicles. This paper presents a listening experiment assessing the accuracy and time of localization using a concealed array of 24 loudspeakers. Static single- and multiple-vehicle scenarios were compared using combustion engine noise, a two-tone AVAS, a multi-tone AVAS, and a narrowband noise AVAS. The results of 52 participants show a significant effect of the sound type on localization accuracy and time for all evaluated scenarios ($p < 0.001$). Post-hoc tests revealed that the two-tone AVAS is localized significantly worse than the other signals, especially when simultaneously presenting two or three vehicles with the same type of sound. The multi-tone and noise AVAS are generally on par but localized worse than combustion noise for multi-vehicle scenarios. For multiple vehicles, the percentage of failed localizations drastically increased for all three AVAS signals, with the two-tone AVAS performing worst. These results indicate that signals typically performing well in a single-vehicle detection task are not necessarily easy to localize, especially not in multi-vehicle scenarios.
 https://doi.org/10.1121/10.0036248

(Received 11 December 2024; revised 4 March 2025; accepted 6 March 2025; published online 24 March 2025)

[Editor: Francesco Aletta] Pages: 2029–2041

I. INTRODUCTION

Today's urban acoustic environments are, to a large part, coined by road traffic noise (EAA, 2020). With the current transition to electro-mobility, the character of this noise may change to some extent (International Energy Agency, 2024). Even though electric vehicles (EVs) are often considered a solution for the quieter cities of tomorrow, tire-road noise is, at medium to high driving speeds, more relevant for overall sound levels than propulsion noise (Pallas *et al.*, 2016) and changing the propulsion type does not necessarily decrease sound levels for all situations. However, tire-road noise often plays a minor role at low driving speeds. Hence, slowly driving EVs typically radiate less sound than slowly driving internal combustion engine vehicles (ICEVs) (Garay-Vega *et al.*, 2010; Pallas *et al.*, 2016). While this may be considered beneficial from an environmental noise perspective, the downside of this, generally speaking, low sound emission of EVs at slow driving speeds, is a lack of acoustic localization cues for pedestrians, cyclists, and other vulnerable road users such as the visually impaired. This drawback has been associated with an increased risk of accidents involving slow-driving electric and hybrid-EVs in urban environments, shown by analyzing accident statistics in different nations (Hanna, 2009; Wu *et al.*, 2011; Morgan *et al.*, 2011; Hou *et al.*, 2023; Edwards *et al.*, 2024). Even though this relation and the underlying statistical assumptions were, during the early stage of electric vehicle (EV) market introduction, questioned by some (Sandberg *et al.*, 2010; Stelling-Kończak *et al.*, 2015), consecutively performed human subject experiments confirmed that without any countermeasures or visual cues and assuming low background noise levels, slow-driving electric and hybrid-EVs approaching a test person are often detected later than ICEVs (Goodes *et al.*, 2009; Garay-Vega *et al.*, 2010; Kim *et al.*, 2012). As a countermeasure, regulations in many nations now mandate that all newly produced EVs are equipped with an acoustic vehicle alerting system (AVAS), i.e., a loudspeaker radiating artificial warning sounds that shall indicate the location and driving behavior of the vehicle.

In the United States, this regulation was implemented based on research commissioned by the National Highway Traffic Safety Administration (NHTSA), for which researchers first measured sound levels of ICEVs and hybrid-EVs under different safety-critical operating conditions and then conducted a human subject experiment evaluating how well these sounds are detected. The measure used to assess detection efficiency was time-to-vehicle arrival, i.e., the time between the moment a participant first noticed an approaching vehicle and the moment the vehicle reached the participant's position. The longer this time-to-vehicle arrival, the lower the risk of a potential accident. The results, published as technical reports, showed that participants detected electrically propelled vehicles later than ICEVs for most operating modes (Garay-Vega *et al.*, 2010). The agency then considered adding recorded ICEV sounds, synthesized ICEV-like sounds, or synthetic sounds designed according

[a)]Email: leon.mueller@chalmers.se

to psychoacoustic principles as possible countermeasures. The results of a second human subject experiment showed that the latter synthetic sounds are, under most driving conditions, detected earlier than ICEV sounds. ICEV-like sounds produced similar results as ICEVs (Hastings *et al.*, 2011). A third experiment finally indicated that warning signals with energy in a large number of non-adjacent third-octave bands seem to be most efficient and that pure tones are more detectable in terms of time-to-vehicle arrival than third-octave band noise (Hastings and McInnis, 2015). Apart from these human subject experiments, sound level calculation procedures based on different psychoacoustic models were examined, and it was concluded that a velocity-dependent pitch shift would be beneficial (Hastings *et al.*, 2012).

While the original NHTSA proposal suggested a quite extensive AVAS warning signal consisting of up to eight different third-octave bands and a pitch shift of at least 1%/km/h (National Highway Traffic Safety Administration, 2013), the finally implemented regulation FMVSS No. 141 was somewhat toned down, allowing for either using four non-adjacent third-octave bands spanning no fewer than nine bands in total or using two non-adjacent one-third octave bands between 315 and 3150 Hz, with one band below 1000 Hz and the other band at or above 1000 Hz. Therein, minimum third-octave band levels for reversing, stationary, and pass-by conditions from 0 to 10, 10 to 20, 20 to 30, and 30 to 32 km/h are specified (National Highway Traffic Safety Administration, 2016a). Due to concerns regarding measurement reproducibility, the initially proposed pitch shift for vehicle acceleration was replaced by demanding a relative volume increase in 3 dB for each 10 km/h increment of driving speed (National Highway Traffic Safety Administration, 2016b).

In the EU and other countries, such as China and Japan, the corresponding electric vehicle warning sound regulations are based on or aligned with United Nations Economic Commission for Europe (UNECE) Regulation No. 138 (United Nations Economic Commission for Europe, 2017), which was developed in parallel with the US regulations (Fiebig, 2020). This regulation demands that the AVAS signal covers at least two third-octave bands, of which one shall be within or below the 1600 Hz band. Minimum sound pressure levels for those bands and a minimum overall level are defined for 10 and 20 km/h passages. Additionally, at least one tone within the specified frequency range should be shifted in frequency by at least 0.8% per 1 km/h speed change. Only a minimum overall sound pressure level is specified for reversing vehicles, and an AVAS sound for still-standing vehicles is optional but not required. Unlike the US regulations, UNECE Regulation No. 138 also limits the maximum overall sound pressure level for AVAS-equipped vehicles to 75 dBA measured at a 2 m distance.

Apart from the previously mentioned NHTSA-issued research, several other studies found that slowly driving EVs are often harder to detect than ICEVs (Mendonça *et al.*, 2013), adding warning sounds can increase detection probability (Parizet *et al.*, 2014; Roan *et al.*, 2021) and help with estimating the trajectory of accelerating vehicles (Wessels *et al.*, 2022), synthetic sounds can be more efficient in terms of detection than ICE sounds (Poveda-Martínez *et al.*, 2017), a velocity-dependent pitch shift and amplitude modulation may benefit detectability (Fleury *et al.*, 2016; Emerson *et al.*, 2013), and experiments performed in virtual environments can predict real-world EV detection (Singh *et al.*, 2015). However, most of these studies are limited to evaluating the detection of single vehicles. While this might be the most critical measure for overall traffic safety, some scenarios, such as a busy parking lot, may also require a certain level of localization accuracy, especially in the presence of multiple vehicles.

To our knowledge, the only study explicitly investigating the localization accuracy of EVs was performed by Stelling-Kończak *et al.* (2016), which found that ICEV sounds yield more correct localizations than the evaluated EV sound, especially at low driving speeds. However, the spatial resolution of this experiment was limited to discrete steps of 45°, and only the number of correct localizations was evaluated, not the size of the localization error. Additionally, only one type of EV sound, presumably without any AVAS, was compared to three combustion engine vehicle sounds. While studies on ICEV sounds confirmed that competing noise from a second vehicle significantly affects detection (Ulrich *et al.*, 2014), the presence of multiple EVs with the same or different AVAS sounds seems to not have been investigated yet.

This study aims to close this gap by performing a listening experiment evaluating the high-resolution localization accuracy for ICEVs and EVs with different AVAS sounds, including scenarios where multiple vehicles with the same or different sounds are present. We thereby limited the scope of this experiment to static scenarios with all vehicles standing still, mimicking a busy parking lot where vehicles are slowly maneuvering rather than the often studied pass-by or street crossing situations.

