THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

AURALIZATION, PERCEPTION AND DETECTION OF TYRE-ROAD NOISE

ALICE HOFFMANN

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers University of Technology
Gothenburg, Sweden, 2016

AURALIZATION, PERCEPTION AND DETECTION OF TYRE-ROAD NOISE
ALICE HOFFMANN
ISBN 978-91-7597-469-9

Doktorsavhandlingar vid Chalmers tekniska högskola
Chalmers University of Technology
Ny serie nr 4150
ISSN 0346-718X

Department of Civil and Environmental Engineering
Division of Applied Acoustics, Vibroacoustics Group
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone: +46 (0)31-772 2200

Printed by
Chalmers Reproservice
Gothenburg, Sweden 2016

AURALIZATION, PERCEPTION AND DETECTION OF TYRE-ROAD NOISE

ALICE HOFFMANN

Department of Civil and Environmental Engineering
Division of Applied Acoustics, Vibroacoustics Group
Chalmers University of Technology

Abstract

Due to improvements in combustion engines and electric engines for cars, tyre noise has become the prominent noise source at low and medium speeds. Models exist that simulate the noise produced by a rolling tyre, as do models that auralize specific traffic situations from a basic data set. A model that combines both could assist in the planning stage of a tyre by delivering not only estimates of the physical behaviour of the tyre, but also by further making the resulting sound perceivable. Further, such a model could help to design acoustic traffic situations with full control of all parameters. Focusing on that, this thesis has three aims. All focus is on the perception of the sound of a car from the outside, perceived by, for example, a pedestrian. The first aim is to combine an established model for tyre noise (SPERoN) with an auralization tool. The combined model can predict the spectrum of a car pass-by at 7.5 m, as well as reproduce the sound at a given listener position. Psychoacoustic judgements are used to compare the modelled signals with recorded signals. It was found that responses for simulated and recorded signals correlated for all cases, but the ranked orders differed slightly. The second aim focuses on the perception of tyre–road noise and whether it can be differentiated and characterized by its perceptual qualities. When designing tyre sounds, the main aim should be to increase the pleasantness of the total vehicle sound while maintaining the carried information and reducing the sound level. Achieving this requires an understanding of how physical changes in a tyre are reflected in the perception of that tyre. Listeners were asked to judge different tyre–road combinations and their perception in terms of their emotional and psychoacoustic responses. The results confirmed that rolling noise can be perceptually differentiated. The third aim in this thesis was to increase understanding of the parameters that influence the detection of a single car in background traffic noise. For this, both variations in the sound of the test car and in the background (e.g. distance, traffic amount, speed, tyre/engine noise) were investigated and found to influence the reaction time. The introduced auralization method was utilized to generate the sound files for the different traffic situations.

Keywords: *Psychoacoustics, Auralization, Perception, Tyre/Road Noise, Rolling Noise, Traffic Detection, Reaction Time*

List of publications

This thesis is based on the work contained in the following appended papers:

Paper I
Auralisation model for the perceptual evaluation of tyre-road noise
J. Forssén, A. Hoffmann, W. Kropp
to be submitted, 2016.

Paper II
Auralization of simulated tyre noise: psychoacoustic validation of a combined model.
A. Hoffmann, W. Kropp
submitted to Applied Acoustics, 2016.

Paper III
Perception of tyre noise: Can tyre noise be differentiated and characterized by the perception of a listener outside the car?
A. Hoffmann, P. Bergman, W. Kropp
submitted to Acta Acustica united with Acustica, 2016.

Paper IV
Psychometric function for car pass-by in background noise based on simulated data
A. Hoffmann, W. Kropp
submitted to Acta Acustica united with Acustica, 2016.

The following publications are not included in the thesis due to an overlap in content or due to being published in another language:

Auralization of tyre/road noise based on the SPERoN prediction tool.
A. Hoffmann, J. Forssén, W. Kropp
Proceedings of AIA-DAGA 2013 Conference on Acoustics, Meran, Italy, March 18-21, 2013.

Optimierung der Auralisierung von Reifengeräuschen basierend auf dem Modellierungs-Tool SPERoN
A. Hoffmann, J. Forssén, W. Kropp
DAGA 2014, Fortschritte der Akustik, 40. Deutsche Jahrestagung für Akustik, 10.–13. März 2014, Oldenburg.

There's a car coming? - Psychometric function for car pass-by in background noise based on simulated data.
A. Hoffmann, P.Bergman, W. Kropp
Proceedings of Euronoise 2015, 10th European Congress on Noise Control Engineering p. 2417-2421., Maastricht, Netherlands, 2015

Perception of rolling noise
A. Hoffmann
Licenciate thesis, Chalmers University of Technology, Göteborg, Sweden, 2015

Acknowledgements

The work included in this thesis was financially supported by the German Federal Highway Research Institute (BAST) within the Project 'Silent road traffic noise 3 (Leiser Straßenverkehr 3)'.
First and foremost I would like to thank my supervisor, Professor Wolfgang Kropp, for his support and guidance during the process leading to this thesis and for the freedom and support to follow own ideas and interests.

I would also like to thank Jens Forssén for the help with the auralization, Penny Bergman for the guidance and support during listening tests and for all the helpful discussions and ideas and Lars Hansson for the technical support during the detection experiments. Thanks to all the PhD students and colleagues at the Department of Applied Acoustics for the good atmosphere, an open ear when needed and a great time together. Special thanks to Gunilla and Börje for all the help with technical and administrative matters.
Thanks to my friends both in Sweden and in Gemany for the support and kind words of encouragement.
Thanks to my parents and grandparents for the support throughout the years, for feeding my curiosity with piles of books and for letting my interest in science grow.
Last but not least endless thanks to my dearest husband Jona for the support during the last years, the hours of prove reading and a cup of tea when needed.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Road traffic noise has historically been dominated by engine noise. Due to improvements of combustion engines and especially the new electric alternatives for car engines, modern car engines have become increasingly silent. This has led to tyre–road noise being the most prominent noise source nowadays for the speed range from about 30 km/h up to 100 km/h [121, 130]. Consequently, today tyre–road noise is the main source of road traffic noise.

For car manufacturers, the main focus concerning perception of noise and vibration properties is on how the customer perceives the interior sound quality of the car. However, focusing on the problem of environmental noise, the area of noise perception has to be expanded to include the noise radiated by the car to the environment. A Word Health Organization (WHO) report [48] estimates environmental noise to be responsible for the loss of about 1 million healthy life years in Europe. Environmental noise has been shown to increase the risk of cardiovascular disease, cognitive impairment in children and sleep disturbances. As a main contributor to road traffic noise, tyre–road noise is one of the main sources of environmental noise and thus a major factor in its health effects. As a consequence, in order to reduce the severe health effects of road traffic noise, it is essential to reduce the negative impact of tyre–road noise.

Most of the work being carried out today focuses on reducing tyre/road noise at the source by using low-noise road surfaces [120, 121] and low-noise tyres [132, 114]. Noise barriers and a series of innovative noise-control measures can also be used to reduce the propagation of traffic noise from source to receiver [13]. Finally, increased façade and window insulation is a last, desperate attempt to cope with the noise problem. However, practice has shown that this work is cumbersome and its progress is slow.

A design tool that can auralize the noise produced by the designed tyre or road would be of great help in the process. This requires a model for the sound generation process, as well as a method to generate a realistic audio signal for e.g. a pass-by situation under desired conditions. Such a tool would help to make tyre and road design more cost efficient.

On the other side, there are discussions about individual cars becoming too silent and thus increasing the risk of accidents, especially for visually impaired people. This was brought forward by the US National Highway Traffic Safety Administration in a study on accidents involving electric vehicles [111] and led to to the enactment of legislation in the USA [99] that demands a minimum sound level from new electric and hybrid vehicles. This is done by adding alert sounds to help pedestrians perceive the presence, direction, location and operation of the vehicle. These differing viewpoints on noise requirements may lead to a conflict of interests, where on the one hand there are health risks due to noise emitted by passenger cars, on the other hand there is an increasing risk of accidents due to non-detection if this noise is reduced. Further knowledge in the field of perception of vehicle noise is needed to overcome this conflict of interests.

1.2 Aim

The overall hypothesis behind the first two parts of this work is that tyre–road noise varies in the perception of a listener outside the car depending on road and tyre selection, and that this could be utilized as a complement to the ongoing work to reduce the negative consequences of tyre–road noise without losing the information the sound provides to identify and detect the sound source.

This work pursues three aims. All of its studies focus solely on passenger cars. The first aim is to model an authentic pass-by situation of a passenger car from the point of view of a listener on the roadside. The concept of the applied model is to combine the SPERoN prediction model [83] for tyre noise and the Listen Demonstrator [108]. This work combines both approaches to create a tool for auralizing tyre–road noise. The combined model can predict the spectrum of the sound close to the tyre, as well as reproduce the sound at a given listener position. This tool needs a set of parameters that define the properties of the tyre and the road as input data.

The second aim in this work is to verify that there is a perceptual difference for different tyre–road combinations and to show how these differences can be described. Investigating the perceptual space of tyre–road noise provides information about the main parameters that influence the perception and the possible spread in those parameters for tyre–road noise. This will provide a framework for possible changes and improvements.

In the last part, the focus shifts to traffic noise. The third aim in this work is to investigate the reaction time for detecting a car with a combustion engine (test vehicle) passing by in the presence of background noise from a road with high traffic flow under a set of conditions specific to the background traffic and test car. All parameters apart from the tested condition (i.e. car type, road surface, speed, etc.) are constant. One parameter that is known to commonly have a strong effect on the reaction time to a masked stimulus is the level difference between the two sounds [67]. In this study, this signal-to-noise ratio (SNR) is defined as the difference between the equivalent sound pressure level from the test vehicle and the equivalent sound pressure level from the road with high traffic flow. For each of the study parameters, an investigation will be made on the underlying effect through changes in SNR. One parameter that might affect the reaction time is the similarity between the signals. This study is based on the hypothesis that changes in similarity reflect on the reaction time independent of changes in SNR. To test the hypothesis, the similarity between the signals will be modified with additional tonal components to the test car in the first study. As an additional test, the similarity is increased again in the second study by adding the tonal component to both test car and background.

1.3 Outline

The thesis is divided into three parts. It starts with a general introduction and methodology (Chapters 1, 2 and 3). The second part deals with the perception of tyre–road noise, from two different angles. The first angle is the development of an auralization model, its validation and extensions (Chapter 4) and the second angle is the psychoacoustic investigation of tyre–road noise (Chapter 5). Both aspects are studied based on the same listening tests, but using different viewpoints and methods. In the third part, the focus is moved from individual cars and tyre–road noise to traffic noise (Chapter 6). The question of detectability of a single car in traffic noise is investigated utilizing auralized car pass-by.

Chapter 2 features a literature review on the perception of rolling noise. The fields of tyre–road noise and psychoacoustics are described in more detail. The general information about the methodology of the presented studies is collected in Chapter 3, which presents the psychoacoustic methodology applied in the studies and describes the applied statistical analysis model ANOVA. The utilized tyres and roads and their sources are also described in this chapter.

In Chapter 4 the auralization method is introduced and validated utilizing listening tests. This corresponds to the content of the first and second papers. The introduced auralization method is a combination of the SPERoN prediction model [83] for tyre noise and the Listen Demonstrator [108]. To verify the model, a set of listeners rated the perceived sound quality of the modelled signals as well as real recordings. The recorded and simulated signals incorporated the same tyres, roads and speeds. The listeners were asked to rate the pleasantness, loudness, roughness and sharpness of each signal. These properties are commonly seen to be important for sound quality perception in general [144]. The tests show that the combined model provides an estimate of how the real situation is perceived. If the pass-by situations are similar, the modelling uncertainties lead to differences in assessments.

In Chapter 5 the focus of the work shifts to the perception of rolling noise. A listening test and its statistical analysis are presented. This corresponds to the third paper. One basis for this work is the psychoacoustic annoyance defined by Zwicker and Fastl [144], as well as the components defining it: loudness, roughness, sharpness and fluctuation strength. These parameters are commonly used in comparable studies for different kinds of sound sources. Additionally, emotional responses were tested, since research [48] indicates that emotional reactions to sounds are related to health effects. Västfjäll et al. [136] evaluated the use of emotional measures on interior and exterior vehicle auditory quality, and found them well fitting. This motivates the inclusion of emotional measures in the evaluation of tyre–road noise.

In Chapter 6 the focus is on the detection of a single car in background noise. This corresponds to the fourth paper. For this, the auralization method developed in Chapter 4 is extended to full cars and traffic noise, and used to investigate the influence of individual parameters on the reaction time to a car in background traffic noise. Both parameters affecting the background (distance, traffic amount etc.) and parameters affecting the test car (engine type, speed etc.) were investigated.

A summary of the conclusions and an outlook on possible future work are presented in Chapter 7 and 8.

Chapter 2

Overview of the involved fields

This thesis is positioned at an intersection of different fields, namely tyre–road noise generation and modelling, auralization and psychoacoustics. Thus, it includes a literature review investigating the state of the art in in each of these fields concerning the presented studies and an introduction to the field of tyre–road noise, the field of auralization and to the field of psychoacoustics.

2.1 Literature review

The field of perception of tyre–road noise interacts with different areas. One set of studies focuses on the perception of noise inside the car. An example is the study by Bergeron et al. [21], who developed a method to describe the perception of internal automotive road noise by designing a sensory grid and a predictive tool in which calculated sound metrics are related to perceptual dimensions.

If the focus is more on noise and noise pollution in urban areas, the interest is not in the perception of interior noise in a car, but mainly on exterior noise. One field of ongoing studies is to apply psychoacoustics on car pass-by and traffic noise. Pörschmann et al. [109] evaluated the velocity and distance perception of moving sources to find important cues for moving sources, such as temporal loudness changes, binaural cues and Doppler shift. Park and Lee [102] analysed booming sounds and investigated which psychoacoustic parameters are related to this sound characteristic. Gärtner et al. [50] modified car pass-by sounds to investigate the perception of sound quality and to find basic parameters affecting the sound quality. Lee et al. [87] used another approach to investigate sound quality. They measured the electrical field that is induced by brain activity as a reaction to the sound of an accelerating car. This electrical field was then related to the perceived sound quality. Still another method was developed by Cik [33], in which the health effects of traffic noise are related to sleep disturbance and annoyance, and a method was developed to investigate these factors in situ.

A quite new field is to combine the different approaches of tyre–road noise research and psychoacoustics. Keulen and Duskow suggest in an inventory study of basic knowledge on tyre–road noise [130] that the use of psychoacoustics might give new insights into tyre noise. He used this approach to find explanations for disagreements in measurements and subjective ratings of the noisiness of silent roads. The use of the psychoacoustic concept of coloration helped to explain the emerging differences. Buss [30] follows this idea and investigates pattern noise as a part of tyre–road noise by evaluating sound quality assessments of professional subjective testers.

2.2 Tyre–road noise

Tyre–road noise describes the noise produced by a tyre rolling on a surface. It is the most prominent noise source for the speed range from typically 30 km/h up to 100 km/h [121, 130] for passenger cars, and it depends strongly on the speed of the car.

The start of the research on tyre–road noise generation dates back to the sixties and seventies. The very first workshop on tyre–road noise generation was held in Stockholm in 1979 [98]. At that time two mechanisms were identified as main contributors to tyre–road noise: tyre vibrations and air pumping. This view has been extended over time. The most important effects will be discussed in more detail as follows:

Tyre vibrations Tyre vibrations are caused by the interaction between road and tyre.The contact geometry between road and tyre varies with both tread pattern and roughness of the road surface. Consequently, the contact forces vary over time, as do the tyre vibrations. The time-varying tyre vibrations lead to sound radiation.

The response of the tyre to the varying contact forces depends strongly on the tyre properties, i.e. their geometry and material properties. An extensive study of the influence of the different design parameters on the vibration behaviour of the tyre can be found in [64].

Research showed that tyre vibrations can be responsible for the radiated sound within a wide frequency range. In the range of 1 kHz, the radiation is mainly determined by the motion of low-order modes on the tyre structure due to the time-varying contact shape [81]. These modes have high radiation efficiency that counterbalances a lower excitation energy.

Flow-related processes Traditional literature tends to use the term air pumping, which was introduced by Hayden [62]. Air pumping refers to a time-varying air flow, which creates monopole sources at the leading and the trailing edge of the tyre during rolling. However, the mechanisms leading to the time-varying flow are not very well understood and are largely speculative. The following phenomenological mechanisms are suggested:

In 1971, Hayden [62] suggested that, as the tread enters the leading edge of the road contact area, the tread is compressed and penetrates the road surface. This leads to the air being squeezed out of the void. At the trailing edge, the tread is decompressed and lifts up from the road surface, with the result that air flows back to fill the void.

Deffayet and Hamet [59] assumed that the opening and closing of cavities in the contact leads to sound generation. They measured the pressure in cylindrical cavities of different dimensions as a slick tyre rolled over the opening.

Ronneberger [113] assumed that when the tread was deformed by roughness asperities intruding into the rubber, air is displaced due to the changing gap between rubber and road surface. He considered this flow to be a monopole source and estimated the radiated sound.

None of these suggestions is able to explain the air pumping measured in the field in a satisfactory way. The meagre result might be due to the fact that experimental investigation is very difficult, since it is hard to directly observe the exact process in the contact between tyre and road during rolling without disturbing the process.

Less speculative was the work by Conte [34], who actually made the very first complete non-linear model of the flow in between tyre and road when the tyre passes over cavities in the road surface.

Much of the literature claims that air pumping is responsible for tyre–road noise above 1 kHz (see e.g. [121]). This is based on the observation that below 1 kHz the speed dependency is normally velocity squared (U^2), while above the dependency it is around U^4, which indicates that the sound-generation mechanism is airflow-related (monopole source). Winroth and Kropp [141] showed, however, that this U^4 dependency can even be obtained when only taking into account tyre vibrations in the simulations. Further, they observed this dependency even at lower frequencies, a fact which can hardly be explained by traditional air-pumping models. Therefore they introduced the expression air-flow-related mechanism, which is broader and in this context may be more accurate.

Other processes Further processes can be found in the literature, but they play only a minor role in very specialized situations. For very smooth surfaces, one might find stick-slip and stick-snap processes where friction and adhesion are involved.

Stick-slips occur when individual parts of the tread (e.g. a block) come into contact. Tangential stresses are then built up while this section travels through the contact.

At the end of the contact patch, the tangential stresses will exceed the frictional force and the section will start to slide (snap out) as described in [93], among others. This process is mainly related to high-frequency sound radiation, but is normally not so important.

Stick-snap describes forces occurring when the adhesive bonds between rubber and road change. If the tyre is sticky, the adhesive bonds may break up and cause vibrations. It can also happen that the adhesive bonds are increased, which leads to an increase of the excitation at the trailing edge of the tyre footprint [93]. Also, stick-snap is not one of the major effects contributing to the sound radiation.

Propagation processes The propagation of the tyre–road noise is influenced by various circumstances. The cavity formed by tyre and road in and opposing the rolling direction is shaped like a horn. It grows exponentially and thus gives a smooth impedance matching from the contact area to the surroundings. This leads to very effective sound radiation, especially in the region of 2 kHz [122]. The strength of the horn effect depends strongly on tyre width and the structure of the road surface. The narrower the tyre width and the more porous the road surface, the smaller the horn effect becomes ([78] and [56]).

The sound radiation is dependent on the absorptive parameters of the different media along the radiation path. One of the main considerations being the parts of the car surrounding the tyre, and the road surface. The more sound-absorbent these parts are, the less sound is radiated.

Another aspect is that the tyre cannot be seen as a spherical source. The levels in front of the tyre are higher than at the back of the tyre. These differences are presumably due to stick-snap phenomena [121]. The radiation to the side is even less than to the back, due to the horn effect mainly working in the forward and backward directions.

2.2.1 Model approaches

Today, independent models of varying complexity exist for predicting different phenomena involved in tyre–road interaction. In most of the models, tyre–road contact is substantially simplified, considering e.g. only stationary contact or rather crude contact models. In many cases tyre dynamics, tyre profile and road surface profiles are neglected, i.e. a smooth static tyre rolling on smooth surface is modelled. The quality of a tyre–road interaction model depends on the quality of the tyre model, the road model and the contact model used for the simulation. The following section gives an overview of the state of the art for these models.

Contact between tyre and road Typically a third-body approach is introduced for calculating the contact between two surfaces. This approach aims to characterize the local deformation of the involved bodies. One of the most common methods is to model the tread as a set of uncoupled springs with constant stiffness (e.g. [76]). A more advanced method is to use an elastic half-space formulation (e.g. [142]). Kalker's classic model for 3D contact between rolling bodies [71] follows this approach. It is applied in many practical applications, especially for wheel/rail contact. However, both approaches are far from sufficient to describe the contact between tyre and road. Uncoupled springs neglect the coupling of displacements within the tread. The elastic half-space includes this coupling, but demands that the contact area is small in comparison to a typical dimension of the body (e.g. the diameter of the minimum curvature radius of the involved bodies). Instead of a third-body model, the two bodies in contact could be modelled by taking into account the local elasticity of the structure.

In most of these approaches, the solution to the contact problem is carried out in two steps [28]. First, the contact problem for stationary rolling, i.e. a smooth tyre rolling on a smooth surface, is solved. After obtaining the contact geometry this way, the influence of road roughness is then applied as an external force acting on the deformed tyre structure. This only works for very small roughnesses, however, where the roughness does not alter the contact area – an assumption that is hardly correct for the tyre–road interaction.

Tyre models Tyres are composite structures with frequency-, temperature-, load- and strain-dependent properties. From a modelling point of view, inflation adds a dimension of complexity, as it alters the undeformed tyre shape and is the origin of additional pre-tension forces acting on the sidewalls and belt. Acoustic pressure fluctuations within the enclosed air cavity also induce vibrations on the tyre structure and vice versa, yielding a coupled fluid/structure problem. Models for simulating tyre dynamics range from analytical approaches, based on rough ([31] and [101]) simplifications in the description of the physical tyre properties, to highly elaborate numerical models that take into account the very details of the complicated tyre features [28]. In the 1980s and 1990s, two-dimensional models were developed that focused on noise generation from tyres. These also covered the medium and high frequencies (e.g. [75] and [92]). A further step forward was the model by Larsson and Kropp [85] consisting of two elastic layers. It is one of the few tyre models that have been demonstrated to capture the high-frequency response of the tread and especially its local deformation. A similar model, but with curvature in the circumference was later suggested by O'Boy and Dowling [97]. In parallel to these models, from the mid-1980s on, finite element models have been used to describe tyre dynamics (e.g. [112, 106, 88, 84]). Although capturing more geometrical details, they have been limited to the low-frequency region due to the computational cost at the time. Often, they are directly combined with a contact model (see [42, 28, 60, 89]). Nilsson [94] presented a model based on Waveguide Finite Element Method (WFEM). Fraggstedt

[47] developed this model further. The WFEM is one of the most advanced tyre models. Sabiniarz and Kropp used this model to discuss the vibration properties of tyres [119].

Sound radiation Typically, the Boundary Element Method (BEM) is utilized to calculate the radiated sound from tyre vibrations. For example, in [27] and [56], the amplification due to the horn effect was modelled with BEM, while in [118] and [139] the calculated radiation from a rolling tyre was simulated. An interesting alternative to BEM is based on so-called 'infinite elements' as described in e.g. [28]. The idea is to model the near-field close to the tyre in terms of conventional finite elements, which are coupled to 'infinite elements' describing wave propagation outside the near-field zone. The infinite elements are easily incorporated into existing Finite Element software and preserve the banded matrix structure of standard FE models. The latter feature is a great advantage over BEM, as it allows the use of efficient numerical algorithms developed for standard FE problems [134]. In addition, the very narrow gap building the horn might create problems in standard BEM packages. Most of the calculations were made for smooth tread patterns. In the case of normal tread patterns, the geometry between tyre and road is given by a complex system of narrow channels and cavities, creating viscous losses and groove resonances. Work related to this subject includes [55], where Graf modelled the transfer function from a point source in a groove to the far field radiation, including the horn effect. Recently Hoever and Kropp [65] published calculated sound pressure spectra for truck tyres rolling on different road surfaces (simulated with the model developed at Chalmers) and found very good agreement with measurements.

Hybrid models In addition to typical statistical models not discussed here, there are so-called hybrid models. These models combine deterministic calculations (e.g. contact forces) based on physics with a statistical approach, e.g. correlation of calculated contact forces with measured sound pressure levels. This approach has been used by Beckenbauer and Kuijpers [19] and was used in the form of the SPERoN Prediction model [49] in the presented studies. This model will be described in detail in Chapter 4.1.1.

2.3 Auralization

The term auralization was introduced by Kleiner et al. in 1991 [74]. In [73] they define auralization as "*the process of rendering audible, by physical or mathematical modelling, the sound field of a source in a space, in such a way as to simulate the binaural listening experience at a given position in the modelled space*". Another, newer definition by Vorländer defines auralization as the technique of creating audible sound files from numeric (simulated, measured, or synthesized) data [135]. In general, one can say that auralization describes different techniques that recreate sounds in such a way that they represent different acoustic situations either from basic recordings, or by the use of information about the sound source and the sound propagation in the surrounding environment.

There have been some efforts to develop models that auralize different traffic situations from a basic data set. Eerden et al. [129] used a monitoring system to investigate the sound level distribution in different urban areas. These measurements were used to create dynamic noise maps of different areas. Another method was used in the Swedish Listen Project [108]), which aimed to develop a demonstrator that simulates and auralizes the sound environment in urban areas [103]. The demonstrator is based on a set of single vehicle passages [90].

A useful step that has not been accomplished yet, is to to combine tyre-noise modelling with the auralization approach. This combination would create a design tool that can play back the noise produced by the designed tyre or road. In this way, the acoustic effects of a newly developed tyre or road can be analysed even prior to the start of production.

2.4 Psychoacoustics

The Encyclopaedia Britannica defines psychoacoustics as “the study of the physical effects of sound on biological systems” [29]. This means that psychoacoustics deals with the physical properties of stimuli like sounds and vibrations and their effect on the human body as well as the subjective experience of those stimuli.

Thus, psychoacoustics studies the relationship between the physical signals and their interaction with the auditory system all the way to the interpretation of and reaction to them (perception). Models are derived for different stages of this relationship. Psychoacoustics is an interdisciplinary field, so different models follow different approaches and concepts.

Some models focus on the relationship between stimuli and percepts, like Zwicker and Fastl [144] or Moore [91]. Zwicker and Fastl focus on the way in which sound is perceived or described. They collected general descriptive parameters for sounds like pitch, loudness, roughness, fluctuation strength, pleasantness/annoyance, subjective duration and rhythm, and tried to develop mathematical models to describe them. For this, they related the perceptions to sound properties and linked them together in models. The descriptive parameters will be discussed in detail in a separate section. Moore uses a slightly different approach, focusing more heavily on the process of sound processing. He analyses how the sound is processed from the ear to the brain and how this leads to the perception of loudness, as well as how frequency selectivity and temporal processing explain pitch, space, and object and speech perception. To do so, he follows the neuronal responses on the auditory path-way and describes them in models. These two approaches have resulted in complex models of auditory signal processing. Examples are Dau's models, which focus on masking effects [38, 39] and detecting changes in amplitude [36, 37]. Another model is PEMO [61, 66], which utilizes a model of auditory signal processing to estimate the perceived similarity between two audio signals.

Another approach is to focus on the emotional reaction to the sound [24]. This approach focuses on the subjective experience of a feeling (emotion) induced by a stimulus like a sound. . This approach requires the definition of common parameters for different emotional reactions. The most common such terms are valence (pleasantness) and activation but, depending on the question, many other parameters can be defined that represent different stages of interaction between the two, such as annoyance, stress, happiness, relaxation and so on [115]. This was the approach Västfjäll et al. used to investigate the affective evaluation of vehicle auditory quality [136].

In product design, the focus is often on sound quality. Here the different approaches become intermingled, depending on the purpose of the product and the type of sound. Quite often, a mix of emotional measures and psychoacoustic measures is used that is suitable for the sound and understandable for the evaluating person. Depending on the sound, specific measures are used, such as comfort, booming and many more. Finding a suitable measure is often a challenge and often approached by free verbalization interviews [12].

2.4.1 Psychoacoustic measures

Psychoacoustic measures describe the perception of a sound's characteristics by a listener. They are properties of an acoustic signal that are rated and recognized in a consistent way by most listeners and can be used to characterize a sound.

The psychoacoustic measures used in a certain situation should reflect properties that are relevant to the problem at hand. One example is dieselness, which describes the specific sound characteristics of a diesel engine and is often related to poor engine quality and thus often undesired. Dieselness can be an important measure for car engines, but it would not be a well-chosen measure for aeroplane cabin noise. Other examples are that booming would not fit to describe birdsong, whereas sharpness would not always fit to describe the noises of huge engines.

The measures used in this study are described in more detail in the next paragraphs.

Loudness is defined in the ANSI [2] standard as "*the intensive attribute of an auditory sensation in terms of which sounds may be ordered on a scale from soft to loud*". A distinction must be made between loudness and loudness level. The unit related to loudness is called *sone* and the unit related to loudness level is called *phon*.

The loudness level of a specific tone is defined as the level in dB SPL of a 1 kHz reference tone that sounds equally loud as the specific tone [3]. Thus a 1 kHz tone at 60 dB SPL has 60 phon. The loudness is defined as the estimate between the strength of a sound compared to a sound with a loudness level of 40 phon [5]. A pure 1 kHz tone at 40 dB SPL has a loudness of 1 sone if presented binaurally from the front in free field. Loudness is additive. A sound of 2 sone is twice as loud as one of 1 sone. Loudness does not depend solely on the magnitude of the signal, but also on its frequency, bandwidth and duration.

The relationship between loudness and loudness level is linear for levels over 20 phon, but non-linear for lower levels. This was first described by Fletcher [45].

Calculation methods for loudness have been developed and are defined in different standards [6, 9]. Both are based on the model introduced by Zwicker and Fastl [144]. Here the sound is split up in bark bands or third octave bands and the specific loudness is first calculated for the individual bands before summing them up. The specific loudness depends on the level in the band and the centre frequency of the band.

Roughness is a perception of sounds that describes fast envelope fluctuations. It occurs for amplitude and/or frequency modulations of signals with modulation frequencies between 15 and 300 Hz. The unit for roughness is *asper*. One asper is defined as a 1 kHz tone with 100% modulation at 70 Hz modulation frequency and at a level of 60 dB.

Roughness depends on the level of the sound and increases with increasing level. There are a few calculation models, e.g. [35, 124]. Most are based on the model by Aures [14]), in which the signal is split into overlapping bands. Modulation frequencies are estimated for each band and used to estimate the partial roughness. Out of these, the total roughness can be estimated.

Fluctuation strength is a perception of slow envelope fluctuations in sounds [125]. It occurs for amplitude and/or frequency modulations of signals with modulation frequencies between 1 and 20 Hz.

The unit of fluctuation strength is *vacil*. It is defined by a 1 kHz tone with 100% modulation at 4 Hz modulation frequency and at a level of 60 dB having a roughness of 1 asper.

Sharpness is defined in the DIN standard [8] as that aspect of timbre that is related to the frequency distribution of the spectral envelope of sounds. It gives a description of the relative amount of high-frequency components in the sound. Sounds with a high sharpness are described as high-pitched, sharp or bright.

Sharpness mainly depends on the spectral content and the centre frequency of the sound (for narrow band sounds). The unit is *acum*, 1 acum being the sharpness of a narrow-band noise with a centre frequency of 1 kHz (920 Hz - 1080 Hz) and a level of 60 dB.

A more objective method of analysing the effects behind sharpness is to perform a frequency analysis. Aures [15] and Bismarck [133] have derived different methods of doing this. Both derive the sharpness from the specific loudness, the loudness and the frequency in Bark.

Pitch is defined in the ANSI standard [2] as "*that attribute of auditory sensation in terms of which sound can be ordered on a scale from low to high*". Pitch relates to the repetition rate of the waveform of a sound. For pure tones (sinusoid) it corresponds to the frequency and for complex tones to the fundamental frequency. The unit of pitch is *mel*. It is defined by a frequency of 1 kHz with 40 dB above the hearing threshold, producing a pitch of 1,000 mel.

2.4.2 Emotional measures

Emotional measures try to capture the subjective experience of a feeling (emotion) and describe them in a commonly understandable way.

The emotional space has two main dimensions: valence (pleasantness) and arousal (activation), as discussed by Barrett [16] or Bradley [26]. These two dimensions are pan-cultural. The concept is that all other emotions can be expressed by combinations of valence and arousal. Russell [116] uses these two parameters to define the core affect as the primitive, universal and ubiquitous base of emotions.

Pleasantness There are multiple theories that use pleasantness. It is used both as a psychoacoustic measure and as an emotional measure. In emotional theory, pleasantness or valence is one of the dimensions for measuring emotions. It ranks from negative valence, or unpleasant, to positive valence, or pleasant.

As a psychoacoustic measure, there are approaches to find other parameters that form the perception of pleasantness. One suggested model for this is *sensory pleasantness*, suggested by Zwicker and Fastl [144]. They describe pleasantness as an emotional response to a sound that is influenced by elementary auditory sensations and the relationship of the listener to the sound. Their model is based on the basic auditory sensations of loudness, roughness, sharpness and tonality.

Activation or arousal is the second basic dimension in measuring emotions [16]. It is ranked from low, or passive, to high, or active.

Activation describes how strong the reaction to a stimulus is, and how much the human body and brain react to it. A high activation is often coupled with a high heart rate and increased blood pressure, whereas a low activation means that the body is relaxed [17]. Activation is the key reaction to attention and alertness.

