DOI: 10.71568/daga2026.321

Comparing different simulation approaches of a moving vehicle in a Wave Field Synthesis laboratory

Jonas Krautwurm[1], Ercan Altinsoy[1]

[1] *Institute of Acoustics and Speech Communication, Chair of Acoustics and Haptics, TU Dresden*

Introduction

In traffic scenarios, pedestrians must be aware of their surroundings. Besides actively avoiding objects, they must make important decisions, such as when to cross the street. Due to the higher electrification of vehicles, the sounds of approaching vehicles are becoming more and more unconventional. This is why acoustic cues and how people react to them must be investigated.

To quantify the behavioral patterns of pedestrians, perception experiments were conducted in virtual environments. The Time to Collision (TTC) point is especially a good indicator of whether pedestrians over- or underestimate an approaching vehicle. [1] [2] [3] In these investigations, single source recordings of electric and combustion vehicles were used to create simulations in two different playback systems (HOA and WFS). The different perceptional behavior in these laboratories has already been investigated regarding different subjective attributes. [4] In addition, approaches were developed for synthesizing electric vehicle sounds, particularly AVAS and tire noise. Müller and Kropp created a useful toolbox for synthesizing AVAS and tire noise that also includes directivity patterns. [5] Other studies have conducted TTC experiments with a simplified version of the sub-source sounds but experienced problems with sound level calibration and large TTC estimation offsets as a consequence. [6]

Considering these calibration issues, this study carries out a new experiment with a new synthesis approach. Additionally, driving scenes of a vehicle are simulated using dummy head recordings of real traffic scenarios. This paper presents initial results from a complex listening test that provide insights into how participants rate different simulation approaches based on certain subjective attributes.

Methods

Before creating different vehicle simulations, source sounds must be recorded. In addition to choosing the car and test track, this section discusses other preliminary work.

Test track/ car selection

A Cupra Born EV with 170 kW of power was used as a test car. It was fitted with 215/50 R19 AllSeasonContact tires. Cruise control could be activated at speeds of 20 and 30 kph in 1 kph increments. Although the AVAS sound for EVs only has to be present at low speeds [7], the Cupra Born AVAS sound is present at speeds above 40 kph. [8] For recording the vehicle at all relevant speeds over a sufficient distance, a street at an industrial site was chosen. It has a long, straight road with light traffic.

Recording Procedure

Vehicle sounds come from three main sources. Tire and engine sounds (or AVAS sounds) are dominant at lower speeds, while wind noise becomes relevant at speeds of 60 kph and above. [9] For this study, which focuses on speeds up to 30 kph, AVAS and tire noises were recorded using four Brüel & Kjaer free-field measurement microphones (Type 4188) and preamps (Type 2671), which were placed around the vehicle. (Figure 1) A previous study is recommended for analyzing the exact positioning of the microphones. [10] For recording the audio files, a SQuadriga III interface from Head Acoustics was used. The interface is equipped with a GPS antenna that collects position data in the WGS format for the vehicle and its speed for later simulation. The microphones were calibrated before recording, and the calibration value was saved in the software.

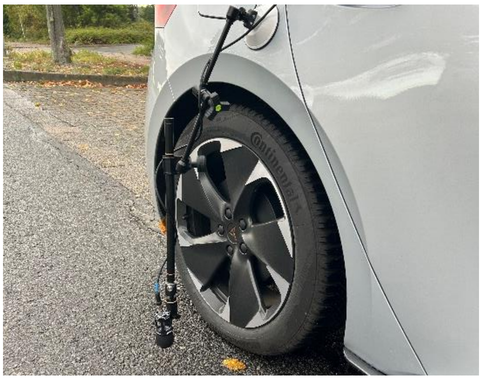


Figure 1: Free field microphone placement besides the tire.

In addition to the close-up microphones, a dummy head (HSU III, Head Acoustics) was placed 1.5 meters away from the passing vehicle. It was equipped with windshields, and the Independent of Direction (ID) equalization from Head Acoustics was used. The dummy head microphones were 1.6 meters high, which is the height approximation of ears of a pedestrian.

Signal Analysis

For a first overview, a FFT analysis is performed for the audio recordings of the close-up microphones of the back tire and engine hood in Figure 2.