II. METHODS

A. Experiment setup and procedure

A circular loudspeaker array with a 3 m radius, consisting of 24 Genelec 8020 loudspeakers (Genelec OY, Iisalmi, Finland), was installed in an anechoic chamber with the acoustic center being fixed at a height of approximately 1.6 m as shown in Fig. 1. The loudspeakers were covered by an acoustically transparent curtain so that the participants, placed in the center of the array, could not see them. In front of this curtain, a strip comprising 695 individually addressable light-emitting diodes (LEDs) provided visual feedback for the direction estimation, and a small monitor mounted at the 0° position displayed experiment instructions. The participants were handed a custom-built motion controller, i.e., a water blaster modified with a microprocessor, motion sensors, and an electronic trigger switch. The horizontal movement of this controller was translated to a red dot on the LED strip so that the subjects always had visual feedback indicating in which direction they were aiming. The motion controller and LED strip were controlled using Arduino

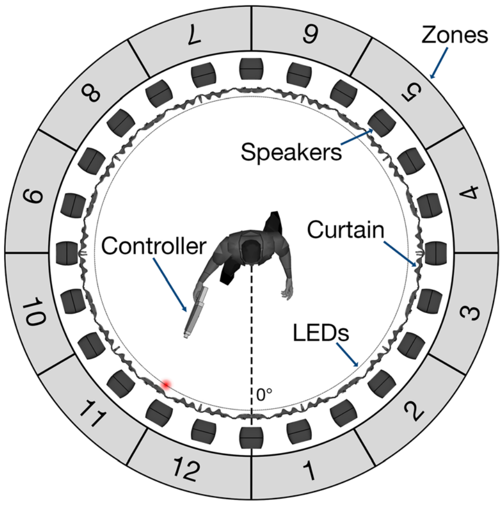

FIG. 1. (a) Schematic illustration of experiment setup, (b) and picture of a participant performing the experiment. The zone numbers in (a) are used to describe the distribution of stimuli (cf. Sec. II D).

microprocessors (Arduino LLC, Somerville, MA) microprocessors. The stimuli playback and overall experiment sequence were implemented in Matlab R2024b (The Mathworks Inc., Natick, MA).

For the experiment, the participants were instructed to aim the controller as quickly as possible in all directions from which they perceive a vehicle and to lock in each perceived vehicle position by pulling the trigger switch or, in simpler terms, to shoot in the direction of all vehicles they perceive in arbitrary order. These locked vehicle positions were visualized by green lights on the LED strip and reset after each trial. Thereby, the participants used the red light shown on the LED strip to indicate their answer position, which means that these perceived positions were measured with an angular resolution of $360^\circ/695\,\text{LEDs} \approx 0.5^\circ$. After each trial, the synchronization between the motion controller and the red LED position was recalibrated by instructing the participants to aim the controller toward the 0° point and pull the trigger. This trigger initiated a 3 s countdown, after which the subsequent trial started. All stimuli were played back for 10 s with a countdown displayed on the screen indicating the remaining time in each trial. After 10 s, the trial stopped and the subjects could not give any more answers. Additionally, the participants were told that a maximum of three vehicles were presented simultaneously without specifying the exact number of cars to expect in each trial. The subjects were allowed to move their heads and bodies freely, with the only limitation being to stay in the center of the loudspeaker array. The repository (Müller *et al.*, 2024) contains a video illustrating the experiment procedure.

B. Stimuli

The stimuli presented in the experiment comprised three different types of AVAS signals (band noise, two-tone, and multi-tone) and one combustion engine sound. The AVAS signals were generated using a recently developed electric vehicle auralization toolbox (Müller and Kropp, 2024) simulating a 7.5 m distance between the vehicle and the listener position, with the car frontally facing the observer. Apart from ground reflections included in the simulated AVAS radiation. directivity, the virtual acoustic environment was completely anechoic. In addition to the vehicle stimuli, a constant background noise recording of an empty parking lot was played back. All stimuli are described in the following and can be openly accessed at Müller *et al.* (2024).

1. Variations in velocity

Since evaluating multiple vehicles with exactly the same AVAS signals is not very realistic and can, especially for tonal signals, lead to strong interferences, four different versions of each stimulus type that are slightly but perceivably different were rendered by using different velocities for the AVAS signal generation and different idling speeds for the combustion noise recording. Thereby, it is important to distinguish between the velocity used for the signal generation and the velocity of the virtual sound source in the listening experiment. For the AVAS stimuli generation, constant velocities between 2 and 8 km/h were used to generate sounds typical for the different AVAS types. In the experiment, however, the position of the sound sources was static, i.e., the velocities used to generate the AVAS signals do not match the fact that the vehicles are standing still in the experiment. Hence, the experiment does not evaluate the localization accuracy for the exact vehicle models used as a reference, which might radiate a different or no sound when standing still, but instead compares three different types of AVAS sounds that are similar to currently implemented solutions and comply with the existing regulations.

2. Combustion engine recording

A 2014 Volkswagen Golf VII (Volkswagen AG, Wolfsburg, Germany) with a 77 kW four-cylinder petrol engine was recorded outdoors on a quiet parking lot, idling at

https://doi.org/10.1121/10.0036248

700, 800, 1000, and 1200 rpm using two GRAS 46AE free-field equalized 1/2 in. microphones (GRAS Acoustics, Holte, Denmark) mounted in front of the grill and above the exhaust. Additionally, a free-field equalized HeadAcoustics HMS-V (HeadAcoustic GmbH, Herzogenrath, Germany) artificial head was placed at a 7.5 m distance, facing the front of the vehicle. The front and exhaust microphone signals were then mixed to match the spectral balance of the reference artificial head recording, resulting in stimuli as shown in Fig. 2(a).

3. Noise AVAS

The noise AVAS signal was synthesized based on measurements of a forward-driving Tesla Model Y 2019 (Tesla Inc., Austin, TX) (Müller and Kropp, 2024). The signal contains two narrowband noise components, as shown in Fig. 2(b), whose center frequency increases with vehicle velocity. The lower band has a center frequency of around 580 Hz with a 3 dB bandwidth of 45 Hz, and the higher band has a center frequency of 1020 Hz with a 3 dB bandwidth of 150 Hz. Both bands vary in frequency in steps of around 30 cents (≈1.7%) between the four generated stimuli versions.

4. Two-tone AVAS

The two-tone AVAS signal was synthesized based on measurements of a reversing Tesla model Y 2019 (Müller and Kropp, 2024). The signal consists of two pure tones with a frequency of 340 and 1350 Hz, as shown in Fig. 2(c). Both tones are amplitude modulated; the lower with a rate of approximately 4.2 Hz, the higher with a rate of approximately 5.5 Hz. Both the pitch and modulation rate of both tones change with vehicle velocity. The four stimuli versions generated for the experiment vary in pitch in steps of around 30 cents and in modulation rate by around 0.25 Hz.

5. Multi-tone AVAS

The multi-tone AVAS signal was synthesized based on measurements of a Volkswagen ID.3 Pro Performance 2021 (Müller and Kropp, 2024) and contains 25 tones between 70 and 2700 Hz, each amplitude modulated with rates between 0.03 and 2 Hz as shown in Fig. 2(d). The pitch of each tone shifts by approximately 30 cents between the four generated stimuli versions, including slight variations in amplitude modulation rates.

6. Background noise

In addition to the vehicle stimuli, constant background noise was played back during the experiment. The purpose of this noise was not to investigate potential masking effects but rather to achieve a higher level of immersion in the virtual acoustic environment. The background noise was recorded by placing an EigenMike em32 (mh acoustics LLC, Summit, NJ) spherical microphone array on a quiet parking lot. The recording contains natural sounds and far-distant road traffic but no clearly recognizable sound sources or auditory events relevant to the experiment. The recording was filtered with a 12 dB/octave bandpass filter between 100 Hz and 4 kHz to reduce low-frequency wind noise and high-frequency microphone noise and calibrated to match the equivalent continuous background noise level of 40 dBA measured on said parking lot.

7. Loudness and compliance

In order to rule out loudness differences as a possible factor affecting localization accuracy, all stimuli were normalized to the same loudness of 4 soneHMS, according to the Ecma International (Geneva, Switzerland) (ECMA)-418-2 method

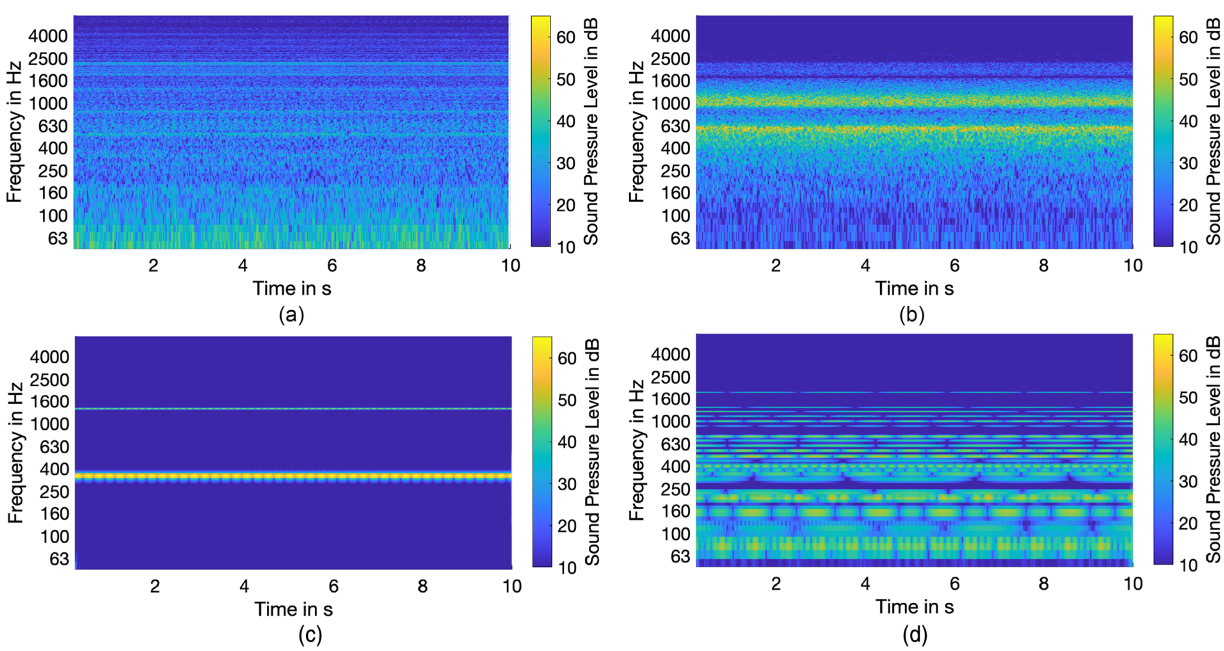

FIG. 2. Spectrograms of (a) combustion engine idling at 700 rpm noise AVAS, (b) two-tone AVAS, (c) multi-tone AVAS, (d) stimuli. Only one of the four different versions generated for each stimulus type is shown.

https://doi.org/10.1121/10.0036248

TABLE I. Equivalent continuous sound pressure level and loudness as determined by ECMA-418-2 for all stimuli recorded at the listening position. The given range of sound pressure levels covers all four generated stimuli variations.