Stress is a combination of valence and arousal. It is seen as a relationship between a person and the environment that is appraised by the person as exceeding their own resources and thereby endangering their well-being [86]. This means that it is an emotional reaction to a stimulus that indicates health risks caused by this stimulus. It can be measured by self-report or by measuring hormones like cortisol. It has been found that stress is an emotional reaction with generally high activation and low pleasantness, as can be seen in studies by Russell [117]. Stress has been utilized to deepen the focus on negative valence and arousal combinations in measures of emotional responses in the first attached paper.

All the above measures have been utilized in the present studies. For complementation, annoyance will be discussed as well, even though it has not been utilized in the presented studies, due to a strong overlap with pleasantness, stress and activation.

Annoyance is a term often used in sound quality evaluations. It is related to activation (arousal) and inverse to pleasantness.

In the WHO guidelines for community noise [23], annoyance is one of the criteria used, and is defined as a feeling of displeasure associated with any agent or condition.

The annoyance of sounds is strongly affected by their loudness [22], but there are other parameters involved as well. It is closely related to other negative responses such as anger, displeasure, exhaustion, and stress-related symptoms [145]. Öhrström [145] recommends annoyance as one of the parameters to be used to indicate health effects caused by noise.

Fastl and Zwicker developed a mathematical approach to calculating psychoacoustic annoyance in their book [144]. They calculate the psychoacoustic annoyance of a sound out of its loudness, roughness, fluctuation strength and sharpness.

Chapter 3

Methodology

3.1 Psychoacoustic methodology and statistics

This chapter introduces the utilized testing methods and the important statistical methods for this study.

Two testing methods were chosen for this study: *categorial scaling* and *paired comparison*. Both methods are capable of providing answers to the posted questions with a reasonable amount of time and effort. They are described in detail in the following section.

The results of the listening tests, utilizing the two introduced methods, were analysed using different statistical methods. In addition to mean values and standard deviations, the ANOVA (ANalysis Of VAriance) method was chosen to test the statistical reliability of the test results. Different methods of applying ANOVA were utilized in these studies, depending on the question at hand, and are described in more detail below.

3.1.1 Semantic differential

The semantic differential was first introduced by Osgood in 1957 [41]. It was developed as a scaling instrument for the measurement of meaning [100]. Thus, it can be used as a method to investigate the connotative meaning and affective qualities of random objects or words [41].
An object can be evaluated by either metaphoric or emotional relations to the object and not by a rating of the object itself [26].

The semantic differential is defined as a set of several semantic scales [100]. Each scale is defined by a bipolar combination of adjectives. Examples for these polar adjectives could be ‘smooth–rough’ or ‘loud–silent’. The universal semantic differential introduced by Osgood [100] is one available method, but it is common to use semantic differentials that are adapted to suit the question at hand, called context-specific semantic differential [41].

It is most common to use bipolar pairs of adjectives, but there is also the possibility to create an artificial bipolar scale by using one adjective and a scale, for example from ‘agree’ to ‘disagree’ [41]. This method is sometimes called categorical scaling test or polarity profile and is the method used in the presented studies. The method is closely related to the Likert scale, where statements expressing a clear opinion are tested, and the respondent can express agreement or disagreement on a rating scale. For Likert scaling, a set of statements should investigate the same question from different angles.

For the semantic differential, the scale for a single attribute is divided into steps. Different lengths of scale are possible, but the most common has seven steps [100]. This arises from the fact that shorter scales have a resolution that is too poor for most questions, whereas longer scales are harder for the participants to handle and might not be used to the fullest extent. The spaces in between the steps are assumed to be equidistant. The studies presented here use a seven-step scale.

For perceptual acoustics, semantic differentials are often used to characterize sounds. There are different kinds of semantic differentials known for this. For example, in the car industry, semantic differentials are used to analyse the sound and vibration impression of cars in different driving situations.

If more than one object is tested with the same semantic differential, the resulting profiles can be tested for their similarity with the help of correlation analysis [41].

3.1.2 Paired comparison

A paired comparison test can be used to define a ranking order of test objects under a chosen attribute. The objects — for example sounds — are compared in pairs. For each individual comparison, the test subject is asked to choose one of the two objects, e.g. the louder of the two. This method was introduced by Thurstone in 1927 [126].

The results are collected in a matrix of predominance, where each object is compared with all others. For each object, the matrix indicates how often it was preferred over the other objects, which can be used to rank the objects.

The disadvantage of paired comparison tests is that they can be very time-consuming [41]. The number of pairs to test can be calculated using the formula $\binom{n}{2} = \frac{n\cdot(n-1)}{2}$,where n is the number of objects to test. Another problem of the paired comparison is that the information gained primarily describes the order of the tested objects for the tested attribute [41]. The distance between them and the overall validity of the attribute for the tested objects can be investigated by additional methods like multi-dimensional scaling or Thurston scaling, but it is information gained indirectly.

3.1.3 Detection and Reaction Time

With environmental noise, the question is quite often whether or not a specific sound is noticeable. There are different ways of measuring such a situation. One can measure the threshold of hearing in the background to determine the level at which a sound becomes audible, or one can measure the reaction time to the sound in background noise. Reaction time means the time difference between the onset of a given signal or stimulus and the reaction to it [138].

The advantage of using reaction time as a measure for detection is that it can be used for both strong and weak signals and thus link the perception of both [57]. Thus, it not only measures whether a sound is noticeable, but further qualifies the detection. Reaction time has a resolution that is as good as other detection methods, such as the proportion of correct responses. The difference is that the reaction time still changes when the proportion of correct responses is saturated. The reaction time to an acoustic stimulus depends on its level: the louder the sound, the shorter the reaction time [32]. Reaction time can thus be used to derive loudness functions. This has been utilized and validated in several experiments [67], especially with animals, e.g. [104, 105, 53].

Humes [67] found that the reaction time to a single acoustic stimulus varies between 150 ms to 450 ms depending on the signal strength. Several experiments showed that the measured reaction time is independent of whether the instruction focuses on speed or accuracy [52, 69, 43]. The reaction time strongly depends on whether it is a simple reaction time with just one stimulus and reaction or if it is a choice or discrimination reaction time, where several stimuli or responses might be involved [40]. Hick [63] postulated that the time to make a decision depends on the possible choices. The equation he found for the reaction time is $T_R = b * log_2(n+1)$ where b can vary between 0.518 and 0.626 depending on participant and situation. Thus, the more complex the situation and the more decisions are included, the higher the reaction time.

Utilizing reaction time (RT) as a measure for detection is a convenient approach that has been used for investigations on vehicles in traffic not only by Grosse [58] and Altinsoy [11], but also in more general investigations like Kerber's study of the perception of exterior vehicle noise [72]. Kerber used reaction times to investigate the level from which on a car is detectable in background traffic. He states that the reaction time for a stimulus with an increasing level is about 560 ms. For this reason, the study presented in Chapter 6 uses reaction time as a measure to investigate the effect different parameters of background traffic have on the detection of a car passing by close to a pedestrian.

3.1.4 ANOVA

ANOVA was developed by Fisher [44] and is an acronym for analysis of variance. It is a statistical test for differences between or among groups of measurement values.

Like most other statistical theories, ANOVA is based on the idea that the collected data of repeated measurements of any kind are random samples from a normal (Gaussian) distribution. The distribution can therefore be characterized by its mean value and variance. If the sample size is large enough, i.e. if the measurement has been repeated often enough, the mean values and variance of the measurement are equal to the mean value and variance of the underlying distribution. Different statistical tests, like the Student t-test and ANOVA, have been developed to test whether groups of values originate from the same distribution. Whereas the t-test only allows testing for two groups, ANOVA allows for multiple groups by comparing group and overall variances to statistical tables. The test is conducted on the null hypothesis that there is no difference or variance between the tested groups.

Different versions of ANOVA with different underlying statistics have been developed for various relationships between the test groups. One main difference is whether the tested relation lies within the test group (within design) or between different groups (between design). The one-way analysis can only test effects between designs, whereas the repeated measures ANOVA can even test within design effects.

A one-way ANOVA tests the effect of an 'independent variable' on a 'dependent variable' [25]. The dependent variable is normally a measurement result or set of test subject responses. The independent variable is part of the test design, e.g. the sound files that are evaluated. ANOVA can be used to test if the variance in the results can be explained by the change of the independent variable. It is possible to include more than one independent variable in an ANOVA test by extending the statistics of a one-way ANOVA, for example, to a two-way or n-way ANOVA.

The repeated measures ANOVA provides a method for analysing variances in data sets where a limited number of test subjects are tested multiple times. It allows detection of not only influences of independent variables between test subjects, but also between different responses from the same test subjects.

An ANOVA calculates the mean squares for the different degrees of freedom in the data set. The ratio between the mean squares is then evaluated using the F-Test. The variance between the data sets in the variable is compared to the variance within the data sets in the variable. The results of such a test are commonly presented depending on the degrees of freedom, as $F(df, d_{error})$. df gives the degrees of freedom of the independent variable and d_{error} gives the degrees of freedom of the rest of the data set. Error is used in this context to designate unexplained variations in the collected observations.

For easier interpretation, the probability p is calculated as well. It calculates the probability of the null hypothesis being correct. Commonly the null hypothesis is that the data sets belong to the same population. Common levels used to decide whether to accept or dismiss the null hypothesis are significance levels of 1% or 5%. If the p value is below the chosen significance level, the null hypothesis is dismissed, meaning that the data differs on the tested variable.

The probability is only an indicator of an existing difference, not a measure of the size of the difference. If more information is desired about the degree of difference, the effect size can be calculated. There are several different methods to calculate an effect size. Generally, the higher the effect size is, the stronger the tested effect. The method used in this thesis is called eta-square (η^2). It tests the amount of variance that is explained by the underlying model.

3.2 Used road surfaces and tyres

The used road surfaces and tyres in this study are based on documented data from the SPERoN Database. This database contains sets of documented data from measurements on a test track in Germany called Sperenberg [83]. From these measurements, three tyres and three roads were chosen for the validation of the auralization of pass-by signals.

3.2.1 Road surfaces

Figure 3.1 shows the pass-by levels of a set of measurements in the SPERoN Database. The three road surfaces that were chosen for the simulation are marked with ellipses. They cover most of the level differences of the road surfaces.

Road surface A04 is an asphalt concrete surface 0/8 produced according to the regulations defined in ISO 10844 [1], and marked by A in the following studies. This surface is mainly used for testing vehicles in Europe, but not used for public roads. A picture of the surface can be seen in Figure 3.2. The grain size used is up to 8mm.

Road surface A07 is a stone-mastic asphalt 0/8. The main grain size is 5 to 8 mm. It is a common road surface. It was produced and treated in the same way as it would be for public roads. A picture of the surface can be seen in Figure 3.3.

Road surface B09 is a concrete surface covered with synthetic resin and gravel, marked by C in the following studies. The concrete surface was polished before a synthetic resin was applied and gritted with gravel with sizes from 5 mm to 8 mm.

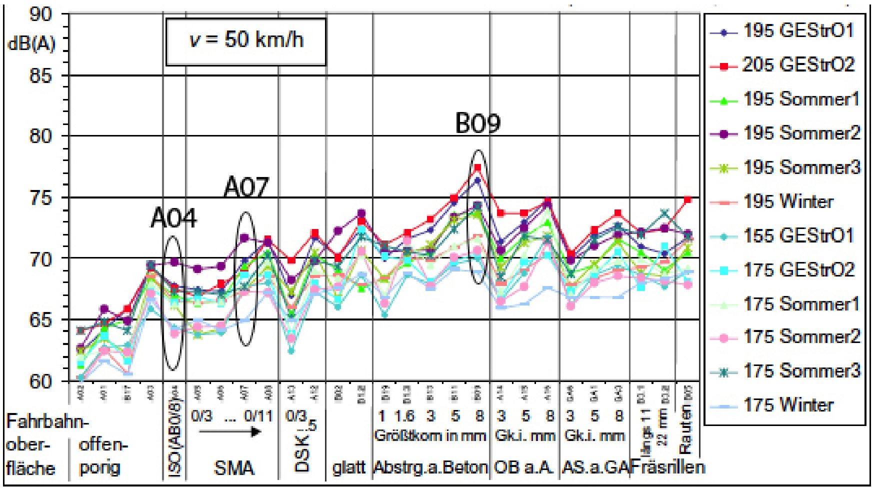


Figure 3.1: The pass-by levels in dB(A) are given for all tyres that were tested in Sperenberg and are displayed over the different road surfaces [18]. Marked are those road surfaces that are utilized in this thesis.

A picture of the surface can be seen in Figure 3.4.

All three surfaces contain mainly gravel of the same size (5 – 8 mm), but differ in the mix of materials and the production methods.

3.2.2 Tyres

Three tyres were chosen from the same data set for the simulations and validations. The selected tyres are marked in the full data set in Figure 3.5. All the used tyres are conventional tyres.

The first and the second used tyre were attached to a Mercedes C280, while the

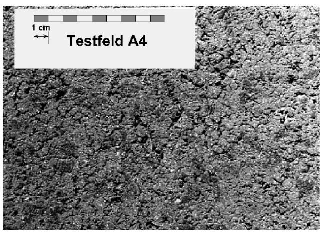


Figure 3.2: Picture of road surface A04 (A) [18]

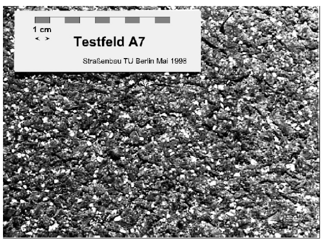


Figure 3.3: Picture of road surface A07 (B) [18]

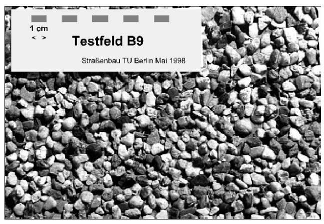


Figure 3.4: Picture of road surface C (B09) [18]

third was attached to a VW Polo. The first tyre (DB3) is one of the most silent tyres from the available set, as can be seen in Figure 3.5, and the third tyre (VW3) is one of the loudest in the data set. The second tyre (DB4) is more in the middle and less varying in sound levels, depending on the road surface.

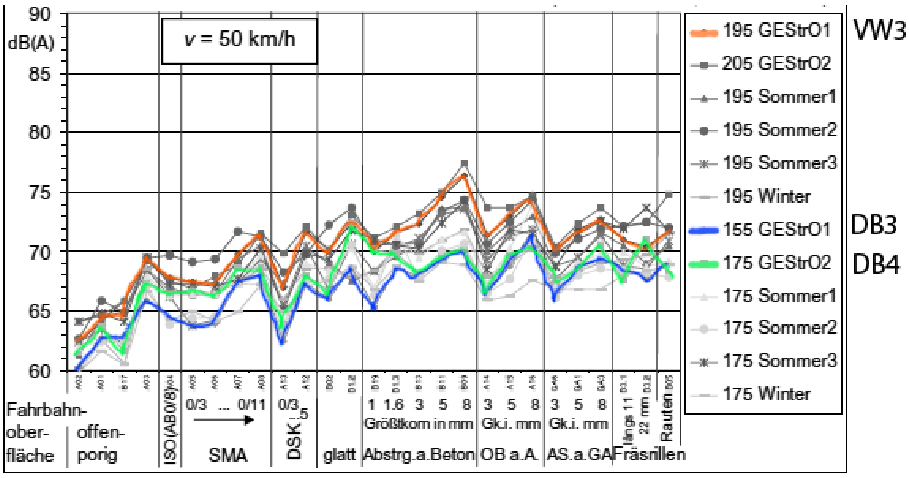


Figure 3.5: The pass-by levels in dB(A) are given for all tyres that were tested in Sperenberg and are displayed over the different road surfaces [18]. Marked are those tyres that are utilized in this thesis.

The first tyre (DB3) was made by Continental: SportContact CH90 (1995/65-R15 90H), the second tyre (DB4) is from Pirelli: P600 (205/60-R15 91V), and the third tyre (VW3) is from Michelin: MXT (155/70-R13 75T). Figures 3.6 to 3.8 show the profiles of the three tyres.

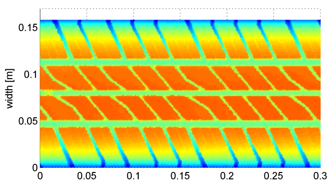


Figure 3.6: Profile of tyre 1 (DB3)

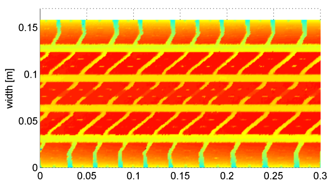


Figure 3.7: Profile of tyre 2 (DB4)

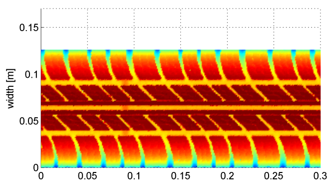


Figure 3.8: Profile of tyre 3 (VW3)

For both cars, tyre pressure and weight were adjusted according to ISO 13325, with some variation in tyre pressure. There were variations between the recommendations by the car manufacturer and the ISO, and the pressures recommended by the manufacturer were chosen over ISO. For the VW Polo, the load estimated by ISO would have led to a total weight at the limit of the total accepted weight for the car.

Thus, the load was reduced to a reasonable amount. The used pressures and loads are given in Table 3.1.

Car	Car 1 (DB3, DB4)	Car 2 (VW3)
total load	1620 kg +80 kg	1060 kg +100 kg
load rear	800 kg + 80kg	400 kg + 100kg
load front	860kg	680kg
pressure rear	2.3 bar	2.1 bar
pressure front	2.1 bar	2.1 bar

Table 3.1: Loads and tyre pressures of the two test cars. The loads are given as the basic weight and driver plus the weight added to achieve a fitting load index

The combination of the three tyres and three roads will lead to a set of nine pass-by signals covering all combinations.

Chapter 4

Auralization of pass-by signals

In this thesis and in the first two papers, an auralization technique is introduced to make simulated tyre–road noise audible, using a new combination of two existing models. Tyre–road noise that was simulated in SPERoN [83] is made audible by combining SPERoN with an auralization software developed for the often called Listen Demonstrator [108]. The models used in this approach are described below.

4.1 Models

4.1.1 SPERoN

SPERoN is an acronym for Statistical Physical Explanation of Rolling Noise. It is a model to estimate the controlled pass-by sound level in one-third octave band spectra emitted by a passenger car for a certain tyre–road combination. The model requires information about the tyre and road. The SPERoN model has been developed and improved in a set of research projects such as 'Silent Road Traffic Noise 1-3' by the German Federal Highway Institute [49].

SPERoN is the combination of a physical model, which calculates he contact forces of the tyre–road contact, with a statistical model, which relates these contact forces to measured pass-by levels. The physical model can further be divided into two parts: a tyre model and a contact model. The interaction of the three sub-models and their needed input and output are depicted in Figure 8.2.

The tyre model was developed by Kropp ([76] and [77]) and simplifies the tyre structure by projecting the tyre on a plane, as shown in Figure 4.2. The resulting plate is characterized by different properties of the tyre: the tyre geometry, the elasticity, the bending stiffnesses for the different directions, the pre-tension due to inflation pressure and loss factors for different movements. The elastic foundation parameters are varied for different modes to compensate for the neglected round shape of the tyre. Validations with measured vibrations on tyres showed good agreement with the tyre model [80], as can be seen on the right side in Figure 4.2. The difference in

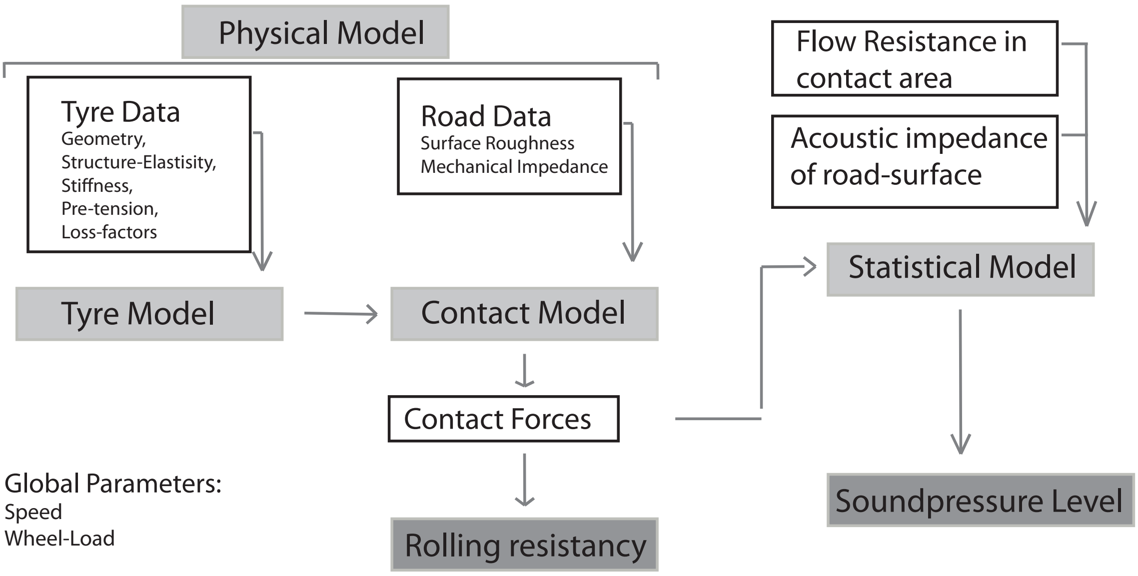


Figure 4.1: Illustration of the SPERoN model with the three sub-models and their needed input and output

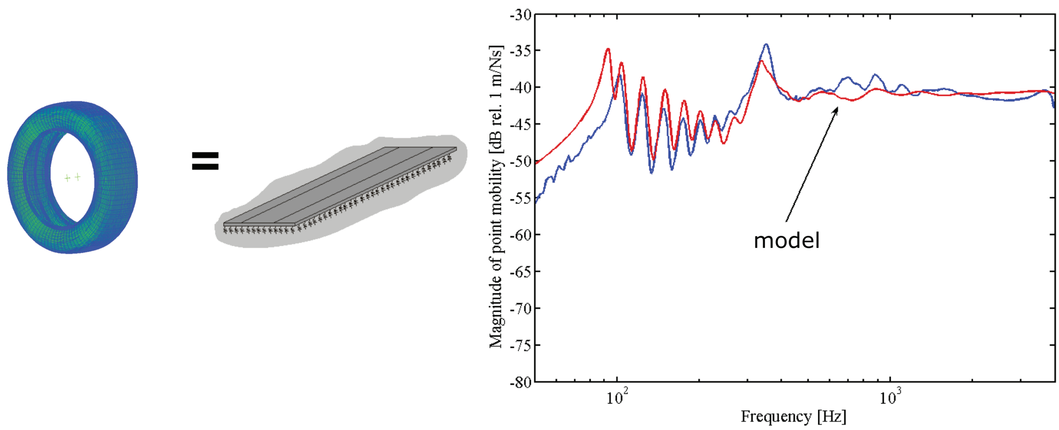


Figure 4.2: Illustration of the tyre modelled as an orthotropic plate, and an example of a resulting vibration response from the modelled tyre and a measurement (the excitement was located radial in the middle). Figure taken from [76]

the figure for the first resonance is related to the fact that the tyre in the measurement was freely suspended whereas the tyre in the model had a fixed rim.

The next phase in the SPERoN model is the contact model. This model needs information from the tyre model, as well as the roughness and flow resistance of the road surface. A version of the 'Chalmers tyre–road interaction' model [82] and [128] is used as a contact model, in which the 3D structures of the tyre and the road are reduced to a two-dimensional problem by translating the roughness of the road surface and the tyre profile into a contact stiffness depending on the intrusion of the road into the tyre surface. For the tyre, this is achieved by dividing the tyre into contact segments. For each of these segments the possible contact is compared to that of a tyre without profile. The relative surface in contact is then used to estimate the stiffness of the segment and the stiffness of a tyre without a profile. To transfer the roughness of the road into stiffness parameters, the contact between tyre and road can be described by a set of springs. The more the tyre presses onto the road, the more springs get in contact and are compressed. This concept is depicted in Figure 4.3. This leads to a non-linear stiffness function. The stiffness parameters for road and tyre structure can then be used to calculate the contact forces for the rolling tyre on the road surface, using a nonlinear algorithm. The resulting total contact force is time varying and is handed over to the statistical model in the form of a one-third octave band spectrum.

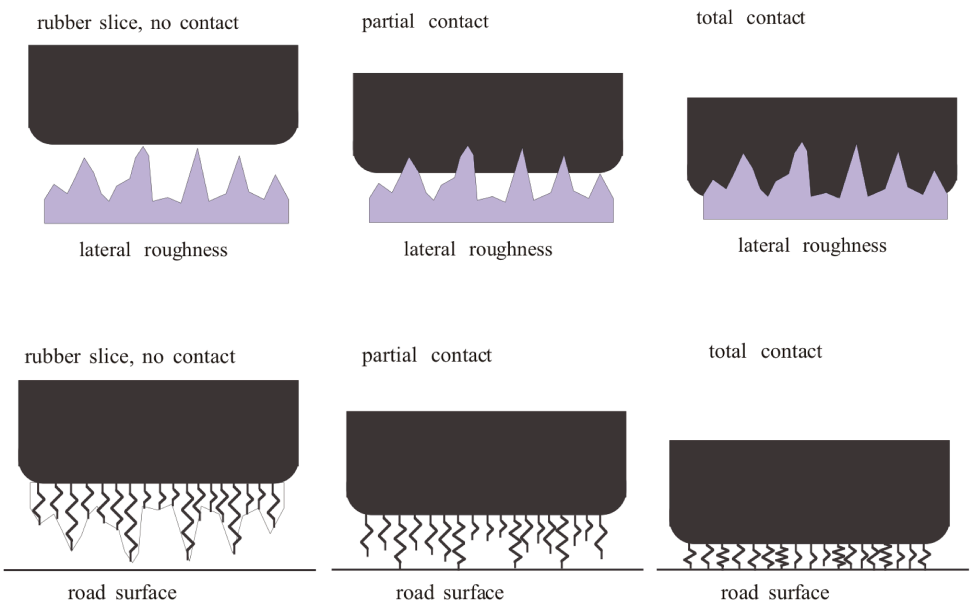


Figure 4.3: Illustration of the contact model; the contact is described in the form of springs. The closer the contact between rubber and road, the more contact with the springs and, therefore, the greater the compression. Figure taken from [140]

The last stage of the SPERoN model is the statistical model, which is based on a documented set of measurements in Sperenberg [83], the SPERoN Database. They consist of controlled pass-by measurements for a variety of tyre–road combinations. The spectra from the measurements are related to the contact forces by a set of physical relations involving four sound-generation mechanisms: (1) sound radiation due to vibrations in the tyre, (2) sound radiation due to airflow-related processes, (3) sound radiation by cavity modes inside the tyre and (4) aerodynamic processes around the tyre and vehicle that contribute to sound radiation. These mechanisms are related to a set of parameters: surface texture, flow resistance in the tyre–road contact area, vibration properties of the tyre, stiffness of the contact patch, tyre profile, size of the tyre, load and rolling speed. [18]. Utilizing these parameters, the four radiation mechanisms can be estimated and summed up to the levels of a pass-by. One example is shown in Figure 4.4, in which estimated sound spectra by SPERoN and measured sound spectra are validated for different cases, demonstrating the high quality of the model [79].

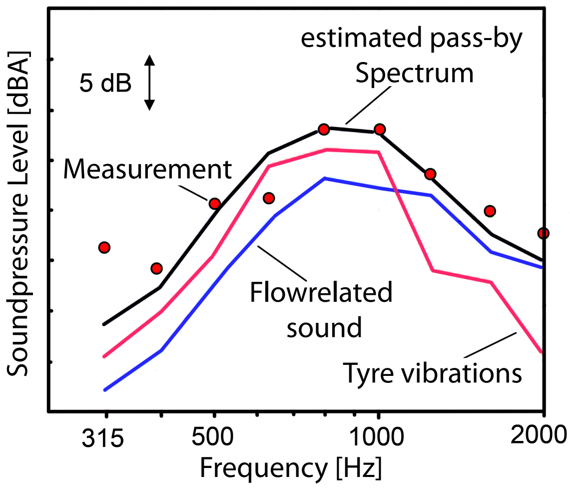


Figure 4.4: Measured one-third octave band spectrum of a controlled pass-by at 50km/h on a rough street in comparison to the calculation by SPERoN [140]

4.1.2 Auralization tool

The auralization approach used in this study is based on the Listen Demonstrator [108]. The main objective of that project was to develop a demonstration software for simulating and auralizing the acoustic environment in urban areas [103]. The tool was intended to enable city planners and stakeholders such as politicians to better understand sound environments in the planning stage [90]. It is based on noise mapping methods described in the Harmonoise methods ([137], [131] and [96]) and the Nord2000 methods ([107] and [70]). The concept of the demonstrator was to separate the source signal and the radiation and propagation effects. To this end, the demonstrator implemented both sound source models and models of sound propagation.

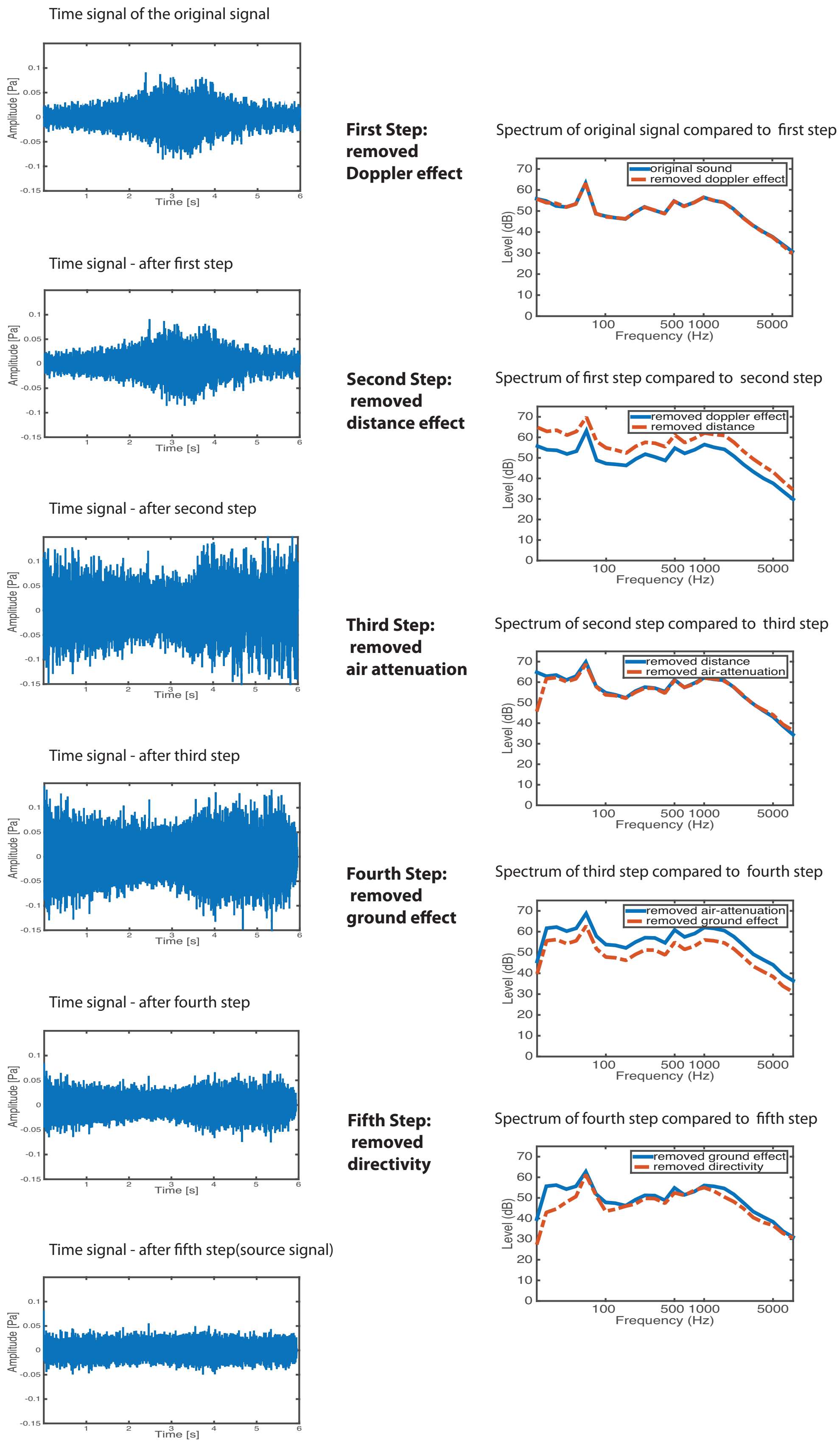


Figure 4.5: The concept of the auralization process from a mono recording to a source signal. Illustration of the changes in the signal in both time and frequency domain

The part of the Listen Demonstrator that was applied in this study is based on an approach by Forssén [46]. The starting point was a recorded monaural pass-by signal of a car with defined parameters such as speed, tyre specifications and road specifications. Applying the inverse propagation effects to that signal creates a stationary signal that can be considered as the source signal. The different stages of this process are illustrated in Figure 4.5. First, frequency shifts due to the Doppler effect are removed (first step in the figure). Then the effect of the varying distance is removed. This can be seen in the second step in the figure as a remarkably constant level increase over frequency, and in the time domain the early and late times are increased in relation to the middle. The third step is to remove the effect of air attenuation. As can be seen in the figure, this mainly affects the low frequencies. The fourth step is to remove the ground reflections, which leads to a decrease in level for all frequencies and mainly the early and late times. As a final step, the influence of the directivity is removed, leaving a very steady time signal as our source signal. This source signal is separated into two terms, by means of engineering models such as Harmonoise [70] and Nord2000 [96]. One source term characterizes the propulsion-related sound sources, such as the engine, air intake, air exhaust etc. (Figure 4.6, dashed line), and the other characterizes the tyre–road noise (Figure 4.6, solid line). Source terms have been generated for 30, 50, 70, 90 110 km/h. To auralize speeds in between those, interpolation methods are applied. Both terms can be modified to create new driving scenarios with differing speeds, road surfaces and tyres. To recreate new pass-by signals, all propagation effects are added back to the source signals. To go back from one-third octave band data to a full spectrum, each band is filled with noise making up the same total level as the corresponding one-third octave band. The information about the shape of the noise in each octave band was stored during the inverse process of the Listening demonstrator.

