NTNU
Norwegian University of
Science and Technology

Psychoacoustic evaluation of noise from metro trains

A case study of the Oslo metros

Andris Broks

Master's thesis in Electronics Systems Design and Innovation
June 2019

Supervisor 1: Guillaume Dutilleux
Supervisor 2: Alice Hoffmann

Department of Electronic Systems
Norwegian University of Science and Technology

Preface

The following assignment is a Master's thesis, which is a final work for the study program 'Electronic Systems Design and Innovation' in the Norwegian University of Science and Technology. The topic for this Master's thesis was proposed by Brekke & Strand Akustikk AS, which is the company that works within the fields of acoustics, noise and vibrations. Thus, the current work is done in a collaboration with Brekke & Strand.

Mainly, the work was done in the city of Oslo, Norway, which includes both the measurement campaigns and listening tests inside the sound proof room at the Oslo's office of Brekke & Strand. Most of the meetings, consultations and writing also took place at the Brekke & Strand office. However, some theoretical preparations and consultations took place at the NTNU campus Gløshaugen in the city of Trondheim, Norway. Also one meeting regarding the post-processing of the obtained results took place in the city of Gothenburg, Sweden. Since the NTNU supervisor Guillaume Dutilleux and supervisor from Brekke & Strand Alice Hoffmann were not located in Oslo, skype meetings and emails were used to share the information and knowledge concerning the thesis work.

I would like to thank my NTNU supervisor Guillaume Dutilleux for sharing his knowledge during the meetings we had, all the helpful words of advice and comments that he gave. I am very thankful to Alice Hoffmann, my main supervisor from Brekke & Strand, who introduced me to the field of psychoacoustics and explained a great deal of concepts and terms that were completely new for me. Her Ph.D work was extremely helpful and served as a huge source of inspiration with many hints for my own thesis. It was a big pleasure to work together with Alice. I would also like to thank my second Brekke & Strand supervisor Sigmund Olafsen for his passion for the research within the field of acoustics, his support during the practical part of work and fast answers to any sorts of questions that came to my mind. I'd like to thank all the Brekke & Strand employees who helped me during this assignment both with the preparations for the measurement campaigns and participation in the listening tests. I am especially thankful to Teresa Fernandez Espejo for her help with the practical part of work. She is a great human who is always ready to give a piece of advice and share her knowledge despite her big workload. I would also like to thank Malene Monslaup, a master student at NTNU, whom I worked with during the practical part of this work. She was extremely helpful during the whole assignment, was always positive about sharing her thoughts and assumptions and it was a big pleasure to work with her throughout the spring semester. I was lucky enough to have around some very smart and friendly students who were all very kind and supportive during my stay in Oslo, so I'd also like to thank Per Christian Olafsson, Therese Öqvist and Gry Bernakiewicz. At last, I'd like to thank my family members and friends for being the biggest source of love, understanding, support and motivation during my studies in Norway.

Oslo, June 16, 2019
Andris Broks

Abstract

The classical approaches used to analyze metro train noise work nicely in most cases, however, sometimes they do not explain why a particular noise source is experienced as very bothersome. The new technological era also brings some new challenges since the cities are rapidly growing and the technological changes lead to the changes in noise that is being produced. Thus, there is a demand for the new ways to look at the old problems from different angle. Psychoacoustical assessment of noise could broaden our understanding of how humans react to noise sources nowadays. This work is a continuation of the Alice Hoffmann's Ph.D thesis on the topic of psychoacoustical assessment of the tyre-road noise [1], this time applied to the metro train noise.

The following paper presents the study of psychoacoustical evaluation of metro train noise in the Oslo city, Norway. The main purpose of this work was to reveal which psychoacoustical and emotional parameters can be used in the assessment of the metro train noise. The data collected from the four different measurement sites in Oslo was processed with a help of software calculations and listening tests in a controlled lab environment. The statistical analysis showed that there is a good agreement between the theoretical software computations and listening test results for the parameters sharpness and loudness. A deeper analysis using one-way ANOVA test proved that sharpness and loudness are the most suitable psychoacoustical parameters to describe the metro train noise. Roughness and fluctuation strength showed rather poor results and it seems that theoretical models proposed for these parameters do not agree with the human perception that was estimated using the set of listening tests. The emotional parameters, stress and pleasantness, can also be valuable indicators of the metro train noise perception. Statistical analysis showed that some of the signals have similar results despite the parameter being used, and thus the nature of the signals can be crucial and should be considered when analyzing metro train noise.

Sammendrag

De klassiske tilnærmingene som blir brukt til å analysere T-bane støy fungerer bra i de fleste tilfellene, men noen ganger forklarer de ikke hvorfor en bestemt støykilde oppleves som svært plagsom. Den nye teknologiske tiden bringer også noen nye utfordringer ettersom byene vokser raskt, og de teknologiske endringene fører til endringer i støy som blir produsert. Dermed finnes det en etterspørsel etter de nye måtene å se på de gamle problemene fra forskjellig vinkel. Psykoakustisk vurdering av støy kan utvide vår forståelse av hvordan mennesker reagerer på støykilder nå til dags. Dette arbeidet er en fortsettelse av Alice Hoffmann's doktorgradsavhandling om psykoakustisk vurdering av trafikkstøyen [1], denne gangen anvendt på T-bane støy.

Følgende oppgave beskriver studiet av psykoakustisk vurdering av T-bane støy i Oslobyen, Norge. Hovedformålet med dette arbeidet var å avdekke hvilke psykoakustiske og emosjonelle parametre som kan brukes ved vurderingen av T-bane støy. Dataene som ble samlet inn fra de fire forskjellige målestedene i Oslo, ble behandlet ved hjelp av programvareberegninger og lyttetester i et kontrollert laboratoriemiljø. Den statistiske analysen viste at de teoretiske programvareberegningene og lyttingstestresultatene korrelerer ganske godt for parametrene skarphet og lydstyrke. En dypere analyse ved hjelp av enveis ANOVA-test viste at skarphet og lydstyrke er de beste blant alle psykoakustiske parametrene under vurdering for å beskrive støy fra T-banen. Ruhet og fluktueringsstyrke viste ganske dårlige resultater, og det ser ut at teoretiske modeller foreslått for disse parametrene ikke stemmer overens med den menneskelige oppfatningen som ble estimert ved hjelp av lyttetester. De emosjonelle parametrene, stress og behagelighet, kan også være verdifulle indikatorer for lydopplevelse av støy fra T-banen. Statistisk analyse viste at noen av signalene har lignende resultater uansett hvilken parameteren blir brukt, og dermed signalenes natur er avgjørende og bør vurderes når T-bane støy blir analysert.

Contents

List of Tables

List of Figures

Abbreviations

AM	=	Amplitude Modulation
ANOVA	=	Analysis of variance
ATP	=	Automatic train protection
EEA	=	European Environment Agency
FM	=	Frequency Modulation
IQR	=	Interquartile range
JND	=	Just-Noticable Differences
NTNU	=	Norwegian University of Science and Technology
Ph.D	=	Philosophiae doctor (Latin)
SPL	=	Sound Pressure Level

Chapter 1

Introduction

There is an ongoing discussion regarding the most suitable transportation means for the city of Oslo. The capital city of Norway is rapidly growing and therefore there occurs a demand for the most efficient and economically profitable way of transportation. Oslo metro lines, while currently being the most rapid transit system, face some problems already in the present times because of the lack of capacity to carry enough passengers and therefore extension of metro lines is one of the obvious solutions to the existing problem [2]. However, one of the highly discussed and debated topics is noise produced by metro trains. Thus, it is necessary to consider how people perceive noise from Oslo metro trains and take their experiences into account along with the classical noise measures.

Even though psychoacoustics has its roots in early science when Pythagoras had studied musical consonance and dissonance with his monochord, it is a new field and many new studies in modern psychoacoustics exist, which are similar to those of the old days but can be realized with a help of new technologies making use of digitalization [3]. One of such recent studies is work on psychoacoustic assessment of railway noise in sensitive areas and times by Lercher who focuses on the effects of railway noise on annoyance, making use of psychoacoustical parameters loudness and sharpness [4]. He found out that psychoacoustical analysis should be applied when considering health effects and annoyance caused by railway noise. Kasess states in his article on relationship between psychoacoustical factors and annoyance for railway noise that 'loudness seems to be a better descriptor of railway noise than A-weighted SPL' [5]. Patsouras presented the study on the psychoacoustical evaluation of tonal component's effects on high-speed train interior noise proving that tonalness has a strong effect on perceived sound quality [6]. These examples clearly show that there is a need to assess psychoacoustical parameters when analyzing effects of railway noise, but they mostly focus on the one or a couple of specific psychoacoustical parameters. The current work in turn aims at the assessment of railway noise based on various psychoacoustical and emotional parameters covering much wider range of possible psychoacoustical effects.

This work will entirely focus on metro trains in the city of Oslo since this location is not properly characterized in terms of the alternative ways, like psychoacoustical approach, that can be used to describe train noise perception. The goal of the assignment is to investigate which psychoacoustical parameters can be used to describe the perception of noise from Oslo metro trains and improve our understanding of the railway noise. Work will be carried with a reference to the fundamental study of psycho-acoustic parameters and models described in a book 'Psychoacoustics' by Fastl and Zwicker [7].

The current thesis is made in collaboration with Brekke & Strand akustikk. The approach that will be used in the project work is similar to the one Alice Hoffmann used in her Ph.D thesis that was applied to the perception of noise from road traffic. The intention is to transfer her knowledge from car noise to metro trains. On the whole, this master assignment is a part of a bigger research work run by Brekke & Strand akustikk concerning noise and vibrations from metro trains and trams called Metronova. Yet another NTNU master student, Malene Monslaup, also took part in the Metronova project and even though the two assignments consist of different steps and have different purposes, some overlaps should be clarified. While the following work focuses entirely on the psychoacoustical parameters, Malene's work is about the comparison between the most commonly used parameter for the metro noise evaluation, i.e. the A-weighted sound pressure level and the objective psychoacoustical parameters like loudness, sharpness, roughness and so on. The raw data obtained from the measurement campaigns is common for both works since measurements were done together by both students. However, the listening tests and the analysis of obtained results was done separately for each student as the final goals for the two thesis works are different. Since both students were allocated an office place in the Oslo's office of Brekke & Strand akustikk, many mutual discussions took place to share some knowledge and ideas about the master's work.

The practical part of the following work will consist of two major parts: outside measurements of noise from metro trains in 4 different positions along the Oslo metro lines and psychoacoustical listening tests in a controlled laboratory setting, which will make use of the measurements. The data obtained from the noise measurements will be used for further analysis based on theoretical models for different psychoacoustical parameters with a help of ArtemiS software. Results from ArtemiS and listening tests will be analyzed both separately and also compared against each other to test if the measured results correlate with perception of train noise by real humans. Thus, the set of psychoacoustical parameters, which can be used to rate train noise will be revealed.

It is neither expected nor required that a reader has a deep understanding of acoustics and psychoacoustics. However, it is expected that a reader has some basic knowledge in mathematics and statistics corresponding to the level of secondary school student or the first year university student. The necessary theoretical background will be explained in details in Chapter 2. Some information regarding the Oslo metro trains, the practical aspects of the measurement procedure as well as the detailed explanation of listening tests and the list of equipment will be given in Chapter 3. Chapter 4 will present the analysis of obtained results. The post processing of the results will be done using Matlab and ArtemiS software. At the end, some important aspects will be highlighted and discussed in Chapter 5 and conclusions will be made in Chapter 6 based on the observations, calculations and discussion.

Chapter 2

Basic Theory

The following chapter describes the theoretical background for understanding, interpreting and calculating the desired values. The formulas for calculation of different parameters will be given as well as the way to apply certain corrections if needed.