Stimulus	Sound pressure level in dBA	Loudness in soneHMS
Combustion	50.65–51.19	4.00
Noise AVAS	63.27–63.77	4.00
Two-tone AVAS	60.34–60.51	4.00
Multi-tone AVAS	54.24–54.89	4.00
Background noise	40.09	1.84

(European Computer Manufacturers Association, 2022), which is assumed to be more accurate in estimating the perceived loudness of strong tonal and multi-tonal signals than other loudness calculation methods (Lobato and Sottek, 2023). Table I lists the resulting *A*-weighted equivalent continuous sound levels for all stimuli types.

All stimuli fulfill the current UNECE and US regulations in at least two third-octave bands, as shown in Fig. 3. For this comparison, minimum values specified for 2 m measurement distance by the regulations were scaled to the 7.5 m distance used for the stimuli generation by applying a reduction of $10\log\left((2^2+1.2^2)/(7.5^2+1.2^2)\right)=-10.26$ dB, assuming a microphone height of 1.2 m, a source near the ground, and geometrical spreading.

C. Acoustic reproduction

While the loudspeaker array setup would have allowed the implementation of sophisticated methods of sound field reproduction, such as wave field synthesis or Ambisonics (Ahrens, 2012), a narrow apparent source width and an artifact-free reproduction were prioritized for this study.

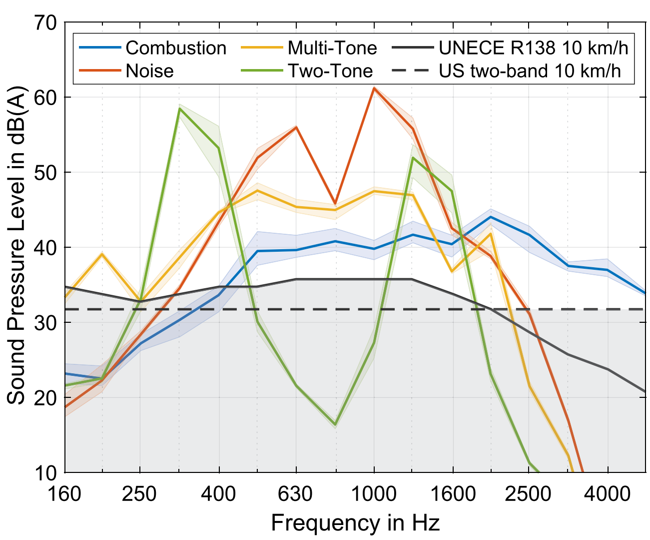

FIG. 3. Third-octave spectra of combustion engine, noise AVAS, two-tone AVAS, and multi-tone AVAS stimuli including corresponding minimum levels of UNECE Regulation 138 and US regulation FMVSS No. 141 for a 10 km/h pass-by, adapted to a measurement distance of 7.5 m. The area around the lines represents the variations between the four different versions of each stimulus.

Therefore, the individual stimuli were played back via single loudspeakers corresponding to the desired horizontal position of the vehicle without any further processing, ensuring a minimum perceived source width without the risk of the reproduction method negatively affecting the localization accuracy.

Additionally, parking lot ambient noise recorded with a 32-channel spherical microphone array as described in Sec. II B 6 was decoded to be played back via the loudspeaker array using fourth-order all-round Ambisonic decoding (Zotter *et al.*, 2012) with max-rE weighting as implemented by the Institute of Electronic Music and Acoustics plug-in suite (Institute of Electronic Music and Acoustics, 2024). This noise was played independently of the vehicle stimuli from the moment the participants entered the room until the end of the experiment to achieve a constant acoustic immersion. Calibrated binaural recordings of all 72 trials, including background noise, can be accessed at Müller *et al.* (2024).

D. Stimuli distribution

In order to keep the evaluated scenarios realistic for a real-life parking lot situation, we decided to simulate a maximum of three vehicles simultaneously. Four different cases were investigated: single vehicles, two vehicles with the same type of sound, three vehicles with the same type of sound, and all possible combinations of two vehicles with different sounds. Each case was repeated four times with different spatial distributions and stimuli variations, resulting in a total of 72 trials as listed in Table II. The order of these 72 trials was randomized for each participant. Prior to the main experiment, all participants completed a short supervised training session of ten trials, representing all four cases and all stimuli types.

Deciding on the location for each sound source in each trial that rules out all possible direction, order, and memory effects, proved challenging. For example, vehicles presented in front of the participant might be easier and faster to localize than lateral sounds, and directly comparing two stimuli presented at completely different horizontal positions might result in wrong conclusions. On the other hand, always using the same vehicle positions might cause participants to memorize those locations. We, therefore, chose a compromise between control and randomness by dividing the 24 loudspeakers into 12 different zones, each comprising three loudspeakers as shown in Fig. 1(a). We then specified playback zones for each stimulus in each trial as described in Table II but randomized which of the three loudspeakers within that specified zone was used for each trial and participant. This means that the coarse stimuli locations for each trial were the same for all participants, but the exact playback locations varied randomly by up to 30°. Overall, care was taken to spatially distribute the stimuli in a meaningful way so that averaging both localization error and localization time over repetitions and participants' balances out all possible influences of order and location.

https://doi.org/10.1121/10.0036248

TABLE II. Stimuli distribution among trials. The numbers after the stimuli names represent the different stimuli variations. For example, in trial number 35, stimulus two-tone 1 was played in zone 3, stimulus two-tone 3 was played in zone 7, and stimulus two-tone 4 was played in zone 11.

	Trial Nr.	Stimulus 1	Stimulus 2	Stimulus 3	Playback zone of stimulus 1	Playback zone of stimulus 2	Playback zone of stimulus 3
Single	01–04	Two-tone [1,2,3,4]			[2,5,9,12]		
	05–08	Multi-tone [1,2,3,4]			[2,5,9,12]		
	09–12	Noise [1,2,3,4]			[2,5,9,12]		
	13–16	Combustion [1,2,3,4]			[2,5,9,12]		
Two same	17–20	Two-tone [1,1,2,3]	Two-tone [2,4,3,4]		[4,2,5,1]	[10,8,7,11]	
	21–24	Multi-tone [1,1,2,3]	Multi-tone [2,4,3,4]		[4,2,5,1]	[10,8,7,11]	
	25–28	Noise [1,1,2,3]	Noise [2,4,3,4]		[4,2,5,1]	[10,8,7,11]	
	29–32	Combustion [1,1,2,3]	Combustion [2,4,3,4]		[4,2,5,1]	[10,8,7,11]	
Three same	33–36	Two-tone [1,1,1,2]	Two-tone [2,2,3,3]	Two-tone [3,4,4,4]	[1,2,3,4]	[5,6,7,8]	[9,10,11,12]
	37–40	Multi-tone [1,1,1,2]	Multi-tone [2,2,3,3]	Multi-tone [3,4,4,4]	[1,2,3,4]	[5,6,7,8]	[9,10,11,12]
	41–44	Noise [1,1,1,2]	Noise [2,2,3,3]	Two-tone [3,4,4,4]	[1,2,3,4]	[5,6,7,8]	[9,10,11,12]
	45–48	Combustion [1,1,1,2]	Combustion [2,2,3,3]	Combustion [3,4,4,4]	[1,2,3,4]	[5,6,7,8]	[9,10,11,12]
Two Different	49–52	Two-tone [1,2,3,4]	Multi-tone [1,2,3,4]		[11,5,6,4]	[5,11,4,6]	
	53–56	Two-tone [1,2,3,4]	Noise [1,2,3,4]		[11,5,6,4]	[5,11,4,6]	
	57–60	Two-tone [1,2,3,4]	Combustion [1,2,3,4]		[11,5,6,4]	[5,11,4,6]	
	61–64	Multi-tone [1,2,3,4]	Noise [1,2,3,4]		[11,5,6,4]	[5,11,4,6]	
	65–68	Multi-tone [1,2,3,4]	Combustion [1,2,3,4]		[11,5,6,4]	[5,11,4,6]	
	69–72	Noise [1,2,3,4]	Combustion [1,2,3,4]		[11,5,6,4]	[5,11,4,6]	

E. Data processing

The data obtained throughout the experiment were pre-processed in MATLAB R2024b, as described in the following. First of all, the number of answers in each trial was limited to the number of simultaneously played-back stimuli for each participant. This means that if a participant marked, for example, three positions but only two stimuli were played, only the first two answers were considered in the analysis. If a participant gave fewer answers than vehicles present in a trial, the given answers were assigned to the vehicle positions that they matched best, and the not localized vehicles were handled as failed localizations.