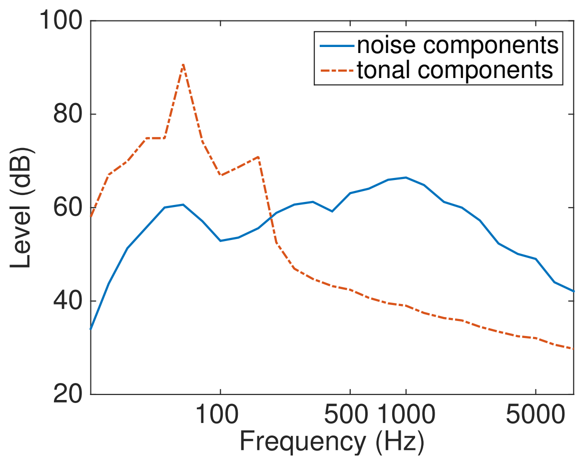


Figure 4.6: The two parts of the source signal: the noisy part is related to tyre–road noise (solid) and the tonal part can be related to the engine etc. (dashed)

4.1.3 Combined model

Listen Demonstrator, the source component in the Listen Demonstrator that is related to tyre–road noise — specified in one-third octave bands — is modified according to the sound pressure levels estimated by SPERoN.

Calculated for the one-third octave bands from 315 Hz to 2000 Hz, the values in SPERoN are now adjusted to match the format in the Listen Demonstrator. This adjusted value is used for all bands where a value from SPERoN (315 Hz to 2000 Hz) exists. For the lower and higher frequency bands, the original values in the Listen Demonstrator are retained. With this approach, the source term in the Listen Demonstrator, which is mainly related to tyre noise, can be shaped by the spectra estimated in SPERoN and synthesized back into a pass-by signal. This adjustment is illustrated in Figure 4.8 for one tyre–road combination (A1).). It shows the one-third octave band values stored in the Listen Demonstrator as source data for the tyre–road noise and the source data after including the values given by SPERoN.

Figure 4.7 illustrated the combined auralization process.

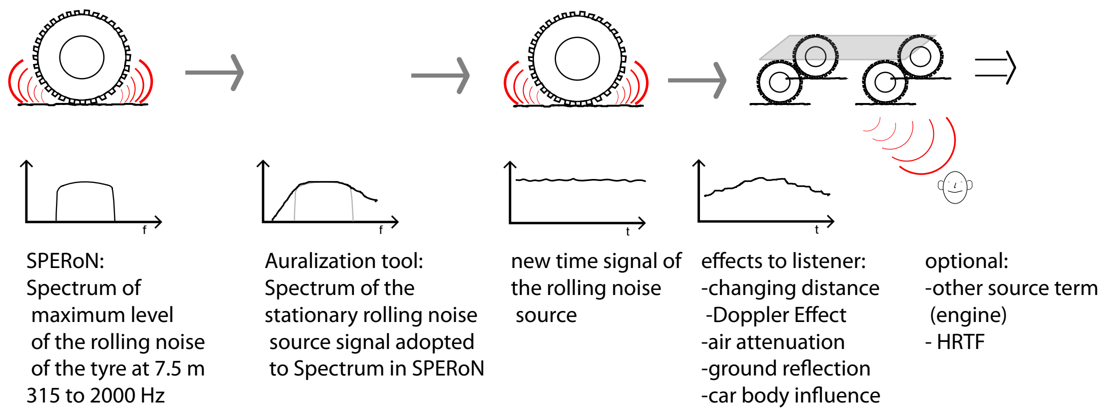


Figure 4.7: Illustration of the Auralization Process: SPERoN estimates the rolling noise spectrum based on basic properties of tyre and road; the source terms in the auralization are compared with the calibrated spectrum and fitted to the new source; propagation effects are added to the source term and a pass-by signal is generated for such desired conditions as Distance, Speed, Surrounding

The SPERoN model was used to calculate the sound pressure levels in the one-third octave bands from the contact between tyre and road for different tyre–road combinations (described in Chapter 3.2). These levels were then transferred to the Listen Demonstrator as described above, and pass-by sounds for the cases were generated. These signals were used in a listening test to validate the combined model.

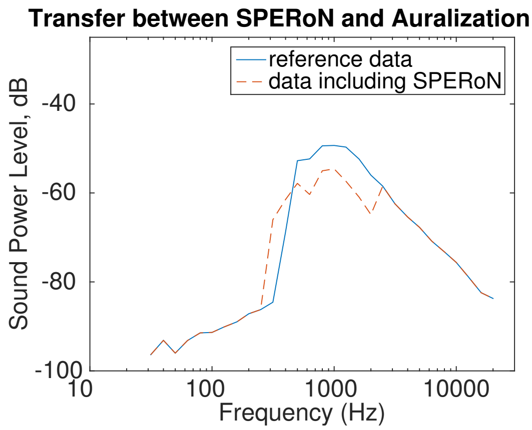


Figure 4.8: One-third octave band values of the source data in the listen demonstrator (reference data) and one-third octave band values after replacing the values from 315 Hz to 2000 Hz by the values given by SPERoN (data including SPERoN) for one tyre–road combination (A1)

4.2 Preliminary study

Two methods were investigated to validate the auralization method: the semantic differential and a paired comparison test. In the paired comparison test, participants were asked to rank the signals in pairs of two for different attributes, to order the signals on an ordinal scale.

Due to time limitations, the test was only done for the attributes 'pleasant', 'loud' and 'rough'. Each participant carried out the experiment twice, and the order of the signals was randomized for each repetition. The question was: "Which signal is more pleasant/loud/rough?", with Signal A and Signal B to choose from.

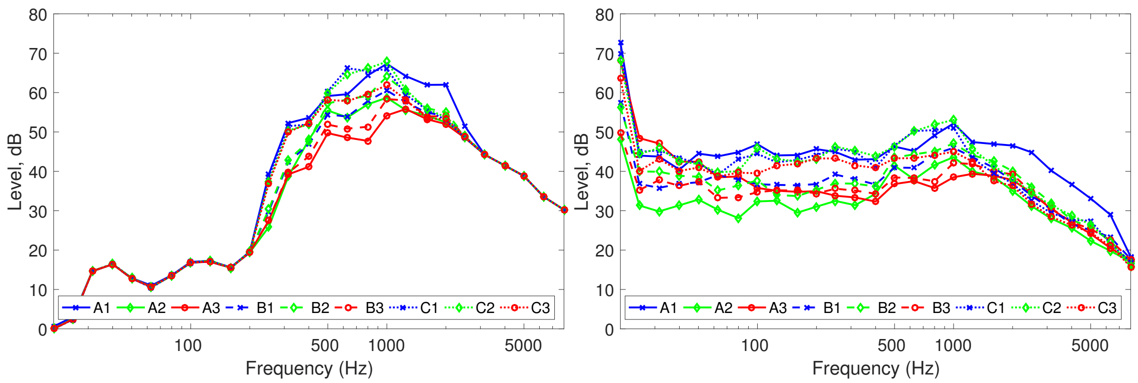
The results correspond well with those of the semantic differential, which were obtained using the same signals and participants. Due to the fact that the results did not show a far better consistency or resolution, but require a longer experiment time, the decision was made to focus on the semantic differential as the method for the performed studies.

To test if the combined auralization method is a valid approach, two listening tests were designed, one with simulated and one with recorded signals. The model was tested for the three different road surfaces and for the three different tyres presented in the previous chapter.

The listening tests were designed as a seven-step semantic differential. The participants were asked to rate the signal according to their impression of pleasantness, sharpness, loudness, and roughness. In the first listening test, they were also asked to rate fluctuation strength. In the second listening test, fluctuation strength was replaced by pitch. For each signal and each attribute, the participants were asked to rate on a scale from 1 to 7 their agreement that the attribute describes the sound.

The first listening test contained only simulated sounds; the second contained only recorded sounds. The spectra of the nine simulated sounds in one-third octave bands are shown in Figure 4.9a and the spectra of the nine recorded sounds are shown in Figure 4.9b. The simulations used only the tyre–road noise, without the engine. The recordings were coast-by measurements. Both measurements and recordings were at a distance of 7.5 m at the moment of pass-by and a speed of 50 km/h. The participants were non-experts in both experiments.

In the first listening test, the signals were played over a loudspeaker in a sound-insulated room furnished as a lecture hall. The participants listened in groups of up to three at a time and received the questions on paper. Only simulated signals (a total of nine) were presented. The signals had a length of 6 seconds, centred round the moment of pass-by. 14 participants (7 male, 7 female) participated in the listening test (age: $mean$ = 28 years, $s.d.$ = 5.1 years). The experiment was



(a) auralized sounds (b) recorded sounds (uncalibrated)

Figure 4.9: Spectrum in third octave bands of the nine auralized and recorded pass-by sounds. Measured are the L_{eq} values for each band. The signals are auralized/recorded at a distance of 7.5 meters to the moment of pass by and a speed of 50 km/h

conducted two times and the order of the signals and questions was randomized.

The second listening test was conducted in a soundproof, neutral room. The test was set up on a computer and the sounds were played via open headphones (Sennheiser HD 650). The signals were adjusted in level according to the peak levels that were measured during the recording process. However, they were not played back at the exact measured levels; rather, the playback level for the listening test was adjusted in such a way that all sounds stayed within a comfortable range for the participants and accounted for the different environment the participants were in, while still sounding realistic. The focus of the experiment was not on the absolute values, but on the relative differences. The signals had a length of 1.5 seconds, centred round the moment of pass-by. The signal length was defined by the shortest test track used for the recordings – 20 m for road surface C (Chapter 3.2). Both signals and questions were presented in randomized orders, with different orders for each participant and each repetition to minimize order and learning effects. Each trial contained one signal and one question. A training exercise on the signals and rating was included in the experiment. The experiment was carried out two times. In total 18 individuals (9 male, 9 female) participated in this second listening test (age: $mean$ = 26 years, $s.d.$ = 3.3 years).

The results of the listening test with the simulated signals and the results of the listening test with the recorded signals will first be investigated separately, and then compared to validate the simulation.

4.2.1 Simulated Signals

In the first experiment, only the simulated signals were presented to the listeners. Figure 4.10 shows the mean value with standard deviation of all participants. The responses are plotted for the different tyre–road combinations. The different markers indicate the different perceptual attributes.

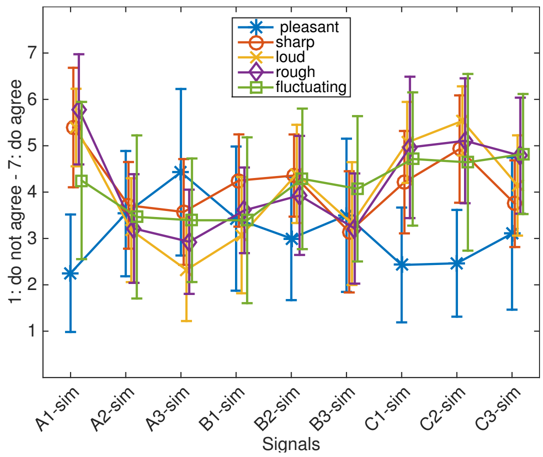


Figure 4.10: Results for simulated signals. Mean values and standard deviations of the responses of all participants are plotted for all tyre–road combinations and for the different percepts (markers)

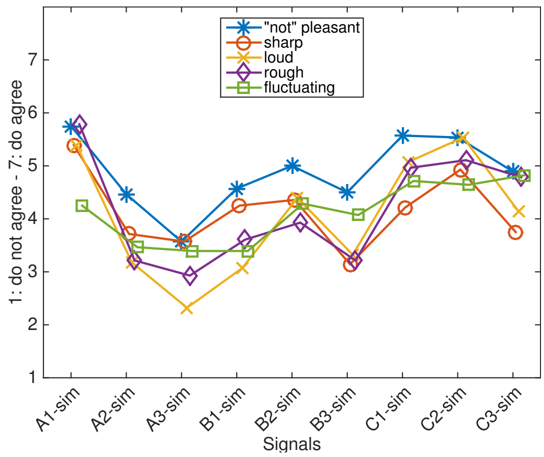


Figure 4.11: Results for simulated signals. Mean values of the responses of all participants are plotted for all tyre–road combinations and for the different percepts (markers). The responses for pleasantness have been inverted for easier comparison

One can see that the pleasantness seems to be approximately inverse to the other percepts. That is why it is plotted inversely in Figure 4.11. The difference in perception between the different tyre–road combinations is small, compared to the standard deviation of the responses.

In addition, the different percepts vary similarly for the different signals. This might be due to the effects of sound generation for rolling noise, which becomes clear in Figure 4.11 where the standard deviations had been removed. One can see that the responses to the different tyres varied strongest on surface A, with A1 showing the highest agreement and A3 showing the lowest agreement in most percepts. This makes sense, considering that A is the ISO Asphalt that is designed for testing tyres and vehicles.

Tests were conducted to see if the signals differed significantly for each percept with an ANOVA, to validate the importance of the high standard deviation. The results for the F-test will be presented together with the degrees of freedom, along with the probability (p value) for accepting the null hypotheses. The meaning of these values is described in Chapter 3.1.4. The results showed significant differences between the signals for all percepts except fluctuation strength (pleasantness: $F(8,117) = 3.13; p < 0.01$, sharpness: $F(8,117) = 5.72; p < 0.01$, loudness: $F(8,117) = 16.23; p < 0.01$, roughness: $F(8,117) = 9.5; p < 0.01$, fluctuation strength: $F(8,117) = 1.81; p = 0.081$). The ANOVA results indicate that the data is interpretable.

Because the perception of fluctuation strength gives no significant difference in variations, and it is only rated around the middle of the scale, it was decided to remove it from the experiment. It seems to have a fairly low significance for the tested signals.

4.2.2 Recorded Signals

In the second experiment, the recorded signals were played for a set of listeners. Figure 4.12 shows the mean value with standard deviation of all participants. The responses are plotted for the different tyre–road combinations. The markers indicate the different perceptual attributes.

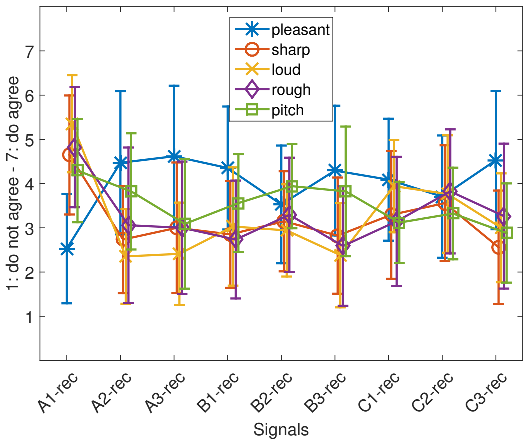


Figure 4.12: Results for recorded signals. Mean values and standard deviations of the responses of all participants are plotted for all tyre–road combinations and for the different percepts (markers)

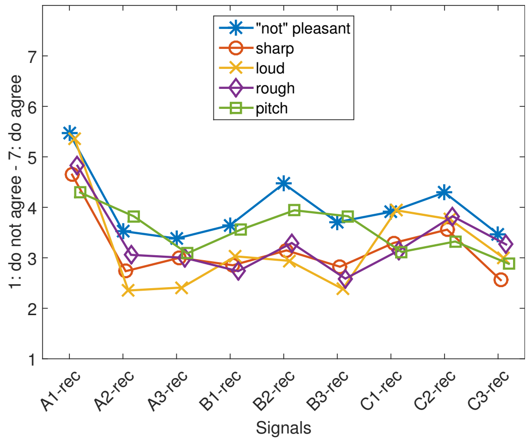


Figure 4.13: Results for recorded signals. Mean values of the responses of all participants are plotted for all tyre–road combinations and for the different percepts (markers). The responses for pleasantness have been inverted for easier comparison.

Due to the high standard deviation tests were conducted to see if the signals differ significantly for each percept with an ANOVA test. The results of the F-test will be presented together with the degrees of freedom. Additionally, the probability (p value) for accepting the null hypotheses will be given. The meanings of these values are described in Chapter 3.1.4. The results indicated significant differences between the signals for all percepts (pleasantness: $F(8, 297) = 7.26; p < 0.01$, sharpness: $F(8, 297) = 7.86; p < 0.01$, loudness: $F(8, 297) = 23.69; p < 0.01$, roughness: $F(8, 297) = 7.15; p < 0.01$, pitch: $F(8, 297) = 5.29; p < 0.01$).

The trends indicated by the results look similar to those of the simulated signals. The results seem somewhat compressed. This might be due to the presence of two simulated signals that were tested additionally during the listening test, but excluded from the evaluation. They were perceived to be stronger than the recorded signals. The comparison between the recorded signals and these simulated signals will not be considered here due to an error in calibration between simulated and recorded signals.

4.2.3 Comparison of the models

In the following the results from the listening tests of recorded and simulated signals are compared in more detail for each percept. For all four cases, the absolute distance between the curves has no meaning, due to different sets of listeners being used and different methods in the experimental set-up.

For inverse *pleasantness*, the mean responses of the simulated and recorded signals are plotted over the different tyre–road combinations in Figure 4.14. A correlation analysis was carried out between the signals, resulting in a correlation coefficient of $R = 0.73$. This leads to a probability of $P = 0.026$ that the null hypothesis of no correlation between the signals is true. Thus, the null hypothesis is dismissed and the responses for the simulated signals correlate with those for the recorded signals at the 5% significance level.

Looking closer at the result, one can see that the highest and lowest responses fall on the same tyre–road combination. However, there are changes in the order of the combinations in between. But those changes only occur between responses that are very close to each other.

For *loudness*, the mean responses of the simulated and recorded signals are plotted over the different tyre–road combinations in Figure 4.14. A correlation analysis was done between the signals, resulting in a correlation coefficient of $R = 0.81$ with a probability of $P = 0.01$. This means that the responses for simulated signals correlate with those for the recorded signals at the 1% significance level. Looking closer at the results, one can see that there are changes in the order of the tyre–road combinations from high to low response. These changes occur between responses that do not differ significantly from each other.

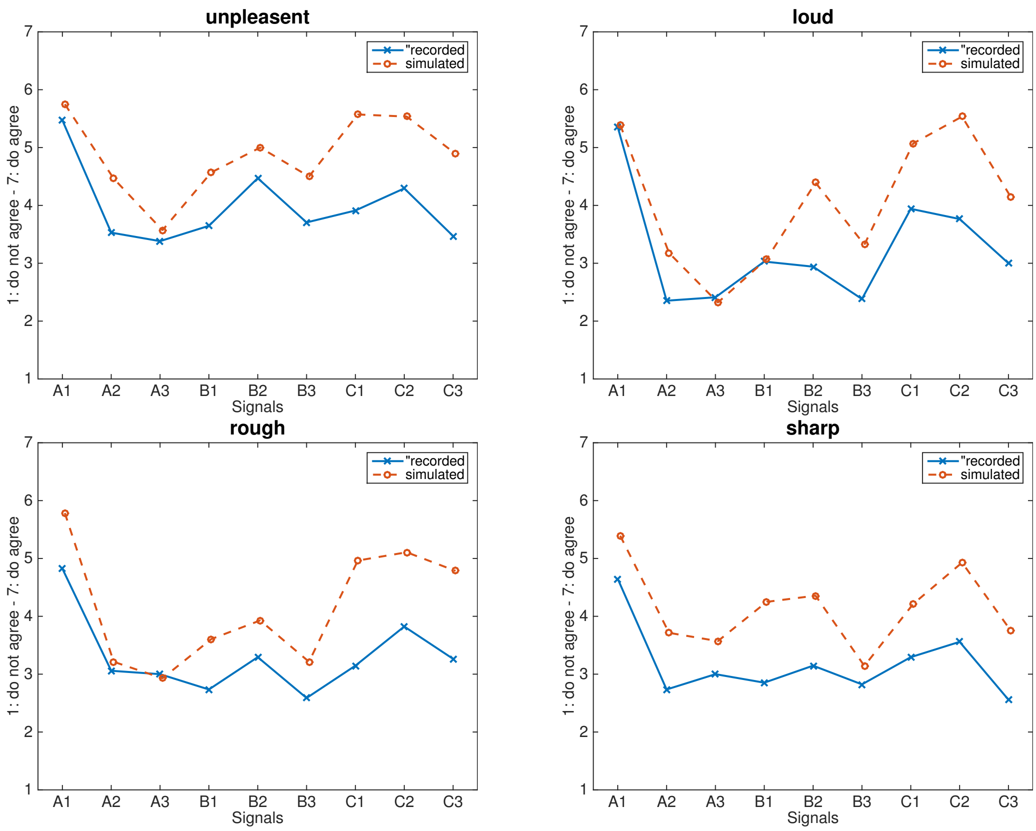


Figure 4.14: Comparison between the responses in the listening tests between recorded and auralized signals for the psychoacoustic variables '(inverse) pleasant', 'rough', 'loud' and 'sharp'.

For *roughness*, the mean responses of the simulated and recorded signals are plotted over the different tyre–road combinations in Figure 4.14. A correlation analysis was done between the signals, resulting in a correlation coefficient of $R = 0.81$ with a probability of $P = 0.01$. This means that the responses for simulated signals correlate with those for the recorded signals at the 1% significance level. Looking closer at the results, one can see that the two highest responses fall on the same tyre–road combinations. However there are changes in the order of the combinations for the lower responses. These changes only occur between responses that do not differ significantly from each other.

For *sharpness*, the mean responses of the simulated and the recorded signals are plotted over the different tyre–road combinations in Figure 4.14. A correlation analysis was done between the signals, resulting in a correlation coefficient of $R = 0.85$ with a probability of $P = 0.01$. This means that the responses for simulated signals correlate with those for the recorded signals at the 1% significance level. Looking closer at the result one can see that the two highest responses fall on the same tyre–road combinations. However there are changes in the order of the combinations for the lower responses. These changes only occur between responses that do not differ significantly from each other.

In general, one can see good correlations between the recorded signals and the signals simulated by the combined auralization tool. However, when looking at the order of the signals from high agreement to the rated percept to low agreement, one notices that the orders differ for all percepts at several positions between the recorded and simulated signals. These changes in order occur mainly between signals that are very close to each other in the responses. Further, one can see that the agreement between recorded and simulated signals is less for surface C than for the other two. Listening to the sounds, one could get the impression that surface C could be more heavily affected by the missing low frequencies.

Although the auralization based on the SPERoN model works fairly well, there is still need for improvement. The main problem with the simulation used is that SPERoN is not analysing frequencies below 315 Hz and those frequencies were assumed to decay linear to the lower frequencies. In the recordings, the very low frequencies are partly influenced by aerodynamic noise around the vehicle. This noise term is not taken into account in the simulations. However, frequencies below 315 can play an important part in the perception of sound and have an influence on different percepts like loudness, roughness and pleasantness. This problem will be discussed further in the next chapter.

Another aspect could be the lack of consideration of tonal components. The spectral information is transferred from SPERoN to the auralization tool in 1/3-octave bands. This leads to a strong suppression of tonal components. If present, tonal components have a strong effect on the perception, as has been shown in various research [123, 20, 7].

Including these effects could significantly improve the simulation and result in an even better estimate of the tyre–road noise. This can then be used to test tyres in different traffic situations.

4.3 Extended Auralization

The preliminary study indicated that SPERoN and the Listen Demonstrator can be combined into an effective auralization tool. A problem with this auralization tool is that the simulations used in SPERoN do not deliver values below 315 Hz. Spectral comparison of the auralized pass-by signal with recordings shows that the signals differ strongly in the low frequency range (Figure 4.17) and that the levels are generally lower. Even if the low frequencies are mainly related to the wind noise, we cannot neglect them, since they are included in the recordings and in every real situation and have an impact on perception.

To better adapt the simulated signals in the auralization process to the recorded signals, the transfer of information between SPERoN and the Listen Demonstrator was changed. Instead of replacing the one-third octave band values from 315 Hz to 2000 Hz with the information from SPERoN and leaving the other octave bands as they were, the data in both models were fitted better with one another. The process of this extended auralization is depicted in Figure 4.15.

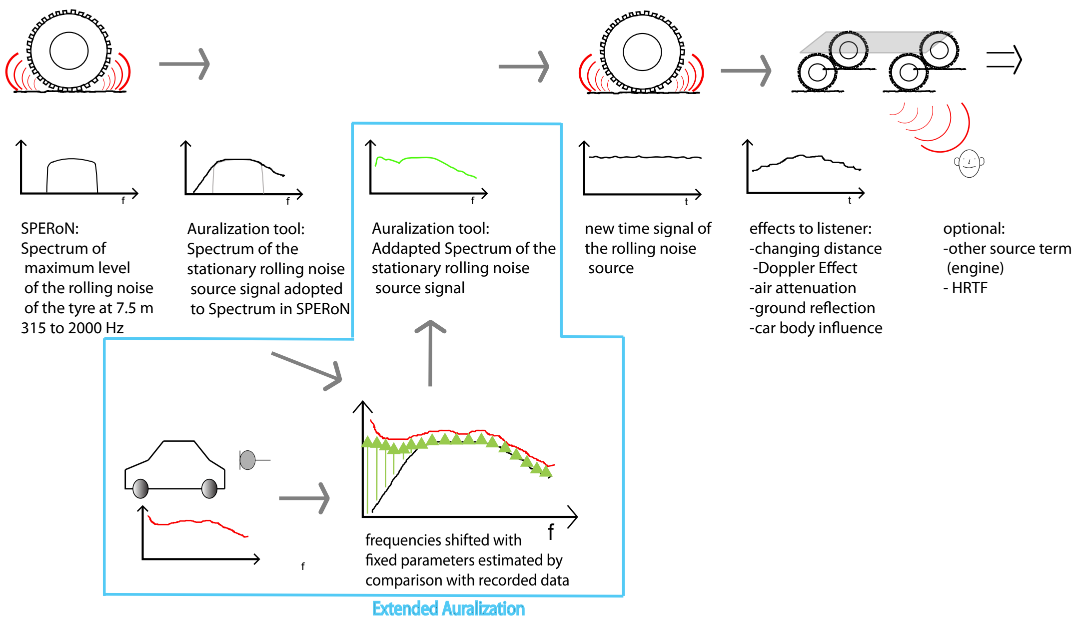


Figure 4.15: Illustration of the extended Auralization Process: SPERoN estimates the rolling noise spectrum from the basic properties of tyre and road;
new: in a comparison with measured pass-by sounds, the rolling noise spectrum estimated by SPERoN is adjusted accordingly; The source term in the auralization is compared with the shifted spectrum and fitted to the new source;
propagation effects are added to the source term and a pass-by signal is generated for desired conditions such as Distance, Speed and Surrounding

The data in the Listen Demonstrator are given in one-third octave bands. Those values are now modified according to the sound pressure levels estimated by SPERoN. To do so, different frequency regions are treated in different ways. The main separation is bands below 315 Hz, bands from 315 Hz to 2000 Hz and values above 2000 Hz.

For the lower frequencies, the preliminary study indicated that the values generated by the Listen Demonstrator, or a linear decay starting downwards from the 315 Hz band value, lead to a hearing impression that lacks in the low frequency range. Thus, starting from the 315 Hz band value calculated by SPERoN, the values for

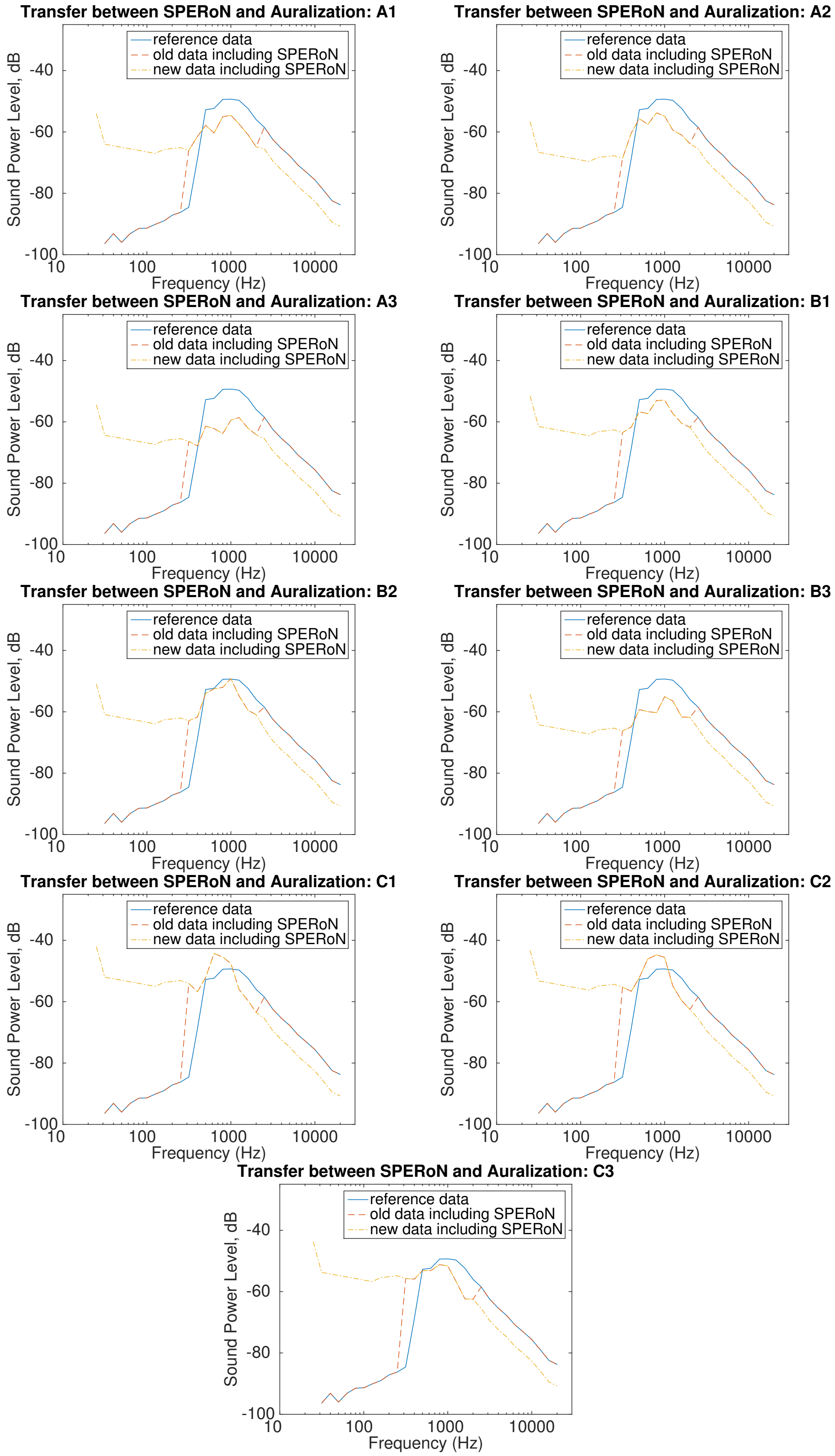


Figure 4.16: Source data in the Listen Demonstrator (reference data),), one-third octave band values after replacing the values from 315 Hz to 2000 Hz with the values given by SPERoN (old data including SPERoN) and), one-third octave band values after the extended auralization method (new data including SPERoN) for nine applied tyre–road combinations. The figure shows the overall sound pressure level.

the one-third octave bands were chosen in such a way that the resulting auralization matches well with corresponding real recordings in the resulting spectrum. This step is needed to consider e.g. wind noise, which will always be present in pass-by recordings. For the frequency bands below 315 Hz, the levels are based on the lowest value from SPERoN. From there, values are decreased until they reach 125 Hz. For the lower frequency bands, the levels are increased again step-by-step. The resulting adjustments are Hz +0.9 dB for 250, +0.6 dB for 200 Hz and +0.3 dB for 160 Hz. From 125 Hz to 31.5 Hz, the values are changed in steps of 0.5 dB from -1 dB to 2 dB. The value for the lowest band (20 Hz) needed an increased value for a stable auralization.

SPERoN delivers values for the one-third octave bands from 315 Hz to 2000 Hz. These values are adjusted to match the format in the Listen Demonstrator. For the bands from 315 Hz to 2000 Hz, where a value from SPERoN exists, these adjusted values are used as input.

For the higher frequency, bands the levels from the source data (Listen Demonstrator) were used, but with a corresponding calibration factor to match the levels in SPERoN. The resulting input data into the auralization tool can be seen in Figure 4.17 for all nine tyre–road combinations that were tested. The Figure shows the one-third octave band values that are stored as source data for each method. The reference data from the Listen Demonstrator is given in the figures for each of the generated signals (reference data) as well as the option of ‘just’ replacing the values given by SPERoN as was done in the preliminary study (old data including SPERoN) and the presented method with adjusted values (new data including SPERoN). The adjustment was made for one tyre–road combination at first. Comparisons between spectra for simulated and recorded sounds of other combinations suggest that it fits quite generally, and so the same adjustment set-up was used for all tested tyre–road combinations.

The changes in the transfer of the sound-pressure level values from SPERoN to the Listen Demonstrator led to a higher agreement of the spectra of the auralized and recorded signals. Figure 4.17 uses one pass-by situation to exemplify the differences in the spectrum of the generated signal to the recording. The spectrum is calculated as the overall sound-pressure level in one-third octave bands. The spectrum of the first auralization is shown as a dashed line, the spectrum of the recording is the solid line and the new auralization is dotted. It is clear that the auralization with the new calibration factors is very similar to the recording, whereas the earlier one differs, especially for the low frequencies.