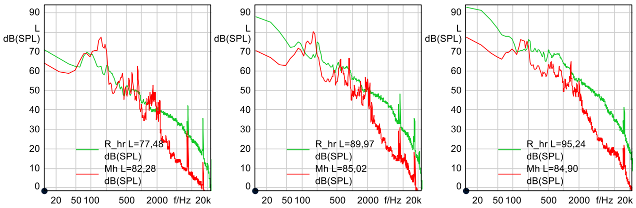
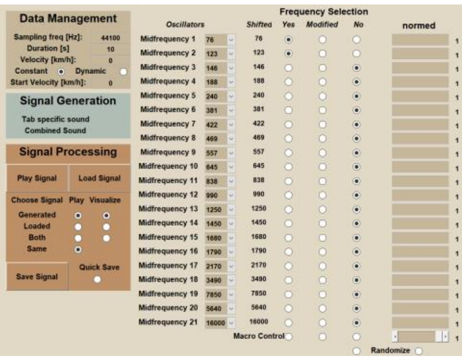


Figure 2: FFT spectrum of the recordings of the back tire and engine hood (abbr.: R_hr and Mh) for three constant speeds. (from left to right: 10, 20 and 30 kph)

The red line represents the FFT spectrum of the recording taken in front of the engine hood. As can be seen, the AVAS sound creates several frequency peaks in the FFT spectrum. For example, there is a peak at about 1000 Hz at a speed of 10 kph. This is caused by the use of frequency- and amplitude-modulated sinusoidal sounds. Electric vehicles sound more futuristic and less conventional because they use a different sound design than combustion engine vehicles, which create sound through the combustion process and the sound of several engine parts. At higher speeds, these peaks shift to higher center frequencies. An increase in those peaks is suggested by an increase in speed. [7] The green line represents the FFT spectrum of the tire sound from the right rear tire. Compared to the AVAS sound, the frequency content of the tire recording is more present under 100 Hz, especially at speeds of 20 kph and 30 kph. Since the influence of wind becomes dominant below that frequency [11], it can be assumed that the microphone in front of the rear tire is influenced more by the vehicle flow field than the microphone in front of the engine hood. One reason for this could be the increased turbulence right beside the tire. Based on the analysis of the original vehicle sounds, a synthetic sound is recreated.

Sound Synthesizer

This simulation uses two different synthesizing approaches. The AVAS sound is characterized by specific frequency peaks. The first simplified synthesizer (Figure 3) uses sinus oscillators, for which the central frequency and normalized amplitude can be entered.



Data Management

Sampling freq [Hz]: 44100
Duration [s] 10
Velocity [km/h]: 0
Constant ● Dynamic ○
Start Velocity [km/h]: 0

Signal Generation

Tab specific sound
Combined Sound

Signal Processing

Play Signal | Load Signal

Choose Signal	Play	Visualize
Generated	●	●
Loaded	○	○
Both	○	○
Same	●	

Save Signal | Quick Save ○

Frequency Selection

Oscillators		Shifted	Yes	Modified	No	normed
Midfrequency 1	76	76	●	○	○	1
Midfrequency 2	123	123	●	○	○	1
Midfrequency 3	146	146	○	○	●	1
Midfrequency 4	188	188	○	○	●	1
Midfrequency 5	240	240	○	○	●	1
Midfrequency 6	381	381	○	○	●	1
Midfrequency 7	422	422	○	○	●	1
Midfrequency 8	469	469	○	○	●	1
Midfrequency 9	557	557	○	○	●	1
Midfrequency 10	645	645	○	○	●	1
Midfrequency 11	838	838	○	○	●	1
Midfrequency 12	990	990	○	○	●	1
Midfrequency 13	1250	1250	○	○	●	1
Midfrequency 14	1450	1450	○	○	●	1
Midfrequency 15	1680	1680	○	○	●	1
Midfrequency 16	1790	1790	○	○	●	1
Midfrequency 17	2170	2170	○	○	●	1
Midfrequency 18	3490	3490	○	○	●	1
Midfrequency 19	7850	7850	○	○	●	1
Midfrequency 20	5640	5640	○	○	●	1
Midfrequency 21	16000	16000	○	○	●	1
		Macro Control	○	○	○	1

○ Randomize ○

Figure 3: Additive AVAS sound synthesizer MATLAB GUI.