In the next step, the absolute error between the exact playback positions and the given response locations was calculated. For cases with multiple vehicles, this was done using a greedy matching algorithm, i.e., by first finding the pair of playback and response location with the lowest absolute error, then removing that pair and repeating for the remaining data. If a single error between the source and answer position was above 90°, it was treated as failed localization.

Additionally, the localization time for each stimulus in each trial was estimated. For single vehicles and the first localized stimulus in a multi-vehicle comparison, this time was defined as the duration from the trial start to the moment the participant marked the vehicle position. For subsequent answers in a multi-vehicle comparison, the previous answer time was used as a starting point for the localization time measurement, including failed localizations. The mean localization error and localization time for each stimulus in each comparison type were then calculated by averaging both measures over the four repetitions for each participant.

While this pre-processing seemed most reasonable to us, we also tested other approaches, such as allowing for more answers than played-back stimuli or including errors above 90° in the mean value calculations. However, none of these variations in processing had a significant effect on the overall results and drawn conclusions, indicating that the obtained data is quite robust. For full transparency, both the raw and the pre-processed data can be accessed at Müller *et al.* (2024).

F. Participants

The experiment was performed by 55 participants, mainly recruited from Chalmers University students and faculty members. Three of these 55 participants were excluded from the following analysis as their overall localization error, averaged over all trials, was more than 1.5 interquartile ranges above the upper quartile. The remaining 52 participants (24 female, 26 male, one non-binary, and one other) were between 20 and 38 years of age, with a median age of 25.5 years. All participants had self-reported normal hearing and gave their written consent for participation as well as collection, processing, and publication of their data.

Twenty-three of the participants stated they had never performed a listening experiment before, 17 seldom, eight several times, and four many times. Prior to the experiment, the participants were also asked: "*How often do you notice electric vehicles (cars/busses/trucks) and the special sounds they emit in your everyday life?*" to which two responded "*never*," nine participants responded "*rarely*," 16 responded "*occasionally*," 17 responded "*frequently*," and eight participants responded "*very frequently*," which indicates that most of the subjects had some prior experience with electric vehicle sounds.

https://doi.org/10.1121/10.0036248

III. RESULTS

After pre-processing the data following the methods described in Sec. II E, the results were statistically analyzed using MATLAB R2024b and assuming a significance level of 0.05. The effect of vehicle sound type on localization time and error was investigated using separate repeated measures analyses of variance (rmANOVAs) with Greenhouse–Geisser correction for lack of sphericity in some of the data; the effect size for these rmANOVAs is given as η_p^2. *Post hoc* paired comparisons were Bonferroni corrected, and Cohen's d_z was calculated as a standardized measure of effect size (Lakens, 2013). The influence of vehicle sound type on the percentage of failed localizations was evaluated using a non-parametric Friedman test in combination with Bonferroni corrected Wilcoxon signed rank paired comparisons.

The following sections present the results of this analysis separately for single vehicles (Sec. III A), two vehicles with the same type of sound (Sec. III B), three vehicles with the same type of sound (Sec. III C), and two vehicles with different sounds (Sec. III D). In addition to the results presented here, the data were analyzed for between-subject effects of gender, age, or self-reported exposure to EV sounds. However, no significant between-subject effects were found, and since these aspects are not considered crucial for the main research question of this study, those results are not further discussed here.

A. Single vehicles

Figure 4 shows the localization error (a) and localization time (b) results for single vehicles, averaged over all evaluated positions and repetitions. A rmANOVA confirmed a significant main effect of the sound type on the localization error for single vehicles [$F(2.48, 126.57) = 14.30$, $p_{gg} < 0.001$, $\eta_p^2 = 0.22$]. *Post hoc* paired comparisons indicated that the two-tone AVAS resulted in a significantly larger localization error than the combustion engine noise ($\Delta_{T-C} = 2.17°$, $d_z = 0.77$), the noise AVAS ($\Delta_{T-N} = 1.54°$, $d_z = 0.65$). and than the multi-tone AVAS ($\Delta_{T-M} = 1.52°$, $d_z = 0.52$). The difference between multi-tone AVAS and combustion noise ($p_{bon} = 0.118$, $\Delta_{M-C} = 0.64°$) and the difference between noise AVAS and combustion noise ($p_{bon} = 0.263$, $\Delta_{N-C} = 0.63°$) were not found significant.

A second rmANOVA confirmed a significant effect of stimulus type on the localization time for single vehicles [$F(2.40, 122.46) = 29.35$, $p_{gg} < 0.001$, $\eta_p^2 = 0.37$] and *post hoc* paired comparisons indicated that all three AVAS signals resulted in a significantly longer localization time than the combustion engine noise ($\Delta_{N-C} = 0.32$s, $d_z = 0.71$; $\Delta_{M-C} = 0.41$ s, $d_z = 0.87$; $\Delta_{T-C} = 0.66$ s, $d_z = 1.14$).

Additionally, the two-tone AVAS was detected significantly more slowly than both the noise AVAS ($\Delta_{T-N} = 0.33$ s, $d_z = 0.52$) and the multi-tone AVAS ($\Delta_{T-M} = 0.25$ s, $d_z = 0.54$). There was no significant difference between multi-tone AVAS and noise AVAS in localization time ($p_{bon} = 0.929$, $\Delta_{M-N} = 0.09$ s).

In summary, participants localized single vehicles with two-tone AVAS less accurately than the combustion noise and the other two AVAS signals, and the localization time of all three AVAS signals was slower than for the combustion engine noise with the two-tone AVAS being localized most slowly. Even though statistically significant, all these differences are relatively small for road-traffic scenarios, with the mean localization errors for all types of sound being below 6° and the differences in mean localization time reaching a maximum of 0.66 s. Except for a small number of outliers, all participants successfully localized all vehicles in the single-vehicle trials, i.e., responded within the time limit of 10 s with an absolute error below 90°. Therefore, the number of failed localizations is not further evaluated in this single-vehicle analysis.

B. Two vehicles with same sound

Adding a second vehicle with the same type of sound increased the influence of the sound type on localization error and localization time, as shown in Fig. 5. Performing two separate rmANOVAs confirmed this increased effect

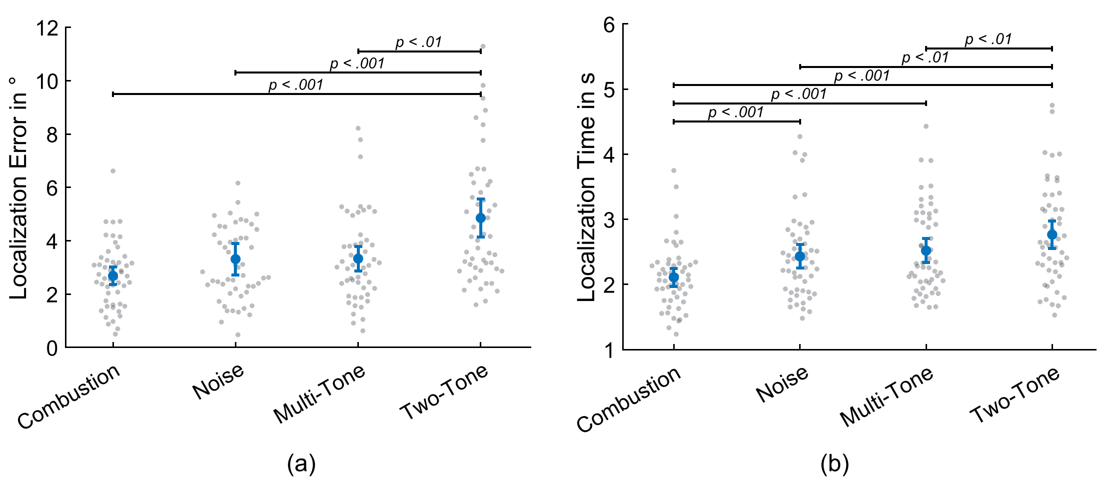

FIG. 4. (a) Mean localization error, (b) localization time for single vehicle condition including Bonferroni corrected p-values for significant paired comparisons. The gray dots show individual subject results. The blue error bars show arithmetic means with 95% confidence intervals.