2.1 Railway noise

This section is aimed at brief explanation of railway noise that is dealt with in this thesis, its main sources, effects and exposure.

Railway noise is an environmental noise, meaning that it is harmful for the environment, i.e. humans and other creatures [8]. It is defined as 'the noise created by the operation of rail-bound vehicles' [9]. According to Vos, the quality of life might be affected by the environmental noise [10]. The most common effect caused by this kind of noise is annoyance, i.e. 'a general feeling of disturbance or discomfort, usually occurring after a long period of continuous or repeated exposure' [10]. The following study deals with the problem of annoyance caused by the railway noise, which can possibly be resolved with a help of psychoacoustics. Moreover, some more serious health problems in a form of diseases like insomnia, cardiovascular diseases or even mortality might occur due to longer and more frequent railway noise exposure [10].

EEA reports that 'railways are the second most dominant source of environmental noise in Europe, with nearly 7 million people exposed to levels above 55 dB L_{den} (Average Noise Level Index total day) in 2012 considering people exposed both inside and outside urban areas, as reported in August 2013'. Moreover, according to the same report, estimation models double the mentioned numbers. It means that huge numbers of people across Europe are affected by railway noise and it should be considered as one of the most significant environmental noise sources [11].

Sources of noise for railway vehicles are highly dependent on the speed of the vehicle. At low speeds, the so-called traction noise or engine noise is dominant [9]. At medium speeds, i.e. between 30 and 200 km/h rolling noise is the dominant noise source, which is also considered the main noise source, which affects most people living near railway track [12]. It occurs due to the contact between the rails and the wheels. At very high train speeds, the aerodynamic noise becomes dominant, however, it becomes vital only for high speed vehicles [12]. The other important railway noise sources include squeal noise in curves, noise caused by braking and accelerating, vibration from rail corrugation and out-of-round wheels, vehicle coupling in shunting yards as well as signalling noise [9] [12].

2.2 Sound pressure level

Human hearing is a unique process that is guided by the auditory system. It is a system integrated in our head that converts acoustical sound waves into certain patterns of neural activity in our brain by the use of array of miniature acoustical detectors. This system is extremely sensitive to sound, that is to vibrations of air molecules in the air that create pressure waves. Even slight displacements of the air molecules can be detected by our ears. Thus, it is the sound pressure changes that we actually are looking for during the noise measurements [13].

Although sound pressure unit is Pascal, it is common in acoustics to work with Decibel scale, so that the sound pressure level L_p can be expressed with formula (2.1) in Decibels. The rms value p_{rms} is the sound pressure in $[Pa]$, whereas p_0 is the reference pressure value in airborne acoustics, i.e. $p_0 = 20\mu Pa$. Results for L_p will be presented along with the psychoacoustical values.

$$L_p = 10\log_{10}\left(\frac{p_{rms}}{p_0}\right)^2 \tag{2.1}$$

The mean value of the square of $p_{real}(t)$ is given by the equation (2.2).

$$p_{rms}^2 = \frac{1}{T}\int_0^T p_{real}^2(t)dt \tag{2.2}$$

2.2.1 The average sound pressure level

In order to compute a sound level as a single value parameter, the average sound pressure level L_{avg} in $[dB]$ should be determined using formula (2.3). The total number of measurement points is denoted by n and L_j is the sound level value in a particular point.

$$L_{avg} = 10\log_{10}(\frac{1}{n}\sum_{j=1}^{n} 10^{0.1L_j}) \tag{2.3}$$

2.3 Background noise

When the outdoor measurements take place, it should be thoroughly checked whether the background noise does not affect and eventually ruin the obtained measurements.

According to the $ISO3095 : 2013(E)$ standard [14], the maximum value of the background sound pressure level $L_{pAeq,T}$ for the time interval $T = 20s$ for all the microphone positions should be at least $10dB$ below the combined level of noise from the train unit and the background noise L_{pAeq,T_p} during the pass-by time T_p. In case of frequency analysis the same rule is applied to each frequency band. This rule is also applied to separate the pass-by, also the noise from the train, from the surrounding background noise.

2.4 Presentation of sounds by loudspeakers and headphones

Distortion factor of only 0.1% is allowed for psychoacoustical applications corresponding to a level difference of $60dB$. In case if loudspeaker sounds are tested in a non-anechoic room, the frequency characteristics of the room is superimposed on the characteristics of the loudspeakers. However, these problems can be solved using headphones instead. Headphones show very little nonlinear distortion (less than 0.1%). Also should likely use combination of headphones and equalizer in order to get a flat frequency response so that the original signal will not be changed [7].

2.5 Psychoacoustical parameters and emotional measures

Psychoacoustics is defined as 'a branch of science dealing with the perception of sound, the sensations produced by sounds, and the problems of communication' [16]. Therefore, the main objective of the current thesis, which is a human perception of metro train noise, can be considered a psycho-acoustical problem. The following section introduces 4 psychoacoustical parameters and 2 emotional measures based on which the sound samples will be evaluated.

2.5.1 Loudness

According to Zwicker and Fastl [7], loudness belongs to the category of intensity sensations, which means that relative sound intensity changes are proportional to relative changes in loudness. Loudness is defined as the intensive attribute of an auditory sensation in terms of which sounds may be ordered on a scale from soft to loud (ANSI) [1].

The reference value for loudness is a 1 kHz pure tone at a level of 40 dB presented binaural from the front in free field. This reference tone is equally loud as 1 [sone], which is a unit for loudness. The loudness level, which sounds equally loud as the reference tone is defined in [phon]. The equal-loudness contours, shown in figure 2.1, depict the relationship between the pure acoustical representation of intensity in dB (vertical scale) and psycho-acoustical parameter loudness level shown as the curved lines. For instance, the sound level of 40 dB at 1 kHz corresponds to the loudness level of 40 phons or loudness of 1 sone [1]. For loudness level higher than 40 phons the increase in the loudness level by 10 dB is equal to the doubling of loudness value, which means that the sound will be perceived twice as loud by human hearing, however, not for all frequencies.

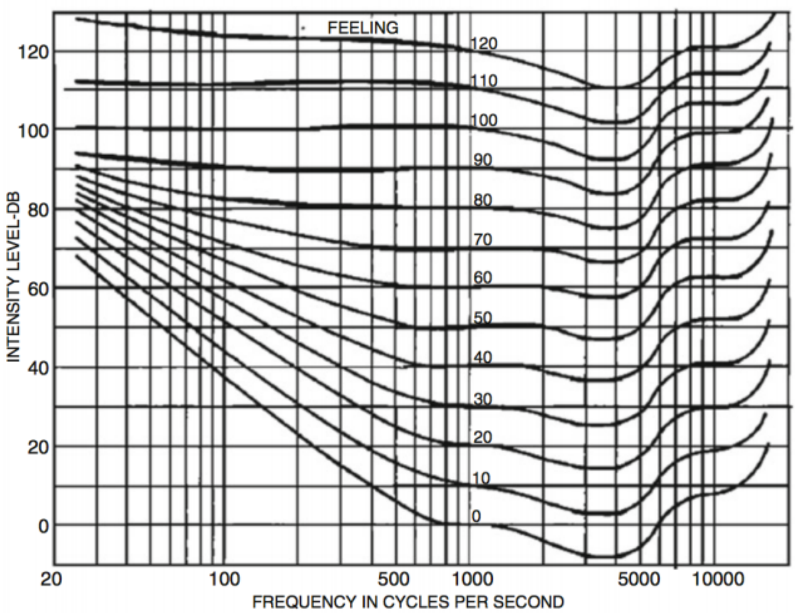


Fig. 10.5 The equal-loudness contours, known as the Fletcher–Munson curves, are taken from [19]. The *solid lines* correspond to the intensity of sound in air, in dB re: 1 pW/m^2, which is required to produce a perceived loudness equal to that of a 1.0 kHz tone with the same intensity level. That "loudness level" is called the "phon." To produce a loudness level of 60 phon, at 1.0 kHz, the sound pressure level would be 60 dB_{SPL}. The "0 dB" curve is often called the "threshold of hearing"

Figure 2.1: Equal loudness curves, [17]

Even though loudness depends highly on frequency, it also depends on bandwidth and duration, meaning that both spectral and temporal effects should be taken into account. It also leads to a thought that broadly used A-weighted SPL is a rather rough estimation of loudness level [7]. Here are two examples, which show the importance of spectral and temporal effects:

- Spectral effect: Uniform exciting noise (broad and equally distributed) is much louder than a $1kHz$ tone (narrow band sinusoid) at the same SPL.
- Temporal effect: Loudness reduces by approximately 10 phon if the signal is shortened by a factor of 10 for durations below 100ms.

In order to deal with different loudness contributions, the model of loudness was proposed by Zwicker [3] [7], which accounts both for spectral and temporal effects. The 3 major steps of the spectral model are depicted in figure 2.2. The first step is a graph in the left part of the figure, where the physical frequency scale is transformed to the psycho-acoustical Bark scale. In the second step given by the graph in the center, the masking effects are taken into account. The loudness pattern is shown as a hatched area in the right graph of the figure, which depicts the fundamental assumption that the total loudness is a sum of the specific loudnesses produced at different critical-band rates. The model can be extended to include temporal effects as well. The following model has been standardized and is a foundation for loudness calculations using acoustical software such as ArtemiS.

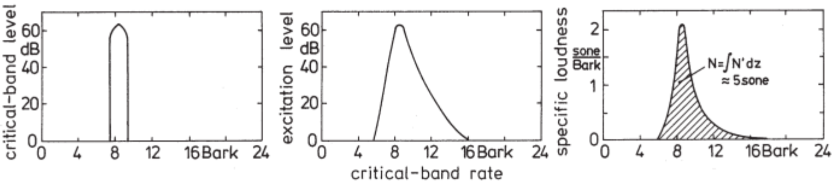


Figure 2.2: Zwicker loudness model, [3]

2.5.2 Sharpness

According to Head Acoustics website [18], sharpness is a sensation value which is caused by high frequency components in a given noise. Sharpness is considered to be a measure of tone color, i.e. it can give a character of powerfulness or aggressiveness depending on how much sharpness is added to the sound [3].

The unit for sharpness is [acum] and the reference value for 1 [acum] is the narrow-band noise one critical band wide with center frequency of 1 kHz and a level of 60 dB [7].

Spectral content and center frequency of narrow-band sounds influence sharpness the most. As a general rule, sharpness increases with high-frequency energy. Since it can also be decreased by adding sound at lower frequency (and vice versa), the influence of bandwidth is undeniable [3] [7].

It is known that sharpness can be estimated from the loudness pattern, which becomes evident from the equation (2.4) used for the model of sharpness. S stands for sharpness, N' is a specific loudness and g is a factor [7].

$$S = 0.11 \frac{\int_0^{24Bark} N'g(z)zdz}{\int_0^{24Bark} N'dz} \tag{2.4}$$

2.5.3 Roughness

Roughness is defined as a fundamental hearing sensation for fast amplitude modulations [19]. It expresses how 'rough' the sound is perceived by our hearing. Roughness is created by rapid temporal variations produced by either FM (frequency modulation) or AM (amplitude modulation) signal in the region between about 15 to 300 Hz with a peak value at around 70 Hz [7].

The reference value for roughness is defined as a 1 kHz tone with 100 % modulation at 70 Hz modulation frequency and having a level of 60 dB, which is equal to 1 [asper] [1]. Asper is a unit for roughness.

The following 3 parameters are vital in determining roughness [7]:

- For AM: degree of modulation and modulation frequency
- For FM: frequency modulation index and modulation frequency

However, it should be noted that FM can influence roughness much stronger than AM.