The extended auralization method was then used to create a new set of pass-by sounds, that were evaluated in a listening test in comparison with the recorded pass-by sounds.

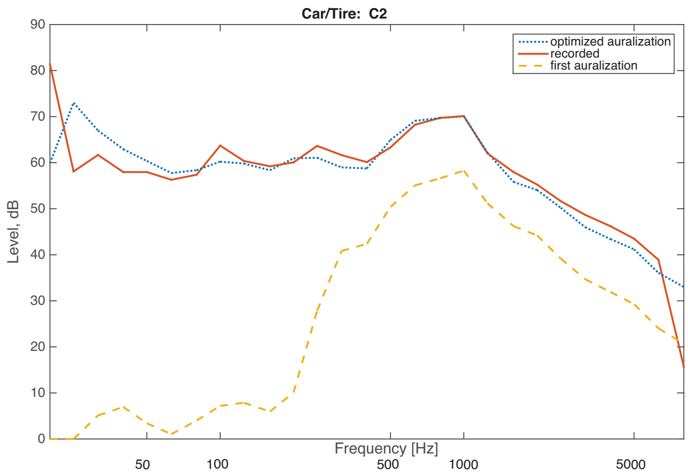


Figure 4.17: Comparison for the smoothed spectra of the recorded signal, the basic auralized signal (without level calibration) and the signal by the extended auralization including a level calibration

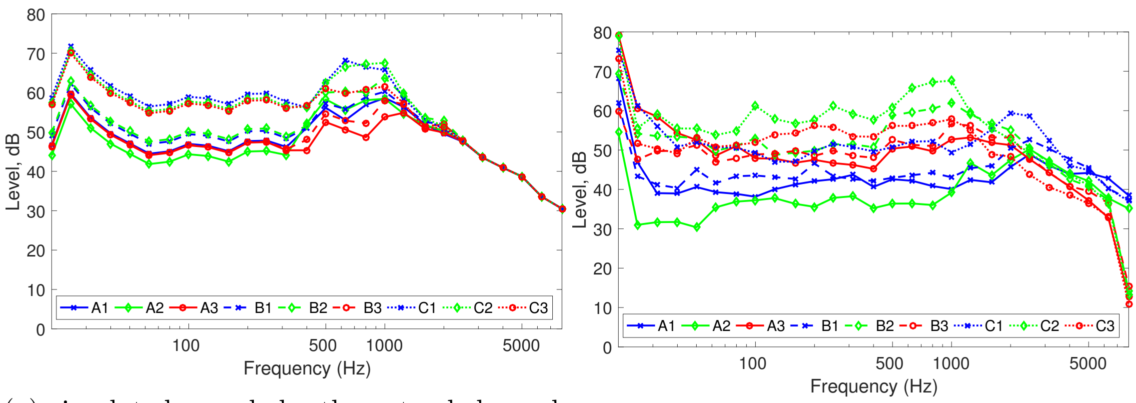
4.4 Validation of the auralization method

4.4.1 Listening Test

A listening test was performed to validate the extended auralization. Both recorded and auralized signals were used in the test, at a speed of 50 km/h. For both cases, nine tyre–road combinations were used. These were the same combinations as in the pre study (three roads: A–C and three tyres: 1–3 as described in Chapter 3.2).

The recorded signals were calibrated at their maximum sound-pressure level (L_{AFmax}). The simulated signals were generated as described above. The spectra of the resulting nine simulated sounds in in one-third octave bands can be seen in Figure 4.18a, and the spectra of the resulting nine recorded sounds can be seen in Figure 4.18b. All signals had a length of 1.5 seconds, centred round the moment of pass-by. The signal length was defined by the shortest test track used for the recordings. This was 20 m for road surface C [18].

The listening test was performed as a semantic differential (Chapter 3.1.1), and the language was Swedish. For each signal and each attribute, the participants were asked to rate on a scale from 1 to 7 how much they agreed that the attribute described the sound. The statements were: The sound is pleasant/sharp/loud/rough. The exact statements and instructions used in Swedish are attached in Appendix A. The response scale was in seven steps and the limits were marked as “do not agree” / “agree”. As common in most scaling design methods (e.g. Likert and Thurston scaling), statements were used instead of questions. The idea is that a statement can be perceived as more clear and more neutral than a question. Furthermore, the



(a) simulated sounds by the extended auralization

(b) recorded and calibrated sounds

Figure 4.18: Spectrum in in one-third octave bands of the nine recorded and simulated (by the extended Auralization model) pass-by sounds. Measured are the L_{eq} values for each band. The signals are auralized/recorded at a distance of 7.5 metres to the moment of pass-by and a speed of 50 km/h

middle of the scale is more neutral when using statements.
The listening test was performed in a soundproof, neutral room. The test was set up on a computer base and the sounds were presented via Sennheiser HD 650 headphones (calibrated via a HEAD Acoustics dummy head). Both signals (recorded and simulated) and questions were presented randomized with different orders for each participant and each repetition to minimize order and learning effects. Appendix A includes a screenshot of the experiment.

Each trial contained one signal and one question. Before the experiment, the participants conducted a practice session to familiarize themselves with the sounds and the attributes. The main session was carried out twice. In total 18 individuals (9 male, 9 female) participated in the listening test (mean age = 26 years, s.d. = 3.3 years). The participants were all non-experts, but most of them had already participated in a number of listening tests. The participants were recruited from students in the ‘Sound and Vibration’ master programme and from a list of potential participants collected over the years, with people of different backgrounds. The participants were asked about their hearing ability. All participants rated their hearing as average or better.

In the beginning of each listening test, the participants were informed about the set-up of the experiment. The statements were introduced to each participant. Furthermore, the participants were informed that they could pause the experiment at any time if questions arose. The first session was a trial run, to get the participants accustomed to both statements and signals. The participants were instructed to go with their first feeling and not reflect too long. However, the experiment was designed in such a way that the participants could listen to each sound as many

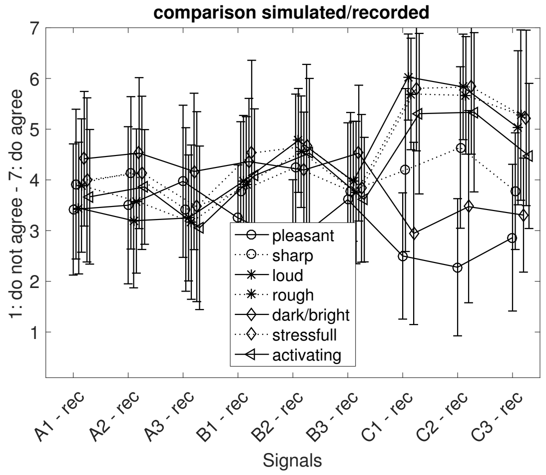


Figure 4.19: Results for simulated signals. Mean values and standard deviations of the responses of all participants are plotted for all tyre–road combinations and for the different percepts (markers)

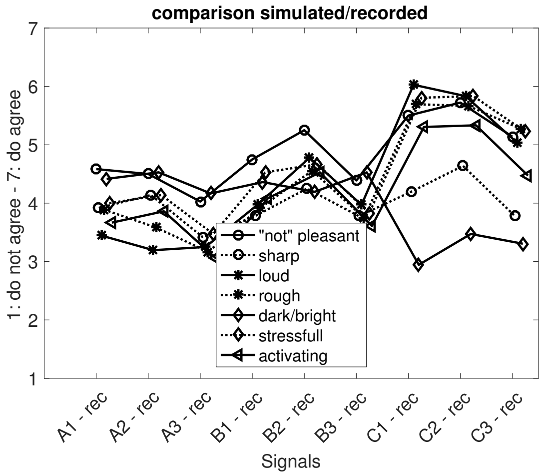


Figure 4.20: Results for simulated signals. Mean values of the responses of all participants are plotted for all tyre–road combinations and for the different percepts (markers). The responses for pleasantness have been inverted for easier comparison

times as they liked and decide for themselves when to move on to the next question. The participants each received a cinema ticket as compensation for their time.

4.4.2 Results and comparison

The results of the listening test were investigated first for simulated and recorded signals individually. Then both were compared to validate the simulation.

4.4.3 Simulated Signals

Figure 4.19 shows the mean value with standard deviation of all participants for the simulated signals. The responses are plotted for the different tyre–road combinations. The different indicate the different perceptual attributes.

As expected, pleasantness seems to be approximately inverse to the other percepts. That is why it is plotted inversely in Figure 4.20. The difference in perception between the different tyre–road combinations is small, compared to the standard deviation of the responses for all percepts. In addition, the different percepts vary similarly for the different signals. That might be due to the effects of sound generation for rolling noise. This becomes clear in Figure 4.20, where the standard deviations are removed. The only percepts that differ clearly from the others are pitch (dark/bright) and sharpness.

Tests were conducted to see if the signals differ significantly for each percept with an ANOVA, to validate the importance of the high standard deviation. The results of the F-test will be presented together with the degrees of freedom. Additionally, the probability (p value) for accepting the null hypotheses is given, and the effect size. The meaning of these values is described in Chapter 3.1.4.

The results of the one-factor analysis revealed that only sharpness could not be differentiated between the different signals. Pleasantness is significant with $F(8,153) = 3,61; p < 0.001$ and an effect size of $\eta_p^2 = 0.16$; stress is significant with $F(8,153) = 7,91; p < 0.001$ and an effect size of $\eta_p^2 = 0.29$; activation is significant with $F(8,153) = 7,30; p < 0.001$ and an effect size of $\eta_p^2 = 0.28$; loudness is significant with $F(8,153) = 21,89; p < 0.001$ and an effect size of $\eta_p^2 = 0.53$; roughness is significant with $F(8,153) = 10,31; p < 0.001$ and an effect size of $\eta_p^2 = 0.35$; pitch is significant with $F(8,153) = 3,86; p < 0.001$ and an effect size of $\eta_p^2 = 0.17$; and sharpness is not significant with $F(8,153) = 1,45; p = 0.18$ and an effect size of $\eta_p^2 = 0.07$.

All emotional responses show significant differences in the perception of the nine signals and all show medium effects. For the psychoacoustic parameters, loudness, roughness and pitch show significant differences in the perception of the nine signals, but pitch only has a minor effect. Loudness and roughness show a large effect. Sharpness has no significant difference in the perception of the nine signals and no effect either.

4.4.4 Recorded Signals

Figure 4.21 shows the mean value with standard deviation of all participants. The responses are plotted for the different tyre–road combinations. The markers indicate the different perceptual attributes.

As for the simulated signals, ANOVA tests were done to see if the signals differ significantly for each percept, to validate the importance of the high standard deviation. The results of the F-test will be presented together with the degrees of freedom. Additionally, the probability (p value) of accepting the null hypotheses is given, and the effect size. The meaning of these values is described in Chapter 3.1.4.

Pleasantness is significant with $F(8,153) = 4,55; p < 0.001$ and an effect size of $\eta_p^2 = 0.19$; stress is significant with $F(8,153) = 5,4; p < 0.001$ and an effect size of $\eta_p^2 = 0.22$; activation is significant with $F(8,153) = 7,95; p < 0.001$ and an effect size of $\eta_p^2 = 0.28$; loudness is significant with $F(8,153) = 21,38; p < 0.001$ and an effect size of $\eta_p^2 = 0.53$; roughness is significant with $F(8,153) = 7,16; p < 0.001$

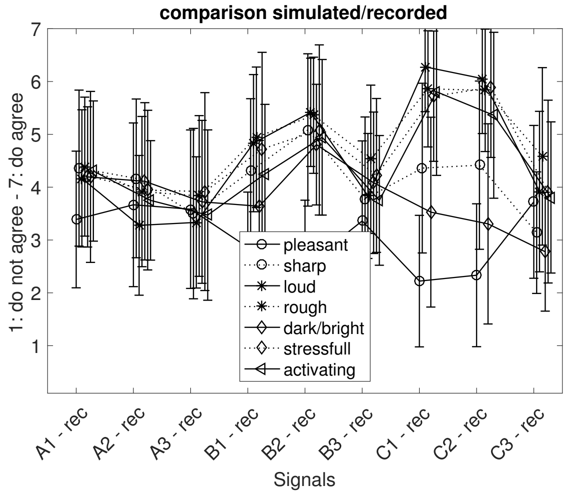


Figure 4.21: Results for recorded signals. Mean values and standard deviations of the responses of all participants are plotted for all tyre–road combinations and for the different percepts (markers)

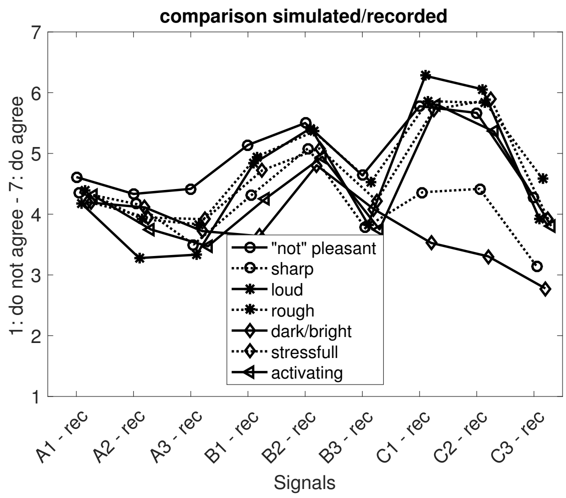


Figure 4.22: Results for recorded signals. Mean values of the responses of all participants are plotted for all tyre–road combinations and for the different percepts (markers). The responses for pleasantness hve been inverted for easier comparison.

and an effect size of $\eta_p^2 = 0.27$; pitch is significant with $F(8, 153) = 3,82; p < 0.001$ and an effect size of $\eta_p^2 = 0.17$; and sharpness is not significant with $F(8, 153) = 3,65; p < 0.001$ and an effect size of $\eta_p^2 = 0.16$.

All responses show significant differences in the perception of the nine signals. Loudness shows a large effect. Pleasantness, stress, activation and roughness show a medium effect and pitch and sharpness only have a minor effect.

4.4.5 Comparison

To be able to compare how similarly recorded and auralized signals are perceived, Figure 4.23 shows the combined results for each psychoacoustic variable. It is clear that the ratings became both closer and more similar in variations between the signals than in the preliminary study (Figure 4.14). This might be explained to a small degree by the calibration of the recorded signals (Chapter 4.3.2).

To investigate only the effects of the extended auralization, a correlation analysis was done. Table 4.1 shows the corresponding correlations, which indicate a correlation between the judgements of inverse pleasantness, stress and pitch at the 5% limit, and for loudness, roughness and activation at the 1% limit. So for those, both auralized and recorded signals are perceived the same. For sharpness, however, the ratings of the auralized signal and the recorded signal do not correlate at the 5% limit.

	pleasant	sharp	loud	rough	pitch	activating	stress
R_{ext}	0.769	0.658	0.86	0.843	0.672	0.8766	0.7824
P_{ext}	0.015	0.054	0.003	0.004	0.048	0.002	0.013
R_{pre}	0.729	0.852	0.813	0.811			
P_{pre}	0.026	0.004	0.008	0.008			

Table 4.1: Table of the correlation coefficients between simulated and recorded signals for the preliminary study and the extended auralization

Table 4.1 also includes the correlation values for the earlier auralization. With the new auralization method, the correlations increased for pleasantness, loudness and roughness, but decreased for sharpness. It must be considered, though, that sharpness was rated as very neutral and in neither of the two auralization methods did it correlate between recordings and simulation. Increasing the low frequencies might have increased the suppression of this percept, since it depends on the relative amount of high frequencies.

Investigations in the form of listening tests showed that a good agreement can be reached for the simulated signals compared to recordings under the same conditions. The agreement is best with an extra inclusion of the low frequencies. The analysis of the listening tests showed that the auralization is improved by the extension. Recordings and simulations are perceived as being very similar. Thus the simulation method can be used instead of recordings and it can be expected that the simulated acoustic situation will represent the real situation well.

Figure 4.14shows the results of the preliminary study. There is a noticeably greater difference between the recorded and simulated signals for road surface C. This difference disappeared with the use of the extended auralization. Thus, the lower frequencies in the auralization had an extra strong effect on these signals. Listening to the signals, one could also get the impression that all tyres on road surface 3 contained a higher amount of low frequencies than on the other surfaces. This is even confirmed by the question about pitch, where the signals on surface C are rated lowest/darkest.

4.5 Conclusions about the auralization method

The aim of this part of the thesis was to create a tool that makes simulated tyre–road noise audible. A combination of the SPERoN prediction model [79] and the Auralization tool developed by Forssén [46] were used to achieve this.

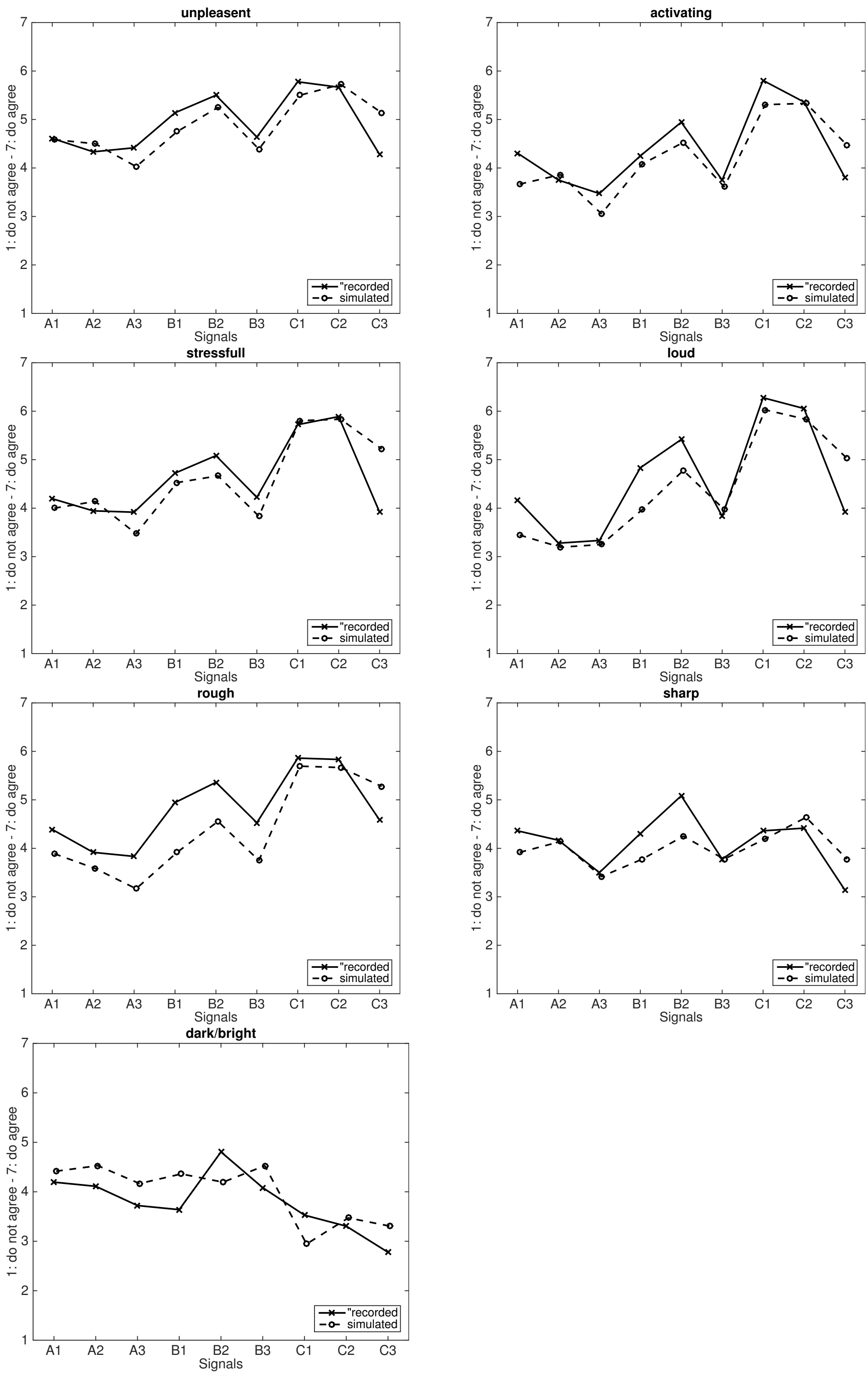


Figure 4.23: Comparison between the responses in the listening tests between recorded and auralized signals (adopted for $f < 315$ Hz) for the psychoacoustic variables 'pleasant', 'activating', 'stress', 'rough', 'loud', 'sharp' and 'pitch (dark-/bright)'

The auralization tool has been validated with the help of listening tests, which showed that the simulated signals are perceived as being very similar to recordings under the same conditions. One limitation of the auralization was caused by the lower frequency limit in SPERoN. It is not possible to simulate for frequencies below 315 Hz. Noise in these lower frequency bands is mainly related to wind. It cannot be excluded in recordings and thus might have an effect in comparisons between recorded and simulated sounds. A method was included in the auralization tool to consider those frequencies and to improve the data transfer between the two tools. Comparisons with recordings showed that the agreement between simulations and recordings improved with this extra consideration of the low frequencies. Good results could be gained with the auralization for the perception of pleasantness, activation, stress, loudness and roughness, for all of which the simulated and recorded signals were rated very similar. For sharpness and pitch, the results were less good. However, the tested signals had a small range in variation for those particular parameters, both for the simulated and the recorded signals. In general, the listening test shows that the pass-by signals generated by the tested auralization method are perceived very similarly to the recorded equivalents.

Chapter 5

Perception of rolling noise

The aim of the third paper and this chapter is to evaluate how changes to the tyre or roads could affect perceptual responses to rolling noise. Is it possible to measure differences in the perception of rolling noise? And is it possible to distinguish between the influence of the sound of tyre and road on the perception? To investigate these questions, a parametric study was designed using the three tyres and three roads described in Chapter 3.2 to simulate the rolling noise of a driving car. The approach is, to use a set of psychoacoustic parameters to characterize and assess the perception. This study was part of the same listening test that was also used to validate the extended auralization tool in Chapter 4.4.1. The listening test performed to validate the first version of the auralization tool (Chapter 4) was also conducted as a preliminary study, investigating the nine signals by psychoacoustic means to find a set of fitting psychoacoustic parameters.

5.1 Preliminary study

To determine which psychoacoustic parameters might be useful to characterize tyre–road noise, a preliminary study using only a small set of participants was conducted. The participants were asked to rate the 9 pass-by signals based on the tyres and roads described in Chapter 3.2 and at a speed of 50 km/h by a set of psychoacoustic parameters. The signals were generated with a previously developed auralization method 4. These parameters were chosen based on the psychoacoustic annoyance defined by Zwicker and Fastl in [144]. The chosen parameters were pleasantness, loudness, roughness, fluctuation strength and sharpness. The results of the preliminary study showed that fluctuation strength was not adequate to describe the car pass-by sounds that were used. The most prominent fluctuation in level is due to the car's passing by, and that might be interpreted more as information about the sound source than as fluctuation strength. It is also the same for all cars, since all sounds were simulated at the same speed (50 km/h). For all other psychoacoustic parameters, the participants appeared to be able to distinguish between the rolling noises, so all those were included in the main experiment. The statistical analysis of the preliminary study further indicated that it is possible to differentiate between the influence of the street and of the tyre with regards to the perception of the tyre–road noise. The analyses also indicated an interaction between the influence of the tyre and the influence of the road on the perception.

Interviews with the participants indicated that the sound pitch is not perceived as being the same for the different pass-by signals. This led to the decision to include pitch as another psychoacoustic parameter in the main experiment. Due to the promising results in the preliminary study, the decision was made to extend the evaluation of the emotional response to the sounds with activation as well as pleasantness. The reason for this was to be able to locate the sounds in the two main emotional dimensions of valence and activation [117]. Additional stress was added as a perceptual attribute, since it is related to high activation and negative valence as well as to negative health effects [68]. The emotional response was extended to allow investigation of whether there are any particular physical aspects that are connected to higher activation and negative valence, and thus potential negative health effects. For the main experiment, a total of seven attributes were tested in the listening test: pleasantness, loudness, roughness, sharpness, pitch, stress and activation. The utilized psychoacoustic parameters were chosen to characterize the emotional response to the signals and to determine psychoacoustic parameters that describe the sound characteristics of rolling noise. Calculation models exist for some of the utilized psychoacoustic parameters (loudness, roughness, sharpness).

Based on the preliminary study, the first hypothesis was that physical differences between tyres and between roads affect the perceptual responses to rolling noise for both psychoacoustic and emotional parameters. The second hypothesis is that the influence by the street and the road on the perception can be separated and that tyre and road interact in their influence on the perception.

5.2 Method

To investigate the stated hypothesis, a listening test was designed using nine different car pass-by sounds. The sounds were all synthetic monaural car-pass-by signals, generated by the method described in Chapter 4. Table 5.1 shows the sound-pressure levels estimated by the SPERoN prediction tool for the nine sounds.

Signal	A1	A2	A3	B1	B2	B3	C1	C2	C3
L_{tot}, dB	67.43	67.24	65.94	68.45	68.68	66.71	68.90	69.29	67.24
$L_{tot,A}, dBA$	67.04	66.91	65.59	67.95	68.19	66.31	68.22	68.60	66.59

Table 5.1: Levels of the nine signals used, in dB and dB(A), estimated by the SPERoN prediction model

Signal	A1	A2	A3	B1	B2	B3	C1	C2	C3
loudness [sone]	34.7	34.9	31.1	37.5	41.4	33.8	52.5	52.2	44.5
roughness [asper]	3.36	3.36	3.09	3.49	3.75	3.32	4.13	4.16	3.75
sharpness [acum]	2.61	2.62	2.69	2.57	2.55	2.65	2.38	2.41	2.43
f. strength [vacil]	0.29	0.29	0.29	0.30	0.31	0.28	0.38	0.36	0.39

Table 5.2: Calculated psychoacoustic parameters for the nine signals, using Artemis from Head acoustics

To characterize the nine signals utilized in this study, they were analysed with Artemis from HEAD Acoustics [10] and the maximum values for the psychoacoustic parameters loudness, roughness, sharpness and fluctuation strength were calculated (Table 5.2). In a study on just-noticeable differences (JND) for different psychoacoustic parameters using refrigerator noise [143], the reported JND were 0.5 sone for loudness, 0.08 acum for sharpness, 0.04 asper for roughness and 0.012 vacil for fluctuation strength. Using these as a reference to evaluate the calculated psychoacoustic parameters for the nine sounds in this study, it can be expected that the signals will be perceived differently in loudness and roughness for both tyre and road variations in an experiment. For sharpness it should be possible to differentiate at least between some signals. The variations between roads seem to have more effect than the variations between tyres. Fluctuation strength only has a few variations stronger than JND.

Apart from one signal (C3), the calculated values for loudness follow the order from more silent to louder as expected from the levels of the signals (Table 5.1).

5.2.1 Listening Test Design

A seven-step semantic differential was utilized to determine whether the participants could differentiate between the different tyres and roads on their acoustic and emotional parameters (the same as used in the study described in Chapter 4.4.1). For each signal and each attribute, the participants were asked to rate on a scale from 1 to 7 how much they agreed that the attribute described the sound. The emotional responses were tested with the sentences: The sound is 'pleasant'/'stressful'/'activating'. For the psychoacoustic response the test sentences were: The sound is 'sharp'/'loud'/'rough'/'dark/bright'. The listening test was conducted in a soundproof, neutral room. The test was set up on a computer and the sounds were played via open headphones (Sennheiser HD 650). The relative relation between the levels of the recorded signals was adjusted according to the data set of the recordings. The levels were adjusted using a HEAD Acoustics dummy head to the same maximum levels as measured during the recordings. The simulations were adjusted to match the same relative levels. Both signals and questions were presented randomized with different orders for each participant and each repetition to minimize order and learning effects.

Each trial contained one signal and one statement. In total the session had 63 trials. Before the experiment, the participants conducted a practice session to familiarize themselves with all the sounds and all the attributes. The main session was repeated twice. In total 18 participants (9 male, 9 female) participated in the listening test (mean age = 26 years, s.d. = 3.3 years).

5.3 Results

The results of this listening test were analysed with a repeated measures ANOVA (ANalysis Of VAriance) to allow statistical validation of the results and to be able to investigate effects by embedded parameters such as tyres and roads. This method also allows measuring the interaction of embedded parameters. The results confirmed the hypothesis that tyre noise can be differentiated by its perception, and they further indicate that it is possible to differentiate between streets and tyres in the perception of rolling noise.

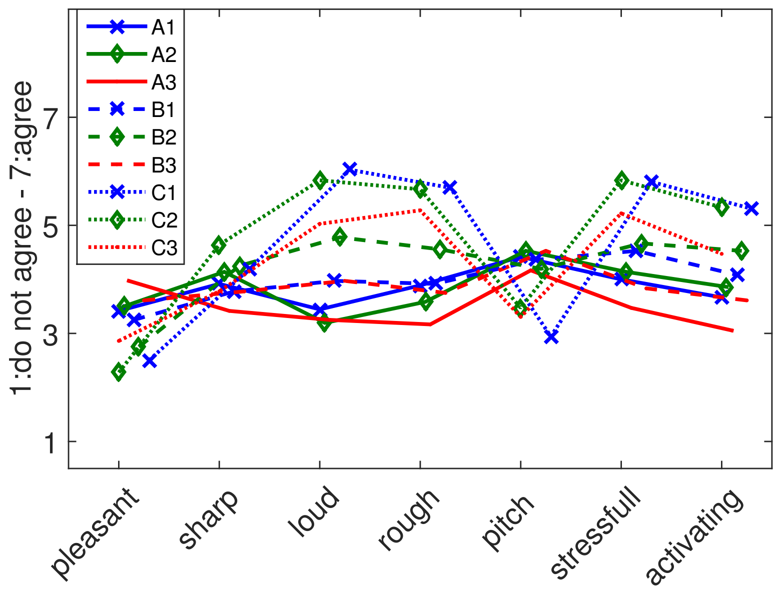


Figure 5.1: The results of the listening test are displayed: the means of the nine different tested sounds are shown for the tested percepts. The lines are marked corresponding to the different streets (A: solid, B: dashed and C: dotted line) and the different tyres (1: cross, 2: diamond and 3: dot). The judgement ranges from 1: no participant agreed to 7: all participants agreed that the percept describes the sound; for pitch the range was 1: dark to 7: bright

Figure 5.1 shows the mean values of the responses of the participants for the nine signals over the emotional and psychoacoustic parameters. There are clear differences in the responses for the different rolling noise cases. The results are compared to the calculated sound pressure levels and to the calculated psychoacoustic parameters and are analysed statistically.

Comparing the results for loudness (Figure 5.1) with the estimated sound pressure levels (Table 5.1), one can see similarities, but the loudest signal from Figure 5.1 (C1) has only the second-highest sound pressure level in dB or dB(A). This is even stronger for the least loud signals. A2 and A3 are rated the same in loudness, but the level of A2 is about 1.2 dB higher, and there are signals on levels between them that are rated higher in loudness. A2 and A3 have the same level of 67.24 dB and only half a dB difference in dB(A) but they were rated very differently in loudness.

Comparing the results from the listening tests (Figure 5.1) with the values calculated in Artemis by HEAD acoustics (Table 5.2), one can see that the loudness results lead to similar ranking in both methods. Only the signals A2 and B3 change more than 2 places between the 2 methods. For roughness the basic order remains similar for the two methods. There are only changes in order between signals that are rated closely to each other. For sharpness, the orders differ completely. This can be due to the fact that they are all very similar and almost not differentiable at all. These findings can be confirmed by plotting the experimental results over the calculated results for each parameter (Figure 5.2).

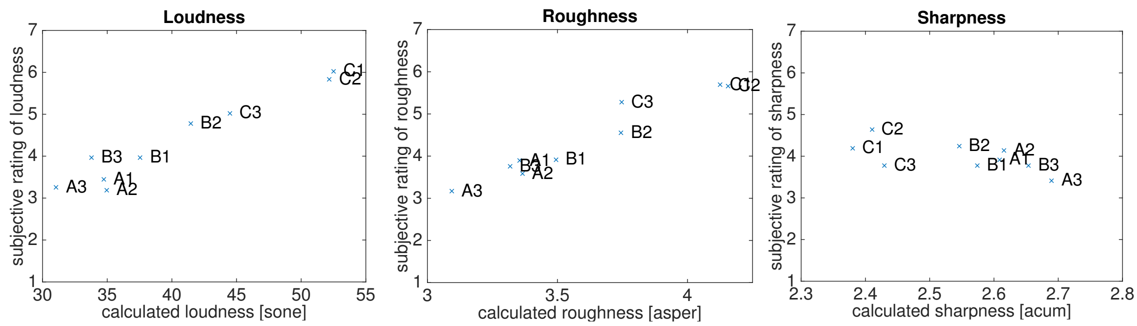


Figure 5.2: Comparison between the results from the listening tests and calculations for the nine signals for loudness, roughness and sharpness

To analyse if significant differences in the perception of the nine signals can be found, a one-way analysis was conducted for the responses of each psychoacoustic and emotional parameter. For each analysis, the results for the F-test will be given together with the degrees of freedom. Additionally, the probability (p value) for accepting the null hypotheses will be given, and the effect size. The meaning of those values is described in Chapter 3.1.4.

The results for the one-way analysis revealed that only sharpness could not be differentiated between the different signals. Pleasantness is significant with $F(8, 153) = 3,61; p < 0.001$ and an effect size of $\eta_p^2 = 0.16$; stress is significant with $F(8, 153) = 7,91; p < 0.001$ and an effect size of $\eta_p^2 = 0.29$; and activation is significant with $F(8, 153) = 7,30; p < 0.001$ and an effect size of $\eta_p^2 = 0.28$.

All emotional responses show significant differences in the perception of the nine signals and all show medium effects.