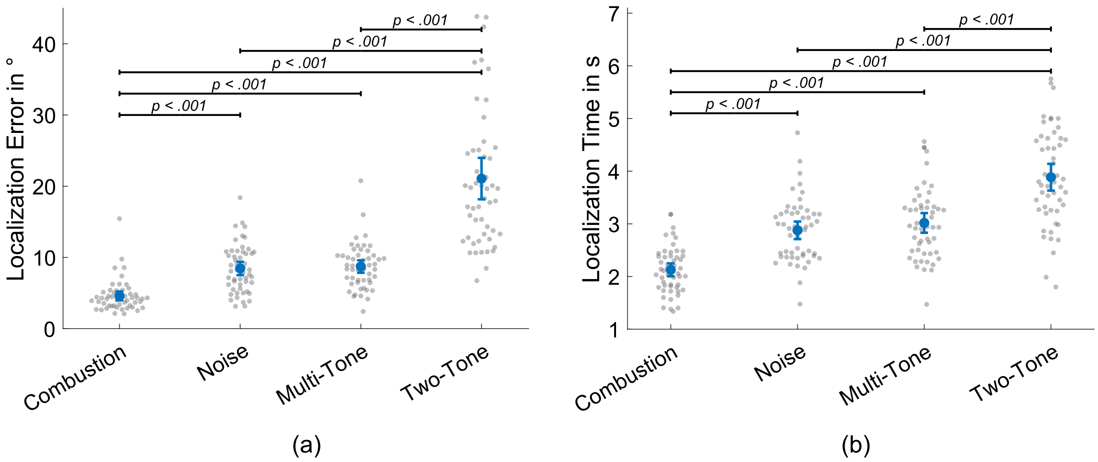

FIG. 5. (a) Mean localization error, (b) localization time for two vehicles with the same type of sound, including Bonferroni corrected *p*-values for significant paired comparisons. The gray dots show individual subject results. The blue error bars show arithmetic means with 95% confidence intervals.

size of stimulus type on localization error [$F(1.29, 65.54) = 95.45$, $p_{\mathrm{gg}} < 0.001$, $\eta_{\mathrm{p}}^2 = 0.65$)] and on localization time [$F(2.41, 122.95) = 125.11$, $p_{\mathrm{gg}} < 0.001$, $\eta_{\mathrm{p}}^2 = 0.71$] compared to the single-vehicle scenario. Unlike for single vehicles, *post hoc* paired comparisons indicated that all three AVAS signals caused a significantly larger localization error than the combustion engine noise ($\Delta_{\mathrm{N-C}} = 3.85°$, $d_z = 1.16$; $\Delta_{\mathrm{M-C}} = 4.13°$, $d_z = 1.21$; $\Delta_{\mathrm{T-C}} = 16.47°$, $d_z = 1.57$) and were detected significantly more slowly than the combustion engine noise ($\Delta_{\mathrm{N-C}} = 0.75$ s, $d_z = 1.40$; $\Delta_{\mathrm{M-C}} = 0.89$ s, $d_z = 1.62$; $\Delta_{\mathrm{T-C}} = 1.76$ s, $d_z = 2.21$). As for the single-vehicle scenario, the two-tone AVAS resulted in significantly larger errors than the two other AVAS signals ($\Delta_{\mathrm{T-N}} = 12.62°$, $d_z = 1.29$; $\Delta_{\mathrm{T-M}} = 12.34°$, $d_z = 1.25$) and is detected significantly slower than the two other AVAS signals ($\Delta_{\mathrm{T-N}} = 1.00$ s, $d_z = 1.33$; $\Delta_{\mathrm{T-M}} = 0.87$ s, $d_z = 1.19$). There was no significant difference in localization error ($p_{\mathrm{bon}} > 0.999$, $\Delta_{\mathrm{M-N}} = 0.27°$) and in localization time ($p_{\mathrm{bon}} = 0.321$, $\Delta_{\mathrm{M-N}} = 0.14$ s) between the noise AVAS and the multi-tone AVAS.

In addition to localization time and error, the percentage of failed localizations per participant, i.e., the percentage of presented vehicles that a participant did not localize within the time limit of 10 s or localized with an error of more than 90°, was evaluated. As shown in Fig. 6(a), 51 out of 52 participants achieved 0% of failed localizations for the combustion noise while, for the AVAS sounds, the amount of failed localizations is much larger, with some participants even failing to detect more than 50% of the two-tone AVAS vehicles when presented two at a time. A non-parametric Friedman test confirmed the significance of this effect of sound type on failed localizations [$p < 0.001, \chi^2(3) = 102.54$] and Wilcoxon signed rank paired comparisons revealed a significant difference between all three AVAS signals and the combustion engine sound ($p_{\mathrm{bon}} < 0.001$, $Z_{\mathrm{C-T}} = -6.02$, $Z_{\mathrm{C-M}} = -4.52$, $Z_{\mathrm{C-N}} = -4.28$) as well as a significant difference between the two-tone AVAS and the multi-tone and noise AVAS ($p_{\mathrm{bon}} < 0.001, Z_{\mathrm{M-T}} = -5.75$, $Z_{\mathrm{N-T}} = -5.62$). However, the paired comparisons showed no significant difference in failed localizations between the multi-tone AVAS and the noise AVAS ($p_{\mathrm{bon}} > 0.999$, $Z_{\mathrm{M-N}} = 1.06$).

Overall, simultaneously presenting two vehicles with the same sound enhanced the differences that were already observed for the single-vehicle case. In all three evaluated metrics, the AVAS signals performed worse than the

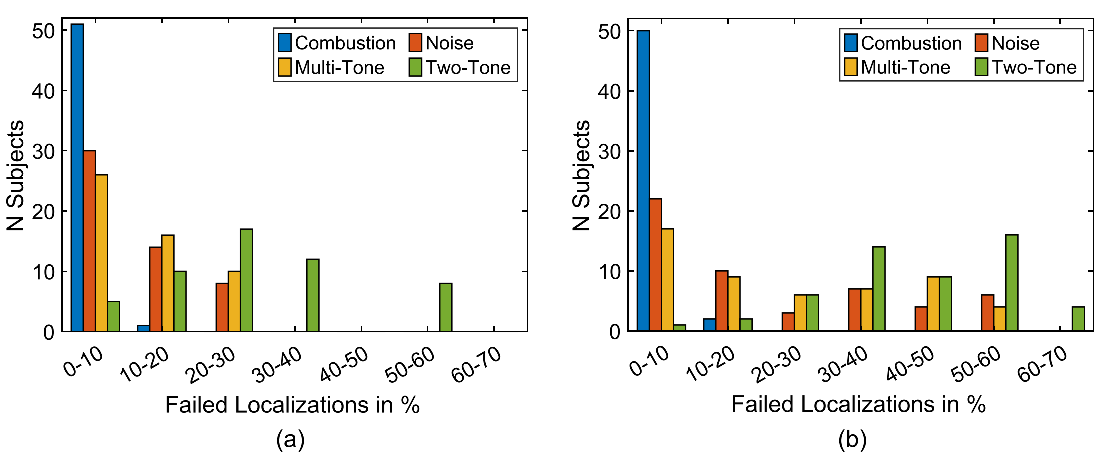

FIG. 6. Distribution of failed localizations for two vehicles with the (a) same sound, (b) for three vehicles with the same sound.

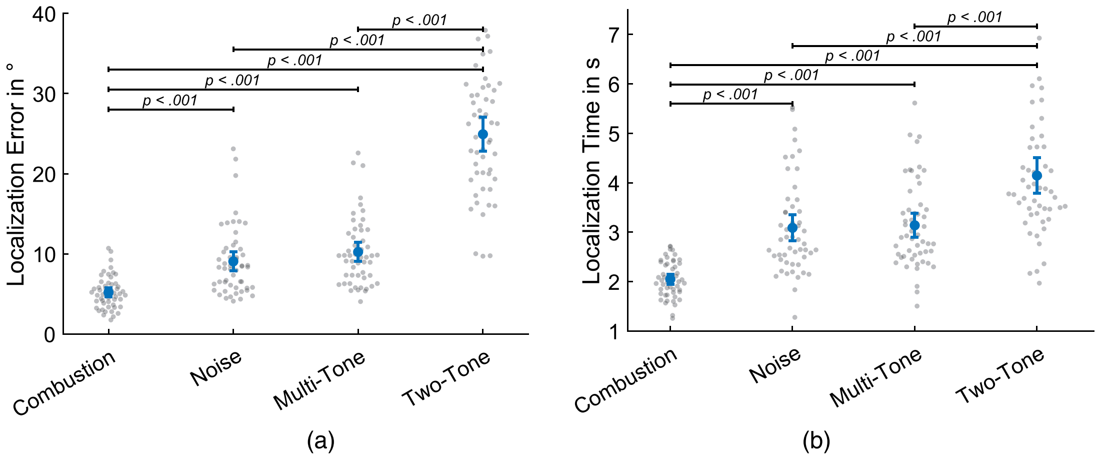

FIG. 7. (a) Mean localization error, (b) localization time for three vehicles with the same type of sound, including Bonferroni corrected p-values for significant paired comparisons. The gray dots show individual subject results. The blue error bars show arithmetic means with 95% confidence intervals.

combustion noise, with the two-tone AVAS being less localizable than the noise AVAS and than the multi-tone AVAS, reaching a mean localization error of up to 21° and a mean localization time of up to 3.9 s. The number of failed localizations underlines that participants had no problems localizing the combustion noise while struggling with the AVAS signals.