Roughness model can be described by temporal-masking pattern of sounds [3]. It is based on the differences in excitation levels produced by the modulation. According to formula (2.5), roughness R is proportional to the speed of change of the temporal masking pattern, i.e. f_{mod}, where ΔL is a modulation depth.

$$R \approx \Delta L \cdot f_{mod} \tag{2.5}$$

2.5.4 Fluctuation strength

Fluctuation strength is another psychoacoustical parameter, very much similar to roughness as it is also created by frequency or amplitude modulation of a signal (temporal modulation). However, unlike roughness, fluctuation strength indicates the perception of fluctuation, i.e. slow modulation of sound, and reaches its peak value at around 4 Hz modulation frequency.

The reference value for fluctuation strength is defined as a 1 kHz tone with 100% modulation at 4 Hz modulation frequency and at a level of 60 dB, which is equal to 1 [vacil] [1]. Vacil is a unit for fluctuation strength.

FS depends on the SPL. The higher is the SPL, the higher is the value for FS and vice versa. It also depends on modulation depth and modulation factor as well as center frequency and frequency deviation [7].

Model for the FS has the same input as the roughness model, which is shown by the formula (2.6), where FS is the fluctuation strength [3].

$$FS \approx \frac{\Delta L}{\frac{4Hz}{f_{mod}} + \frac{f_{mod}}{4Hz}} \tag{2.6}$$

2.5.5 Annoyance

According to the WHO guidelines for community noise, annoyance is defined as a feeling of displeasure associated with any agent or condition [1].

Zwicker and Fastl [7] claim that annoyance can asses the psychoacoustic elements of sound quality. They have worked out a model for psychoacoustic annoyance (PA), which is a combination of hearing sensations and is given by the formula (2.7). The following equation shows that PA is a combination of psychacoustic quantities such as loudness N, sharpness S, fluctuation strength FS and roughness R. It follows that PA mostly depends on loudness, the tone colour and the temporal structure of sounds.

$$PA \approx N(1 + \sqrt{[g_1(S)]^2 + [g_2(FS, R)]^2}) \tag{2.7}$$

2.5.6 Stress

Stress is a combination of valence(pleasantness) and activation(arousal), which are two main dimensions of emotional space [1].

According to The Merriam-Webster Online Dictionary (2009) [20], stress is defined as 'a physical, chemical, or emotional factor that causes bodily or mental tension and may be a factor in disease causation'. Sound is one of such physical factors that can negative affect our physical and/or mental condition and cause stress.

2.6 Type of listening test

According to the Nordtest Method [21], a listening test is defined as 'a test where one or more persons in a systematic way are presented to samples of sound and requested to give their evaluations/response in a prescribed manner'. Listening tests can be divided into subjective and objective tests. The current work focuses on objective or perceptive tests, which is about what the test participant hears and not about what the participant prefers of dislikes. Thus, the perceived stimulus is rated in objective terms, for instance the well-known psychoacoustical and emotional parameters.

Objective listening tests may be in turn subdivided into two main testing methods, namely semantic differential and paired comparison. The focus of this work will be on semantic differential test [1] [21].

2.6.1 Semantic differential

The word 'semantic' is defined in Merriam-Webster online dictionary as 'of or relating to meaning in language' [22]. It means that semantic differential method is a descriptive method when the sound samples can be evaluated or described with a help of certain words that can characterize a sound. It is vital to note that no comparison with other sound samples is implied. Perceptual acoustics is one of the fields where semantic differential test is often used for sound characterization [21].

It is required to decide the type of semantic scale to use for the listening test. Susini defines semantic scales as 'category scales defined either by a single semantic descriptor (unipolar scale) or by a pair of antonymic descriptors (bipolar scale)' [23]. It is typical to use numerical scales, i.e. seven- or nine-point scales, for the evaluation of psychoacoustical parameters like loudness. The problem with such scales is that the test participant is obliged to choose the exact number, which might affect the precision of the results. The number of points on a scale can be odd or even depending on the study.

In the current work, the category scaling is used with a seven-point bipolar scale, where the test participant has to evaluate a statement, which includes an adjective describing a sound, on a scale from 'disagree' to 'agree'. Advantage of using a seven-point scale is that it has an optimal length meaning that it is not too short, which would limit the possible answering options for a test participant, and also not too long, which could make it harder to choose a point out of large number of options [1].

2.7 Statistical model

The mean values, standard deviation, box plots, least square method and correlation will be used in the basic statistical analysis. The more detailed statistical analysis of the result will be made using one-way ANOVA model, which stands for the analysis of variance.

2.7.1 One-Way ANOVA

A one-way ANOVA statistical test is used to find whether there exist any significant differences between the two (usually at least three) or more independent population groups (levels) by comparing the variances of the group means. Note that One-Way ANOVA works only for one independent variable unlike two-way ANOVA, which has two independent variables. For instance, different sound samples can serve as independent variables for each population group. Consequently, such a test can be run separately for each psyachoacoustical and emotional parameter (dependent variables) under consideration to find out whether test participants can differentiate signals within each parameter. It is about the evaluation of an effect of independent variables on dependent variables. [26] [27]

One-way ANOVA is a hypothesis-based test, meaning that it checks if the null-hypothesis should be accepted or rejected. If rejected, then alternate hypothesis should be accepted. The null-hypothesis claims that there are no significant differences between the different population means, whereas the alternate hypothesis declares that at least two means belonging to different groups have significant variations between each other. The so-called F-test, i.e. the ratio of the variance calculated among the means of all groups to the variance within each group, is used to clarify whether null-hypothesis should be accepted or rejected. Along with the F-test, the p-value, which stands for probability, is also computed and tested against the predefined limit, which is typically defined as $\alpha = 0.05$ (a 5% limit). Although, the p-value limit can be defined as low as $\alpha = 0.001$ (a 0.1% limit). [26]

F follows an F-distribution if certain ANOVA-assumptions are followed. These assumptions include [26] [28]:

- Normality: normally distributed population for each sample;
- Sample independence: samples are drawn independently from one another;
- Continuousness: dependent variable should be continuous
- Homogeneity: equality of variances for each group.

However, the problem occurs when the one wants to find out which particular group means are different. A one-way ANOVA just states that at least two groups are different. So-called post-hoc test should be used to tell which particular groups have different means. ANOVA test can be run using Matlab software, which also allows to perform a multiple comparison test, which in turn lets the one discover which particular means are significantly different. [29] [30]

Meanwhile the F-test and the p-value tell us about the statistical significance of the obtained results and simply state whether the means are significantly different, they do not give any information on how different the means are. An effect size measure, which is denoted by η_p^2, can give us this crucial piece of information. In order to tell whether the obtained effect size value is substantial, the following rules can be applied [28]:

- $\eta_p^2 > 0.01$ indicates a small effect;
- $\eta_p^2 > 0.06$ indicates a medium effect;
- $\eta_p^2 > 0.14$ indicates a large effect.

Chapter 3

Experiment

The experimental part of work consists of outside measurements of train noise and listening tests in the controlled lab environment. The brief specification of the Oslo metro trains will be given first, followed by the detailed description of measurement procedure and listening test experiment. The list of the equipment used for this work will be given at the end of this chapter.

3.1 Metro trains

There are in total 345 metro trains produced by Siemens of the type $MX3000$ in Oslo, which replaced the old $T2000$ trains during the years $2005 - 2009$. Figure 3.1 shows the picture of such metro train. A typical trainset consists of 3 cars (wagons) like the one shown on the picture or 6 cars [24].

Some technical characteristics of the $MX3000$ trains are listed in table 3.1. $MC1$, $MC2$ and M notations denote 3 different car types, which when coupled together form the whole module or simply a train. The most important difference between $MC1$ and $MC2$ is that air compressor can be only found in $MC1$, whereas ATP (Automatic train protection) controller only in $MC2$. $MC1$ and $MC2$ have similar length, whereas weight of $MC2$ is is bigger than both both $MC1$ and M. Relevant technical drawings of train cars as well as drawing of the frontal view of the train are given in the Appendix 6.7.

Figure 3.1: Oslo metro train, [25]

Table 3.1: MX3000 characteristics

Length of MC1	18210mm
Length of MC2	18210mm
Length of M	17920mm
Whole module length	54340mm
Max width	3160mm
Max height	3800mm
Bogie distance	11000mm
Weight of MC1	ca. 30.7t
Weight of MC2	ca. 31.6t
Weight of M	ca. 30.6t
Whole module weight	ca. 92.9t
Maximum speed	70km/h
Electric system	750 V DC third rail

3.2 Measurements

3.2.1 Measurement protocol

Some peculiarities of Norwegian nature should be taken into account while discussing the train noise measurements. Mountain landscapes with various slopes are quite typical, although there are many plane surfaces in the Oslo region as well. The time of the year also plays vital role. The 4 measurement campaigns took place at the end of February, and each of these campaigns will be thoroughly described in the following sections. Different tracks and metro trains have different acoustical properties. Therefore, the 4 measurement campaigns took place during the daytime in 4 different locations along the metro lines in Oslo. Each of these 4 locations is special in a way that Brekke & Strand company experienced that A-weighted SPL measurements are not sufficient to fully describe the nature of noise at these particular locations. Psychoacoustical analysis was chosen as an attempt to find some other parameters that could fill this gap. Calibration of the microphones was checked both before and after the measurements and was found to be within the $ISO3095 : 2013(E)$ [14] limits, i.e. the difference between the two consecutive calibrations is within $\pm 0.5dB$.

Minimum of 10 pass-bys were recorded by measuring sound pressure level along with the background noise measurements at each site, including minimum of 5 pass-bys for each train direction. The sound pressure level L_p was measured with a help of free-field 1/2" microphone along with the two microphones placed in the artificial head. No frequency weighting was applied. The measurement time interval for each pass-by was controlled manually. Typical vehicle speed and distance from the fence to the microphone were chosen to be as similar as possible. Speed was recorded using a stopwatch whereas the distances were measured with a help of laser distance measure. Since neither thermometer nor anemometer was available during the measurement procedure, the rough estimates of temperature and wind speed were taken from the internet website [15]. No instructions were given by the traffic manager to the tramway drivers, trains were in a regular traffic. The serial number of each tram was noted.

3.2.2 The first measurement campaign at Voksenlia

Measurement setup and acoustical environment

The first campaign took place at Voksenlia on 11.02.19. The side view of the measurement setup is shown in figure 3.2. The two pressure microphones, mic #1 (top image) and mic #2 (central image) as well as an artificial head denoted as mic #3(head), were set perpendicular to the rails. The distances between the fence (vertical grey bar) and the centres of the closest (out) and furthest (in) rails was 5.2m and 10.1m respectively. The train passages in both directions (in and out of the city) were recorded. Vegetation of rather low density was mainly present on a very steep slope, on the opposite side of the track with respect to the microphone position, hence its damping was negligible. The top view is shown in figure 3.3. The track part was straight in front of the measurement point, but there were turnabouts both on the left and on the right sides of the rails, which might be the cause of so-called squeal noise. Sometimes the trains came simultaneously. There was a rather steep slope at the positions of the microphones as shown in figure 3.4. None of the big objects that could reflect the sound waves have been detected in the nearby area. The ground was covered with a rather thick layer of fresh snow, as shown in figure 3.4, which might have affected the sound propagation due to the absorptive properties of snow layer since it has rather low flow resistivity [31]. The table with characterization of different ground surfaces with respect to relevant flow resistivity values is given in Appendix 6.1.

Weather conditions

The weather was sunny with clear blue sky and no precipitation (see figure 3.4). Observed temperature was around $2°C$ with 52% humidity and the wind speed of around $5km/h$. Snowdrifts covered the ground in the area between the rails and the measurement points.