The results for the psychoacoustic parameters show that sharpness is not significant with $F(8, 153) = 1,45; p = 0.18$ and an effect size of $\eta_p^2 = 0.07$; loudness is significant with $F(8, 153) = 21,89; p < 0.001$ and an effect size of $\eta_p^2 = 0.53$; roughness is significant with $F(8, 153) = 10,31; p < 0.001$ and an effect size of $\eta_p^2 = 0.35$; and pitch is significant with $F(8, 153) = 3,86; p < 0.001$ and an effect size of $\eta_p^2 = 0.17$.

For the psychoacoustic parameters, loudness, roughness and pitch show significant differences in the perception of the nine signals, but pitch only has a minor effect. Loudness and roughness show a large effect. Sharpness has no significant difference in the perception of the nine signals and no effect either. This shows that the hypothesis of a difference in perception of rolling noise is true.

Further, the results in Figure 5.1 indicate that the different street surfaces (solid, dashed and dotted line) modulate the responses. Road C (dotted) is rated for all tyres as the least pleasant, the lowest pitch, and the loudest, roughest and most stressful. The difference between the other 2 roads is not as great, but road A (solid) is rated more pleasant, less loud, rough, stressful and activating than road B (dashed). For the tyres, the third one (dot) is the most pleasant on all roads. It is also the least rough, stressful and activating. The first tyre (cross) and the second tyre (diamond) are very similar in their perception and change order on the different streets for most psychoacoustic parameters.

For all the psychoacoustic parameters that showed significant results in the one-way ANOVA, a multivariate ANOVA was conducted to analyse the effects within the signal parameters road ($F_{road} = F_r/F(r)$) and tyre ($F_{tyre} = F_t/F(t)$) and their interaction.

For the emotional parameters, the results are significant both for the road and the tyre and the effects are medium to large. There was no significant result for the interaction between road and tyre for the emotional responses. The effect of the roads was significant for pleasantness with $F_{road}(2, 34) = 5.92; p < 0,01$ and an effect size of $\eta_p^2 = 0.26$, for activation with $F_{road}(2, 34) = 17.02; p < 0,001$ and $\eta_p^2 = 0.5$ and for stress with $F_{road}(2, 34) = 36.05; p < 0.001$ and $\eta_p^2 = 0.7$. The effect of the tyres was significant for pleasantness with $F_{tyre}(2, 34) = 12.22; p < 0.001$ and an effect size of $\eta_p^2 = 0.42$, for activation with $F_{tyre}(2, 34) = 18.68; p < 0.001$

and $\eta_p^2 = 0.52$ and for stress with $F_{tyre}(2,34) = 17.71; p < 0.001$ and $\eta_p^2 = 0.5$. The interactions between tyre and road were not significant (pleasantness with $F_{road} \times F_{tyre}(4,68) = 1.6; p = n.s.$, activation with $F_{road} \times F_{tyre}(4,68) = 0.71; p = n.s.$ and stress with $F_{road} \times F_{tyre}(4,68) = 0.12; p = n.s.$).

For the psychoacoustic parameters, loudness and roughness show the strongest effects. Both road and tyre have significant main effects on the perception – large for the road and medium for the tyre. Pitch gives a significant difference only for the road with a minor effect, but no significant result for the tyre. The interaction is significant for loudness and pitch, but not for roughness. The effect of the roads was significant for loudness with $F_{road}(2,34) = 51.21; p < 0.001$ and an effect size of $\eta_p^2 = 0.75$, for roughness with $F_{road}(2,34) = 39.12; p < 0.001$ and $\eta_p^2 = 0.7$ and for pitch with $F_{road}(2,34) = 5.2; p < 0.05$ and $\eta_p^2 = 0.23$. The effect of the tyres was significant for loudness with $F_{tyre}(2,34) = 11.71; p < 0.001$ and an effect size of $\eta_p^2 = 0.41$, for roughness with $F_{tyre}(2,34) = 8.97; p < 0.001$ $\eta_p^2 = 0.35$. But not for pitch ($F_{tyre}(2,34) = 0.72; p = n.s.$ and $\eta_p^2 = 0.04$). The interactions between tyre and road was significant for loudness: $F_{road} \times F_{tyre}(4,68) = 7.41; p < 0.001$ and for pitch: $F_{road} \times F_{tyre}(4,68) = 3.06; p < 0.05$ but not for roughness: $F_{road} \times F_{tyre}(4,68) = 2.03; p = n.s.$.

5.4 Discussion

The results confirm the hypothesis that tyre noise can be differentiated by the perception. This is valid both for the emotional psychoacoustic parameters of pleasantness, stress and activation and for the psychoacoustic parameters of loudness, roughness and pitch. Only in the perception of sharpness were the used signals not rated as significantly different.

The experiment also indicates that it is possible to distinguish between the influence of the tyres and the roads on the resulting rolling noise. Our experiment was evaluated using a multivariate ANOVA to analyse the influence of tyre and road. The effects of the tyres and roads are significant to emotional responses, with a medium to large effect size. For the psychoacoustic parameters, the responses show significant effects for the road as an influencing parameter for loudness, roughness and pitch, with large effect size for loudness and roughness and a medium effect size for pitch. For the perception of the tyre, the response differed significantly for loudness and roughness with medium effect sizes, but not for pitch. In general the effect sizes indicate a higher difference in the perception of the roads than of the tyres. This can also be seen in Figure 5.1.

The analysis of the interaction between the perception of the tyre and the road showed no interaction for pleasantness, stress, activation, and roughness but for loudness and pitch, whereas the preliminary study indicated an interaction for all tested psychoacoustic parameters. One reason could be that the signals used in the preliminary study were based on an older version of the auralization tool, so the signals in the two studies were not exactly the same. The signals in the main study consider low frequencies stronger, to include wind noise. This might mask the interaction effect between the perception of the tyre and the road. This difference must be further investigated. It would be helpful to know if the effects of the tyre and of the road on the rolling noise are independent from each other or if they interact in the perception.

The spread in the rating of the signals was largest for loudness and roughness, and also showed the highest effect sizes. This leads to the assumption that they are suited to characterize rolling noise. Pitch and sharpness seem to be of less importance due to their lower effect sizes. Stress and activation show a large variance and are thus of importance to the signals. Pleasantness showed a smaller effect size, with a smaller spread in the ratings than for the other emotional parameters. One reason may be that the sound in general was not seen as pleasant overall. This can limit the scale to half. It could be better in further experiments to test annoyance instead, and to focus more on that aspect of the perceptual space.

When comparing listening test results for loudness with the calculated sound pressure levels of the signals, it can be concluded that the levels do not properly indicate how strongly all signals will be perceived. This is also valid when applying A-weighting.

The calculation models in Artemis led to good agreement with the experiment for loudness and roughness. For sharpness, no agreement between model and experiment could be found.

The performed experiment improves the understanding of the perception of rolling noise. It indicates that rolling noise evokes psychoacoustic and emotional responses in a measurable range. This study confirms that a distinction can be made between the contribution of the road and the tyre on the perceptual rating of tyre–road noise. The emotional responses have primary effects on all tested parameters (pleasantness, activation and stress) for the tested sounds. The psychoacoustic parameters of loudness and roughness are of primary importance to the perception of rolling noise and have the largest effects. In relation to these, pitch and sharpness show a much smaller effect on the perception of the tested signals.

Due to contradictory results, no final conclusion about the interaction of the two could be drawn. An interaction between the effects by road and tyre could only be shown for loudness and pitch, and not for the other tested parameters, as the preliminary experiment suggested.

5.5 Interaction of psychoacoustic parameters

Another matter that could be followed up with the performed listening tests was to investigate the interaction of the psychoacoustic parameters with the emotional parameters. Which parameters influence the perception of pleasantness, stress and activation the most, and do they interact with each other? To follow up on that question, a correlation analysis between the parameters was conducted.

	sharp	loud	rough	pitch	stress	activation
pleasant	-0.81^{**}	-0.93^{***}	-0.95^{***}	0.75^{*}	-0.96^{***}	-0.98^{***}
sharp		0.64	0.67	-0.35	0.71^{*}	0.81^{**}
loud			0.97^{***}	-0.87^{**}	0.94^{***}	0.94^{***}
rough				-0.89^{**}	0.97^{***}	0.95^{***}
pitch					-0.86^{**}	-0.78^{*}
stress						0.98^{***}

Table 5.3: Analysis of the correlation between the different percepts over the mean value responses of the participants for the different sounds. The displayed values are the $\mathbf{R}$ values for each cross-correlation. $*$ marks the correlations at the 5% limit, $**$ marks correlations at the 1% limit and $***$ marks correlations at the 0.1% limit

The results are given in Table 5.3. Pleasantness is inverse to all attributes but pitch. The emotional responses pleasantness, stress and activation correlate with each other, as expected for the investigated type of sounds. Pleasantness and activation correlate with roughness, loudness and sharpness, but not with pitch on a 1% limit. Stress correlates with roughness, loudness and pitch, but not with sharpness on a 1% limit.

Due to the fact that there is no significant difference in the ratings of sharpness (Chapter 5.3), even though sharpness correlates with activation and pleasantness at the limit of 1%, it cannot be interpreted as a parameter that influences the emotional reaction to the tested signals. This leaves loudness and roughness as the main parameters that influence both pleasantness and activation. Stress is additionally affected by the pitch of the sound.

Chapter 6

Psychometric function for car pass-by in background noise

6.1 Introduction

For the traffic safety of pedestrians, it is important that approaching cars are detected in the background noise of overall traffic. This question has led to increasing interest in more silent vehicles, such as hybrid or electrical vehicles (see e.g. [111, 99]). The addition of artificial sound is discussed at least for low speeds where these vehicles are quietest. In this context it is essential to know the parameters that a listener uses to identify an approaching car.. It might be possible to use these parameters to lower the emitted sound levels from vehicles, while maintaining the information relevant for identifying them in time when approaching. Recent studies have focused on the perception of electric vehicles (see e.g. [58, 11]). In [58]), Grosse evaluated the differences in audibility between cars using a combustion engine and cars using electric engines. The sounds of vehicles passing by were recorded binaurally and presented with either recorded traffic noise or pink noise as background noise. The results indicate that electric cars are less audible than cars with conventional combustion engines when approaching at low speeds. The study also indicated that pink noise is not a suitable substitute for the recorded traffic noise, as the reaction times differed substantially when using pink noise as background noise. A study by Altinsoy [11] measured reaction time differences between cars with combustion engines and cars with electric engines. The results showed that the participants detected the electric cars later than those with combustion engines. This study also utilized recordings for the different stimuli.

The drawback when using recordings, however, is that it is difficult to vary the contents of the sound signals systematically as needed for a thorough parametric study, because among other things, the composition of the traffic responsible for the background noise cannot be controlled. Therefore, simulated data were used instead of recordings in the present study. This allows for systematically varying parameters such as speed, traffic composition, tyres, road surfaces or vehicle type individually.

6.2 Generation of traffic noise

Traffic noise is composed out of all the different noise sources that are in hearing distance of a listener. In this study, the focus is only on road traffic noise, so no structures such as houses or trees are considered. Only sound shields between the listener and a street were included. A sound environment might involve several streets, each of which can be split up into lanes. Thus, traffic in our case is a composition of vehicles driving in a number of lanes. Each lane is characterized by the distance and angle to the listener, measured at the moment of pass-by (closest distance). Further, the lane has a direction of traffic and properties for traffic in the lane: number of cars per second, regularity of the cars passing, speed of the cars, and vehicle types involved.

The auralization tool introduced in the previous chapters allows individual car pass-bys to be auralized for a given speed at a given distance with a given tyre–road combination. In addition, the auralization can include engine noise and a noise barrier. The pre-set signal length is 6 seconds, but it is possible to choose longer times as well. For a binaural impression, head-related transfer functions (HRTF) were added to the signal using the open source KEMAR dummy head recordings by Gardner and Martin at MIT [51]. These recordings provide head-related transfer functions at an angular resolution of 5 degrees on the horizontal plane.

To get a realistic impression of the traffic in a lane, the chosen speed is the average speed in the lane. A speed was selected for individual cars from a 'normal' distribution around the average speed. The range of possible speeds is distributed around the average in 2 km/h intervals from -6 km/h to +6 km/h with declining probability on both sides.

To distribute the chosen amount of cars per second over the total time of the traffic noise signal, the time was split into equal windows for the total number of cars. Since traffic is never completely regular, each car is shifted by a random factor within its possible window, which prevents cars from driving unrealistically close to each other. By default, the possible variation is set to 0.4 of the window length, to create a variation in time while keeping the cars from 'running into each other'.

The amount of traffic per lane can be set as desired. However, traffic amounts higher than two cars per second per lane are quite unrealistic. At 50 km/h, a car travels 13.9 m/s. With an average car length of 4-5 m, this leaves about 4 m, or one car length as a safety distance between two cars. The maximum possible physical density at 50 km/h would be three cars per second, with a car every 4.6 m. The higher the speed, the more cars per second could be on the road, but with increasing safety distances needed, two cars per second was found to be a reasonable limit. The basic scenario for the presented set of studies is described in the following. Some part of this will be varied in each study, but most details will stay the same. For most of the auralizations in the present studies, a Pirelli tyre (type P600, size 205/60-R15 91V) on an asphalt concrete surface 0/8 constructed to fulfil the requirements of ISO 10844-1994 was used. A combustion-engine car was chosen.

The test car had a driving speed of 50 km/h. The test car always approached from the left side. The pass-by duration of the test car was 6 seconds, where the car was directly in front of the listener after three seconds. The test car always passed by at a distance of 7.5 m. Figure 6.1 shows the position of the listener in relation to the track of the test car and road with high traffic responsible for the background noise.

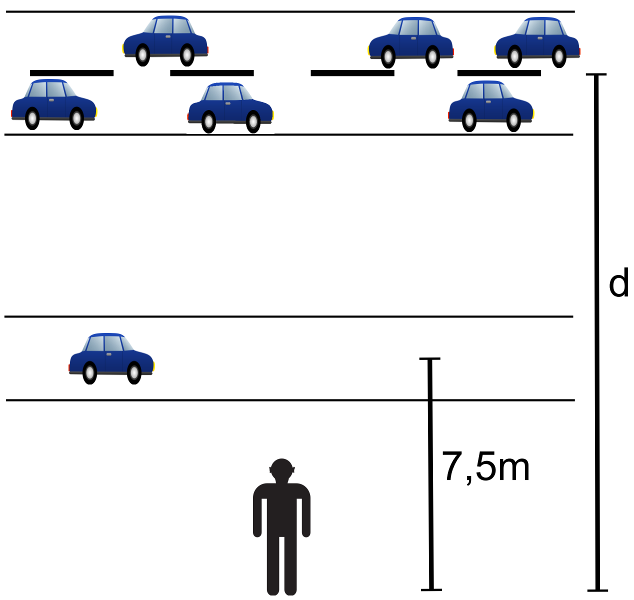


Figure 6.1: Illustration of the auralized traffic situation. The distance d varied between 7.5 m and 100 m.

For the background noise, the traffic flow was created by superposing single-car pass-by events based on the same vehicle properties as the test vehicle. A traffic flow of 3,600 vehicles per hour and lane was assumed (i.e. 1 car per second and lane on average). The sequence of the vehicles was varied on a band of ±0.4 seconds to allow traffic density fluctuations without overlapping the virtual positions of the cars. The speeds were distributed with an approximated normal distribution around 50 km/h (from 44 km/h to 56 km/h in 2 km/h increments) for each lane.

6.3 Detecting a car in background traffic noise

An increasing proportion of electric vehicles in everyday traffic has raised the question of the acoustic detectability of a car in the surrounding traffic. Experiments have been done to test certain electric vehicles against combustion vehicles in recorded traffic noise [11, 58] or in pink noise. The introduced tool to generate car pass-by noise and traffic noise allows this question to be followed up further, without the need to rely on sets of recordings. In the following section, the possibility of using validated simulated pass-by sound is used to follow up on the question of the detectability of a single car in background traffic noise. The study investigated which characteristics of the situation affect detectability and in what way.

Several aspects that might affect detectability were investigated individually and in interaction. To start with a reasonable number of options, a simple scenario was chosen with just two streets going in parallel, frontal to the listener. The street where single cars were to be detected was set at 7.5 m from the listener in all cases. Parameters that might affect detectability include distance from the second street, traffic amount, traffic regularity, speed, variation in speed, similarity of the involved cars and more. The parameters that were assumed to be of highest significance were tested in a set of listening tests. They are discussed below.

The listening tests were conducted in a soundproof, neutral room. The tests were set up on a computer and the sounds were played via open headphones (Sennheiser HD 650). The focus of the experiments was to investigate the influence of the distance and subsequently the relative changes in sound levels on reaction time; hence the sound levels were adjusted to never be too loud, but still audible, without changing the relationship between the signals. The exact levels are given for each experiment. The participants were compensated for their participation and gave their informed consent prior to their inclusion in the study.

The participants were given the information that they were standing in front of two roads (see 6.1)), where the more distant one was described as a road with heavy traffic and the closest road was described as a street with few individual cars passing by from time to time. The participants were asked to listen for an approaching car on the nearer road and press the space bar on a normal keyboard as soon as they detected the test car. They were further told that this car could only appear when a green arrow appeared on left side of the screen. No arrow was shown between the trials, and if the participant missed a car, a red arrow was shown on the right side of the screen. The word-by-word instructions and a screenshot of the experiment can be found in Appendix B.

The reaction time for detection of the test car was measured between the onset of the test car sound and the time the participant pressed the space bar. If no response was given, the trial was counted as a miss. The trial ended and a new trial started when the participant pressed the key.

For each trial in the experiment, a random background signal was presented from the set of background noises. With a random delay varying between 0.3 and 0.5 seconds, the test car was presented during the playback after the onset of the background signal. If more than one test car was used in the experiment, the order of test cars was randomized as well. To avoid accidental keyboard operations, reaction times were only registered 0.1 seconds after the onset of the test car. The test car needed three seconds to reach the frontal position to the listener. The Inter-trial interval was 1 second containing silence. Each participant conducted two sessions with all tested combinations and a short break in between. To ensure that the participant actually responded when they heard a signal, one trial of each condition only contained the background signal (a false positive test).

6.4 Distance

The aim of the study was to investigate the reaction time for detecting a car with a combustion engine (test vehicle) passing by in the presence of background noise from a road with high traffic flow. The distance between the path of the test vehicle and the road with high traffic volume is varied. All other parameters (i.e. car type, road surface, speed, etc.) were kept constant.

The hypothesis was that the closer the distance between the lane of the test vehicle and the road with high traffic volume, the more difficult it would be for the listener to detect the single car. The change in distance will strongly affect the signal-to-noise ratio (SNR) and thus have a strong effect on the detection of the test car. Here, the SNR is defined as the difference between the maximum level of the test vehicle $L_{max,fast}$ (fast time weighted) and the equivalent sound pressure level L_{eq} from the road in the background.

The distance d between the road with high traffic flow and the listener varied from 10 m to 100 m in 10 m increments. Additionally, another 7.5 m was included in order to have a case where the background street and the test street are at the same position. Further on, the distances were varied.

6.4.1 Method of the distance experiment

Twenty-five participants (15 male, 10 female) participated in the listening test (M = 29.2 years old, $s.d.$ = 9.7 years). All participants reported normal hearing.

To evaluate the impact of distance in the recognition of a car in a traffic noise background, a set of stimuli were created as described in the auralization method. The background traffic noise was created for a varying distance d from the listener as described above. Six different sound files were generated for each distance, with a total of 66 different background noises. Table 6.1shows the levels of the resulting background traffic noises in the simulation without calibration. The test car had the level of $L_{max,fast} = 50.40 dB(A)$ in the simulation.

d	$L_{eq}, dB(A)$
7.5m	61.25
10m	59.69
20m	55.27
30m	52.32
40m	49.98
50m	48.18
60m	46.85
70m	45.71
80m	44.80
90m	43.82
100m	43.04

Table 6.1: A-weighted levels of the simulated background sound files. The mean values over the 6 cases of the same distance are shown

The table reveals that the equivalent sound pressure levels do not decay at 3 dB per distance doubling as expected. This is due to the length of the single pass-by events of 6 seconds, which only represents a road of about 80 m length for a driving speed of 50 km/h.

6.4.2 Results of the distance experiment

Due to an error in the software, three participants had to be removed from the analysis of the results. Figure 6.2 shows the responses of all other participants. As can be seen in the figure, one participant's response stood out. This participant had not understood the task and was therefore removed from further analysis.

Participants who gave a false positive response (i.e. responded to hearing a car even when there was no car present) in more than 15% of the cases were removed. The two nearest conditions (7.5 m and 10 m) had a higher risk of leading to false positive responses and were not included in the 15% limit. This resulted in another three participants being removed from the analysis, since their responses were not seen to be consistent enough. For the remaining participants, the false positive responses varied between 0 and 11% ($M = 3.7\%$). In total 18 participants were included in the further analysis.

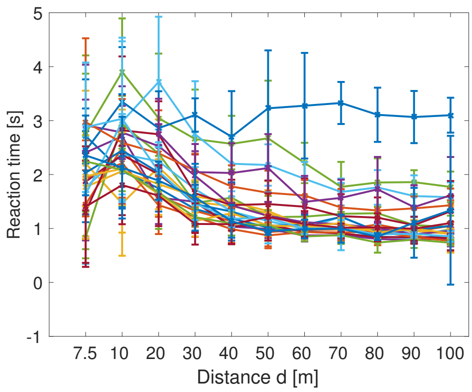


Figure 6.2: The figure shows the average reaction time and standard deviation over the distances d (as defined in Figure 6.1) for each of the participants in the experiment.

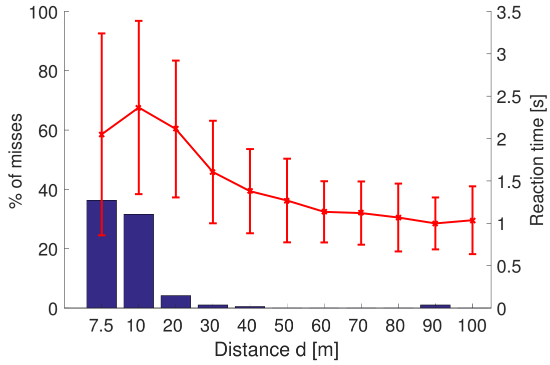


Figure 6.3: The figure shows the average reaction time and standard deviation over the distances d (as defined in Figure 6.1). The barplot shows the relative number of misses for the different background signals.

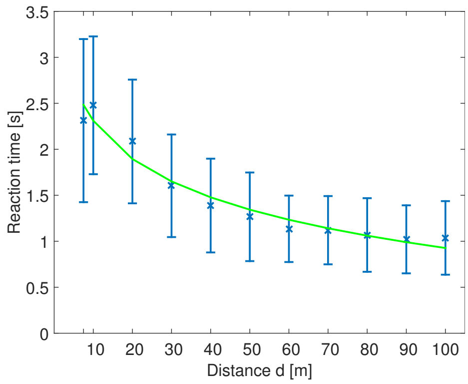


Figure 6.4: Average reaction time and standard deviations of the reaction times over the distances d (as defined in Figure 6.1) and the fitted line from the regression analysis.

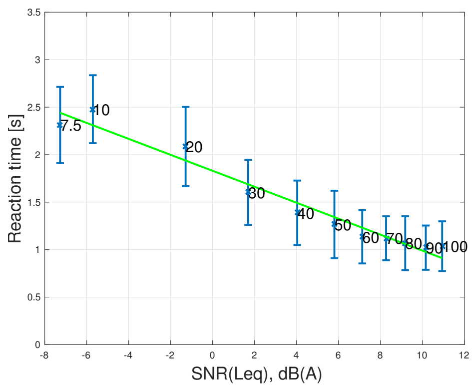


Figure 6.5: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 11 cases.

The maximum latency was set to 3,000 milliseconds after the onset of the test car; latencies longer than this were replaced by 3,000 milliseconds. Figure 6.3 shows the average response time for the different distances, along with the percentage of misses. A repeated measure analysis of variance (ANOVA) with distance as main factor determined that there was a main effect of distance ($F(10, 170) = 111.04$, $p < .001$). Bonferroni post-hoc tests revealed that when the background sounds were presented at >50 metres there was no significant difference between the adjacent distances and that there was no significant difference between the 7.5 m and the 10 m distance. A regression analysis was conducted to determine the reaction time for detecting a car as a function of the distance from the background. This resulted in a logarithmic regression that explained a significant proportion of the variance in reaction time ($b = 601.3, t(9) = -13.77, p < .001$, $R^2 = .96, F(1, 9) = 189.67, p < .001$) as is illustrated in 6.4.

Since the change in SNR for the different cases is expected to have a strong effect on the reaction time, a regression analysis was conducted using reaction time as a function of the SNR. The SNR significantly predicted the reaction time ($b = .084, t(9) = -14.66, p < .001$). It also explained a significant proportion of variance in the RT ($R^2 = .96, F(1, 9) = 214.77$, $p < .001$). This is illustrated in Figure 6.5, where the regression and the average reaction time are plotted over the SNR between the $L_{max,fast}$ of the test signal and the L_{eq} of the background noise for the 11 distances.

6.4.3 Discussion of the distance experiment

The aim of the present study was to investigate the ability of a listener to detect a car in the presence of background noise of traffic. The varied parameter in the study is distance between the listener and the road with high traffic flow responsible for the background noise. In contrast to most of the published work, simulated sound files for both the test vehicle and the background noise were used. The results showed that there is a logarithmic relationship between the response times and the distance d between the listener and the road with high traffic flow. The logarithmic relationship indicates that the sound pressure levels, or better, the SNR governs the reaction time as expected. This was confirmed by a second regression analysis, relating the reaction time to the SNR. The SNR has a linear relation to the reaction time (Figure 6.5). The regression analysis predicts the reaction time with statistically significant agreement. Thus, the change in reaction time with distance is mainly explained by the change in SNR.

The results also show that the method of using simulated data instead of recorded data seems to be a feasible approach as it is demonstrated here. Grosse's [58] study indicates detections of cars in background noise from traffic with reaction times of 1 s before the pass-by until 1.5 s after the pass-by for different cars with combustion engines. The background noise was recorded for a road 50 m away from the road where the car is assumed to pass by. The results are in agreement with the findings of the presented study, where the reaction time was around 1 second before the pass-by for 50 m distance. The presented study only studied one car with a combustion engine.

6.5 Traffic amount

The aim of this part of the study was to investigate the reaction time for detecting a car with a combustion engine (test vehicle) passing by in the presence of background noise from a road with a varying traffic amount. It was tested for three different distances of the background street: 20 m, 40 m and 80 m. All other parameters (i.e. car type, road surface, speed, etc.) were kept constant.

The hypothesis was that the higher the traffic amount on the street, the higher the total level and thus the more difficult it would be for the listener to detect the individual car. On the other hand, if the traffic flow on the background street is so low that each car is separate, the confusion might increase as to whether it is a close or a distant car. Thus detection might become poorer for cases with a very low traffic flow.

From the previous experiment, three distances were chosen, to test whether or not the effect of the traffic amount was independent of distance. The closest distance was 20 m for changes to be detected in both directions. The furthest distance was 80 m. After that, no significant change in effect by distance could be seen in the previous experiment.

To test the effect of traffic amount, 7 traffic amounts were chosen. The highest physical possible traffic amount for 50 km/h was estimated to be 2 cars per second and lane (7200 vehicles per hour and lane), resulting in the upper limit. The lowest traffic amount tested was 0.1 cars per second and lane (360 vehicles per hour and lane). Low traffic amounts on streets are much lower, but not feasible for testing in this experiment. Outgoing from those limits, the tested traffic amounts were 0.1, 0.25, 0.5, 0.75, 1, 1.5 and 2 cars per second and lane.

6.5.1 Method of the traffic amount experiment

Twenty-six participants (16 male, 10 female) participated in the listening test ($M =$ 26.4 years old, $s.d. = 4.5$ years). All participants reported normal hearing.

To evaluate the impact of traffic amount in the recognition of a car in a traffic noise background, a set of stimuli were created as described in the auralization method. The background traffic noise was created for the described three distances and 7 traffic amounts. Six different sound files were generated for each background condition, to increase the statistical stability of the results. In total 126 different background noises were created. The levels of the resulting background traffic noises in the simulation without calibration are given in Table 6.2. As one would expect from theory, a doubling of the number of cars, and thus a doubling of the number of sound sources, leads to a level increase of 3 dB. The standard deviation for the levels of the 6 repetitions was always below 0.06 dB. The test car had the level of $L_{max,fast} = 54.5\ dB(A)$ in the simulation.

	traffic condition	$L_{eq},\ dB(A)$	Label
Case 1	80m, 0.1 cars/sec	37.1	F1
Case 2	80m, 0.25 cars/sec	41.1	F2
Case 3	80m, 0.5 cars/sec	44.1	F3
Case 4	80m, 0.75 cars/sec	45.8	F4
Case 5	80m, 1 cars/sec	47.1	F5
Case 6	80m, 1.5 cars/sec	48.8	F6
Case 7	80m, 2 cars/sec	50.1	F7
Case 8	40m, 0.1 cars/sec	43.5	M1
Case 9	40m, 0.25 cars/sec	47.5	M2
Case 10	40m, 0.5 cars/sec	50.5	M3
Case 11	40m, 0.75 cars/sec	52.3	M4
Case 12	40m, 1 cars/sec	53.5	M5
Case 13	40m, 1.5 cars/sec	55.3	M6
Case 14	40m, 2 cars/sec	56.5	M7
Case 15	20m, 0.1 cars/sec	49.1	N1
Case 16	20m, 0.25 cars/sec	53.1	N2
Case 17	20m, 0.5 cars/sec	56.1	N3
Case 18	20m, 0.75 cars/sec	57.9	N4
Case 19	20m, 1 cars/sec	59.1	N5
Case 20	20m, 1.5 cars/sec	60.9	N6
Case 21	20m, 2 cars/sec	62.1	N7

Table 6.2: A-weighted levels of the simulated background sound files. The mean values over the 6 signals are given for each case. Additionally, the labels used for each case in the figures are given.

6.5.2 Results of the traffic amount experiment

Figure 6.6 shows the responses by all participants. One participant had to be excluded due to a computational error during the experiment. Participants who made a false positive response (i.e. responded to hearing a car even when there was no car

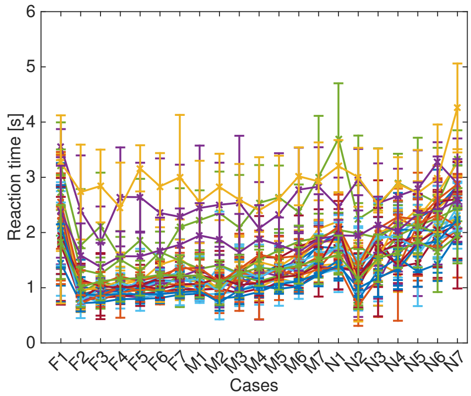


Figure 6.6: The figure shows the average reaction time and standard deviation over the cases for each of the participants in the experiment. The cases are named according to Table 6.2, where the speeds are marked by letters (F: 80 m, M: 40 m, N:20 m) and the traffic amounts are numbered with irising numbers for increasing traffic amount (1: 0.1 cars per second, 7: 2 cars per second).

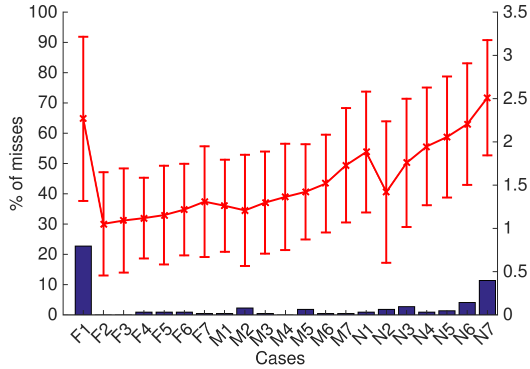


Figure 6.7: The figure shows the average reaction time and standard deviation over the cases. The barplot shows the relative number of misses for the different background signals. The cases are named according to Table 6.2.

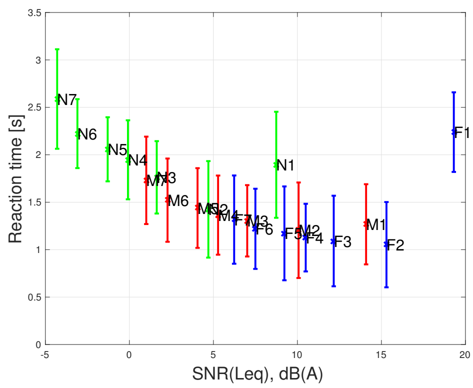


Figure 6.8: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 21 cases. The cases are named according to Table 6.2.

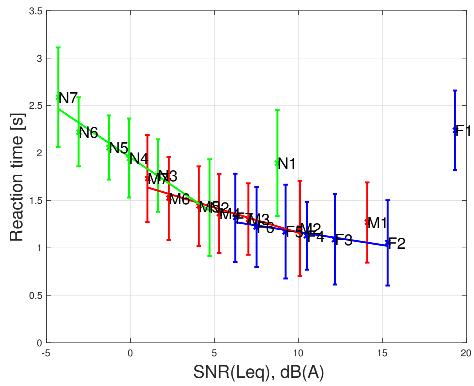


Figure 6.9: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 21 cases. The estimated regression lines are plotted for each distance. The cases are named according to Table 6.2.

present) in more than 15% of the cases were removed. This resulted in another three participants being removed from the analysis, since their responses were not seen to be consistent enough. For the remaining participants, the false positive responses varied between 0 and 11.9% ($M = 5.17\%$). In total 22 participants were included in the further analysis.

Figure 6.7 shows the average response time for the different distances are presented along with the percentage of misses. A repeated measure analysis of variance (ANOVA) with distance as main factor determined that there was a main effect for the tested cases ($F(20, 441) = 20.78$, $p < .001$).