C. Three vehicles with the same sound

Compared to the two-vehicle results, simultaneously presenting three vehicles with the same sound resulted in an even larger effect of the sound type on localization error $[F(1.68, 85.93) = 177.94,\ p_{gg} < 0.001,\ \eta_p^2 = 0.78]$ and a similar effect size of sound type on localization time $[F(1.97, 100.66) = 99.24,\ p_{gg} < 0.001,\ \eta_p^2 = 0.66]$ as shown in Fig. 7. *Post hoc* paired comparisons indicated that all AVAS signals performed significantly worse than the combustion noise both in terms of localization error ($\Delta_{N-C} = 3.86°$, $d_z = 1.00$; $\Delta_{M-C} = 5.05°$, $d_z = 1.18$; $\Delta_{T-C} = 19.74°$, $d_z = 2.45$) and localization time ($\Delta_{N-C} = 1.04$ s, $d_z = 1.31$; $\Delta_{M-C} = 1.09$ s, $d_z = 1.52$; $\Delta_{T-C} = 2.10$ s, $d_z = 1.72$). As in the single-vehicle and two-vehicle scenarios, the participants localized the two-tone AVAS with a significantly larger error ($\Delta_{T-N} = 15.88°$, $d_z = 1.78$; $\Delta_{T-M} = 14.69°$, $d_z = 1.78$) and a significantly slower time ($\Delta_{T-N} = 1.06$ s, $d_z = 1.09$; $\Delta_{T-M} = 1.01$ s, $d_z = 1.13$) than the other two AVAS signals. The multi-tone AVAS and the noise AVAS did not significantly differ in localization error ($p_{bon} = 0.299$, $\Delta_{M-N} = 1.18°$) or localization time ($p_{bon} > 0.999, \Delta_{M-N} = 0.05$ s).

For three vehicles with the same type of sound, the number of participants with a high percentage of failed localizations increased for all three AVAS signals as shown in Fig. 6(b) with 20 out of 52 participants having more than 50% of failed localizations for the two-tone AVAS. Not a single participant successfully localized all two-tone AVAS vehicles when presented with three at a time. Conversely, 45 out of 52 participants achieved 0% of failed localizations for the combustion engine noise. The significance of this effect of sound type on the percentage of failed localizations was confirmed by a Friedman test $[p < 0.001, \chi^2(3) = 116.99]$, and *post hoc* paired comparisons revealed a significant difference between all three AVAS signals and the combustion engine sound ($p_{bon} < 0.001$, $Z_{C-T} = -6.30$, $Z_{C-M} = -5.67$, $Z_{C-N} = -5.39$) as well as a significant difference between the tonal AVAS and the multi-tone AVAS and noise AVAS ($p_{bon} < 0.001, Z_{M-T} = -5.73$, $Z_{N-T} = -5.78$). No significant difference in the percentage of failed localizations was found between multi-tone and noise AVAS ($p_{bon} = 0.118, Z_{M-N} = 1.57$).

Overall, the three-vehicle results show similar trends as the two-vehicle results but with even larger localization errors and more failed localizations. All AVAS signals were performing worse than the combustion noise, and the two-tone AVAS stands out as the least localizable with a mean error of 27° and a mean localization time of 4.1 s.

D. Two vehicles with different sound

The purpose of simultaneously evaluating two vehicles with different warning sounds was to investigate whether certain AVAS sounds are more prone to being affected by other sounds. In order to keep this paper concise, it was decided to present only localization error results as shown in Fig. 8, which were found to be more conclusive than localization time or the percentage of failed localizations. Each sub-figure shows the localization error for each of the four evaluated signals in the presence of a second vehicle, a so-called masker. The *none* case equals the single-vehicle results, i.e., no second vehicle sound being present. Performing multiple rmANOVAs confirmed that the masker type had a significant effect with varying effect sizes on all four evaluated sounds (cf. Table III). Thereby, it is evident that the localization error for each sound was most affected by a second vehicle that radiates the same type of sound. For example, Fig. 8(a) shows the localization error for a combustion engine vehicle in the presence of either no

https://doi.org/10.1121/10.0036248

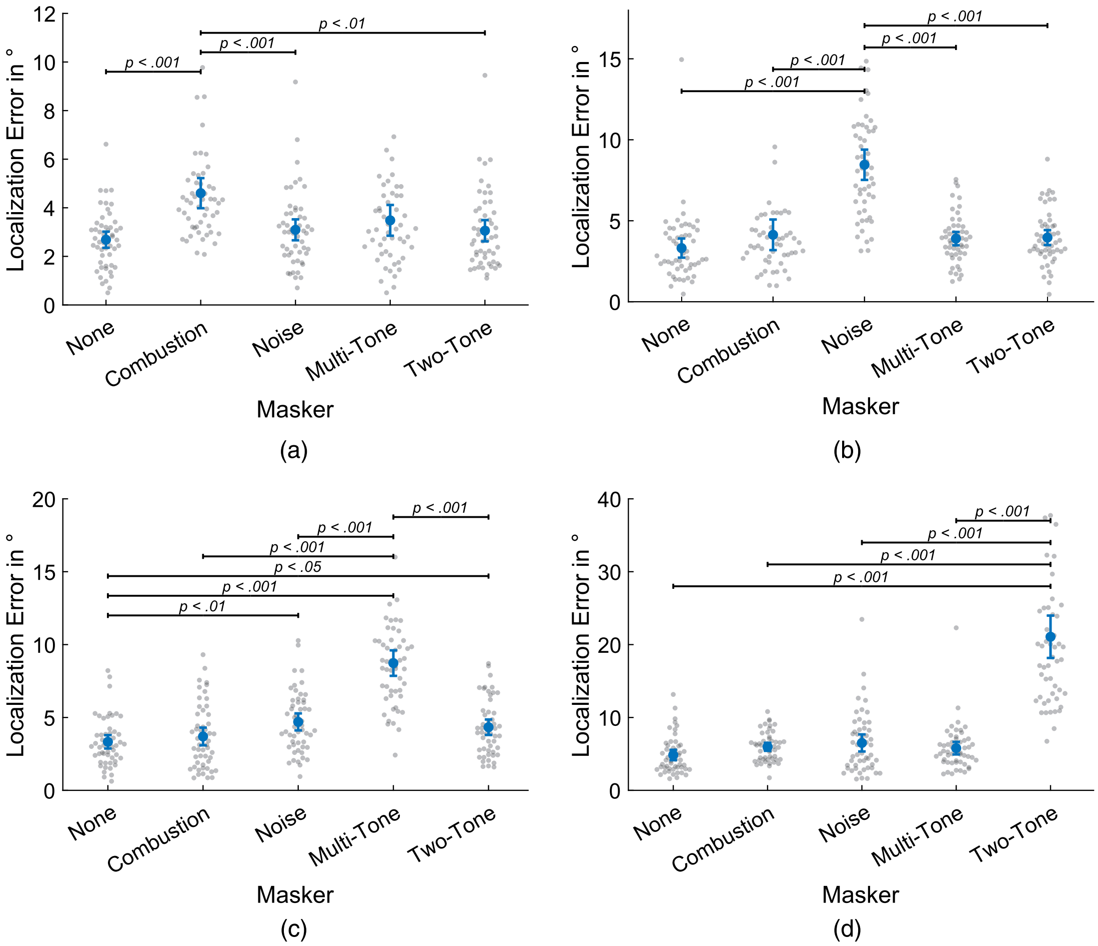

FIG. 8. (a) Mean localization error for combustion, (b) noise AVAS, (c) multi-tone AVAS, (d) two-tone AVAS in the presence of a second masker sound including Bonferroni corrected p-values for significant paired comparisons. The *none* case corresponds to the single-vehicle scenario with no second sound present. The cases where both sounds are the same correspond to the results shown in Fig. 5(a). The gray dots show individual subject results. The blue error bars show arithmetic means with 95% confidence intervals.

second vehicle (*none* data) or a second vehicle with any of the four evaluated sounds. The highest mean localization error for this case was obtained when introducing a second combustion engine vehicle, i.e., the combustion noise was masked most by a second combustion noise. *Post hoc* paired comparisons, as reported by the p-values in Fig. 8, confirmed that this error in the presence of a second combustion engine vehicle was significantly higher compared to no second vehicle being present ($\Delta = 1.92°$, $d_z = 0.85$). This effect is also apparent for the other signal types, i.e., the noise AVAS was masked by a second noise AVAS [cf. Fig. 8(b), $\Delta = 5.14°$, $d_z = 1.54$], the multi-tone AVAS was masked by a second multi-tone AVAS [cf. Fig. 8(c), $\Delta = 5.40°$, $d_z = 1.57$], and the two-tone AVAS was masked by a second two-tone AVAS [Fig. 8(d), $\Delta = 16.22°$, $d_z = 1.53$]. However, the only signal that seems to be significantly affected by other types of sound is the multi-tone AVAS, which also shows a significantly increased localization error in the presence of a noise AVAS ($\Delta = 1.37°$, $d_z = 0.58$) and in the presence of a two-tone AVAS ($\Delta = 1.00°$, $d_z = 0.48$) compared to no masker being present. Nevertheless, even though statistically significant, these mean differences are so small that they could be considered irrelevant in the traffic safety context of this study.

TABLE III. Repeated measures ANOVA results for two vehicles with different types of sound (cf. Fig. 8).

Maskee	df	F	p_{gg}	η_p^2
Combustion	3.32, 169.49	10.20	<0.001	.17
Noise	2.96, 151.11	42.12	<0.001	.45
Multi-tone	2.93, 149.18	59.95	<0.001	.54
Two-tone	1.59, 81.19	88.48	<0.001	.63

IV. DISCUSSION

This study resulted in two primary outcomes: First, the two-tone AVAS was localized significantly worse than combustion engine noise and than the other two AVAS signals under all evaluated scenarios. The multi-tone AVAS and the filtered noise AVAS showed a similar level of localizability. Second, introducing multiple vehicles with the same type of warning sound negatively affected the localizability of all evaluated sounds, whereby the two-tone AVAS was affected the most. In those multi-vehicle cases, all AVAS sounds perform worse than the combustion noise.