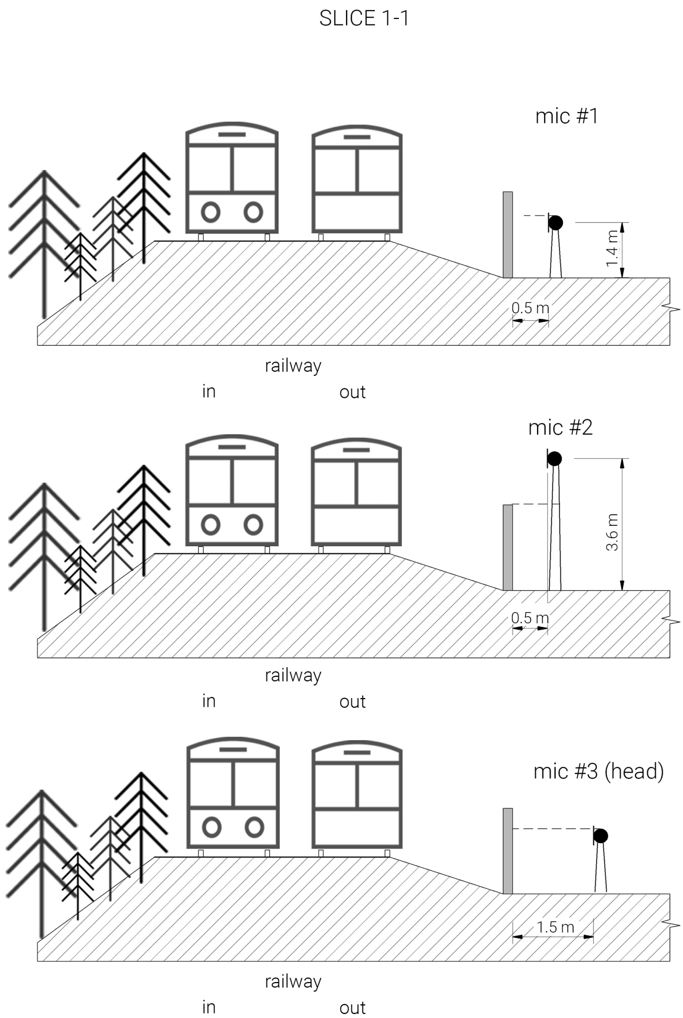


Figure 3.2: Voksenlia,side view

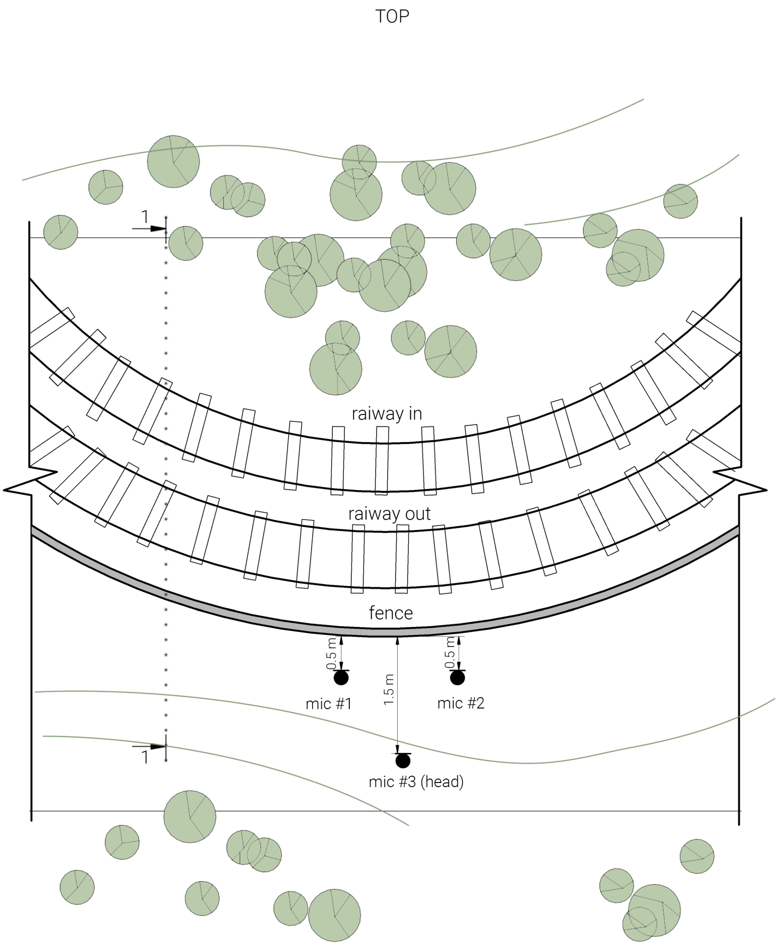


Figure 3.3: Voksenlia,top view

Figure 3.4: Voksenlia,picture from the measurement site

3.2.3 The second measurement campaign at Dalbakkveien

Measurement setup and acoustical environment

The second campaign took place at Dalbakkveien on 12.02.19. The measurement setup, shown on the figure 3.5, consisted of the two pressure microphones, mic #1 (top image) and mic #2 (central image). The distances between the fence (vertical grey bar) and the centres of the closest (out) and furthest (in) rails was 5.85m and 10.07m respectively. The problem was that the lowest mic was clearly lower than the rail height unlike the highest one. An artificial head (mic #3(head), bottom image) was set on the snowdrift further away from the microphones. The train passages in both directions were recorded, however, only the direction when train drives to the city (in) will be used in further analysis. There was a rather steep slope at the positions of the microphones as shown in the side view of figure 3.5, which might be the cause of some unwanted reflections. None of the big objects that could reflect the sound waves have been detected in the nearby area. Figure 3.6 shows the top view of the measurement site with some vegetation present behind the microphone positions, which does not affect sound propagation. The ground at the measurement site was flat and both tracks were straight with with only turnabout on the left side quite far away from the measurement point. The trains that came from the left were sometimes detected a bit too late because of the turnabout and presence of little vegetation close to the fence.

Weather conditions

The weather was chilly with no precipitation. It was dull with cloudy sky (see figure 3.7). Observed temperature was around $0°C$ with 75% humidity and the wind speed of around $2km/h$. Snowdrifts covered the ground in the area between the rails and measurement points.

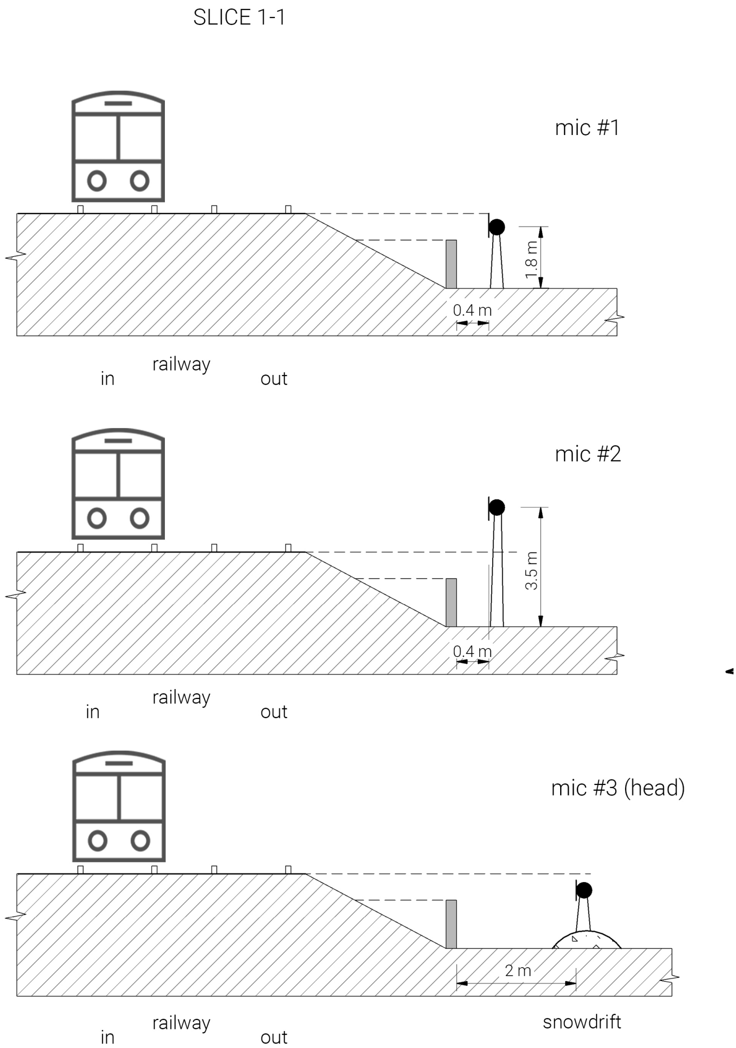


Figure 3.5: Dalbakkveien,side view

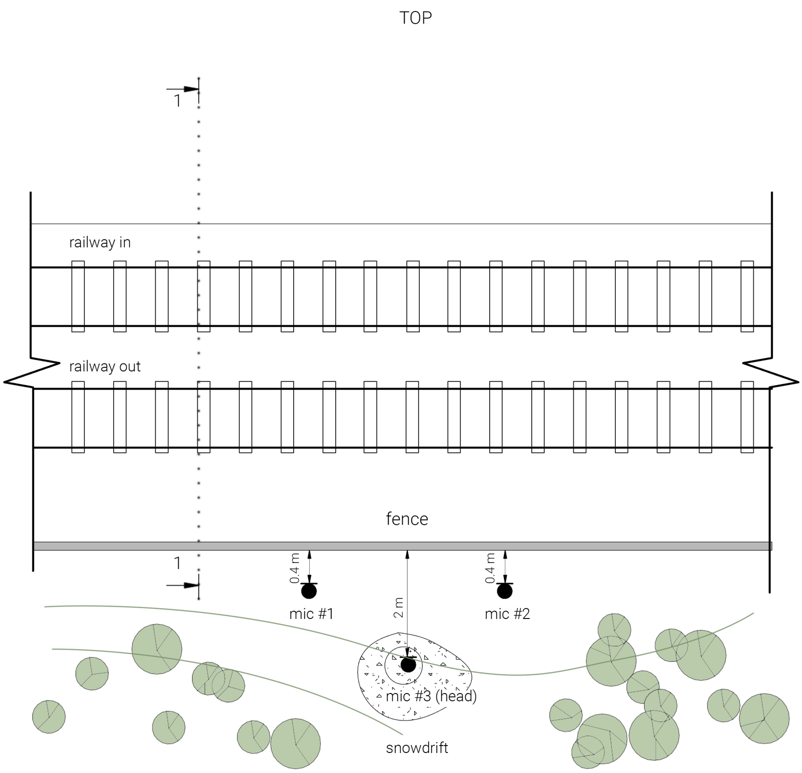


Figure 3.6: Dalbakkveien,top view

Figure 3.7: Dalbakkveien,picture from the measurement site

3.2.4 Third measurement campaign at Voksenlia and Borgen

The third campaign took place at Voksenlia and Borgen on 25.02.19.

Borgen

Measurement setup and acoustical environment

The measurement setup, shown on the figure 3.8, consisted of the two microphones, mic #1 (top image) and mic #2 (central image) as well as an artificial head mic #3(head). The distances between the fence (vertical grey bar) and the centres of the closest (in) and furthest (out) rails was 4.93m and 9.57m respectively. The train passages in both directions were recorded, however, only the direction when train drives from the city (out) will be used in further analysis. None of the big objects that could reflect the sound waves have been detected in the nearby area. Figure 3.9 shows the top view of the measurement site with some vegetation present behind the microphone positions, which does not affect sound propagation. The ground at the measurement site was flat and both tracks were straight with a train running from the tunnel on the right side, which made its appearance spontaneous, and therefore it was sometimes detected too late. The fence also blocked the vision of the approaching trains. Microphones were placed perpendicular to the track switches, which can be the cause of additional noise when the train runs over them. The picture from the measurement site is shown in figure 3.10.