The effect of the SNR on reaction time was investigated, to see if the SNR explains the differences between the cases. A regression analysis was carried out for the three different tested distances using reaction time as a function of the SNR. The first case (traffic amount of 0.1 cars per second) was excluded for all distances. The reaction times for that case demonstrate a different behaviour than the rest.

The SNR significantly predicted the reaction time for all three distances individually (for 20 m $b = -.12,\ t(4) = -10.35,\ p < .01$, for 40 m $b = -.05,\ t(4) = -5.98,\ p < .01$ and for 80 m $b = -.03,\ t(4) = -5.96,\ p < .01$). It also explained a significant proportion of variance in the RT for each of the three distances (for 20 m $R^2 = .86,\ F(1, 4) = 107.13,\ p < .01$; for 40 m $R^2 = .89,\ F(1, 4) = 35.75,\ p < .01$ and for 80 m $R^2 = .89,\ F(1, 4) = 35.51,\ p < .01$). This calculation was done excluding the data points for 0.1 cars per second. The results are illustrated in Figure 6.9, where the regression and average reaction time are plotted over the SNR between the $L_{max.fast}$ of the test signal and the L_{eq} of the background noise for the three distances and seven tested traffic amounts.

6.5.3 Discussion of the traffic amount experiment

The varied parameter in this part of the study is the amount of traffic in the background. The reaction time that a listener needed to detect a closer car in the presence of an additional street in the background was measured for seven different traffic amounts on the background street and three different distances of background traffic. The car to detect always passed at 7.5 m and all cars had an average speed of 50 km/h. Simulated sound files were used for all cars. Figure 6.6 shows the results for all participants. They generally demonstrate similar behaviour. The results show the same effect for distance as the previous experiment. Figure 6.7 shows the mean reaction times and standard deviations over all participants for all tested cases. For all three distances, the detection time increases with increasing traffic amount, as expected. However, for all distances, the lowest traffic amount has a strong increase in detection time. Surprisingly, this effect is strongest for the distance of 80 m. Even the number of misses is significantly larger here. The increase might be explained by the fact that with only 0.1 cars per second it was hard to decide whether or not the test car was part of the background. However, this does not explain the strong effect on 80 m distance.

For each distance, the SNR between background traffic and test car seems to have a strong effect on the reaction time. But it can also be seen that the SNR is not the only factor governing detection time (Figures 6.8 and 6.9). One thing that can be noted is that distance has a somewhat compressing effect on the changes in reaction time due to the traffic amount. An increase in traffic amount will lead to a longer reaction time, but the size of the effect also depends on the distance.

6.6 Traffic regularity

The aim of this part of the study was to investigate the reaction time for detecting a car with combustion engine (test vehicle) passing by in the presence of background noise from a road with high traffic flow. The regularity of the traffic on the road with high traffic volume varied. This was tested for two different traffic amounts. All other parameters (i.e. car type, road surface, speed, etc.) were kept constant. The hypothesis was that a traffic flow where the cars appear regularly, with the same time in between all cars, might sound unrealistic and make it easier to detect an additional car as a 'rhythm breaker'. A very high irregularity, on the other hand, might create unrealistic overlaps, and a higher computational effort. Thus, the question was whether traffic regularity has an effect on the detection of a closer car.

The traffic flow for background noise was created by superposing single car pass-by events that are based on the same vehicle properties as the test vehicle. The background road was placed at 20 m frontal to the listener. A traffic flow of 1,800 and 900 vehicles per hour and lane was assumed (i.e. 0.5 and 0.25 cars per second and lane on average). Those numbers of vehicles were chosen to create traffic amounts where the single cars are still distinguishable, but there are no longer pauses between following vehicles. The sequence of the vehicles was varied in a band of ± 0, ± 0.4 ± 0.9 seconds to investigate the effect of the regularity of the traffic. The speeds were distributed with an approximated normal distribution around 50 km/h (from 44 km/h to 56 km/h in 2 km/h steps) for each lane.

6.6.1 Method of the traffic regularity experiment

Twenty-six participants (16 male, 10 female) participated in the listening test ($M =$ 26.4 years old, $s.d. = 4.5$ years). All participants reported normal hearing.

To evaluate the impact of traffic amount on the recognition of a car in a traffic noise

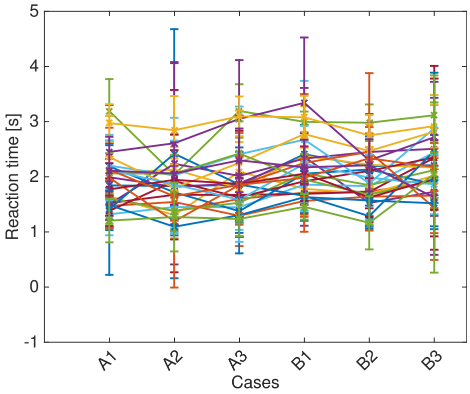


Figure 6.10: The figure shows the average reaction time and standard deviation over the cases for each of the participants in the experiment. The cases are named according to Table 6.3.

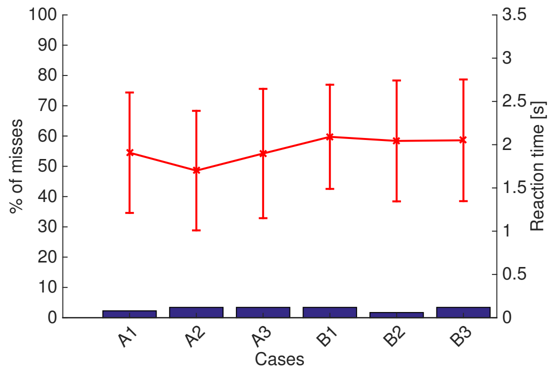


Figure 6.11: The figure shows the average reaction time and standard deviation over the cases. The barplot shows the relative number of misses for the different background signals. The cases are named according to Table 6.3.

background, a set of stimuli were created as described in the auralization method. The background traffic noise was created for varying traffic amount and traffic regularity as described above. For each background condition, six different sound files were generated, to increase the statistical stability of the results. In total 36 different background noises were created. Table 6.3 shows the levels of the resulting background traffic noises in the simulation without calibration. The test car had the level of $L_{max,fast} = 54.5\ dB(A)$ in the simulation.

	traffic condition	$L_{eq},\ dB(A)$	Label
Case 1	0.25 cars/sec, variation 0.0	53.11	A1
Case 2	0.25 cars/sec, variation 0.4	53.10	A2
Case 3	0.25 cars/sec, variation 0.9	53.10	A3
Case 4	0.50 cars/sec, variation 0.0	56.52	B1
Case 5	0.50 cars/sec, variation 0.4	56.10	B2
Case 6	0.50 cars/sec, variation 0.9	56.10	B3

Table 6.3: A-weighted levels of the simulated background sound files. The mean values over the 6 versions of the same case are given. Additionally, the labels used for each case in the figures are given.

6.6.2 Results of the traffic regularity experiment

Figure 6.10 shows the responses of all participants. One participant had to be excluded due to a computational error during the experiment. Participants who made a false positive response (i.e. responded to hearing a car even when there was no car present) in more than about 15% of the cases were removed. This resulted in another three participants being removed from the analysis. Due to the small amount of data and rounding, the false positive responses varied between 0 and 16.66% ($M = 4.35\%$) for the participants included in the further analysis. In total 22 participants were included in the further analysis.

Figure 6.11 shows the average response time for the different distances, along with the percentage of misses. A repeated measure analysis of variance (ANOVA) with distance as main factor determined that there is no main effect for the tested cases ($F(5, 126) = 2.17$, $p = .06$).

6.6.3 Discussion of the traffic regularity experiment

The varied parameter in this part of the study is the regularity of the traffic in the background. The reaction time needed for a listener to detect a closer car in the presence of an additional street in the background was measured for three different traffic regularities and two different traffic amounts. The car to detect always passed at 7.5 m and all cars had an average speed of 50 km/h. Simulated sound files were used for all cars. Figure 6.10 shows the results for all participants. Most participants show a similar response behaviour.

The results indicate that there is no significant effect on reaction time due to the amount of regularity in the background traffic flow. However, a certain amount of irregularity in the traffic might make the hearing impression more realistic. The hypothesis was that a regular traffic flow would make it easier to detect an additional car appearing at a random time. However, if there is such an effect, the presented study shows that its influence on reaction time is too small to be detected in an experiment with such an individual spread of reaction time of over one second.

6.7 Speed

The aim of this part of the study was to investigate the reaction time for detecting a car with combustion engine (test vehicle) passing by in the presence of background noise from a road with high traffic flow. The speed on the road with high traffic volume varied. The effect was tested for three traffic amounts on the background road and two speeds of the test car. All other parameters (i.e. car type, road surface, etc.) were kept constant.

The hypothesis was that the higher the speed, the louder the traffic noise gets. Thus the test car was expected to be harder to detect the higher the speed of the background traffic was in relation to the speed of the test car. On the other hand, with increasing speed the frequency range of a car pass-by shifts towards higher frequencies. If the difference in speed is high enough, the difference in frequency range might make it easier to differentiate between the cars in the background and the test car.

The test car had a speed of 30 km/h in one half of the experiment and 50 km/h in the other half. The distance of the background traffic was chosen to be 20 m as in the previous experiment.Traffic flows of 1,800, 3,600 and 7,200 vehicles per hour and lane were assumed (i.e. 0.5, 1 and 2 car per second and lane on average).

To investigate the effect of the speed of vehicles in the background traffic on the detection of a test vehicle, four different common speeds in traffic were chosen, namely 30 km/h, 50 km/h, 70 km/h and 90 km/h. The speeds were distributed with an approximated normal distribution around the centre speed (from −6 km/h to +6 km/h in 2 km/h increments) for each lane.

6.7.1 Method of the speed experiment

Twenty-six participants (19 male,7 female) participated in the listening test ($M = 28$ years old, $s.d. = 4.69$ years). All participants reported normal hearing.
To evaluate the impact of traffic amount on the recognition of a car in a traffic noise background, a set of stimuli were created as described in the auralization method. Five different sound files were generated for each background condition, to increase the statistical stability of the results. In total 72 different background noises were created. The experiment was done with the test car driving at 30 km/h and with the test car driving at 50 km/h. All participants rated both speeds.

Table 6.4 shows the levels of the resulting background traffic noises in the simulation without calibration. The standard deviation for the levels of the 6 repetitions was always below 0.06 dB. As one would expect from theory, a doubling of the number of cars, and thus a doubling in the number of sound sources, leads to an level increase of 3 dB. The test car had the level of $L_{max,fast} = 54.5\ dB(A)$ for the scenario with a speed of 50 km/h and the level of $L_{max,fast} = 49.6\ dB(A)$ for the scenario with a speed of 30 km/h.

6.7.2 Results of the speed experiment

Figures 6.12 and 6.16 show the responses of all participants. Participants who made a false positive response (i.e. responded to hearing a car even when there was no car present) in more than about 15% of the cases were removed. However, in this experiment the 70 km/h and the 90 km/h sample showed a higher risk of leading to a false response. Thus, they were excluded from the false positive test. This meant that it only took one false positive response to exceed the 15% limit, which resulted

	traffic condition	$L_{eq},\ dB(A)$	Label
Case 1	30 km/h, 0.5 cars/sec	50.9	30A
Case 2	30 km/h, 1.0 cars/sec	53.9	30B
Case 3	30 km/h, 2.0 cars/sec	56.9	30C
Case 4	50 km/h, 0.5 cars/sec	56.1	50A
Case 5	50 km/h, 1.0 cars/sec	59.1	50B
Case 6	50 km/h, 2.0 cars/sec	62.1	50C
Case 7	70 km/h, 0.5 cars/sec	60.4	70A
Case 8	70 km/h, 1.0 cars/sec	63.4	70B
Case 9	70 km/h, 2.0 cars/sec	66.4	70C
Case 10	90 km/h, 0.5 cars/sec	63.4	90A
Case 11	90 km/h, 1.0 cars/sec	66.4	90B
Case 12	90 km/h, 2.0 cars/sec	69.5	90C

Table 6.4: A-weighted levels of the simulated background sound files. The mean values over the 5 signals are given for each case. Additionally, the labels used for each case in the figures are given.

in 6 participants being removed from the analysis, since their responses were not seen to be consistent enough. For the remaining participants, the false positive responses varied between 0 and 20% ($M = 12.25\%$). A total of 20 participants were included in the further analysis.

Figure 6.13 shows the average response time for the different cases with a test car driving at 50 km/h, along with the percentage of misses. A repeated measure analysis of variance (ANOVA) with the cases as main factor determined that there was a main effect for the tested cases ($F(11, 228) = 13.78$, $p < .001$).

Figure 6.17 shows the average response time for the different cases with a test car driving at 30 km/h, along with the percentage of misses. A repeated measure analysis of variance (ANOVA) with the cases as main factor determined that there was a main effect for the tested cases ($F(11, 228) = 4.36$, $p < .001$).

The effect of the SNR on reaction time was investigated for both speed, to see if the SNR explains the differences between the cases. A regression analysis was done for the four different tested speeds of background traffic, using reaction time as a function of the SNR. The results of the regression analysis are quite weak, since only three traffic amounts were measured per speed. Still, it gives an indication of the effect due to speed.

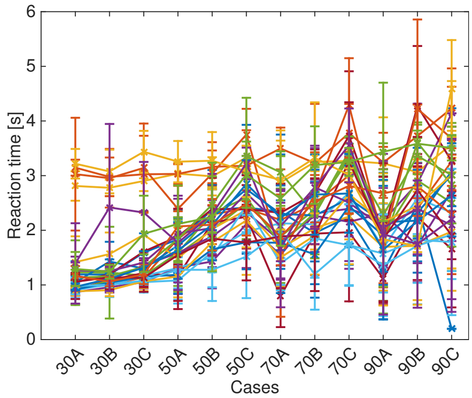


Figure 6.12: The figure shows the average reaction time and standard deviation over the cases for each of the participants in the experiment. Different speeds of the background traffic were tested for a test car at 50 km/h. The cases are named according to Table 6.4. The cases are marked with the tested speed. The following letters indicate the tested traffic amount (A: 0.5 cars per second, B: 1 cars per second, C: 2 cars per second).

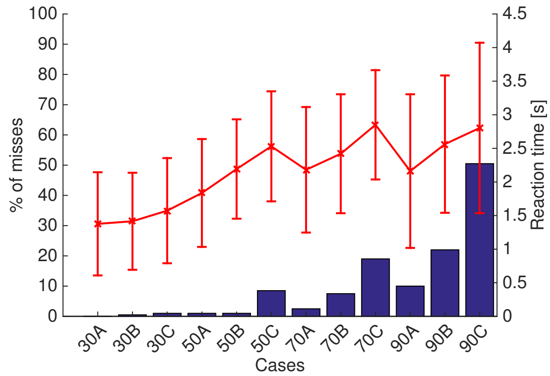


Figure 6.13: The figure shows the average reaction time and standard deviation over the cases. Different speeds of the background traffic were tested for a test car at 50 km/h. The barplot shows the relative number of misses for the different background signals. The cases are named according to Table 6.4.

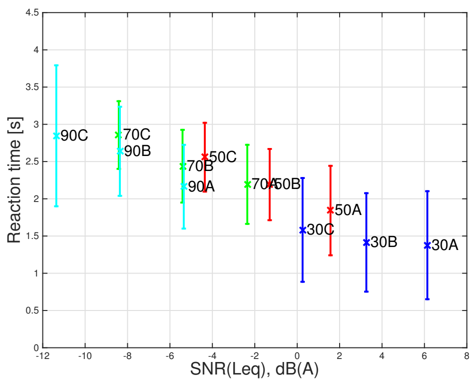


Figure 6.14: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 12 cases and for a test car at 50 km/h. The cases are named according to Table 6.4.

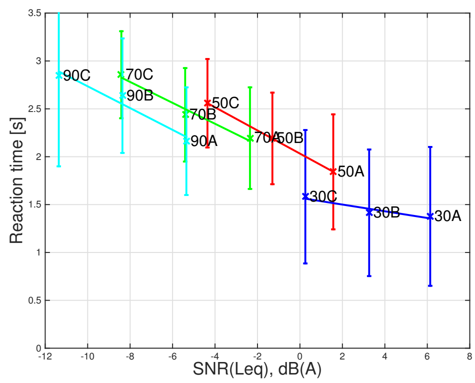


Figure 6.15: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 12 cases and for a test car at 50 km/h. The estimated regression lines are plotted for each speed. The cases are named according to Table 6.4.

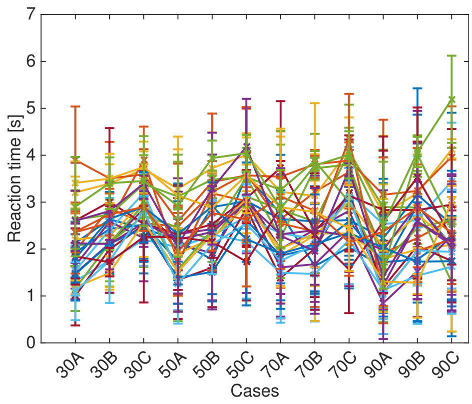


Figure 6.16: The figure shows the average reaction time and standard deviation over the cases for each of the participants in the experiment. Different speeds of the background traffic were tested for a test car at 30 km/h. The cases are named according to Table 6.4. The cases are marked with the tested speed. The following letters indicate the tested traffic amount (A: 0.5 cars per second, B: 1 cars per second, C: 2 cars per second).

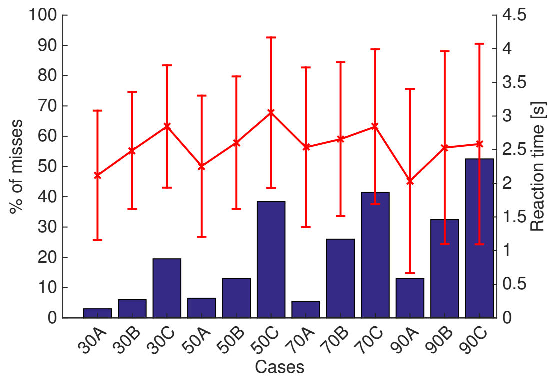


Figure 6.17: The figure shows the average reaction time and standard deviation over the cases. Different speeds of the background traffic were tested for a test car at 30 km/h. The barplot shows the relative number of misses for the different background signals. The cases are named according to Table 6.4.

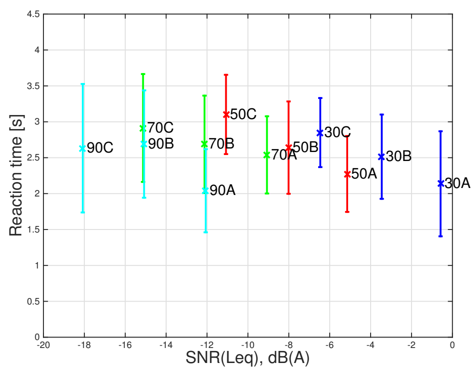


Figure 6.18: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 12 cases and a test car at 30 km/h. The cases are named according to Table 6.4.

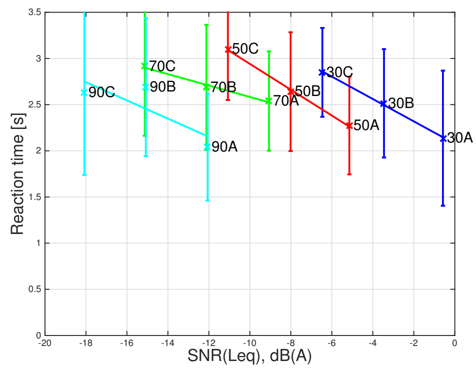


Figure 6.19: SNR between test signal and background noise over the average reaction time for the 12 cases and a test car at 30 km/h. The estimated regression lines are plotted for each speed. The cases are named according to Table 6.4.

For the test car driving 50 km/h, the results of the regression analysis are $b = -.04, t(1) = -2.84, p = .22$ for a background of 30 km/h, $b = -.12,\ t(1) = -254.88,\ p < .01$ for a background of 50 km/h, $b = -.11,\ t(1) = -6.36,\ p < .1$ for a background of 70 km/h and $b = -.11,\ t(1) = -4.46,\ p = .14$ for a background of 90 km/h. The SNR explained a significant proportion of the variance in the RT for a background of 50 km/h $R^2 = 1,\ F(1,1) = 6.4e^4,\ p < .01$ and 70 km/h $R^2 = .98,\ F(1,1) = 40.43,\ p < .1$, but not for 30 km/h $R^2 = .89,\ F(1,1) = 8.07,\ p = .22$ and 90 km/h $R^2 = .95,\ F(1,1) = 19.91,\ p = .14$. The results are illustrated in Figure 6.15, where the regression and the average reaction time are plotted over the SNR between the $L_{max,fast}$ of the test signal and the L_{eq} of the background noise for the four speeds and three tested traffic amounts.

For the test car driving 30 km/h, the results of the regression analysis are $b = -.12, t(1) = -21.59, p < .05$ for a background of 30 km/h, $b = -.14,\ t(1) = -21.17,\ p < .05$ for a background of 50 km/h, $b = -.06,\ t(1) = -7.81,\ p < .1$ for a background of 70 km/h and $b = -.1,\ t(1) = -1.45,\ p = .38$ for a background of 90 km/h. The SNR explained a significant proportion of variance in the RT for a background of 30 km/h $R^2 = 1,\ F(1,1) = 466.14,\ p < .05$, 50 km/h $R^2 = 1,\ F(1,1) = 448.47,\ p < .05$ and 70 km/h $R^2 = .98,\ F(1,1) = 61.07,\ p < .1$, but not for 90 km/h $R^2 = .68,\ F(1,1) = 2.11,\ p = .38$. The results are illustrated in Figure 6.19, where the regression and the average reaction time are plotted over the SNR between the $L_{max,fast}$ of the test signal and the L_{eq} of the background noise for the four speeds and three tested traffic amounts.

6.7.3 Discussion of the speed experiment

The varied parameter in this part of the study is the speed of the traffic in the background. The reaction time needed for a listener to detect a closer car in the presence of an additional street in the background was measured for four different speeds of the traffic on the background street and for three different traffic amounts of the background traffic. The car to detect always passed at 7.5 m and all cars had an average speed of 50 km/h or 30 km/h. Simulated sound files were used for all cars. Figure 6.12 shows the results for all participants for the cases with the test car driving at 50 km/h. Most of them show a very similar behaviour. Figure 6.16 shows the results for all participants for the cases with the test car driving at 30 km/h. Generally they show a similar behaviour.

Figures 6.13 and 6.17 show the mean reaction times and standard deviations over all participants for all tested cases. The results show the same effect for the traffic amount as in the previous experiment; the higher the traffic amount, the higher the reaction time is as well. Not surprisingly, the best reaction times can be found for the case where the test car was driving faster (50 km/h) than the cars in the background traffic (30 km/h) (Figure 6.13). There is no significant difference between the test car driving at 50 km/h with background traffic of the same speed, the test car driving at 30 km/h with background traffic of the same speed and the test car driving at 30 km/h with background traffic of 50 km/h. However looking at the SNR, the cases differ. Interestingly, the reaction time drops a little for the case with the difference in speed between test car (30 km/h) and background traffic (90 km/h). But this case also shows the highest numbers of misses for each of the three traffic amounts.

Looking at the relationship between SNR and reaction time, there seems to be some effect for each speed, but no overall effect. Even though a regression analysis will not show very significant results with only three data points, one was carried out for each speed. The previously discussed experiments already confirmed a linear relationship for the traffic amount. Thus, comparable relationships can be assumed even for this experiment.

6.8 Tyre–road noise and additional tonal components

The aim of this part of the study was to investigate whether differences in the sound characteristics of a test car affect reaction time. The test was done in varying amounts of background traffic. Other parameters, such as distance, speed, traffic regularity etc. were kept constant. The question was whether differences in the perception of tyre–road noise that were discussed in Chapter 5 even have an impact on detection. Two road surfaces were chosen from that study. The resulting tyre–road noise differed most strongly in perceptual parameters such as loudness, roughness and pleasantness (Chapter 5).

The first tyre–road combination was the same as in the previous studies (Chapter 6.2). For the second tyre–road combination, the same tyre (Pirelli, type P600, size 205/60-R15 91V) was driving on a concrete surface covered with synthetic resin and gravel. The concrete surface was polished before a synthetic resin was applied and gritted with gravel with sizes from 5 mm to 8 mm. For the car, a combustion engine was chosen. Each tyre–road noise was tested in background noise composed of the same or the other tyre–road noise.

Further, it was investigated how much the detection of a test car can be improved by adding tonal components. Two tonal components were tested (315 Hz and 2500 Hz), chosen according to the guidelines in [4]. The tonal components were designed in such a way that they were audible, but did not have a strong effect on the the overall car noise level. The level increase was 0.49 dB for both frequencies in relation to the signal without an additional tonal component and for a single pass-by event. The tonal components were added to the second tyre–road combination and tested in all backgrounds.

The experiment was done for the distance d of 20 m to the background traffic at pass-by. Each background was further tested for three traffic amounts (1, 1.5 and 2 cars per second), which were chosen from the previous studies (Chapter 6.5).

6.8.1 Method of the experiment investigating the effect of the tyre and additional tonal components

Twenty-four participants (16 male, 8 female) participated in the listening test ($M =$ 26.5 years old, $s.d. = 4.8$ years). All participants reported normal hearing.

To evaluate the impact of changes in the noise of the test car on the recognition of a car in a traffic noise background, a set of stimuli were created as described in the auralization method. Background traffic noise was created for the two tyre–road combinations and the three chosen traffic amounts. Five different sound files were generated for each background scenario, to increase the statistical stability of the results. A total of 15 different background noises were created and tested with the four different test cars and silence (false positive test) for each tyre–road combination. The two different tyre–road combinations in the background were tested in different blocks of the experiment. One block contained a total of 75 test conditions. Table 6.5 shows the levels of the resulting background traffic noises in the simulation without calibration. The standard deviation for the levels of the five repetitions was always below 0.06 dB. As one would expect from theory, a doubling of the number of cars, and thus a doubling of the number of sound sources, leads to a level increase of 3 dB. The first test car condition (first tyre–road combination) had the level of $L_{max,fast} = 54.5dB(A)$ in the simulation, the second test car condition (second tyre–road combination) had the level of $L_{max,fast} = 55.5dB(A)$, the third test car condition (second tyre–road combination and tonal component of 2500 Hz) had the level of $L_{max,fast} = 58.5dB(A)$ and the fourth test car condition (second tyre–road combination and tonal component of 350 Hz) had the level of $L_{max,fast} = 55.7dB(A)$. In the following section, the test car conditions are labelled from A to D.

6.8.2 Results of the experiment investigating the effect of the tyre and additional tonal components

Figures 6.20 and 6.23 show the responses of all participants. The cases are named from A to D for the four test car conditions. The numbers denote the amount of traffic in the background per second. In Figure 6.20 the results are shown for the background with the first tyre–road combination, and in Figure 6.23 the results are shown for the second tyre–road combination. Participants who made a false positive response (i.e. responded to hearing a car even when there was none) in more than

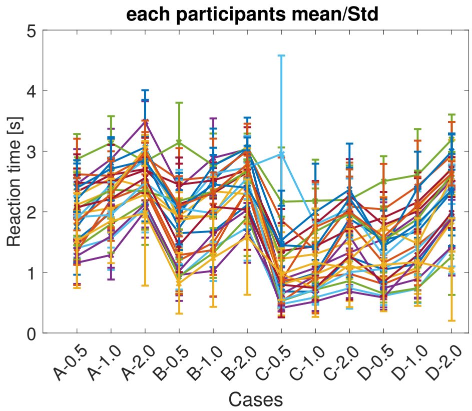


Figure 6.20: The figure shows the average reaction time and standard deviation over the cases for each of the participants. All cases are tested against the background traffic noise on the asphalt surface. A marks the cases with the asphalt surface, B marks the cases with the concrete surface, C marks the cases with the tonal component of 2500 Hz and D marks the cases with the tonal component of 315 Hz. The numbers indicate the traffic amount in the background traffic in cars per second.

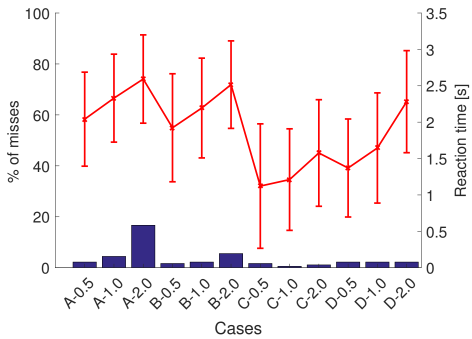


Figure 6.21: The figure shows the average reaction time and standard deviation over the cases. All cases are tested in the background traffic on the asphalt surface. The barplot shows the relative number of misses for the different background signals. The cases are labelled according to Figure 6.20.

traffic condition	$L_{eq}, dB(A)$
first tyre–road combination (asphalt), 0.5 cars/sec	56.1
first tyre–road combination (asphalt), 1.0 cars/sec	59.1
first tyre–road combination (asphalt), 2.0 cars/sec	62.1
second tyre–road combination (concrete), 0.5 cars/sec	56.9
second tyre–road combination (concrete), 1.0 cars/sec	59.9
second tyre–road combination (concrete), 2.0 cars/sec	62.9

Table 6.5: A-weighted levels of the used simulated background sound files. The mean values over the 5 conditions for the background traffic noise are given.

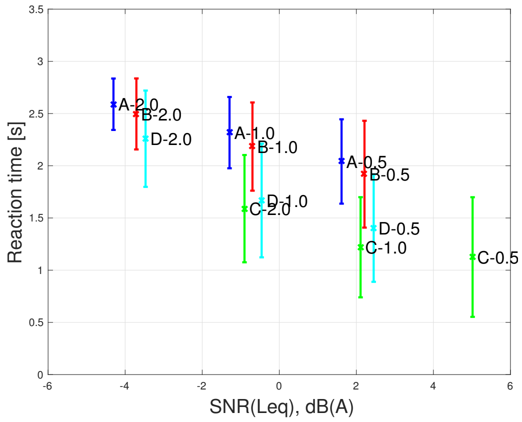


Figure 6.22: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 12 cases. All cases are tested in the background traffic on the asphalt surface. The cases are labelled according to Figure 6.20.

about 15% of the cases were removed. This resulted in 6 participants being removed from the analysis, since their responses were not seen to be consistent enough. For the remaining participants the false positive responses varied between 0 and 13.3% ($M = 5.74\%$). In total 18 participants were included in the further analysis.

Figure 6.21 shows the average response time for the different cases in the background based on the first road surface (concrete asphalt), along with the percentage of misses. A repeated measure analysis of variance (ANOVA) with the cases as main factor determined that there was a main effect for the tested cases ($F(11, 204) = 18.90$, $p < .001$).

Figure 6.24 shows the average response time for the different cases in the background based on the second road surface (treated concrete), along with the percentage of misses. A repeated measure analysis of variance (ANOVA) with the cases as main factor determined that there was a main effect for the tested cases ($F(11, 204) = 30.98$, $p < .001$).

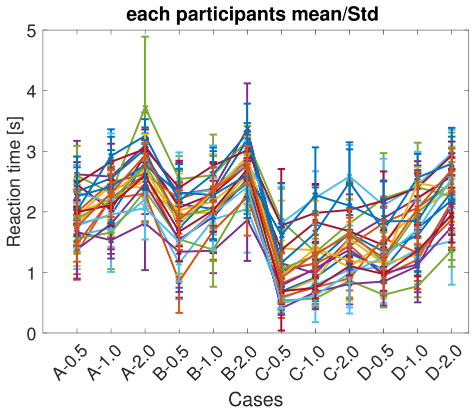


Figure 6.23: The figure shows the average reaction time and standard deviation over the cases for each of the participants in the experiment. All cases are tested in the background traffic on the concrete surface. A marks the cases with the asphalt surface, B marks the cases with the concrete surface, C marks the cases with the tonal component of 2500 Hz and D marks the cases with the tonal component of 315 Hz. The numbers indicate the traffic amount in the background traffic in cars per second.

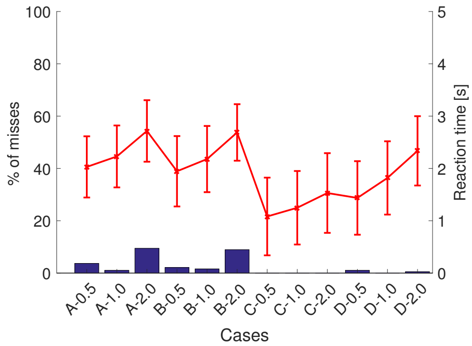


Figure 6.24: The figure shows the average reaction time and standard deviation over the cases. All cases are tested in the background traffic on the concrete surface. The barplot shows the relative number of misses for the different background signals. The cases are labelled according to Figure 6.23.

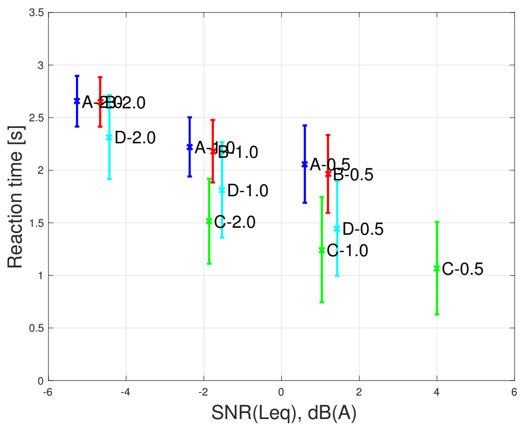


Figure 6.25: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 12 cases. All cases are tested in the background traffic on the concrete surface. The cases are labelled according to Figure 6.23.

For both backgrounds, the effect of the SNR on reaction time was investigated, to see if the SNR explains the differences between the cases. The results are shown in Figures 6.22 and 6.25 as mean values and standard deviations of the reaction time over the calculated SNR between test car and background noise.