To generate the vehicle's specific AVAS sounds, the maximum levels and center frequencies of the dominant frequency peaks are analyzed and normalized according to the peak with the highest level, which corresponds to a normalized value of 1. Then, the level of the rendered synthetic sound is compared to that of the original close-up recording of the vehicle and adjusted accordingly. A different synthetic approach is required for the tire sound. First, a reference tire sound and white noise are loaded into the program. The FFT spectrum of both sounds is analyzed, and a level difference is calculated over specific frequencies. Using this difference, a FIR filter is created, and the filter coefficients are saved. This process uses recordings of the right rear tire of the vehicle at speeds of 10, 20, and 30 kph. Ultimately, the filtered white noise replaces the individual sound sources in the simulation approach. Additionally, the synthetic tire sounds were filtered with a high-pass butterworth filter (6th order, cut off frequency: 120 Hz), to not simulate wind induced sounds in the low frequency range. Figure 4 shows the FFT analysis of certain sounds.

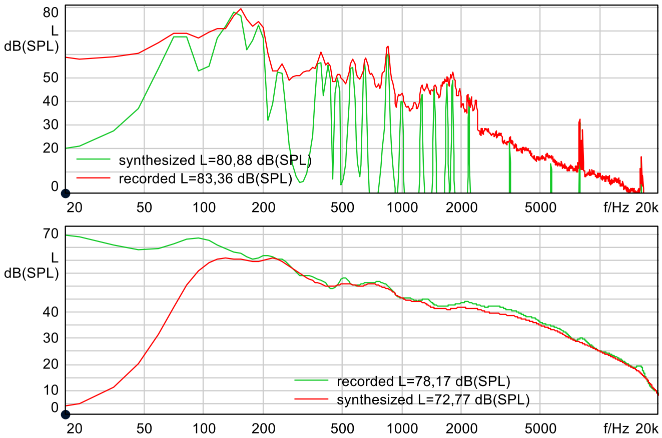


Figure 4: FFT spectrum of the synthesized and recorded close-up vehicle sounds from the AVAS sound (above) and the tire sound (below) for a speed of 10 kph.

WFS Simulation

A special audio rendering system featuring Wave Field Synthesis is used for the traffic scene simulation. In total, there are 464 built-in loudspeakers and four subwoofers. Eight rendering PCs allow up to 32 individual point sources to be played simultaneously. The laboratory is acoustically treated with perforated metal sheets and mineral wool, and the reverberation time is only about 0.2 seconds. [12] The laboratory was also calibrated and equalized, as detailed stated in a previous study. [10] While the recording-based simulation and synthetic approach uses a TASCAR-based toolbox [4], the static approach uses a different method. When using direction-independent recording equalization, the dummy head acts as a stereo microphone with a mixed stereophony concept. [13] In consequence, the two microphone recordings were placed as point sources in an equilateral triangle relative to the listener position in the laboratory's virtual environment, and the static approach is a simple simulation in a stereo playback setup.

Comparison Simulation Approaches

A calibrated free-field microphone (MG MMS 212) was placed at the listener position in the laboratory. The driving scenes of the Cupra were simulated within a 15-second time interval (10 seconds before and 5 seconds after the collision point) and recorded. For the FFT analyses in Figure 5, the left ear recording from the dummy head at the test track was also added. (blue line).

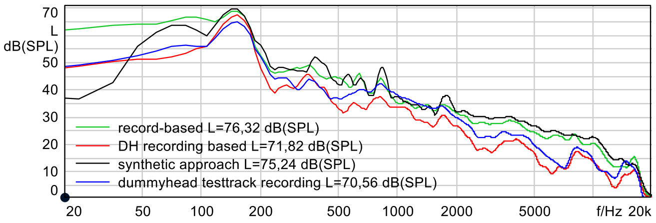


Figure 5: FFT spectra of the different driving by simulations compared to the real test track dummy head recording at a speed of 10 kph.

The figure shows that the close-up and synthetic recordings are generally louder than the real driving scenario. First, the calibration of the single-sourced recordings could be set too high. Additionally, wind-induced noise dominant under 100 Hz leads to a higher low-frequency level and, consequently, a higher overall level. As Figure 5 shows, the frequency content under 50 Hz decreases due to high-passing the synthetic tire noise. This allows the wind rumble to be regulated for the synthetic approach. When comparing the real test track recording to the static simulation approach, the frequency content is nearly identical, though the static simulation approach is about 2 dB louder.

Experimental design

Parameters

This paper is based on an experiment that has two parts. The first part determined TTC times, but only the results of the second part are analyzed in this study. Three different vehicle speeds and three different simulation approaches were set as parameters. (Table 1)

Table 1: Parameters of the scenes for the listening experiment.