Weather conditions

The weather was quite cool with no precipitation. It was rather dull (see figure 3.10). Observed temperature was around $6°C$ with 77% humidity and the wind speed of around $2km/h$. Ice and tiny snowdrifts covered the ground in the area between the rails and measurement points.

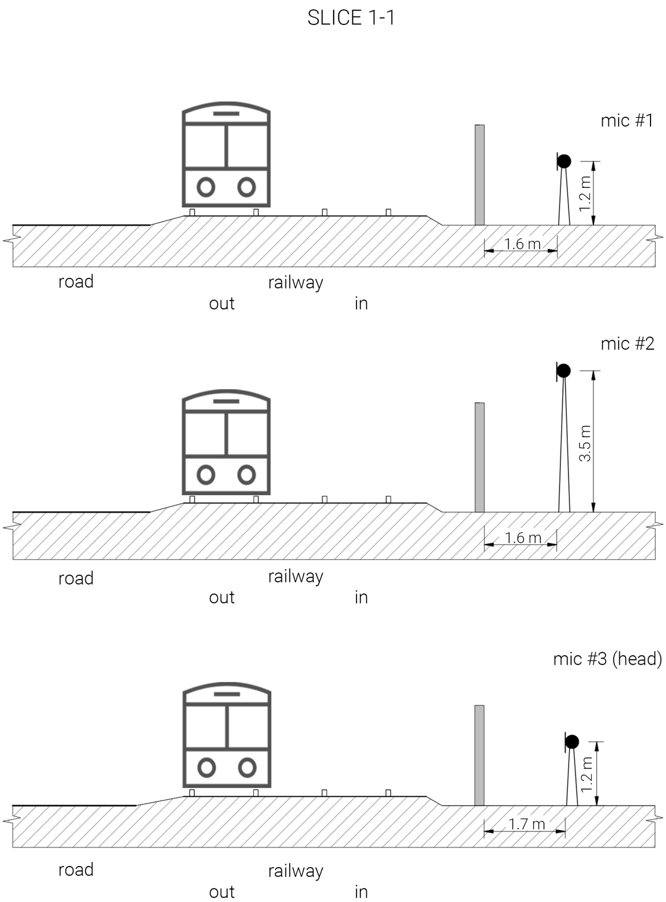


Figure 3.8: Borgen,side view

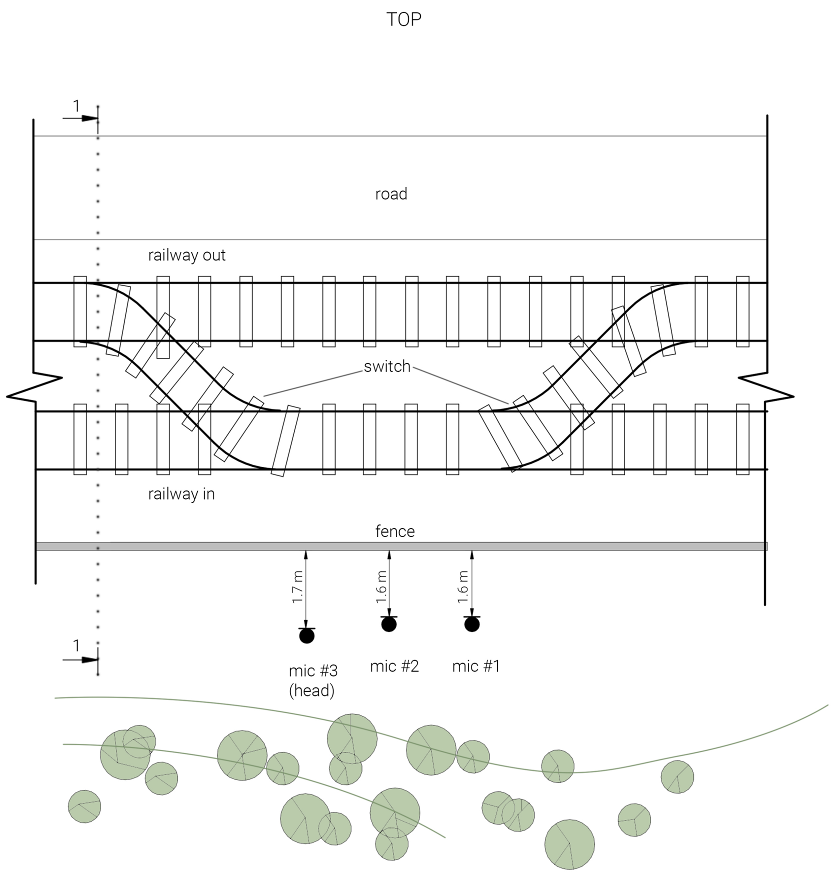


Figure 3.9: Borgen,top view

Figure 3.10: Borgen,picture from the measurement site

Voksenlia

Measurement setup and acoustical environment

The measurement setup with side and top view is shown in figures 3.11 and 3.12 respectively. It is exactly similar to the one used in the first measurement campaign at Voksenlia. The only differences were that this time mic #1 was set at the height of $1.2m$ and mic #2 at $3.5m$ above the top of the ground and both at the distance of $0.3m$ from the fence. Figure 3.13 shows the picture from the measurement site.

Weather conditions

The weather was a bit chilly with no precipitation. It was dull with cloudy sky (see figure 3.13). Observed temperature was around $5°C$ with 69% humidity and the wind speed of around $2km/h$. Ice and snowdrifts covered the ground in the area between the rails and measurement points.

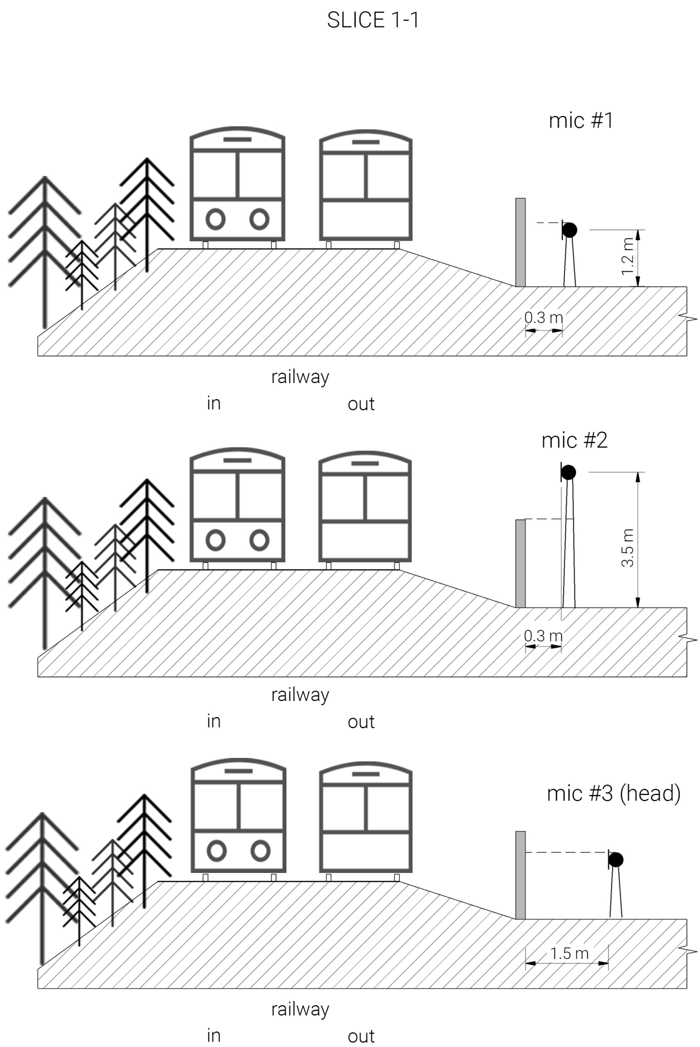


Figure 3.11: Voksenlia,side view

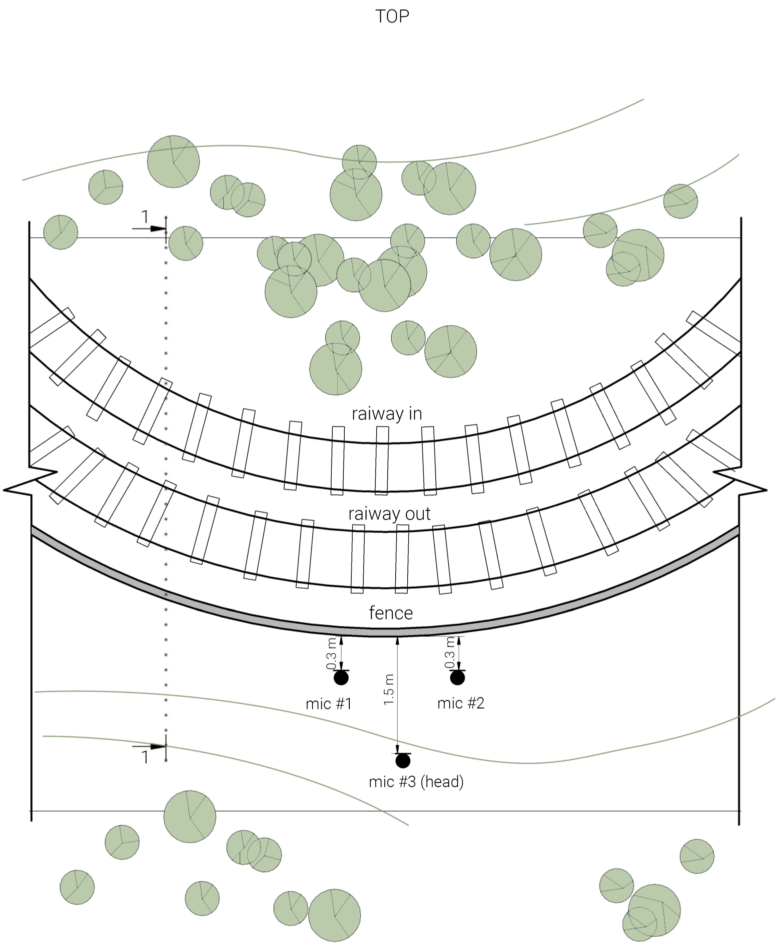


Figure 3.12: Voksenlia,top view

Figure 3.13: Voksenlia,picture from the measurement site

3.2.5 Fourth measurement campaign at Tjensrud(Jar)

Measurement setup and acoustical environment

The fourth measurement campaign took place at Tjensrud(Jar) on 28.02.19. As shown in figure 3.14, the measurement setup consisted of the two microphones, mic #1 (top image) and mic #2 (central image) as well as mic #3(head) set behind the snowdrift. The distances between the fence and the centres of the closest (in) and furthest (out) rails was 5.51m and 10.21m respectively. The train passages in both directions were recorded, however, only the direction when train drives from the city (out) will be used in further analysis. None of the big objects that could reflect the sound waves have been detected in the nearby area. Figure 3.15 shows the top view of the measurement site with some vegetation present on the other side of the rails with respect to the microphone placement, which does not affect sound propagation. On the left from the microphones and also on the opposite side there was detected a concrete obstacle (wall, denoted as a horizontal orange bar), which could cause some unwanted reflections. The fence is once again denoted by the horizontal grey bar. The ground at the measurement site was flat and both tracks were straight in front of the measurement position, but there were two turnabouts from both sides. It should be noted that both the trams and trains run on these rails both ways. There is also some kind of noise from the electrical system. Very unpleasant sound is heard when the trams and trains run on the closest track, high pitched and extremely unpleasant, though not so high.