That the two-tone AVAS performed worst in terms of localization accuracy is, as such, not surprising, given that it has been known since the early days of psychoacoustic research that wideband sound sources are easier to localize than narrowband or tonal sounds (Stevens and Newman, 1936; Blauert, 1996; Yost and Zhong, 2014). However, most of these psychoacoustic studies were performed on pure tones and under synthetic laboratory conditions. Our results confirm that two-tone signals with added amplitude modulation in the presence of outdoor background noise are localized worse than other, more broadband, sounds. While localization accuracy and localization time might generally be affected by a speed-accuracy trade-off, i.e., a faster response tends to be less accurate than a slower response (Heitz, 2014), the results show that combustion noise is not only localized more accurately but also faster while the two-tone AVAS is localized slower and less accurately.

Nevertheless, considering only single-vehicle cases, one could argue that differences in localization accuracy of less than 3° are, even though statistically significant, hardly relevant for real-life traffic scenarios. While the corresponding variations in localization time are larger, it is hard to predict whether, e.g., a 0.66 s slower mean localization time matters in the context of EV traffic safety. Instead, early vehicle detection could be considered more important, for which research, such as Hastings and McInnis (2015), showed that tonal signals perform better than noise bands with the same frequency and energy, especially in the presence of urban background noise. On the other hand, the same research indicated that detectability generally increases with the number of different third-octave band components. While some specific two-band sounds were found to be equally detectable as some four-band sounds, this was not the case for all two-band stimuli. Combined with the present study findings, a sound that contains energy in multiple different third-octave bands is presumably easier to detect and localize than a single- or two-tone sound and could be better suited as an AVAS signal.

The second main finding of this study is that simultaneously introducing multiple EVs with the same type of warning sound drastically increases the localization error, the localization time, and the amount of failed localizations. Perhaps most importantly, when presenting three vehicles with the same sound at a time, the number of failed localizations increases to a level where many participants failed to localize a considerable amount of the presented EVs. While this effect is worst for the two-tone AVAS, for which more than half of the participants localized less than 60% of the presented vehicles, the other two AVAS sounds also yielded a significantly higher number of failed localizations than the combustion noise. Even though this measure of failed localizations does not differentiate whether participants did not notice the second or third vehicle at all, localized it with a too-large error or did not manage to mark all vehicle positions within 10 s, all three explanations could be considered hazardous in a traffic context. The fact that the localizability of the evaluated combustion noise is much less affected by the presence of multiple similar-sounding vehicles shows that this outcome is not caused by an improper experiment setup, e.g., asking for an impossible task, but that it is possible to quickly and accurately localize three different vehicles in the evaluated scenario. It remains to be an open question whether the combustion noise performed best solely because of its acoustic properties or whether the fact that most participants have been exposed to it throughout their entire lives and hence might have trained its localization plays a role as well. Even though the subjects were interviewed regarding their self-reported electric vehicle sound exposure (cf. Sec. II F), the data presented in this study do not allow us to conclude how personal experience and long-term training might affect localizability, and further experiments systematically investigating this aspect are necessary.

Independent of the cause and the potential impact on real-life traffic safety, the presented results show that participants struggle to accurately localize multiple simultaneously presented EVs with the same type of warning sound. This finding is especially relevant in the context of the "sameness requirement" of the current US regulations, which states that vehicles of the same model shall be designed to have the same AVAS sound (National Highway Traffic Safety Administration, 2016a). In combination with not requiring a pitch shift, the scenario that multiple vehicles with similar AVAS sounds approach a pedestrian is not unrealistic. Based on our findings, it might be beneficial to implement a certain level of randomness so that two vehicles never radiate exactly the same AVAS sound. However, while the presented results show the benefit of combining fundamentally different AVAS types, they do not allow us to predict which degree of difference between two sounds would be sufficient to increase localizability. For example, it is possible that two AVAS signals of the same type but with different parameters, e.g., 2, two-tone AVAS sounds with a large difference in pitch and modulation rate, do not significantly improve localizability compared to two vehicles with exactly the same sound. Additionally, this suggestion is only based on localizability, and further human subject studies are required to assess whether the detectability of multiple vehicles with the same type of sound is affected by a similar degree.

The scope of this study was limited to investigating differences in the localizability of AVAS signals that fulfill the current regulations without assessing which specific aspects of the evaluated signals might have contributed to this outcome. Therefore, a more systematic study on how factors such as amplitude or frequency modulation, impulsiveness, and frequency range affect localization accuracy in the presence of other vehicles is necessary. We simulated free-field cases of stationary vehicles, aiming to reproduce a parking lot scenario that might not apply to, for example, intersections with reflecting buildings and moving cars. However, other studies suggest that the presence of sound-reflecting surfaces either does not affect horizontal localization accuracy at all (Guski, 1990) or, when resulting in perceivable reverberation, rather reduces the human ability to localize a sound source (Giguère and Abel, 1993). Since the participants were able to rotate their heads freely and the experiment procedure required them to constantly turn around their own axis, adding vehicle movements would not have introduced any additional binaural localization cues. Hence, we do not expect that introducing environmental reflections or adding vehicle movement would improve the localizability of the signals that showed a bad performance in this stationary free-field experiment. On the contrary, literature regarding representational momentum in spatial hearing suggests that the final position of moving sound sources may be localized as displaced in the direction of motion (Getzmann and Lewald, 2007), i.e., moving AVAS sounds might be localized worse than our findings for stationary sounds suggest. However, this assumption should be further investigated in a follow-up experiment.

As a final limitation, one could argue that normalizing all signals to the same loudness removes the advantage of a tonal signal being perceived as louder than a bandpassed noise with the same energy. Since current regulations only specify sound pressure levels, normalizing the sound pressure might be more relevant from a regulatory point of view. Such a normalization would have resulted in the combustion noise being perceived as much louder than the AVAS signals and the multi-tone AVAS being perceived as louder than the noise and two-tone AVAS. However, as discussed in Sec. II B 7, all three AVAS signals were presented with sound pressure levels at least 10 dB above the minimum US and UNECE requirements and under relatively calm background noise conditions. This means that all signals were loud enough to be audible, and literature indicates that, as long as a signal is audible, the sound level does not significantly influence localizability (Yost, 2016). The fact that the evaluated two-tone AVAS had an *A*-weighted equivalent continuous sound pressure level 10 dB above the combustion noise but still was localized significantly worse supports this assumption that sound pressure level alone is not decisive for localizability. Nevertheless, it would be meaningful to include different types and levels of urban background noise in a potential follow-up experiment to gain more insights into the relation between localizability, detectability, loudness, and sound pressure level.

As a final remark, it is essential to underline that this study only investigated the auditory localization of different AVAS signals. In order to judge which sounds might be the most suitable solution overall, their environmental noise impact needs to be considered as well, and further research regarding the possible effects of AVAS sounds on public health and well-being is needed.

V. CONCLUSION

This study presented a laboratory experiment on the auditory localization of electric vehicle alerting sounds, comparing three common AVAS signal types, i.e., bandpass filtered noise, two amplitude-modulated tones, and a multitude of amplitude-modulated tones, to a combustion engine recording. The results show that the combustion engine noise consistently achieved the highest localization accuracy and fastest localization time, while the two-tone AVAS showed the worst localizability under all evaluated conditions. There was no significant difference between the multi-8tone AVAS and the noise AVAS. While there was barely an interaction between two vehicles with different sounds, introducing multiple vehicles with the same type of sound drastically increased the differences between the evaluated sound types. Under this multi-vehicle condition, all three AVAS signals performed significantly worse than the combustion engine noise, with the two-tone AVAS being the least localizable, reaching mean localization errors up to 20° larger and localization times up to 2 s longer than the combustion noise. Additionally, the number of failed localizations increased drastically for this scenario, with most participants localizing less than 60% of the presented vehicles when simultaneously hearing 3, two-tone AVAS sounds. Even though early detection might be considered more relevant for traffic safety than accurate localization, this large amount of failed localizations may indicate that the participants did not always detect all simultaneously presented vehicles. This assumption should be further investigated by performing detectability experiments involving multiple EVs with the same type of AVAS sound.

ACKNOWLEDGMENTS

This research was funded by FORMAS, a Swedish research council for sustainable development, under Grant No. FR-2020-01931. The authors would like to thank all study participants and are particularly grateful for the expert advice from Marlene Wessels (Johannes Gutenberg University Mainz, Germany).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Ethics Approval

The study was conducted in accordance with the Declaration of Helsinki and approved by the Swedish ethical review authority (Etikprövningsmyndigheten, 2024–04880, September 12, 2024). Informed consent was obtained from all subjects involved in the listening experiment.

DATA AVAILABILITY

All stimuli recordings, a video showcasing the experiment procedure, and the data that support the findings of this study are openly available on Zenodo at http://doi.org/10.5281/zenodo.14261299.

Ahrens, J. (**2012**). *Analytic Methods of Sound Field Synthesis* (Springer, Berlin Heidelberg).