Speeds [kph]	Simulation approaches
10	Recording based
20	Synthetic based
30	Static approach

The sections before describe how the different simulation approaches were created and which recordings had to be done.

Procedure

The participants stood at the listener position on a pedestal in the laboratory. Nine scenes were played in total (three caused by the speed parameter and three caused by the simulation approach parameter). Before the actual experiment, the participants listened to three randomly selected driving by scenes. They could play each scene as many times as they wanted. Afterwards, they had to rate four subjective attributes loudness, realism, localization ability, and timbre on a scale from zero to one hundred percent.

Participants

In total 25 participants (13m, 12f) took part in the experiment. The age ranged from 18 to 62 years with an average age of 33 years. They have shown no hearing damage and understood the procedure well.

Results and discussions

At first, a graph of the average rating results of the subjective attributes is shown in Figure 6.

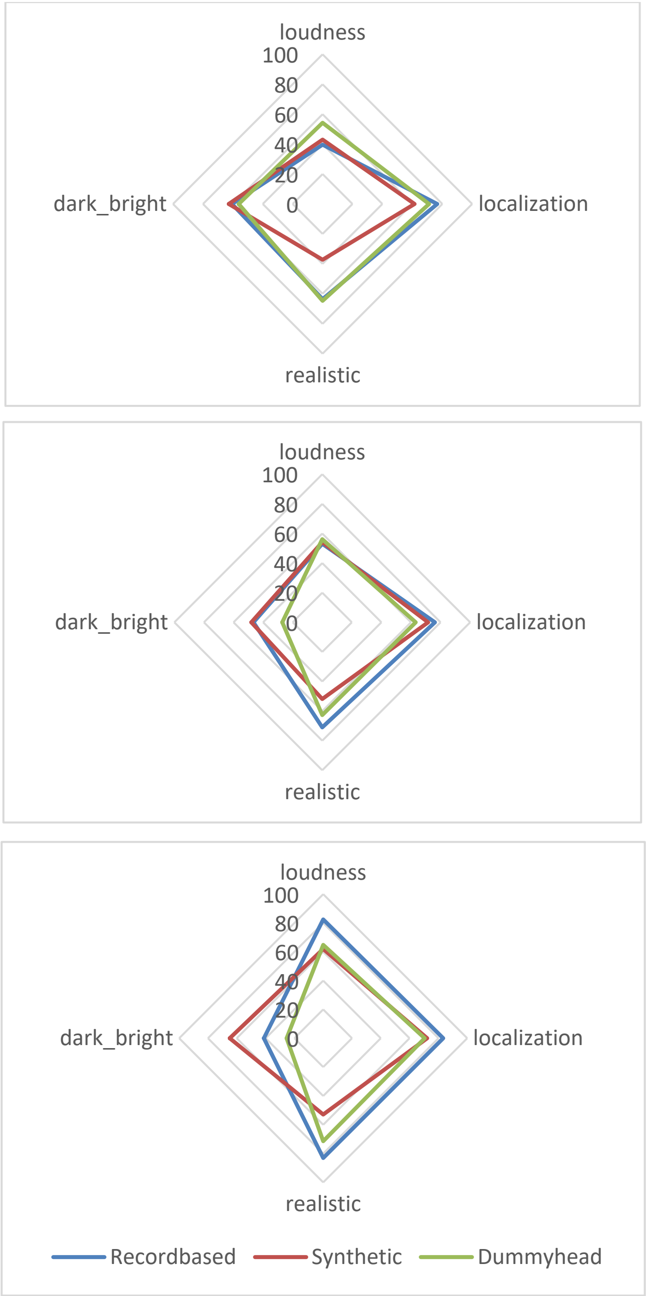


Figure 6: Attribute rating results for vehicle speeds of 10, 20 and 30 kph (from top to bottom).

Some average values catch the eye. The static approach, in which a stereo playback was rendered in a laboratory, was perceived as darker at vehicle speeds of 20 and 30 kph. Probably the background noise, that was present at the industrial site induced such a low frequency content. At 30 kph, the synthetic approach was perceived as much brighter. Wind noise rumble was recorded with the close-up microphones. In the simulation with synthetic sub-source sounds, this frequency content was cut with a high pass filter, which avoided this frequency range. Consequently, the synthetic vehicle scene was perceived as brighter as the original record-based approach. For the static approach, two point sources were placed in a stereo setup in the VR. Although the width of the vehicle scenes is limited to the range of stereo panorama, participants rated the localization ability similarly to that of the single-source simulation approach. Regarding the attribute "realism", a clear trend emerged. The synthetic approach was consistently rated as more unrealistic. This could be due to the simpler additive and subtractive synthesis of the sub-source sounds. In contrast, both the static approach and the recording-based simulations were rated as more realistic. The recording-based scene at 30 kph was rated as louder than the other vehicle scenes. One reason for this could be the low-frequency noise induced by the wind at this speed, which is picked up by the close-up microphones and makes the VR simulation generally louder.