Weather conditions

The weather was chilly but very sunny, clear blue sky and no precipitation (see figure 3.16). Observed temperature was around $8°C$ with 66% humidity and the wind speed of around $3km/h$. Snowdrifts covered the ground in the area between the rails and measurement points. The snow was crystallized, like small transparent caviar, which might have affected reflection of sound. It was rather solid, could easily walk and stand on it.

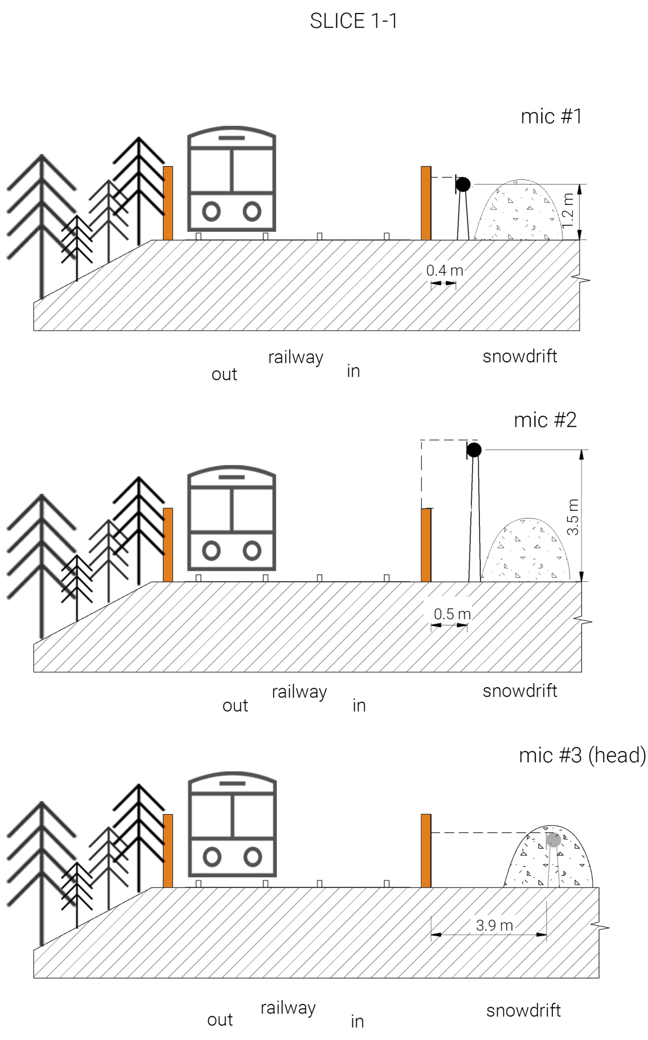


Figure 3.14: Tjensrud,side view

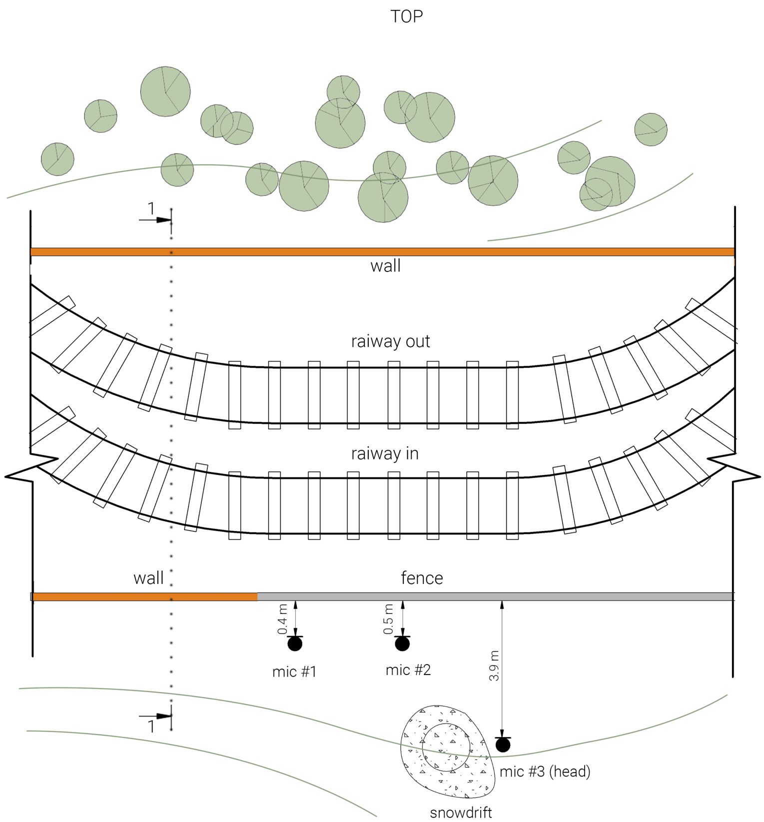


Figure 3.15: Tjensrud,top view

Figure 3.16: Tjensrud(Jar),picture from the measurement site

3.3 Listening tests

Equipment and software used

The test took place in a sound-insulated room. It was a computer-based test realized using laptop connected to the over-ear headphones Audio-Technica ATH-M40x. The headphone frequency response is given in Appendix 6.4. The test itself was programmed using Matlab computer software.

Form of the test

The test was run with a test leader and the focus was on individual listening. The order of stimuli was randomized. The test leader estimated the duration of the test to be ca. 30 - 40 minutes including pauses. Maximum duration of the listening round was no longer than 20 minutes and each round was followed by a short pause. The experimenter decided repetitions of test signal and pauses between test signals. Test participants could decide the time for answering. Non-acoustic sources like pictures from the measurement sites, video presentations and sketches of the measurement setup, weren't included in the test since they are considered as unwanted bias in psychoacoustic test. The test was planned according to the Nordtest method [21].

Introduction before the test

Test persons had a short presentation of the test organization, test leader, purpose, listening conditions, safety, ethics, their rights, what is expected and that they have the right to terminate their participation whenever they want, the expected duration and pauses, and other practical information before the test. It was stressed that there are no correct answers, but hence 'their answer' that is wanted. The written form of presentation was used. All the necessary information about the test procedure including test situation, test conditions, sounds and the scenario and test persons' task was given in the introduction, so that test participants could familiarize themselves with the upcoming test. Introduction to the listening test can be found in the Appendix 6.2 in the Norwegian language. The main session consisted of 30 signals and was repeated twice for statistical reasons. No pretest was done.

Type of the test

A semantic differential experiment with numeric seven-point scale was used as the main test type. Numeric scale's advantage is that it is relatively fast to read after the test, but the drawback is that it does not allow the test persons to choose numbers between the fixed points.

The following table shows the statements used for the test with respect to each psychoacoustic and emotional parameter in the Norwegian language along with the scale that was used:

Table 3.2: Original table of statements

Psychoacoustic parameter	Statement	Scale
Loudness	Lydnivået er 'høyt'	uenig/enig
Sharpness	Lyden er 'skarp'	uenig/enig
Roughness	Lyden har 'ruhet'i seg	uenig/enig
Fluctuation Strength	Lyden er 'fluktuerende'	uenig/enig
Stress	Lyden er 'stressende'	uenig/enig
Annoyance/Pleasantness	Lyden er 'trivelig'	uenig/enig

English translation of the above table can be found in the Appendix 6.3.

Subjects used for the test

The purpose was to include at least 18 persons in the hearing test. The test participants were consulting acoustic company workers as well as university students. Most of the students had rich background in acoustics. Participants filled some background information before the test, including their hearing ability, gender and age.

Playback

It was important to give soft starts and ends (rise time of $200 - 500ms$) to the continuous sound samples that were used for this test, so that no unintended transient sound could occur. The sound was played through the headphones inside the sound-insulated room, which can change the sound depending on the frequency response of the headphones. The headphones were calibrated to the same level as the originally recorded sound with a help of artificial(dummy) head. Frequencies were not adjusted with a help of an equalizer since the frequency response of the headphones used was considered to be flat enough in the frequency range of interest.

3.4 Equipment

Table 3.3: List of equipment

Type of tool	Manufacturer	Serial/licence number	Amount (#)
Alkaline battery, 9V	Philips	-	2
Artificial head	Bruel & Kjaer	1899850	1
Free-field 1/2” microphone, 46AE	GRAS	259983	1
Free-field 1/2” microphone, 46AE	GRAS	260058	1
Free-field 1/2” microphone, 1206	Norsonic	-	2
Headphones ATH-M40x	Audio-Technica	-	1
Laptop Macbook Air	Apple	FVFV2RWNJ1WK	1
Laser distance measure DISTO A3	Leica	2061132653	1
Matlab programming software, R2017a	MathWorks	833468	1
Microphone Calibrator, type 4231	Bruel & Kjaer	-	1
Power Supplier/Amplifier, type 336	Norsonic	20578	1
Software ArtemiS suite	HEAD acoustics	-	1
SQuadriga II	HEAD acoustics	33221268	1
Stopwatch WT035	Asaklitt	20180570743	1
Tape	-	-	1
Tripod	-	-	3
Windscreen	-	-	2

Chapter 4

Analysis

Analysis of the results obtained from the measurements will be presented in following chapter. Both the results from the ArtemiS software, which is based on the mathematical calculation models, and the listening test results will be given separately and also will be compared against each other. At the end of the chapter a deeper statistical analysis using ANOVA-1 test and correlation analysis will be described.

4.1 Train measurement data

The data obtained from the measurements had to be properly sorted in order to use it in the listening tests. The total number of 5 signals was picked from the 4 measurement sites. Only the limited number of signals could be chosen due to the time limitations of listening tests. Two signals were picked from Voksenlia (third measurement campaign) since the two trains moving in opposite directions at this site seem to produce different sounds: one uphill and one downhill. The results obtained from the first measurement campaign at Voksenlia were not used for the data processing since the left channel, i.e. the left ear of the artificial head, did not work and there was also no signal at the right channel, so the only valid measurement data that was obtained was the data from the two microphones on the tripods. One signal was picked from each of the remaining 3 sites by choosing the pass-by which was furthest from the measurement position. This was necessary to ensure that the SPL of the signals is not too high and there won't be any damage to the hearing system of people taking the listening tests. Otherwise, essential volume adjustments had to be made. Also it should be kept in mind that listening test participants are exposed to metro train noise during a longer time period than the one that is typical in the real life situation. Since the focus was to find out whether people could differ between different signals, some small volume adjustments of the signals played via headphones and made similarly for each of the signals were accepted. All the signals had the pattern, which seemed to be representative of each site.

The following names will be used from now on to denote the five signals (the raw data can be found in the Appendix 6.5):

- Sig 1 (Signal Nr. 374 from Borgen, train moving out)
- Sig 2 (Signal Nr. 366 from Dalbakkveien, train moving in)
- Sig 3 (Signal Nr. 358 from Jar, train moving out)
- Sig 4 (Signal Nr. 363 from Voksenlia, train moving out)
- Sig 5 (Signal Nr. 367 from Voksenlia, train moving in)

Afterwards, all the signals had to be bounded in a similar way. The signals were cut in ArtemiS program such that only the part of the signal 2.5 sec from both sides of the pass-by itself was taken into account. Later these signals were stored as stereo wav files to be able to use them in the listening tests. Channels 3 and 4 represented right and left ear respectively. The head files were transferred to wav with 16-bit resolution, original range from HDF and 200ms In Out fading mode.

As for the ArtemiS, the two-channel stereo sounds would give two different values for each parameter. According to head acoustics manual, the signal that is 'the worst', i.e. has the highest values with respect to different parameters, should be chosen among the two channels to be able to compare it with values obtained from the listening tests. It is wrong to compute the average between the two channel values in order to obtain a single value [32].