6.8.3 Discussion of the experiment investigating the effect of the tyre and additional tonal components

This study examined the effect of different sound characteristics of a test car on the detection of the car in background traffic. Four different cases were studied in a set of background traffic conditions. The car to detect always passed at 7.5 m and all cars had an average speed of 50 km/h. Simulated sound files were used for all cars. Figure 6.20 shows the results for all participants for the background that was composed of cars with the first tyre-road combination (asphalt). Most participants show similar behaviour in reaction time. Figure 6.23 shows the results for all participants for the background that was composed of cars with the second tyre–road combination (concrete). Here too, they generally show similar behaviour.

Figures 6.21 and 6.24 show the mean reaction times and standard deviations over all participants for all tested cases. The results show the same effect for the traffic amount as previous studies; the higher the traffic amount, the higher the reaction time as well. The different road surfaces do not lead to a change in reaction time. The measured reaction times do not show significant differences between different tyre–road noises for the test car or for the background under the same traffic amount conditions (Cases A and B in Figure 6.21 and Figure 6.24). The additional tonal components clearly improve reaction time. This is especially valid for the tonal component of 2500 Hz.

Looking at the relationship between SNR and reaction time, one can see that the increase in reaction time with growing traffic amount is reflected in a decrease in SNR. However, the SNR does not show a decrease that corresponds with the changes in reaction time for the four cases. In particular the signals with the tonal components have a smaller reaction time despite having almost the same SNR as the other cases.

6.9 Combustion vs. electric engines

The aim of this part of the study was to investigate reaction time for detecting a car with different types of engine noise (test vehicle) passing by in the presence of background noise from a road with varying traffic amount. It was investigated whether detection depends only on the test car type or also on the number of similar cars in the background traffic. All parameters apart from the tested ones (i.e. road surface, speed, distance etc.) were kept constant for all sound files used. To simulate the most common risk scenarios, the background traffic was set to the speed of a normal inner city street, 50 km/h, while the test car had a lower speed of 30 km/h.

The question was how the detection time changes for an electric car (no engine noise) in relation to a car with a combustion engine. This question has also been investigated in studies by Grosse [58] and Altinsoy [11]. In this study the question was extended by also investigating the effect the proportion of electric vehicles in the background traffic has on detection time. The studies by Altinsoy and Grosse measured longer reaction times for electric cars than for combustion cars. One hypothesis is that a larger proportion of electric cars in the background traffic will lead to a lower background noise level, and thus the reaction time to a car passing by closer might decrease.

A second question that is followed up in this study is based on the results of the study in the previous chapter. Those results indicated that additional tonal components improve the detection of a car. But what happens if those tonal components are present in the background traffic as well? Thus, for the two tested tonal components from the previous study (315 Hz and 2500 Hz), the case was studied where not only the test car, but also all cars in the background traffic have the additional tonal component. The hypothesis was that the tonal component of the test car would be partly masked by the background traffic, and thus the benefit would decrease.

For the auralization, a combustion engine was chosen as one option. The alternative option was to use only the rolling noise and wind noise, to simulate a very silent engine (electric engine).

The test car had a driving speed of 30 km/h, while the background traffic had an average speed of 50 km/h. The experiment was done for two different distances (20 m and 40 m) of the background traffic in order to test once in a range where detection might be difficult (20 m) and once in a range where the test car is expected to be audible in all cases (40 m). The distances were chosen based on the study in Chapter 6.4). From the previous studies (Chapter 6.5), two traffic amounts were chosen (1 and 2 cars per second), to test the effect of the acoustic parameters of the test car in different background conditions. The limitation to two instead of three traffic amounts (as in the previous studies) was due to time constraints for the experiment.

To test the first hypothesis about the effect of electric engine versus combustion engine depending on their respective proportions in the background traffic, five test scenarios were created. The combustion engine was tested both in background traffic with 100% combustion-engine driven cars and in a background of 100% electric cars. For the electric test car, the scenarios were background traffic with 100% combustion-engine driven cars, background traffic with a 50% mix of both car types and a background of 100% electric cars.

Two scenarios were tested regarding the second hypothesis for the tonal components. For each tonal component (315 Hz and 2500 Hz), the test car and all cars in the background traffic were based on the same car with the tonal component. Only the pass-by distance was different.

6.9.1 Method for the experiment testing combustion vs. electric engines

Twenty-five participants (18 male,7 female) participated in the listening test (M = 27.4 years old, $s.d.$ = 3.9 years). All participants reported normal hearing.

To evaluate the impact of changes in the noise of the test car on the recognition of a car in a traffic noise background a set of stimuli were created as described in the auralization method. Background traffic noise was created for the chosen engines, the additional tonal components, two tested distances and the two chosen traffic amounts (Table 6.6). Five different sound files were generated for each background scenario, to increase the statistical stability of the results. A total of 80 different test conditions were created and tested, including 10 false positive tests with no test car.

Table 6.6 shows the levels of the resulting background traffic noises in the simulation without calibration. The standard deviation for the levels of the 6 repetitions was always below 0.06 dB. The first test car condition (combustion engine) had the level of $L_{max,fast} = 53.6dB(A)$ in the simulation, the second test car condition (silent engine) had the level of $L_{max,fast} = 49.7dB(A)$, the third test car condition (silent engine and tonal component of 350 Hz) had the level of $L_{max,fast} = 51.6dB(A)$ and the fourth test car condition (silent engine and tonal component of 2500 Hz) had the level of $L_{max,fast} = 53.9dB(A)$.

traffic condition	$L_{eq}, dB(A)$ at 20 m	$L_{eq}, dB(A)$ at 40 m	Label
100% combustion, 0.5 cars/sec	56.1	50.5	C.5
100% combustion, 2.0 cars/sec	62.1	56.5	C2.
50% combustion, 0.5 cars/sec	55.9	50.3	M.5
50% combustion, 2.0 cars/sec	61.9	56.3	M2.
100% electric, 0.5 cars/sec	55.7	50.1	E.5
100% electric, 2.0 cars/sec	61.7	56.1	E2.
tonal component of 350 Hz, 0.5 cars/sec	56.0	50.4	Tlo.5
tonal component of 350 Hz, 2.0 cars/sec	62.0	56.4	Tlo2.
tonal component of 2500 Hz, 0.5 cars/sec	59.1	53.2	Thi.5
tonal component of 2500 Hz, 2.0 cars/sec	65.1	59.2	Thi2.

Table 6.6: A-weighted levels of the simulated background sound files. The mean values over the 5 conditions for the background traffic noise are given.

6.9.2 Results of the experiment testing combustion vs. electric engines

Figures 6.26 and 6.29 show the responses of all participants. Participants who made a false positive response (i.e. responded to hearing a car even when there was no car present) in more than about 15% of the cases were removed. However, due to the small data set in this experiment, 15% was not a possible limit. Rounding led to a maximum of 20% for an excepted false positive test in this experiment. This resulted in three participants being removed from the analysis, since their responses were not seen to be consistent enough. For the remaining participants, the false positive responses varied between 0 and 20% ($M = 8.18\%$). A total of 22 participants were included in the further analysis.

Figure 6.27 shows the average response time for the different cases with the background traffic 20 m away, along with the percentage of misses. A repeated measure analysis of variance (ANOVA) with the cases as main factor determine that there was a main effect for the tested cases ($F(13, 294) = 32.02$, $p < .001$).

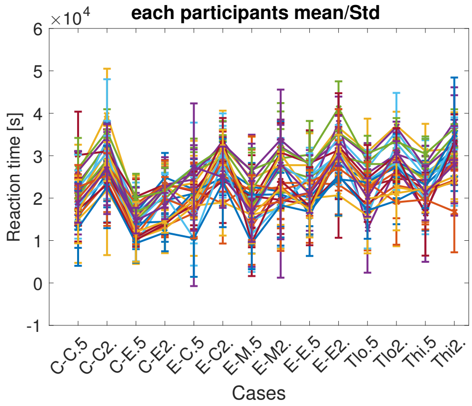


Figure 6.26: The figure shows the average reaction time and standard deviation over the cases for each of the participants in the experiment at a distance of 20 m for the background street at pass-by. The first part of the labels for the different cases defines the test car, with C for combustion engine and E for electric engine. The second part of the labelling defines the background according to Table 6.6.

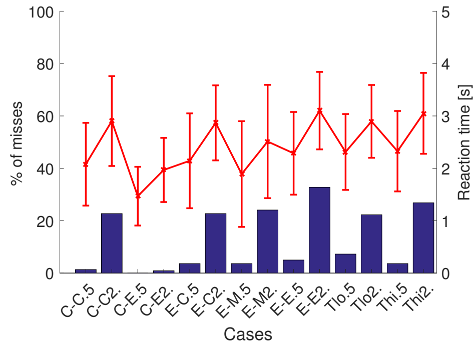


Figure 6.27: The figure shows the average reaction time and standard deviation over the cases at a distance of 20 m for the background street at pass-by. The barplot shows the relative number of misses for the different background signals. The cases are labelled according to Figure 6.26.

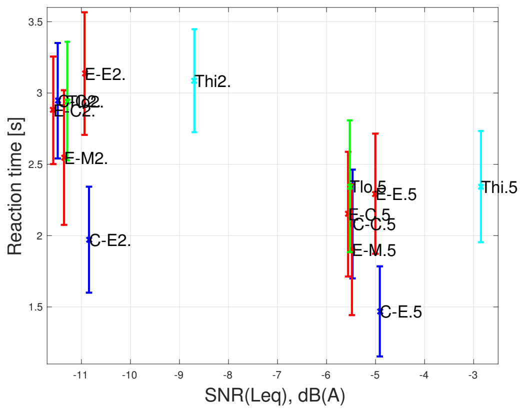


Figure 6.28: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 14 cases at a distance of 20 m for the background street at pass-by. The cases are labelled according to Figure 6.26.

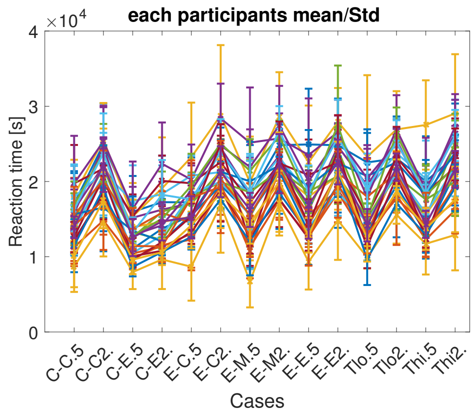


Figure 6.29: The figure shows the average reaction time and standard deviation over the cases for each of the participants in the experiment at a distance of 40 m for the background street at pass-by. The first part of the labels for the different cases defines the test car, with C for combustion engine and E for electric engine. The second part of the labelling defines the background according to Table 6.6.

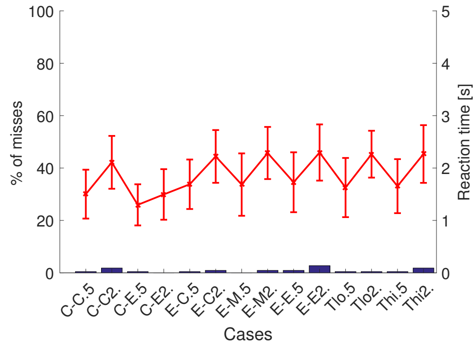


Figure 6.30: The figure shows the average reaction time and standard deviation over the cases at a distance of 40 m for the background street at pass-by. The barplot shows the relative number of misses for the different background signals. The cases are labelled according to Figure 6.29.

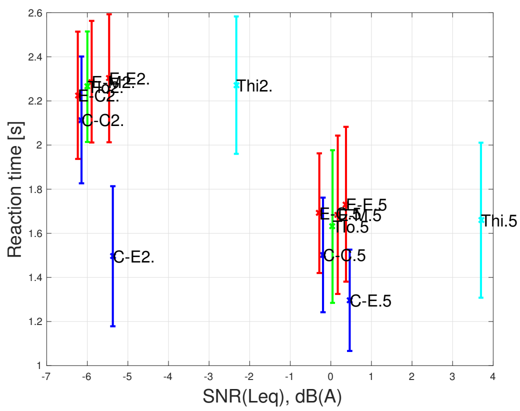


Figure 6.31: Average reaction time and standard deviations of the reaction times over the SNR between test signal and background noise for the 14 cases at a distance of 40 m for the background street at pass-by. The cases are labelled according to Figure 6.29.

Figure 6.30 shows the average response time for the different cases with the background traffic 40 m away, along with the percentage of misses. A repeated measure analysis of variance (ANOVA) with the cases as main factor determine that there was a main effect for the tested cases ($F(13, 294) = 29.37$, $p < .001$).

For both distances, the effect of the SNR on reaction time was investigated, to see if the SNR explains the differences between the cases. The results are shown in Figures 6.28 and 6.31 as mean values and standard deviations of the reaction time over the calculated SNR between test car and background noise.

6.9.3 Discussion of the experiment testing combustion vs. electric engines

This study examined the effect of different engine sounds of a test car on the detection of the car in background traffic. Four different cases were studied in a set of background traffic conditions. The car to detect always passed at 7.5 m at a speed of 30 km/h. In the background traffic, all cars had an average speed of 50 km/h. Simulated sound files were used for all cars. Figure 6.20 shows the results for all participants for background traffic at a distance of 20 m during the pass-by. Overall, the participants show similar behaviour in reaction time. Figure 6.23 shows the results for all participants for background traffic at a distance of 40 m during the pass-by. Here as well, the participants generally show similar behaviour in their reaction time.

Figures 6.21 and 6.24 show the mean reaction times and standard deviations over all participants for all tested cases. The results show the same effect for the traffic amount as previous studies: the higher the traffic amount, the higher the reaction time as well and the number of misses increases for the distance of 20 m. As expected, almost all cars were detected with background traffic at a distance of 40 m (Figure 6.24). Both distances show similar behaviour for the reaction time over all tested cases. The only case that clearly differs from all others in reaction time is the combustion engine against a background of electric cars (C-E). As expected, the reaction time is much lower than for all other cases. One unexpected result is that the detection times for all other cases do not differ much. No significant increase in reaction time or number of misses can be found in the cases with the electric car against a background of cars with combustion engines as opposed to the test car having a combustion engine. The differences are even less between the different cases at 40 m than at 20 m, even though the effect of the traffic amount can be seen for all cases.

The second question that was followed up in this experiment was how the benefit of tonal components is affected when the same tonal component is present in the background traffic as well. The results clearly indicate that all improvement created by the addition of a tonal component disappears when the same tonal component is included for all cars in the background traffic as well, for both distances and both traffic amounts. An interesting question would be how many cars in the background traffic have to have the same tonal component to destroy the beneficial effect of the tonal component on reaction time.

Looking at the relationship between SNR and reaction time, one can see that the increase in reaction time with growing traffic amount is reflected in a decrease in SNR. However, in the case of the combustion car against a background of electric cars (CE.5 and CE2.), the reaction time is much lower than the SNR would suggest. On the other hand, the SNR is much higher in the case with the tonal component of 2500 Hz (Th.5 and Th2.), but the reaction time is the same as for most other cases for each traffic amount. Those results can be seen for both tested distances.

6.10 Discussion

All studies in this chapter had at least one situation that utilized the exact same conditions. This allowed the studies to be linked together, even though they were carried out at different times and with a changing set of participants. The reaction times for the same conditions varied by a maximum of 100 ms, which falls within the range of standard deviations.

The aim of the study was to investigate the reaction time for detecting a test car passing by in the presence of background noise from a road with high traffic flow under a set of specific conditions regarding background traffic and test car. The tested conditions regarding the effects of background traffic were distance to the background traffic, traffic amount, traffic regularity and average speed of the background traffic. Of these parameters, distance, traffic amount and speed showed effects on detection time. The only parameter that did not show any significant effect on reaction time for the tested traffic amounts was variation in traffic regularity (Chapter 6.6). For the other parameters, it was also found that their effects on reaction time interact with each other (Chapter 6.5 and 6.7).

The later experiments in this study (Chapters 6.8 and 6.9) examined in greater detail the effect of the similarity between test car and background traffic, by modifying both signals.
The results of these studies confirm that additional tonal components can improve the detection time of a car by up to a second (Figure 6.21 and 6.24). However, this has to be undertaken carefully. The effect will decline as more tonal components of the same type are present in the sound environment. In the extreme case where all cars in the background have the same tonal component as the test car, the positive effect was lost completely (Figures 6.27 and 6.30).

One parameter that was investigated in all studies was the effect of the SNR on reaction time. It was found that the SNR correlates with distance and also with speed and traffic amount as individual parameters in the effect on reaction time. However, looking at the overall changes in reaction time, the SNR alone is not always a sufficient indicator without further knowledge about the test car and background signal. There are more parameters that can affect reaction time, such as differences in the amplitude envelope of the signal over time due to the pass-by, or differences in the frequency content. Those effects can occur due to differences in both distance and speed. In the presented study, these parameters cannot be clearly differentiated.

For the distance change alone, the SNR seems to be the major effect that determines reaction time (Chapter 6.4). For the changes in traffic amount at different distances, the SNR still has a strong effect and gives a rough estimate of the reaction time (Chapter 6.5). However, the slopes for the decay in reaction time with increasing SNR change with distance, decreasing as distance increases.

Regarding variations in speed and traffic amount, the SNR predicts an estimate of reaction time only within one speed (Chapter 6.7). At different speeds, reaction time seems to depend strongly on the similarity between background and test car (Figures 6.14 and 6.18).

Later experiments (Chapter 6.8 and 6.9) also showed that the similarity between test car and background traffic can have an effect of the same magnitude than the SNR. Adding the tonal component to the test car alone decreased the similarity between the signals while insignificantly changing the SNR. Reaction time dropped significantly. Adding the tonal component to both test car and the cars in the background traffic does not change the similarity; the same reaction times were measured as in the case without any tonal components.

Further, the data suggest that outgoing from identical conditions for test car and background traffic, the reaction time based on changes to the SNR will establish a lower limit for possible reaction time results. Changes to the background traffic noise that affect reaction time beyond the SNR (similarity between the signals) seem to always result in a better reaction time.

Another aspect that can be seen in these studies is that decreasing the level of background traffic noise can lead to a strong improvement in detection time independent of test car condition. In this study, a 6 dB decrease in background level (by changing the distance between the roads) to levels around 50 to 55 dB, significantly improved the detection of vehicles (Chapter 6.9). This effect was shown for both combustion cars and electric cars, and was far stronger than differences between car types. Reaction time improved by about 700 ms and the rate of misses decreased significantly to almost none (Figures 6.27 and 6.30).

6.10.1 SNR calculation

In the experiments presented in this chapter, the SNR between background noise and test car was used as a parameter in the interpretation of the result. The SNR was calculated by relating the maximum value of the L_{fast} of the test signal to the L_{eq} of the background signal. This relation was chosen because it seemed to best represent the properties of the signal. Other methods are also possible. For example, one alternative would be to relate the L_{peak} of a pass-by signal. The results of the experiment from Chapter 6.4 are plotted in Figure 6.32 over both the SNR_{leq} and the SNR_{peak} to compare both methods. The differences between the two are small in this experiment. During the course of this study, the decision was made to analyse the data based on the SNR_{leq}. Even though the results did not differ much, the SNR_{leq} seems to represent the signals better. We do not know where the peak appears in the background. It might even be outside of the time where the test car is present. The L_{eq} is a better representative of the whole signal. And for the test car, the focus on the peak might overrepresent the moment of pass-by, so an average around the pass by might be a better representation. To see if that decision had an impact on the later results, a comparison was done for all experiments. The differences between SNR_{leq} and SNR_{peak} are similar in most cases, but some stand out. The differences between the two methods can vary up to 3 dB and thus affect the interpretation of the result. One example is given in Figure 6.33. For most signals, the difference between SNR_{leq} and SNR_{peak} is similar, resulting in a shift of the results on the SNR axis. However, for some data points the SNR_{leq} is actually bigger than the SNR_{peak}. This will change some data points, but will not change the overall interpretation of the results.

Both methods are valid and lead to reasonable results. However, the SNR_{peak} bears the risk of focusing on small events in the signal. This leaves the SNR_{leq} as the preferred method for this study.

6.10.2 False positive analysis

The results presented include a false positive analysis, which means that empty runs or test runs (often called catch trials [127]) are included to measure if the participants react even if there is no stimulus. The assumption is that participants who too often report a stimulus even if there was none aren't focusing on the right cues and thus deliver less trustworthy results. The assumption is made that false alarms are strongly related to the individual participants' guessing rate [54]. This motivation is then used to remove participants from the data set before further analysis or use statistical methods to adjust the results [127].

The practice of removing data from an experiment is needed in certain cases to be able to receive correct results from statistical methods [110] [95]. This mainly ap-

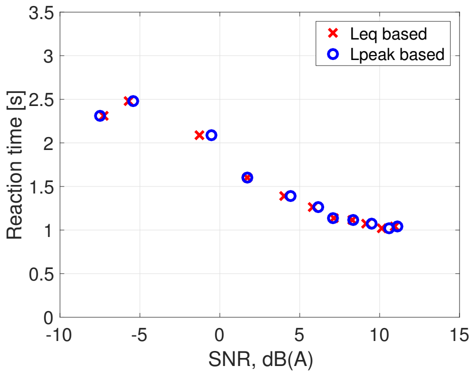


Figure 6.32: Comparison between the SNR based on L_{eq} (x) and the SNR based on L_{peak} (o) in the experiment on the effect of distance on reaction time.

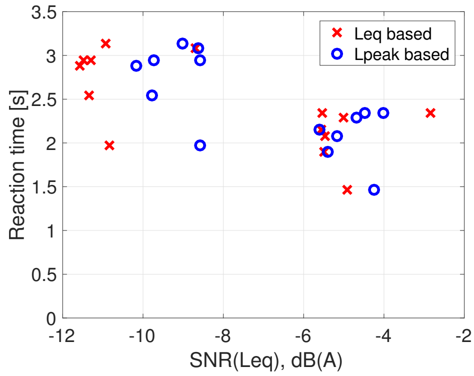


Figure 6.33: Comparison between the SNR based on L_{eq} (x) and the SNR based on L_{peak} (o) in the experiment on the effect of different engine sounds on reaction time (20m distance).

plies to outliers that can have a disproportionate effect in some statistical methods. However, the practice of categorically removing data, be it as outliers or due to false responses, must be undertaken with care [110] [95]. Each removed data point is a change in the true obtained data. In the author's point of view, this should be done if necessary to obtain the better fitting result from a certain method, but not as a general method of data management. The original data should be kept as untouched as possible, even though the results might look nicer with the treated data. However, the focus of presenting and delivering should be to give as true results as possible rather than delivering nice and easily interpretable data.

A false positive test was included in each of the presented experiments. This was done by having each background included without a test car passing by. The empty conditions were repeated five times, and were included in the randomization of the whole set of test conditions. Thus, they appeared at random times during the experiment. Participants who reacted in more than 15% of such cases, even though there was no signal, were excluded from further analysis of the results. One problem with this procedure was that some of the experiments did not have very many different background conditions to test, which limited the number of empty tests. This led to the problem that just two wrong responses put a participant over the 15% limit, making it hard to draw a reasonable line for excluding participants. Another thought is that the number of false hits is assumed to be related to a higher error rate in the detection experiment, but that relationship can also be seen as ambiguous. Certain effects might lead to a different error rate in the false positive test than in the main experiment. For example, the threshold for a participant's detecting a signal can be low for different reasons [54]. One reason can be, that participants have a tendency to react to a stimulus that is below the personal threshold of awareness. Such an effect is in an experiment impossible to differ from the participant being very sensi-

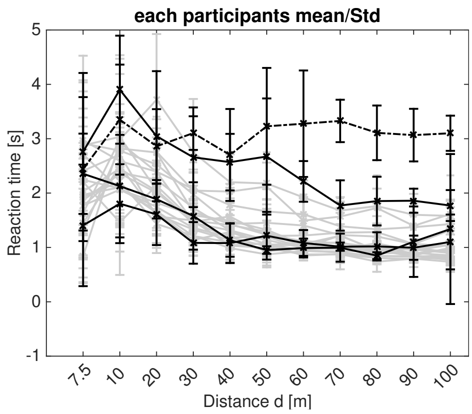


Figure 6.34: The results of the experiment on the effect of distance for all participants individually. The highlight shows participants with more than 15% hits in the false positive experiment. Those were removed in the further analysis presented in Chapter 6.4. The dotted line marks the participant who was removed for having misunderstood the task.

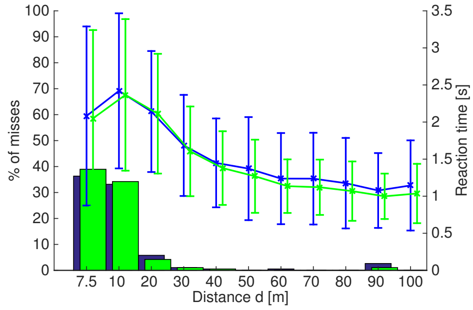


Figure 6.35: The results of the experiment on the effect of distance are compared for the case where all participants are included (dark) and for the case where the participants with more than 15% hits in the false positive experiment were removed (light).

tive. However, both affect the error rate in the false positive test in a different way than the error rate in the main experiment.

Looking at the presented experiments, the removal of participants from any given test does not show a strong effect on the results (Figures 6.35, 6.37 and 6.39). The reaction times actually increase slightly when excluding the data according to the false positive test, but this effect is very small compared to the standard deviation. The percentage of misses is lower if all participants are included than when excluding the false-positive data. However, this does not change the interpretation of the results, since it affects all datasets in a similar way. There is no change in orders or trend, only a small offset for all results.

When looking at the responses by all participants individually (Figures 6.34 and 6.36), the ones to be removed due to the false positive analysis do not show any clear similar behaviour that separates them from the rest of the responses. Only in some experiments (Figure 6.38) is there any impression that the participants with higher hits in the false positive test generally tend to have lower reaction times. In general their reaction behaviour throughout the experiments seems to be similar to the others. A fast reaction time is normally something that is appreciated in participants.

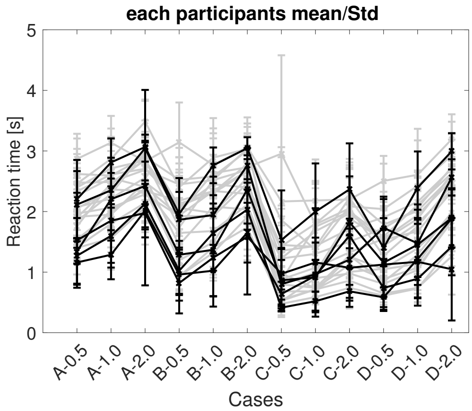


Figure 6.36: The results of the experiment on the effect of speed (test car at 50 km/h) for all participants individually. The highlight shows the participants with more than 15% hits in the false positive experiment. They were removed in the further analysis presented in Chapter 6.7.

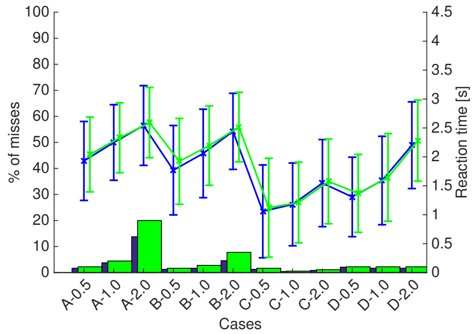


Figure 6.37: The results of the experiment on the effect of speed (test car at 50 km/h) are compared for the case with all participants included (dark) and for the case where the participants with more than 15% hits in the false positive experiment were removed (light).

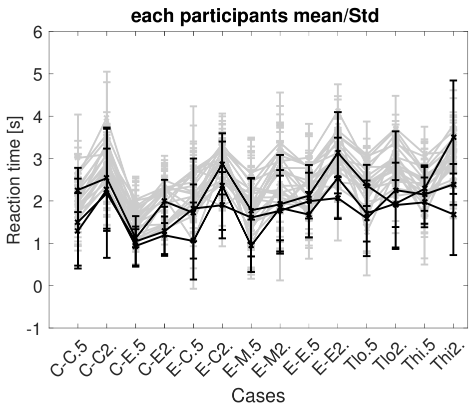


Figure 6.38: The results of the experiment on the effect of different engine sounds on reaction time (20m distance), for all participants individually. The highlight shows the participants with more than 15% hits in the false positive experiment. They were removed in the further analysis presented in Chapter 6.9.

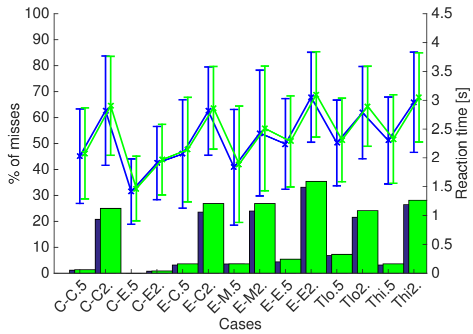


Figure 6.39: The results of the experiment of different engine sounds on reaction time (20 m distance) are compared for the case with all participants included (dark) and for the case where the participants with more than 15% hits in the false positive experiment were removed (light).

A false positive analysis was used in the presented experiments and in the paper, because it seems common practice. However, in this extended content it seemed worthwhile to present the full data and to highlight the issue. The statement by Quinn and Keough [110] that “Sometimes thinking about why particular observations are outliers can stimulate new research questions” is worth keeping in mind. Clean and perfect results do not stimulate much new thinking and ideas. New and interesting research questions might lie in the spread of results and in the differences.

Chapter 7

Conclusion

The first aim of this thesis was to create a tool that makes simulated tyre–road noise audible. In the course of the work, an auralization method was introduced, tested and further developed. The result is an auralization tool that creates pass-by sounds of passenger cars that are perceived comparably to recordings. The tool utilizes a combination of the SPERoN prediction model [79] and the auralization tool developed by Forssén [46].

In the resulting auralization tool, the basic characteristics of the signal are estimated using the SPERoN model. In the simulation process, the SPERoN model specifies tyre and road types and provides source characteristics in the form of one-third octave band spectra to the auralization model. This requires detailed information about the road surfaces (measured roughness, flow resistance), the tyres (mechanical properties and profile), driving speeds and load (i.e. weight of the vehicles). Underlying measurements are included in an extensive database [18], which was created in the Sperenberg project.

The spectrum delivered by SPERoN is then used in the auralization approach based on the Listen Demonstrator, which was developed by Forssén [46] in the Listen Project. The source term that characterizes the tyre–road noise from the reference signal in the auralization is reshaped by the one-third octave band spectrum calculated by SPERoN. To create new pass-by signals according to the source parameters defined in SPERoN, propagation effects are applied to the source signal to create the final signal at a defined receiver position. The considered propagation effects are directivity, ground reflections, air attenuation, distance effect and Doppler effect.

The auralization tool was validated with the help of listening tests, which showed that the simulated signals are perceived as being very similar to recordings under the same conditions. One limitation in the auralization was the lower frequency limit in SPERoN. SPERoN does not provide frequencies below 315 Hz. Noise in these lower frequency bands is mainly related to wind. This cannot be excluded in recordings and thus might have an effect in comparisons between recorded and simulated sounds. A method was included in the auralization tool to consider those frequencies and to improve the data transfer between the two tools. Comparisons with recordings showed that this extra consideration of the low frequencies improves the agreement between simulations and recordings.

The second aim in this thesis was to use synthesized pass-by sounds to investigate and verify the perception of rolling noise. The question was whether tyre noise can be differentiated by human perception. The listening tests and analysis showed that this is true, both for the emotional parameters of *pleasantness*, *stress* and *activation* and for the psychoacoustic parameters of *loudness*, *roughness* and *pitch*. Only in the perception of *sharpness* were the signals not rated significantly different. Fluctuation strength was dismissed as a parameter in most of the performed studies. Both calculations and the experiment indicated that the analysed signals showed no significant variance in fluctuation strength.

The performed studies further indicate that it is possible to differentiate between roads and between tyres in the perception of rolling noise. Statistical analyses to investigate the influence of tyre and road on the perception of rolling noise show significant effects of both road and tyre on most of the tested parameters. For the emotional responses, the effects of the tyres and the roads are significant, with a medium to large effect size. For the psychoacoustic parameters, the responses show significant effects for the road as an influencing parameter on loudness, roughness and pitch. For the perception of the tyre, the response differed significantly for loudness and roughness, but not for pitch. In general, the effect sizes indicate a greater difference in the perception of roads than of tyres.

The effects of the interaction between tyre and road on the perception of rolling noise cannot be answered with the performed studies. The results were partially contradictory. The statistical analysis of the interaction between the perception of the tyre and the road showed no interaction for pleasantness, stress, activation and roughness, but there was for loudness and pitch, whereas the preliminary study indicated an interaction for all tested psychoacoustic parameters.

Furthermore, interactions between the utilized psychoacoustic parameters were analysed. Pleasantness is inverse to most tested attributes, as expected. The emotional responses of pleasantness, stress and activation correlate with each other, as expected for the investigated type of sounds. Due to the fact that there is no significant difference in the ratings of sharpness, even though sharpness correlates with activation and pleasantness at the limit of 1%, it cannot be interpreted as a parameter that influences the emotional reaction to the tested signals. This leaves loudness and roughness as the main parameters that influence both pleasantness and activation. Stress is additionally affected by the pitch of the sound.

To summarize, the performed experiments clearly indicate that variations in rolling noise evoke a variety of psychoacoustic and emotional responses in a measurable range and thus confirm the hypothesis behind this work. This study further confirms that the perception of rolling noise can differentiate between effects in the signals from the road and the tyre. Thus from the viewpoint of perception of rolling noise, both tyre and road contribute to the noise experience.

To pursue the third aim in this work, the auralization tool was further developed. The auralized single pass-by sounds are now combined into the sound of a street with known traffic density, distance, direction of the cars, number of lanes and street surface. For each car the speed, tyre and engine type is known. This traffic auralization was then used to investigate reaction times for detecting a test car passing by in the presence of background noise from a road with high traffic flow. A set of conditions for the background traffic and test car were tested individually and in interaction.