Conclusions

This study compared state-of-the-art simulation approaches for vehicle scenes. First, synthetic recreations of source-based vehicle recordings were built using simplified concepts. Then, a new simulation approach based on dummy head test track recordings was introduced. In a listening test, participants rated the simulation approaches based on different subjective attributes, providing an initial assessment of the quality of the VR simulations. Next, it will be investigated whether differences in attribute ratings influence reaction times, such as TTC. Additionally, future studies will focus on reducing wind noise in simulations because it clearly influences sound pressure levels and loudness perception.

Acknowledgements

This project is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 444809588. The project is part of the priority program AUDICTIVE - SPP2236.

The authors want to thank all subjects that took part in the experiment.

Literatur

[1] D. Oberfeld, M. Wessels und D. Büttner, "Overestimated time-to-collision for quiet vehicles: Evidence from a study using a novel audiovisual virtual-reality system for traffic scenarios," *Accident; analysis and prevention*, Early Access. doi: 10.1016/j.aap.2022.106778.

[2] L. Steinbach, F. Beyer, M. E. Altinsoy, D. Oberfeld-Twistel und M. Wessels, Hg., *Safety Investigation on Traffic Scenarios using Virtual Environments in a Wave Field Synthesis Laboratory*, 2022.

[3] M. Wessels, C. Zähme und D. Oberfeld, "Auditory Information Improves Time-to-collision Estimation for Accelerating Vehicles," *Current Psychology*, Jg. 42, Nr. 27, S. 23195–23205, 2023, doi: 10.1007/s12144-022-03375-6.

[4] J. Krautwurm, D. Oberfeld-Twistel, T. Huisman, M. M. Maravich und E. Altinsoy, "Quality Comparison of Dynamic Auditory Virtual-Reality Simulation Approaches of Approaching Vehicles Regarding Perceptual Behavior and Psychoacoustic Values," *Acoustics*, Jg. 7, Nr. 1, 2025, doi: 10.3390/acoustics7010007.

[5] L. Müller und W. Kropp, "Auralization of electric vehicles for the perceptual evaluation of acoustic vehicle alerting systems," *Acta Acust.*, Jg. 8, S. 27, 2024, doi: 10.1051/aacus/2024025.

[6] F. Beyer, S. Fischer, L. Steinbach und M. E. Altinsoy, Hg., *Comparison of Recorded and Synthesized Stimuli of Traffic Scenarios in an Auditory Virtual Reality Environment Using Wave Field Synthesis*, 2023.

[7] *Regelung Nr. 138 der Wirtschaftskommission für Europa der Vereinten Nationen (UNECE) — Einheitliche Bestimmungen für die Genehmigung geräuscharmer Straßenfahrzeuge hinsichtlich ihrer verringerten Hörbarkeit [2017/ 71],* Amtsblatt der Europäischen Union, 2017.

[8] SEAT S.A., Hg., *How electric sound is Born*, 2021.

[9] R. O. Rasmussen, R. J. Bernhard, U. Sandberg und E. P. Mun, "The Little Book of Quieter Pavements," The Transtec Group Inc., Springfield, 2007.

[10] J. Krautwurm und M. E. Altinsoy, Hg., *Comparing Time To Collision estimations: Experiments in a Wave Field Synthesis Laboratory and the Real World Environment*, 2025.

[11] J. Wuttke, *Mikrofonaufsätze*. Karlsruhe: Schalltechnik Dr.-Ing. Schoeps GmbH, 2000.

[12] M. E. Altinsoy, U. Jekosch, S. Merchel und J. Landgraf, Hg., *Progress in Auditory Perception Research Laboratories - Multimodal Measurement Laboratory of Dresden University of Technology*, 2010.

[13] M. Dickreiter, V. Dittel, W. Hoeg und M. Wöhr, *Handbuch der tonstudiotechnik,* 8. Aufl. (De Gruyter Saur reference). Berlin: De Gruyter Saur, 2014.