The differences between the measured background noise values and train pass-by sounds for each signal both for Channel 3 and Channel 4 are given in table 4.1. The only values that are rather low are values for Sig 1 since the differences for both channels are less than 10dB [14]. The reason might be that there is a highway close to the the measurement position and flow of vehicles could contribute to the increased background noise level.

Table 4.1: Background noise check

Signals	**Ch 3** $\Delta[dB]$	**Ch 4** $\Delta[dB]$
Sig 1	9.41	8.75
Sig 2	21	23.8
Sig 3	25.65	26.19
Sig 4	26.6	29.16
Sig 5	24.25	27.14

4.2 Data from ArtemiS

Table 4.2 shows the SPL values along with psychoacoustical parameter values for each of 5 signals and 2 channels (Channel 3 for the right and Channel 4 for the left ear) from 4 different sites computed in ArtemiS software. According to just noticeable differences, JND, the lowest detectable values for different psychoacoustical parameters are defined as follows [1]:

- 0.5 sone for loudness
- 0.08 acum for sharpness
- 0.04 asper for roughness
- 0.012 vacil for fluctuation strength

Loudness varies significantly for each signal and for both channels within each signal, so that the differences should be easily detectable by the human ear. All the signals seem to differ by more than 0.5 sone from each other. It is curious that the highest SPL value does not necessarily mean highest loudness. For instance, SPL for the signal 'Sig 1' is higher than SPL for 'Sig 5' but still the second signal is higher in terms of loudness values.

Sharpness also varies significantly with more than 0.08 acum for most of the signals and reaches especially high values for the signals 'Sig 4' and 'Sig 5'. It is at least two times higher than sharpness at other sites. Interestingly, the sharpness value for 'Sig 5' is higher than for the 'Sig 4' for channel 3 and almost the same for channel 4 despite the big level and loudness differences in favour of 'Sig 4'. The possible reason might be the difference in the radius length of the rail curvature or the fact that one train runs uphill while the other one runs downhill.

Roughness, at first glance, has the values that are very low and show very little variation for different signals according to the JND value of 0.04 asper.

Fluctuation strength also has very low values but according to JND of 0.012 vacil, people should still be able to differ some of the signals from the table with respect to this parameter, even though the differences are rather small.

In this case, the two channels are very similar and it is hard to tell, which of the them is 'the worst'. Therefore, Channel 3 was randomly picked for the future analysis.

Table 4.2: Psychoacoustic parameters from ArtemiS

Signal	Sig 1		Sig 2		Sig 3		Sig 4		Sig 5	
Channels	Ch 3	Ch 4	Ch 3	Ch 4	Ch 3	Ch 4	Ch 3	Ch 4	Ch 3	Ch 4
Level [dB SPL]	79.62	79.69	76.75	77.09	74.22	74.52	81.24	81.97	73.24	74.71
Loudness [soneGF]	24.2	22.8	26.7	27.7	23.8	25.4	37.7	45.6	29.5	35.2
Sharpness [acum]	1.3	1.28	1.38	1.43	1.4	1.43	2.91	3.36	3.03	3.35
Roughness [asper]	0.0286	0.0272	0.0327	0.0327	0.0305	0.0332	0.0293	0.0306	0.0279	0.0302
Fluctuation strength [vacil]	0.0331	0.0258	0.0183	0.0166	0.0164	0.0165	0.0253	0.0262	0.0268	0.0283

4.3 Listening test results

4.3.1 Mean values

The figure 4.1 shows the results from the listening tests for 21 participants, among which there were 14 men and 7 women. The average age was ca. 36 years. Each participant took the test 2 times, so that the presented data is the average of the total of 42 signals. None of the participants has bad hearing, but 2 claimed that they have hearing that is worse than average. One participant decided not to give information about his hearing abilities. psychoacoustical and emotional parameters are placed along the x-axis whilst y-axis corresponds to the evaluation of each sound based on the mentioned parameters from 1 (the lowest) to 7 (the highest). Different colors represent each of the 5 signals.

People can clearly distinguish between sharpness and stress for the different sounds. However, it seems to be harder to distinguish between loudness of different sounds, although these values are all in the upper part of the graph, and even harder with roughness and fluctuation, which have very little spread. Values for roughness and fluctuation strength are placed in the central part of the graph meaning that test participants struggle to clearly tell whether they agree or disagree that each of the presented sounds is rough or fluctuating. The other explanation can be that these signals are simply middle strong. Possible reason might be that the definitions for psychoacoustical parameters were not presented before the test start, and these terms were explained only if the participants asked for it. It might also be confusing since these terms are not intuitively understandable for many participants. Otherwise, it seems that these parameters are the worst to be used to characterize the metro train noise. In general, the sounds from metro trains do not seem to be pleasant for the test participants. Results for sharpness for the signals 4 and 5 show similar behavior to the ArtemiS results. Also, these two signals are experienced as the least pleasant and the most stressful sounds.

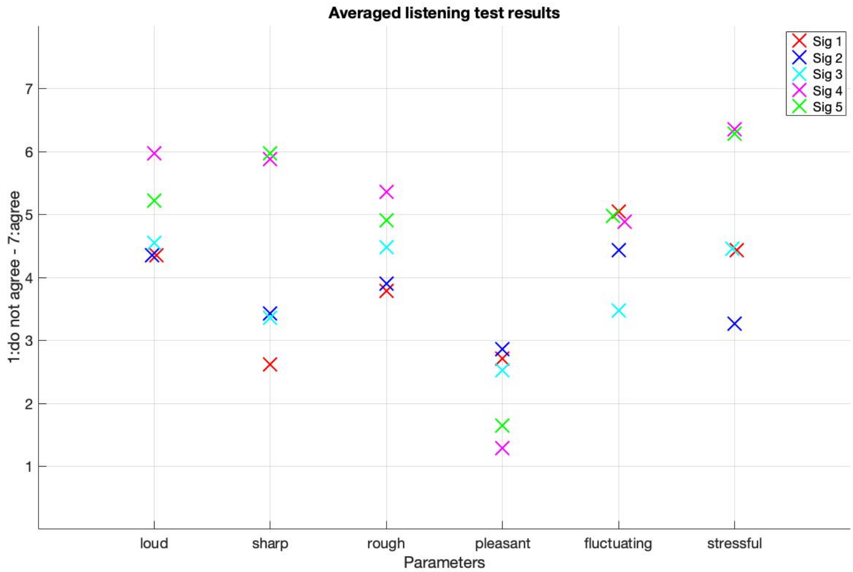


Figure 4.1: The average listening test results

4.3.2 Standard deviation

Figure 4.2 shows the plots of the standard deviation values applied to the results obtained from the listening tests with respect to each of the 6 parameters used (thus, making 6 different plots). The crosses denote the mean values for the signals used, which are placed along the x-axis, and each signal is colored differently. Once again, y-axis denotes the sound evaluation scale ranging from 1 to 7. The vertical lines stand for the range of ± 1 standard deviation.

Results for roughness reveal that there is a big spread in the data, which means that it was rather hard for the participants to evaluate this particular parameter applicable to noise from the metro trains. Results for the parameter fluctuation strength lead to the same conclusion, however participants find it a bit easier to evaluate train noise based on this parameter compared to roughness. Results for other parameters do not have such a big spread, which makes these parameters more reliable with respect to the train noise evaluation. Especially the results for signals 'Sig 4' and 'Sig 5' have rather small spread, which means that most of the participants similarly estimated these particular signals and these results are the most reliable.

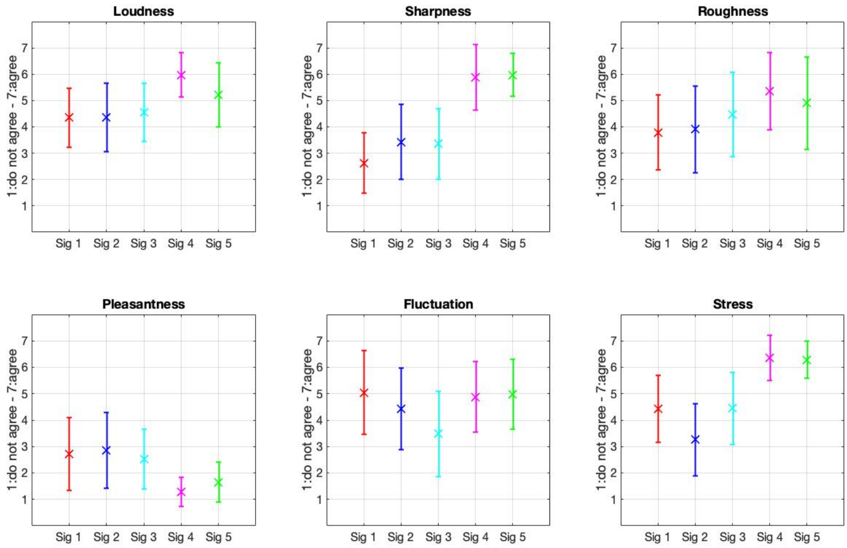


Figure 4.2: Standard deviations of the results from the listening tests

4.3.3 Box plot

Figure 4.3 shows a so-called box plot also known as a box-and-whisker plot for 6 psychoacoustical and emotional parameters under consideration. The x-axis represents the 5 signals mentioned earlier, whereas the y-axis denotes the scale ranging from 1 to 7. The blue boxes cover interquartile range (IQR) meaning that approx. 50% of the data is situated in this range with a red line denoting statistical median of the data (the 2nd quartile). The bottom of the box is called the 1st quartile (or 25th percentile), whereas the top of the box is called the 3rd quartile (or 75th percentile). The bottom and top lines outside of the box are called lower whiskers (denoting minimum value of the data set) and upper whiskers (denoting the maximum value of the data set) respectively. The red crosses are called outliers and denote extreme data points, which are superior than min and max values, and therefore fall outside the plot. [33] [34]

Loudness plot shows that 'Sig 4' has the the highest median and is skewed downwards whereas 'Sig 1' has the lowest median. 'Sig 2' has the biggest IQR as well as the biggest total range, which means that the spread of the data is the largest for this signal. Generally, all the signals have rather low IQR. Moreover, 'Sig 2' is clearly skewed downwards meaning that minimum value(s) of data set is situated rather far from the median value. 'Sig 5' is skewed upwards meaning that its maximum value(s) is situated far from the median value, however this signal also has an outlier around the value of 2.5.

Sharpness plot shows that 'Sig 4' and 'Sig 5' have the highest median values, whereas 'Sig

1' has the lowest median value, which is situated rather close to the 3rd quartile. 'Sig 4' is skewed downwards and has an outlier meaning that certain participant(s) gave abnormally low estimate to this signal. 'Sig 2' and 'Sig 3' are skewed upwards. 'Sig 2' has the highest IQR while 'Sig 3' has the highest total range.

Results for roughness show that 'Sig 1' and 'Sig 2' have the lowest median while 'Sig 4' and 'Sig 5' have the highest median values. Signals 3, 4 and 5 are all skewed downwards while 'Sig 4' and 'Sig 5' also have some outliers. 'Sig 2' has the biggest IQR while both 'Sig 2' and 'Sig 3' have the widest total range.

Pleasantness plot shows that 'Sig 2' has the highest median while 'Sig 4' has the lowest one. Signals 4 and 5 are skewed upwards and 'Sig 4' also has an outlier and a median, which lies on the 1st quartile showing a very small data spread. 'Sig 1' also has an outlier. 'Sig 2' has the highest IQR and the total range as well. The other signals have generally rather small IQR.

Plot for fluctuation strength reveals that 'Sig 1', which also has one outlier, has the highest median value while 'Sig 3' has the lowest median. Signals 1 and 2 are skewed downwards. All the signals have rather big IQR and the total range.