Detection time was found to be affected by distance, traffic amount and speed. Traffic regularity did not show any significant effect. There were no significant differences among the tested road surfaces. For the tests with electric and combustion engines the effects on reaction time were smaller than expected. Perhaps the speed of the test car was still too high at 30 km/h, but that is the lower limit of the model for the time being. Adding additional tonal components to the test car significantly reduced reaction time. However, this effect disappears completely if the same tonal components are also added to the cars in the background.

As expected, the SNR between test car and background has a strong effect on reaction time. However this is not the sole factor governing reaction time. Another aspect that influences reaction time throughout the experiments is the similarity of the signals. This was clearest in the experiment with additional tonal components.

An assumption made on the basis of the performed experiments is that a reaction time based on changes to the SNR will establish a lower limit for the possible results in reaction time. Outgoing from identical conditions for test car and background traffic, changes to the background traffic noise that affect reaction time beyond the SNR (similarity between the signals) seem to always result in a better reaction time.

Another interesting finding from the experiments on reaction time is that decreasing the level of background traffic noise can lead to a strong improvement in detection time independent of test-car condition. This effect was far stronger than the effects of differences between car types (combustion engine/silent engine). Reaction time improved by about 700 ms and the rate of misses decreased significantly to almost none (Figure 6.27 and 6.30). A decrease of 6 dB in background level (in the experiment done by changing the distance between the roads) to levels around 50 to 55 dB significantly improved the detection of all vehicles (Chapter 6.9).

Chapter 8

Future work

In the previous chapters, a method was introduced and validated to auralize tyre–road noise. This method makes it possible to auralize car pass-by sounds under very controlled conditions. Possible parameters under control are distance, speed, type of tyre, type of engine, street surface and presence of a noise barrier. This full control of car pass-by sounds was further used to create more complex acoustic traffic situations, by combining the separate signals into a desired sound. Single auralized pass-by sounds can be combined to the sound of a street or several streets. This method gives full control of the traffic volume, traffic flow, variances in speed and types and numbers of cars. The sounds for the streets can be created for different distances from the listener and in different directions. The last step is to combine different streets into a complex sound impression.

Up to now, listening tests regarding traffic sounds were mainly based on sets of recorded signals representing specific situations. With the introduced method, acoustic impressions of different traffic situations can be produced as very controlled, reproducible and variable in all relevant parameters. This procedure was used in a set of experiments on the detection of a single car in background traffic.

These experiments could be extended further. For example, experiments were done with additional tonal components added to the sound of a car. If only the car in the foreground had those additional tonal components, detection increased significantly. However, if the tonal components were added to the cars in the background as well, the benefit in detection disappeared. This could call for further investigation. How many cars with the same tonal component in the background traffic will destroy the benefit? Would it help to use small variations in frequency? For which frequencies is there a risk of natural maskers, e.g. wind noise between buildings?

A further aspect is that the auralization tool is for the time being limited in terms of lower speeds for the cars. The lower limit at the moment is 30 km/h. For the topic of silent vehicles, auralizations for even lower speeds might be of interest, so it might be of interest to expand the tool for lower frequencies.

In the experiments regarding detection in traffic, the audio files used were limited to 6 seconds. In an updated version, the sound files for a single pass-by sound can now be made as long as desired. The level decay with distance in experiment 6.4 indicated that the streets generated out of the 6-second-long sound files behaved like point sources and not like line sources. This might be changed by using longer audio files. An interesting question would be whether and how the length of the initial pass-by sounds affects detection time.

In the presented studies on detection of a car in background traffic, a lot of other aspects were left out. Some traffic situations are better for the detection of a single car than others, thus saver. But how are they perceived in terms of annoyance and stress by listeners outside and inside a car? How applicable are the saver situations in real life?

Generating a database with a set of auralized traffic situations that can be used by a wider range of research could be of interest. It would make traffic studies more comparable by referring to recordings with fully defined traffic situations.

Appendix A: Semantic Differential experiment

Instructions and used statements

Statement	Scale
Lyssna på ljudet och avgör hur det stämmer in på påståendet nedan	
Ljudet är ”behagligt”	håller inte med / håller med
Ljudet är ”skarpt”	håller inte med / håller med
Ljudnivån är ”hög”	håller inte med / håller med
Ljudet har en ”råhet” i sig	håller inte med / håller med
Ljudet är:	mörkt (lågfrekvent)/ljust (högfrekvent)
Ljudet är ”stressande”	håller inte med / håller med
Ljudet är ”aktiverande”	håller inte med / håller med

Screenshot of the experiment

Figure 8.1: Screenshot of the experiment. Both statements and sounds were ordered at random. Each sound could be played repeatedly. The ‘next’ button was only available after responding.

Appendix B: Reaction time Experiment

Instructions

Welcome to this experiment! Press 'space' to continue.

In the present experiment you will do a reaction time task. Press 'space' to continue.

First, please state your gender. Press 'F' for female and 'M' for male. Press 'enter' when finished.

Please state your age. Press 'enter' when finished.

You will hear a background traffic sound presented from the road further away. Press 'space' to continue.

Once in a while there will be a car passing by on the road closest to you. Press 'space' to continue.

Your task is to press 'space', as quick as possible, whenever you hear the car. Press 'space' to continue.

To help you in when the car might come, a green arrow will show. Press 'space' to continue.

Press 'enter' to start the experiment.

Screenshot of the experiment

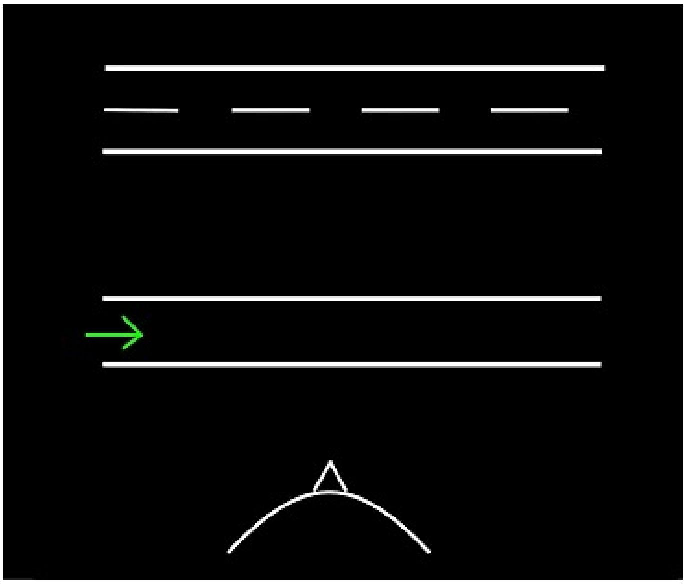


Figure 8.2: Screenshot of the experiment during the time where an additional car might pass by on the closer street. The (green) arrow indicates that there might be a car coming. If the car was not detected the word ‘miss’ was displayed on the screen. If the key was pressed at a wrong time, the word ‘error’ was displayed, and if the key was pressed while the test car sound was played, the word ‘correct’ was displayed

References

[1] ISO 10844-1994. Acoustics - specification of test tracks for the purpose of measuring noise emitted by road vehicles. Standard, International Organization for Standardization, 1994.

[2] ANSI S 1.1-1994. American national standard: Acoustical terminology. Standard, American National Standards Institute, 1994.

[3] ISO 131-1979. Acoustics - expression of physical and subjective magnitudes of sound or noise in air. Standard, International Organization for Standardization, 1979.

[4] NHTSA RIN: 2127-AK93. Federal motor vehicle safety standards; minimum sound requirements for hybrid and electric vehicles [proposal]. Legal Rule or Regulation 2797, Department of Transportation, 2013.

[5] ISO 31/VII-1978. Acoustics - quantities and units of acoustics. Standard, International Organization for Standardization, 1978.

[6] DIN 45631/A1:2010-03. Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum - Verfahren nach E. Zwicker - Änderung 1: Berechnung der Lautheit zeitvarianter Geräusche. Standard, Deutsches Institut für Normen, 2008.

[7] DIN 45681:2006-08. Berichtigung 2: Akustik - Bestimmung der Tonhaltigkeit von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschimmissionen (Entwurf). Standard, Deutsches Institut für Normen, 2006.

[8] DIN 45692:2009-08. Messtechnische Simulation der Hörempfindung Schärfe. Standard, Deutsches Institut für Normen, 2008.

[9] ISO 532-1975. Acoustics - method for calculating loudness level. Standard, International Organization for Standardization, 1975.

[10] HEAD acoustics GmbH. *Artemis 2.0.* Germany, 2000.

[11] M.E. Altinsoy. The detectability of conventional, hybrid and electric vehicle sounds by sighted, visually impaired and blind pedestrians. In *Proceedings*, volume 7, pages 5385–5389, 2013.

[12] M.E. Altinsoy and U. Jekosch. The semantic space of vehicle sounds: Developing a semantic differential with regard to customer perception. *AES: Journal of the Audio Engineering Society*, 60(1-2):13–20, 2012.

[13] K. Attenborough. The green way to reduce rail and road noise. *Acoustics Bulletin*, 38(2):14–17, 2013.

[14] W. Aures. Method for calculating auditory roughness. [Ein Berechnungsverfahren der Rauhigkeit.]. *Acta Acustica united with Acustica*, 58(5):268–281, 1985.

[15] W. Aures. Model for calculating the sensory euphony of arbitrary sounds. [Berechnungsverfahren fuer den sensorischen Wohlklang beliebiger Schallsignale.]. *Acta Acustica united with Acustica*, 59(2):130–141, 1985.

[16] L.F. Barrett. Discrete emotions or dimensions? the role of valence focus and arousal focus. *Cognition and Emotion*, 12(4):579–599, 1998.

[17] M. Basner, W. Babisch, A. Davis, M. Brink, C. Clark, S. Janssen, and S. Stansfeld. Auditory and non-auditory effects of noise on health. *The Lancet*, 383(9925):1325–1332, 2014.

[18] T. Beckenbauer, G. Blokland, and S. Huschek. Effect of pavement texture on tyre/pavement noise [Einfluss der Fahrbahntextur auf das Reifen-Fahrbahn-Geräusch]. Report FE-Nr. 03.293/1995/MRB, German Federal Ministry of Transport, August 2002.

[19] T. Beckenbauer and A. Kuijpers. Prediction of pass-by levels depending on road surface parameters by means of a hybrid model. *InterNoise 2001, The Hague, Netherlands*, 2001.

[20] T. Beckenbauer, I. Stemplinger, and A. Seiter. Basics and use of din 45681 detection of tonal components and determination of a tone adjustment for the noise assessment. In *Proceedings Inter-noise*, volume 96, page 3271, 1996.

[21] F. Bergeron, C. Astruc, A. Berry, and P. Masson. Sound quality assessment of internal automotive road noise using sensory science. *Acta Acustica united with Acustica*, 96(3):580–588, 2010.

[22] B. Berglund, U. Berglund, M. Goldstein, and T. Lindvall. Loudness (or annoyance) summation of combined community noises. *Journal of the Acoustical Society of America*, 70(6), 1981.

[23] B Berglund, T Lindvall, and D.H. Schwela. Guidelines for community noise. Report, World health organization (WHO), 1999.

[24] P. Bergman, A. Sköld, D. Västfjäll, and N. Fransson. Perceptual and emotional categorization of sound. *Journal of the Acoustical Society of America*, 126(6):3156–3167, 2009.

[25] J. Bortz. *Statistik für Human- und Sozialwissenschaftler.* Springer Verlag, 6 edition, 2005.

[26] M.M. Bradley. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1):49–59, 1994.

[27] H. Brick, M. Ochmann, and W. Kropp. *Point-source-scattering from Tyre-like Structures above An Impedance Plane, NAG/DAGA 2009, International Conference on Acoustics,Rotterdam, Netherlands*, 2009.

[28] M. Brinkmeier, U. Nackenhorst, S. Petersen, and O. von Estorff. A finite element approach for the simulation of tire rolling noise. *Journal of Sound and Vibration*, 309(1-2):20–39, 2008.

[29] Encyclopaedia Britannica. Acoustics, 2014. `http://academic.eb.com/EBchecked/topic/4044/acoustics`.

[30] S. Buss. *Subjective perception of pattern noise, a tonal component of the tyre/road noise, and its objective characterisation by spectral analysis and calculating contours.* Phd thesis, Carl von Ossietzky University Oldenburg, Germany, 2006.

[31] F. Böhm. Mechanik des Gürtelreifens. *Ingenieur-Archiv*, 35(2):82–101, 1966.

[32] R. Chocholle. Variation in auditory reaction time as a function of intensity at various frequencies [variation des temps de reaction auditifs en fonction de l'intensite a diverses frequences]. *Annee Psychologique*, 41:65–124, 1940.

[33] M. Cik, K. Fallast, and E. Marth. Free field evaluation of the influence of naturalistic road and rail traffic noise on both psychological and physiological parameters. *42nd International Congress and Exposition on Noise Control Engineering 2013, Inter-noise 2013: Noise Control for Quality of Life*, 6:4562–4568, 2013.

[34] F. Conte. *CFD modelling of the acoustic phenomenon of air pumping in a tyre/road contact patch.* Phd thesis, Institut National Des Sciences Appliquées De Lyon, 2008.

[35] P. Daniel and R. Weber. Psychoacoustical roughness: Implementation of an optimized model. *Acustica*, 83(1):113–123, 1997.

[36] T. Dau, B. Kollmeier, and A. Kohlrausch. Modeling auditory processing of amplitude modulation. i. detection and masking with narrow-band carriers. *Journal of the Acoustical Society of America*, 102(5):2892–2905, 1997.

[37] T. Dau, B. Kollmeier, and A. Kohlrausch. Modeling auditory processing of amplitude modulation. ii. spectral and temporal integration. *Journal of the Acoustical Society of America*, 102(5):2906–2919, 1997.

[38] T. Dau, D. Püschel, and A. Kohlrausch. A quantitative model of the ”effective” signal processing in the auditory system. i. model structure. *Journal of the Acoustical Society of America*, 99(6):3615–3622, 1996.

[39] T. Dau, D. Püschel, and A. Kohlrausch. A quantitative model of the ”effective” signal processing in the auditory system. ii. simulations and measurements. *Journal of the Acoustical Society of America*, 99(6):3623–3631, 1996.

[40] F.C. Donders. On the speed of mental processes. *Acta Psychologica*, 30(C):412–431, 1969.

[41] N. Döring and J. Bortz. *Forschungsmethoden und Evaluation für Sozialwissenschaftler*. Springer Verlag, 4 edition, 2006.

[42] A. Fadavi, D. Duhamel, and H.P. Yin. Tyre/road noise: Finite element modelling of tyre vibrations. *InterNoise 2001, The Hague, Netherlands*, 2001.

[43] L. Festinger. Studies in decision: I. decision-time, relative frequency of judgment and subjective confidence as related to physical stimulus difference. *Journal of Experimental Psychology*, 32:291–306, 1943.

[44] R. A. Fisher. Theory of statistical estimation. *Mathematical Proceedings of the Cambridge Philosophical Society*, 22:700–725, 7 1925.

[45] H. Fletcher and J.C. Steinberg. The dependence of the loudness of a complex sound upon the energy in the various frequency regions of the sound. *Physical Review*, 24(3):306–317, 1924.

[46] J. Forssén, T. Kaczmarek, J. Alvarssonc, P. Lunden, and M. E. Nilsson. Auralization of traffic noise within the listen project - preliminary results for passenger car pass-by. In *8th European Conference on Noise Control 2009*, volume 31 of *Euronoise 2009*, 2009.

[47] M. Fraggstedt. *Vibrations, Damping and Power Dissipation in Car Tyres*. Doctoral Thesis. Royal Institute of Technology, School of Engineering Sciences, Department of Aeronautical and Vehicle Engineering, The Marcus Wallenberg Laboratory for Sound and Vibration Research, 2008.

[48] L. Fritschi, A.L. Brown, R. Kim, D. Schwela, and . Kephalopoulos. Burden of disease from environmental noise - quantification of healthy live years lost in europe. Report, World health organization (WHO), 2011.

[49] Bundesanstalt für Straßenwesen (BAST). Homepage of the silent road traffic project (leistra3), 2014. `http://www.leistra3.de/Leistra-3/Home/home_node.html`.

[50] C. Gaertner, G. Notbohm, and S. Schwarze. Perception of sound quality of vehicle pass-by noises after technical modification. *Euronoise-5th European Conference on Noise Control*, pages 19–21, 2003.

[51] B. Gardner and K. Martin. *HRTF Measurements of a KEMAR Dummy-Head Microphone.* MIT Media Lab, Machine Listening Group, 1994.

[52] H.E. Garrett. A study of the relation of accuracy to speed. *Archives of Psychology, N.Y.*, 56:1–104, 1922.

[53] G.M. Gerken and D. Sandlin. Auditory reaction time and absolute threshold in cat. *Journal of the Acoustical Society of America*, 61(2):602–606, 1977.

[54] G.A. Gescheider. *Psychophysics - The Fundamentals.* Lawrence Erlbaum Associates, 1997.

[55] R.A.G Graf. *Tyre-road Interaction Noise.* Phd thesis, Department of Engineering, University of Cambridge, England, 2002.

[56] R.A.G. Graf, C.-Y. Kuo, A.P. Dowling, and W.R. Graham. On the horn effect of a tyre/road interface, part i: Experiment and computation. *Journal of Sound and Vibration*, 256(3):417–431, 2002.

[57] D.M. Green and J.A. Swets. *Signal detection theory and psychophysics.* John Wiley, 1966.

[58] J. Grosse, R. R. Weber, and S. Van De Par. Comparison of detection threshold measurements and modeling for approaching electric cars and conventional cars presented in traffic and pink noise. In *Proceedings of Meetings on Acoustics*, volume 19, 2013.

[59] J. F. Hamet, C. Deffayet, and M. A. Pallas. Air pumping phenomena in road cavities. In *Proceedings of the International Tire/Road Noise Conference, August 8- 10 1990, Gothenburg, Sweden*, 1990.

[60] J.F. Hamet and P. Klein. Road texture and tire noise. In *Proceedings of the 29th International Congress on Noise Control Engineering*, volume 1, pages 178–183, 2000.

[61] N. Harlander, R. Huber, and S.D. Ewert. Sound quality assessment using auditory models. *AES: Journal of the Audio Engineering Society*, 62(5):324–336, 2014.

[62] R.E. Hayden. Roadside noise from the interaction of a rolling tire with the road surface. In *Proceedings of the Purdue Noise Control Conference, Purdue University, West Lafayette*, pages 62–67, 1972.

[63] W. E. Hick. On the rate of gain of information. *Journal of Experimental Psychology*, 4(1):11–26, 1952.

[64] C. Hoever. *The simulation of car and truck tyre vibrations, rolling resistance and rolling noise*. Doktorsavhandlingar vid Chalmers tekniska högskola. Institutionen för bygg- och miljöteknik, Teknisk akustik, Vibroakustik, Chalmers tekniska högskola,, 2014.

[65] C. Hoever and W. Kropp. The simulation of truck tire rolling noise. In *44nd International Congress and Exposition on Noise Control Engineering 2015: Implementing Noise Control Technology*, 2015.

[66] R. Huber and B. Kollmeier. Pemo-q-a new method for objective audio quality assessment using a model of auditory perception. *IEEE Transactions on Audio, Speech and Language Processing*, 14(6):1902–1911, 2006.

[67] L.E. Humes and J.B. Ahlstrom. Relation between reaction time and loudness. *Journal of Speech and Hearing Research*, 27(2):306–310, 1984.

[68] H. Ising and B. Kruppa. Health effects caused by noise: Evidence in the literature from the past 25 years. *Noise and Health*, 6(22):5–13, 2004.

[69] D.M. Johnson. Confidence and speed in the two-category judgement. *Archives of Psychology, N.Y.*, 34:1–53, 1939.

[70] H. G. Jonasson. Acoustical source modelling of road vehicles. *Acta Acustica united with Acustica*, 93(2):173–184, 2007.

[71] J.J. Kalker. *Three Dimensional Elastic Bodies in Rolling Contact*. Kluwer Academic Publishers, Dordrecht, Netherlands, 1990.

[72] S. Kerber. *Wahrnehmbarkeit von Fahrzeugaußengeräuschen in Hintergrundgeräuschen: Psychoakustische Beurteilung und modellbasierte Prognosen*. Phd thesis, Technische Universität München, Germany, 2008.

[73] M. Kleiner and P. Dalenback, B. I.and Svensson. Auralization - an overview. *AES: Journal of the Audio Engineering Society*, 41(11):861–875, 1993.

[74] M. Kleiner, P. Svensson, and B. I. Dalenbaeck. Influence of auditorium reverberation on the perceived quality of electroacoustic reverberation enhancement systems. experiments in auralization. *Audio Engineering Society Preprint*, 1991.

[75] W. Kropp. Structure-borne sound on a smooth tyre. *Applied Acoustics*, 26(3):181–192, 1989.

[76] W. Kropp. A model for the description of the rolling noise from a smooth tyre on a rough road, ph.d. thesis [Ein Modell zur Beschreibung de Rollgeräusches eines unprofilierten Guertelreifens auf rauher Strassenoberfläche, Doktorarbeit]. Report 166, VDI, Berlin, 1992.

[77] W. Kropp. Mathematical model of tyre noise generation. *Heavy Vehicle Systems*, 6(1):310–329, 1999.

[78] W. Kropp, F.-X. Bécot, and S. Barrelet. On the sound radiation from tyres. *Acta Acustica united with Acustica*, 86(5):769–779, 2000.

[79] W. Kropp, K. Larsson, F. Wullens, P. Andersson, and F.X. Bécot. The generation of tyre/road noise - mechanisms and models. *Proceedings of the Tenth International Congress on Sound and Vibration*, pages 4289–4301, 2003.

[80] W. Kropp, P. Sabiniarz, H. Brick, and T. Beckenbauer. Sound radiation of a rolling tyre. *Proceedings of Forum Acusticum*, pages 795–800, 2011.

[81] W. Kropp, P. Sabiniarz, H. Brick, and T Beckenbauer. On the sound radiation of a rolling tyre. *Journal of Sound and Vibration*, 331(8):1789–1805, 2012.

[82] W. Kropp, J. Winroth, C. Hoever, and T. Beckenbauer. Sound generation and sound radiation from tyres. *42nd International Congress and Exposition on Noise Control Engineering 2013, Inter-noise 2013: Noise Control for Quality of Life*, 1:321–327, 2013.

[83] A. Kuijpers, B. Peeters, T. Beckenbauer, and W. Kropp. Speron and acoustic optimization tool: user manual for version 3.1, 2009. `http://www.speron.net`.

[84] L. E. Kung, W. Soedel, and T. Y. Yang. Free vibration of a pneumatic tire-wheel unit using a ring on an elastic foundation and a finite element model. *Journal of Sound and Vibration*, 107(2):181–194, 1986.

[85] K. Larsson and W. Kropp. A high-frequency three-dimensional tyre model based on two coupled elastic layers. *Journal of Sound and Vibration*, 253(4):889–908, 2003.

[86] R. Lazarus and S Folkman. *Stress, appraisal, and coping*. Springer, New York, 1984.

[87] S. M. Lee and S. K. Lee. Objective evaluation of human perception of automotive sound based on physiological signal of human brain. *International Journal of Automotive Technology*, 15(2):273–282, 2014.

[88] I. Lopez, R.E.A. Blom, N.B. Roozen, and H. Nijmeijer. Modelling vibrations on deformed rolling tyres-a modal approach. *Journal of Sound and Vibration*, 307(3-5):481–494, 2007.

[89] I. Lopez Arteaga. Green's functions for a loaded rolling tyre. *International Journal of Solids and Structures*, 48(25-26):3462–3470, 2011.

[90] P. Lundén, M. Gustin, M.E. Nilsson, J. Forssén, and B. Hellström. Psychoacoustic evaluation as a tool for optimization in the development of an urban soundscape simulator. *Proceedings of the 5th Audio Mostly - A Conference on Interaction With Sound, AM '10*, 2010.

[91] B.C.J. Moore. *An introduction to the psychology of hearing*, volume 2nd edition. Academic Press, London, 2000.

[92] J.M. Muggleton, B.R. Mace, and M.J. Brennan. Vibrational response prediction of a pneumatic tyre using an orthotropic two-plate wave model. *Journal of Sound and Vibration*, 264(4):929–950, 2003.

[93] P.M. Nelson and S. M. Phillips. Quieter road surfaces, ifm research report 6.084.02. Report, TRL Annual Review, 1997.

[94] C.M. Nilsson. *Waveguide finite elements applied on a car tyre*. Doctoral Thesis. Royal Institute of Technology, School of Engineering Sciences, Department of Aeronautical and Vehicle Engineering, 2004.

[95] Nordtest. Nt acou 111, acoustics: Human sound perception - guidelines for listening tests, 2002-05.

[96] R. Nota, R. Barelds, and D. van Maercke. Engineering method for road traffic and railway noise after validation and fine tuning. Report, Harmonoise Project, 2005.

[97] D.J. O'Boy and A.P. Dowling. Tyre/road interaction noise-a 3d viscoelastic multilayer model of a tyre belt. *Journal of Sound and Vibration*, 322(4-5):829–850, 2009.

[98] Proceedings of International Tire Noise Conference. Stockholm. 1979.

[99] Congress of the United States of America. Pedestrian savety enhancement act of 2010. Pub. L. 111-373, Jan. 4, 2011, 124 Stat. 4086 (49 U.S.C. 30111), 2010.

[100] C.E. Osgood, G.J. Suci, and P.H. Tannenbaum. *Measurement of meaning.* Osgood, C.E., 1957.

[101] H.B. Pacejka. The tire as a vehicle component. mechanics of pneumatic tyres. *Monograph 122. National Bureau of Standards, Washington, DC*, 1971.

[102] J. H. Park and S. K. Lee. Identification of vehicle booming sound and its objective evaluation using psychoacoustic parameters. *International Journal of Vehicle Design*, 58(1):46–61, 2012.

[103] C Pendharkar. *Auralization of road vehicles using spectral modeling synthesis.* Master thesis, Chalmers University of Technology, Göteborg, Sweden, 2012.

[104] B. E. Pfingst, R. Hienz, J. Kimm, and J. Miller. Reaction time procedure for measurement of hearing. i. suprathreshold functions. *Journal of the Acoustical Society of America*, 57(2):421–430, 1975.

[105] B. E. Pfingst, R. Hienz, and J. Miller. Reaction time procedure for measurement of hearing. ii. threshold functions. *Journal of the Acoustical Society of America*, 57(2):431–436, 1975.

[106] A. Pietrzyk. Prediction of the dynamic response of a tyre. *InterNoise 2001, The Hague, Netherlands*, 2001.

[107] B. Plovsing and J. Kragh. Nord2000 - comprehensive outdoor sound propagation model. part 1: Propergation in atmosphere without significant refraction. Report, Delte Acoustics and Vibration, 2006.

[108] Listen Project, 2013. `https://www.tii.se/projects/listen`.

[109] C. Pörschmann and C. Störig. Investigations into the velocity and distance perception of moving sound sources. *Acta Acustica united with Acustica*, 95(4):696–706, 2009.

[110] G.P. Quinn and M.J. Keough. *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, 2002.

[111] H. Refaat. Incidence of pedestrian and bicyclist crashes by hybrid electric passenger cars. Report, National Highway Traffic Safety Administration, (NHTSA), 2009.

[112] T.L. Richards. Finite element analysis of structural-acoustic coupling in tyres. *Journal of Sound and Vibration*, 149(2):235–243, 1991.

[113] D. Ronneberger. Experimentelle und theoretische Untersuchung spezieller Mechanismen der Rollgeräuschentstehung und Abstrahlung. Technical report, Institut für Straßen-, Eisenbahn-, und Felsbau der E.T.H. Zürich, 1984.

[114] D. Ronneberger and J. Schewe. Schallabstrahlung von lärmgeminderten Versuchsreifen und von Serienreifen im Vergleich. *Paper within BMFT- Vorhaben TV 8225, Drittes Physikalisches Institut, University of Göttingen, Göttingen, Germany*, 1985.

[115] J.A. Russell. A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6):1161–1178, 1980.

[116] J.A. Russell. Core affect and the psychological construction of emotion. *Psychological Review*, 110(1):145–172, 2003.

[117] J.A. Russell and L.F. Barrett. Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant. *Journal of Personality and Social Psychology*, 76(5):805–819, 1999.

[118] E.A Rustighi, S.J.A Elliott, S.B Finnveden, K.C Gulyás, T.C Mócsai, and M.D Danti. Linear stochastic evaluation of tyre vibration due to tyre/road excitation. *Journal of Sound and Vibration*, 310(4-5):1112–1127, 2008.

[119] P. Sabiniarz and W. Kropp. A waveguide finite element aided analysis of the wave field on a stationary tyre, not in contact with the ground. *Journal of Sound and Vibration*, 329(15):3041–3064, 2010.

[120] U. Sandberg. Design and maintenance of low noise road surfacings. In *International Symposium on pavement surface characteristics, 1996, Christchurch, New Zealand*, volume 1, pages 335–350, 1996.

[121] U. Sandberg and J Ejsmont. *Tyre / road noise reference book*. Informex Handelsbolag, Kisa, Sweden, 2002.

[122] K. Schaff and D. Ronnenberger. Noise radiation from rolling tyres - sound amplification by the "horn effect". *InterNoise 1982, San Francisco, California*, 1982.

[123] A. Seiter, I. Stemplinger, and T. Beckenbauer. Untersuchung zur Tonhaltigkeit von Geräuschen. In *Fortschritte der Akustik, DAGA 96*. Dt. Gesell. für Akustik e. V., Oldenburg, 1996.

[124] R. Sottek. Gehörgerechte Rauhugkeitsberechnung. In *Fortschritte der Akustik, DAGA 94, Dresden*, pages 1197–1200. Dt. Gesell. für Akustik e. V., Oldenburg, 1994.

[125] E. Terhardt. On acoustical roughness and fluctuation strength [Über akustische Rauhigkeit und Schwankungsstärke]. *Acta Acustica united with Acustica*, 20(4)(215-224):37–38, 1968.

[126] L.L. Thurstone. The method of paired comparisons for social values. *Journal of Abnormal and Social Psychology*, 21(4):384–400, 1927.

[127] H. Tiefenau, A.and Neubauer, H. von Specht, and P. Heil. Correcting for false alarms in a simple reaction time task. *Brain Research*, 1122(1):99–115, 2006.

[128] G. Van Blokland and P. The. Test sections for development of a hybrid tyre/road interaction noise model. *Turkish Acoustical Society - 36th International Congress and Exhibition on Noise Control Engineering, Inter-noise 2007 Istanbul*, 1:338–345, 2007.

[129] F. Van Der Eerden, F. Graafland, P. Wessels, and T. Basten. Urban traffic noise assessment by combining measurement and model results. In *21st International Congress on Acoustics, ICA 2013 - 165th Meeting of the Acoustical Society of America*, volume 19, 2013.

[130] W. van Keulen and M. Duskov. Inventory study of basic knowledge on tyre / road noise. Report, Poject report DWW-2005-022, IPG, 2005.

[131] H. Van Leeuwen and R. Nota. The harmonoise engineering model. *Acta Acustica united with Acustica*, 89(SUPP.), 2003.

[132] W.J. Van Vliet and G. Van Blokland. Noise reduction of silent tyres on different road surfaces. *42nd International Congress and Exposition on Noise Control Engineering 2013, Inter-noise 2013: Noise Control for Quality of Life*, 1:200–208, 2013.

[133] G. von Bismarck. Sharpness as an attribute of the timbre of steady sounds. *Acta Acustica united with Acustica*, 30(3):159–172, 1974.

[134] O. Von Estorff. Numerical prediction of noise sources: Facts, fears, future. volume 1, pages 56–64, 2013.

[135] M. Vorländer. *Auralization; Fundamentals of Acoustics, Modelling, Simulation, Algorithms and Acoustic Virtual Reality*. RWTHedition. Springer, 2008.

[136] D. Västfjäll, M. A. Gulbol, M. Kleiner, and T. Gärling. Affective evaluations of and reactions to exterior and interior vehicle auditory quality. *Journal of Sound and Vibration*, 255(3):501–518, 2002.

[137] G. Watts. Harmonoise models for predicting road traffic noise. *Acoustics Bulletin*, 30(5):19–21+23–25, 2005.

[138] C.D. Wickens. *Engineering Psychology and Human Performance*. Charles E. Merrill Publishing Company, Columbus, 1984.

[139] Y.H. Wijnant and A. de Boer. A new approach to model tyre/road contact. In *Proceedings of ISMA 2006*, pages 4453–4462, 2006.

[140] J. Winroth, C. Hoever, and W. Kropp. SPERoN 2020 Teil 1: Erweiterung des Rechenmodells zur akustischen Optimierung von Fahrbahnbelägen (Entwurf). Report FE-Nr. 86.0078/2010, German Federal Ministry of Transport, August 2014.

[141] J. Winroth, C. Hoever, W. Kropp, and T. Beckenbauer. The contribution of air-pumping to tyre/road noise. In *Proceedings of AIA-DAGA 2013*, pages 1594–1597, 2013.

[142] F. Wullens and W. Kropp. A three-dimensional contact model for tyre/road interaction in rolling conditions. *Acta Acustica united with Acustica*, 90(4):702–711, 2004.

[143] J. You and J. Jeon. Just noticeable differences in sound quality metrics for refrigerator noise. *Noise Control Engineering Journal*, 56(6):414–424, 2008.

[144] E. Zwicker and H. Fastl. *Psychoacoustics - Facts and Models, Second Edition.* Springer, Germany, 1999.

[145] E. Öhrström, A. Skånberg, H. Svensson, and A. Gidlöf-Gunnarsson. Effects of road traffic noise and the benefit of access to quietness. *Journal of Sound and Vibration*, 295(1-2):40–59, 2006.