Blauert, J. (**1996**). *Spatial Hearing: The Psychophysics of Human Sound Localization* (The MIT Press, Cambridge, MA).
EAA (**2020**). "Environmental Noise in Europe," European Environment Agency.
Edwards, P. J., Moore, S., and Higgins, C. (**2024**). "Pedestrian safety on the road to net zero: Cross-sectional study of collisions with electric and hybrid-electric cars in Great Britain," J. Epidemiol. Community Health **78**(8), 487–492.
Emerson, R. W., Kim, D. S., Naghshineh, K., Pliskow, J., and Myers, K. (**2013**). "Detection of quiet vehicles by blind pedestrians," J. Transp. Eng. **139**(1), 50–56.
European Computer Manufacturers Association (**2022**). "Psychoacoustic metrics for ITT equipment -Part 2 (models based on human perception)," in *ECMA-418-2*, 2nd ed. (ECMA).
Fiebig, A. (**2020**). "Electric vehicles g*et al*ert signals to be heard by pedestrians: Benefits and drawbacks," Acoust. Today **16**(4), 20–28.
Fleury, S., Jamet, E., Roussarie, V., Bosc, L., and Chamard, J.-C. (**2016**). "Effect of additional warning sounds on pedestrians' detection of electric vehicles: An ecological approach," Accid. Anal. Prev. **97**, 176–185.
Garay-Vega, L., Hastings, A., Pollard, J. K., Zuschlag, M., and Stearns, M. (**2010**). "Quieter cars and the safety of blind Pedestrians: Phase 1," Technical Report No. 811 304 (DOT HS).
Getzmann, S., and Lewald, J. (**2007**). "Localization of moving sound," Percept. Psychophys. **69**(6), 1022–1034.
Giguère, C., and Abel, S. M. (**1993**). "Sound localization: Effects of reverberation time, speaker array, stimulus frequency, and stimulus rise/decay," J. Acoust. Soc. Am. **94**(2), 769–776.
Guski, R. (**1990**). "Auditory localization: Effects of reflecting surfaces," Perception **19**(6), 819–830.
Hanna, R. (**2009**). "Incidence of pedestrian and bicyclist crashes by hybrid electric passenger vehicles," Technical Report (DOT HS), Vol. 811, pp. 1–25.
Hastings, A., Guthy, C., Pollard, J. K., and Garay-Vega, L. (**2012**). "Research on minimum sound specifications for hybrid and electric vehicles - Final report," Technical Report No. 642, Docket no. NHTSA-2011-0148 (U.S. DOT).
Hastings, A., and McInnis, C. (**2015**). "Detectability of alert signals for hybrid and electric vehicles: Acoustic modeling and human subjects experiment," Technical Report Docket No. NHTSA-2016-0125 (U.S. DOT).
Hastings, A., Pollard, J. K., Garay-Vega, L., Stearns, M., and Guthy, C. (**2011**). "Quieter cars and the safety of blind pedestrians, Phase 2: Development of potential specifications for vehicle countermeasure sounds," Technical Report No. 811 496 (DOT HS).
Heitz, R. P. (**2014**). "The speed-accuracy tradeoff: History, physiology, methodology, and behavior," Front. Neurosci. **8**, 86875.
Hou, X., Su, M., Liu, C., Li, Y., and Ma, Q. (**2023**). "Examination of the factors influencing the electric vehicle accident size in Norway (2020–2021)," World Electr. Veh. J. **15**(1), 3–15.
Institute of Electronic Music and Acoustics (**2024**). "IEM Plug-in Suite" (Institute of Electronic Music and Acoustics, Graz), available at https://plugins.iem.at (Last viewed 11/12/2024).
International Energy Agency (**2024**). *Global EV Outlook 2024* (International Energy Agency), available at https://iea.blob.core.windows.net/assets/a9e3544b-0b12-4e15-b407-65f5c8ce1b5f/GlobalEVOutlook2024.pdf.
Kim, D. S., Emerson, R. W., Naghshineh, K., Pliskow, J., and Myers, K. (**2012**). "Impact of adding artificially generated alert sound to hybrid electric vehicles on their detectability by pedestrians who are blind," J. Rehabil. Res. Dev. **49**, 381–393.
Lakens, D. (**2013**). "Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs," Front. Psychol. **4**, 863.
Lobato, T., and Sottek, R. (**2023**). "Modeling the perceived tonal loudness of multiple tonal components," in *Proceedings of the 49th Annual German Conference on Acoustics* [German Acoustical Society (DEGA), Berlin, Germany].
Mendonça, C., Freitas, E., Ferreira, J., Raimundo, I., and Santos, J. (**2013**). "Noise abatement and traffic safety: The trade-off of quieter engines and pavements on vehicle detection," Accid. Anal. Prevent. **51**, 11–17.
Morgan, P. A., Morris, L., Muirhead, M., Walter, L. K., and Martin, J. (**2011**). "Assessing the perceived safety risk from quiet electric and hybrid vehicles to vision-impaired pedestrians," Technical Report No. PPR525.
Müller, L., Forssén, J., and Kropp, W. (**2024**). "Dataset: Auditory localization of multiple electric vehicles," Zenodo. https://doi.org/10.5281/zenodo.14261299 (Last viewed 11/12/2024).
Müller, L., and Kropp, W. (**2024**). "Auralization of electric vehicles for the perceptual evaluation of acoustic vehicle alerting systems," Acta Acust. **8**, 27.
National Highway Traffic Safety Administration (**2013**). "Notice of proposed rulemaking - Minimum sound requirements for hybrid and electric vehicles," Technical Report No. NHTSA-2011-0148 (National Highway Traffic Safety Administration).
National Highway Traffic Safety Administration (**2016a**). "FMVSS No. 141 - Minimum sound requirements for hybrid and electric vehicles" (National Highway Traffic Safety Administration).
National Highway Traffic Safety Administration (**2016b**). "Minimum sound requirements for hybrid and electric vehicles FMVSS No. 141: Final regulatory impact analysis," Technical Report (National Highway Traffic Safety Administration).
Pallas, M.-A., Bérengier, M., Chatagnon, R., Czuka, M., Conter, M., and Muirhead, M. (**2016**). "Towards a model for electric vehicle noise emission in the European prediction method CNOSSOS-EU," Appl. Acoust. **113**, 89–101.
Parizet, E., Ellermeier, W., and Robart, R. (**2014**). "Auditory warnings for electric vehicles: Detectability in normal-vision and visually-impaired listeners," Appl. Acoust. **86**, 50–58.
Poveda-Martínez, P., Peral-Orts, R., Campillo-Davo, N., Nescolarde-Selva, J., Lloret-Climent, M., and Ramis-Soriano, J. (**2017**). "Study of the effectiveness of electric vehicle warning sounds depending on the urban environment," Appl. Acoust. **116**, 317–328.
Roan, M. J., Neurauter, L., Song, M., and Miller, M. (**2021**). "Probability of detection of electric vehicles with and without added warning sounds," J. Acoust. Soc. Am. **149**(1), 599–611.
Sandberg, U., Goubert, L., and Mioduszewski, P. (**2010**). "Are vehicles driven in electric mode so quiet that they need acoustic warning signals?," in *Proceedings of 20th International Congress on Acoustics* [International Commission for Acoustics (ICA), Madrid, Spain].
Singh, S., Payne, S., Mackrill, J., and Jennings, P. (**2015**). "Do experiments in the virtual world effectively predict how pedestrians evaluate electric vehicle sounds in the real world?," Transp. Res. Part F: Traffic Psychol. Behav. **35**, 119–131.
Stelling-Kończak, A., Hagenzieker, M., Commandeur, J. J., Agterberg, M. J., and Wee, B. V. (**2016**). "Auditory localisation of conventional and electric cars: Laboratory results and implications for cycling safety," Transp. Res. Part F: Traffic Psychol. Behav. **41**, 227–242.
Stelling-Kończak, A., Hagenzieker, M., and Wee, B. V. (**2015**). "Traffic sounds and cycling safety: The use of electronic devices by cyclists and the quietness of hybrid and electric cars," Transp. Rev. **35**(4), 422–444.
Stevens, S. S., and Newman, E. B. (**1936**). "The localization of actual sources of sound," Am. J. Psychol. **48**(2), 297–306.
Ulrich, T. A., Barton, B. K., and Lew, R. (**2014**). "Detection and localization of approaching vehicles in the presence of competing vehicle noise," Transp. Res. Part F: Traffic Psychol. Behav. **26**, 151–159.
United Nations Economic Commission for Europe (**2017**). "Regulation No 138 of the Economic Commission for Europe of the United Nations (UNECE) — Uniform provisions concerning the approval of quiet road transport vehicles with regard to their reduced audibility [2017/71]," Technical Report.
Wessels, M., Kröling, S., and Oberfeld, D. (**2022**). "Audiovisual time-to-collision estimation for accelerating vehicles: The acoustic signature of electric vehicles impairs pedestrians' judgments," Transp. Res. Part F: Traffic Psychol. Behav. **91**, 191–212.
Wu, J., Austin, R., and Chen, C.-L. (**2011**). "Incidence rates of pedestrian and bicyclist crashes by hybrid electric passenger vehicles: An update," Technical Report No. 811526 (DOT HS).
Yost, W. A. (**2016**). "Sound source localization identification accuracy: Level and duration dependencies," J. Acoust. Soc. Am. **140**(1), EL14–EL19.
Yost, W. A., and Zhong, X. (**2014**). "Sound source localization identification accuracy: Bandwidth dependencies," J. Acoust. Soc. Am. **136**(5), 2737–2746.
Zotter, F., Pomberger, H., and Noisternig, M. (**2012**). "Energy-preserving Ambisonic decoding," Acta Acust. united Acust. **98**(1), 37–47.