Plot for stress shows that signals 4 and 5 have the highest median values while 'Sig 2' has the lowest median. Median for 'Sig 5' lies very close to the 3rd quartile and this signal also has an outlier. 'Sig 4' is skewed downwards. 'Sig 2' has the highest IQR, whereas 'Sig 3' has the highest total range.

Once again, it seems that results for roughness and fluctuation strength have the largest spread and lead to the biggest uncertainties. Test participants are the most certain with the evaluation of pleasantness. Also results for signals 4 and 5 show the smallest spread while results for signal 2 have very large IQR for for most of the parameters used, which leads to a thought that people seem to give similar evaluations to the same signals regardless of the parameter itself for the mentioned signals.

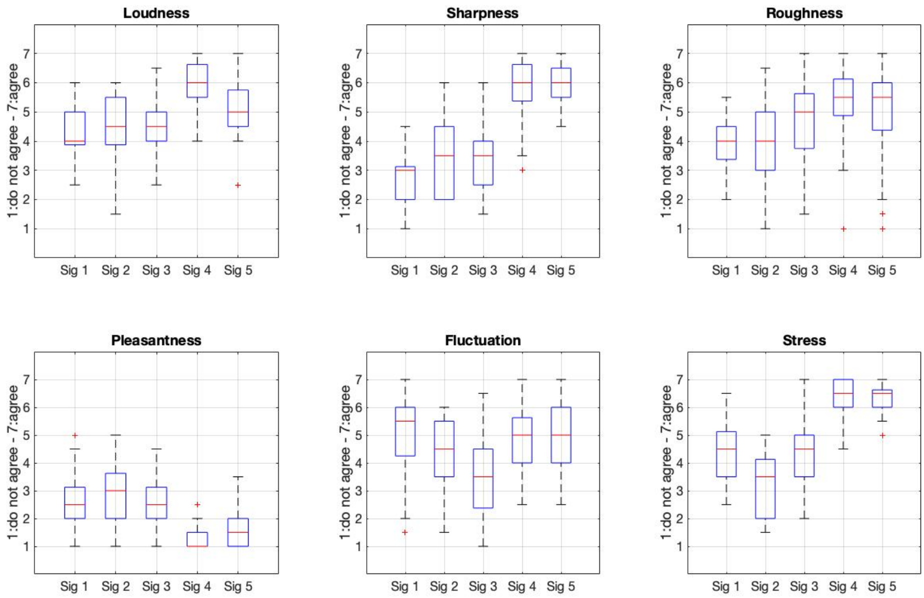


Figure 4.3: Box plot of the results from the listening tests

4.3.4 Comparison

Figure 4.4 shows the comparison between the results from the listening tests for channel 3 (y-axis) and values computed with a help of software (ArtemiS) based on theoretical models (x-axis) for the 4 psychoacoustical parameters. Differently colored markers with corresponding numbers represent the 5 signals used both for the listening tests and software computations. The red line is called regression line, and is based on so-called least squares method [35].

In order to explore the goodness-of-fit, R-squared values should be presented. R-squared is defined as a 'statistical measure of how close the data are to the fitted regression line' [36]. The other value called standard error also tells us how well the actual data fits the estimated values and exactly how large is the distance between these data types [37]. Results for both of these statistical values are summarized in the table 4.3.

Based on the data from the table 4.3 and the figure 4.4 we see that there is a very good correlation for the values of loudness and sharpness, both having very high R^2 values and rather low standard errors, both errors within $\pm 5\%$. Considering loudness, only the two places show mismatch between the listening test and theory, but the difference is very small. As for the fluctuation strength, the R^2 value shows that there is clearly a good fit of the data, but the standard error is slightly higher. The values for roughness, having both

Table 4.3: Regression analysis

Parameter	R^2	Standard error
Loudness	0.9037	0.2515
Sharpness	0.9728	0.2978
Roughness	0.1734	0.6969
Fluctuation strength	0.7230	0.3969

very low R^2 value and quite high standard error compared with other parameters, show that the model for this parameter does not represent the results of the listening test and using of model for this type of noise would need further investigation.

Three graphs show positive relationship since the line increases, whereas graph for roughness shows negative relationship. This points at a disagreement between the listening test results and the calculation model for roughness. Also signals 1 and 4 of the roughness plot are situated further away from the line meaning that errors for these signals are bigger compared to the estimated value (line). Slope for sharpness is quite steep meaning that with increasing x values, y values also increase. Slope for loudness is less steep. Slopes for roughness and fluctuation strength are rather gentle, which means that y values change very little with respect to x values, and therefore listening test participants struggled to differentiate signals based on these psychoacoustical parameters.

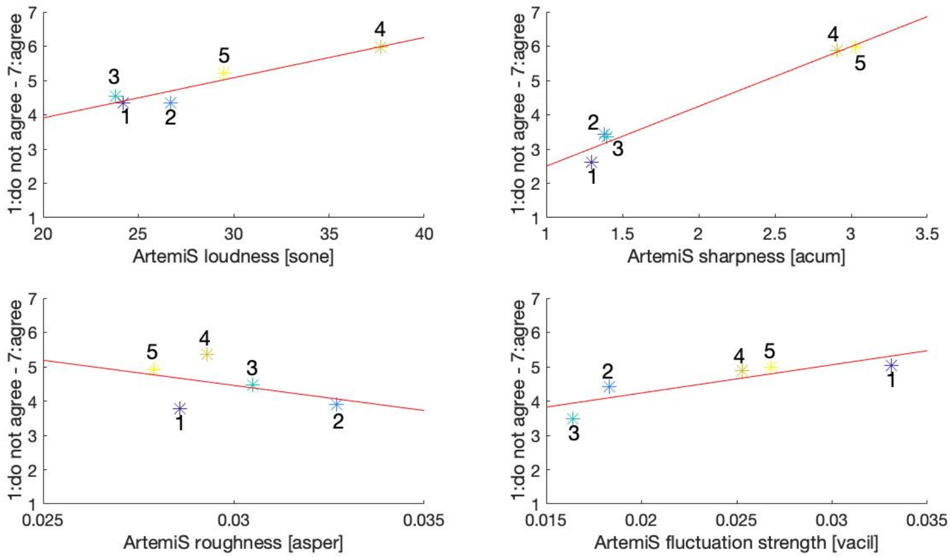


Figure 4.4: Comparison between the results from the listening tests and results obtained using software computations

4.3.5 ANOVA analysis

One-way ANOVA test results for 6 parameters, 4 psychoacoustical and 2 emotional, are summarized in table 4.4. F letter denotes results for so-called F-test, p-value stands for results for probability for accepting the null-hypothesis and η_p^2 denotes the effect size [1]. The F-table can be found in the Appendix 6.6. Null-hypothesis claims a priori that there is no significant difference between the means of the signals of interest for each of the 6 parameters.

Different colors highlight degree of significance of each parameter. Roughness and fluctuation strength are colored in red meaning that these parameters are insignificant on the 0.1% limit, which is the lowest possible limit. Also the F-test value on the 0.1% limit should be higher than 5.02 in order to reject null hypothesis. However, it should be noted that these parameters do fulfill 1% limit requirements, which is good enough and this limit can be satisfying considering that the number of test participants and signals weren't that big. Even though these two parameters have high effect size, it is still the lowest among all the parameters. It means that test participants struggled to differentiate between the test signals based on these two parameters.

Loudness and pleasantness are colored in yellow meaning that these parameters are significant since p-values for both parameters are lower than 0.001 and the F-test values are higher than 5.02. Also the effect size is higher. These parameters can be used to evaluate noise from the metro trains.

Sharpness and stress are colored in green meaning that these parameters are statistically significant since their p-values are much lower than 0.001 and F-test values are much higher than 5.02. Also the values for effect size are the highest. These parameters are the best to use for metro train noise evaluation, and people can easily differentiate signals using these parameters. However, it should be noted that psychoacoustical parameters, unlike emotional parameters, can be measured and modeled.

Table 4.4: ANOVA-1 test results

Parameter	**F(4,100)**	**p**	η_p^2
Loudness	10.49	$3.9 \cdot 10^{-7}$	0.3
Sharpness	47.56	$2.5 \cdot 10^{-22}$	0.66
Roughness	4.58	0.002	0.16
Fluctuation Strength	4.99	0.001	0.17
Stress	39.94	$5.9 \cdot 10^{-20}$	0.62
Pleasantness	13.03	$1.4 \cdot 10^{-8}$	0.34

4.3.6 Correlation between different parameters

Table 4.5 shows data from correlation analysis for the psychoacoustical and emotional parameters compared against each other in order to find out how strong these parameters affect each other. Numbers in the table denote the correlation coefficient R. Cell colors indicate whether the p-value (tests Null-Hypothesis of no correlation) is below 0.05 or 5%(orange), 0.01 or 1%(red) or simply higher than both of these limits(white). [1]

In general, all the parameters except for fluctuation strength show very good correlation at the 5% limit and even at the 1% limit between loudness and pleasantness. P-values for fluctuation strength and a value between sharpness and stress are all above the limits, which means that these parameters do not influence each other significantly.

All the correlation coefficient values for pleasantness are negative meaning that pleasantness is inversely correlated with other parameters. Also absolute values for R between pleasantness and other parameters except fluctuation strength are the highest meaning that pleasantness is highly affected by other psychoacoustical parameters. Absolute R values for stress are lower than absolute values for pleasantness, which shows that even though stress is substantially affected by loudness and roughness, it is generally less influenced by psychoacoustical parameters than pleasantness. Loudness and roughness are the two psychoacoustical parameters, which influence emotional parameters the most. It should be noted that pleasantness and stress as two emotional parameters correlate with each other quite well.

Psychoacoustical parameters loudness, sharpness and roughness also correlate with each other at the 5% limit and have rather high R values. The nature of the signals might explain such unobvious results. There might be some underlying processes that make test participants react similarly to the same signals regardless of the parameter. Of course, it does not apply to all of the parameters, like the already mentioned fluctuation strength, which does not correlate well with other parameters. Even though people might similarly react to the same signals based on different parameters, it does not imply that all these parameters are equally suitable to describe certain type of noise.

Table 4.5: Correlation table

	Loudness	**Sharpness**	**Roughness**	**FS**	**Stress**	**Pleasant.**
Loudness		0.90	0.95	0.36	0.88	-0.98
Sharpness			0.91	0.37	0.87	-0.95
Roughness				0.12	0.89	-0.96
FS					0.43	-0.39
Stress						-0.96
Pleasant.						

Chapter 5

Discussion

The initial task was to discover, which psychoacoustical and emotional parameters can be used to describe the perception of noise from Oslo metro trains and improve our understanding of the railway noise. The total of 6 parameters were analyzed, 4 psychoacoustical, namely loudness, sharpness, roughness and fluctuation strength and 2 emotional including stress and pleasantness. The four measurement campaigns took place in Oslo, and measured values for 5 signals from four measurement sites were used in further analysis. Measured values were analyzed using ArtemiS software and also used for the psychoacoustical listening tests. The measured and computed results were presented in the form of digital values supported by graphical representation. In order to check the reliability of the obtained data, the detailed statistical analysis was carried out. The computed and measured results were also compared against each other.

Results for the psychoacoustical parameters computed using ArtemiS program revealed that loudness and sharpness are the two parameters that have significant variations for different signals. The average results from the listening tests have shown that test participants can differentiate signals for sharpness while roughness and fluctuation strength show the worst performance having the widest spread (standard deviation and box plots). Test participants also seem to be able to differentiate different signals for the emotional parameter stress. It should be noted that people seem to evaluate signals 2, 4 and 5 quite similarly for most of the parameters, which leads to a thought that nature of the signals plays vital role and not only the parameter used. Signals 4 and 5, which represented the two different train directions at the same site - one going uphill while the other going downhill, seem to show different behavior for the parameter sharpness compared with results for SPL and loudness. It might lead to a thought that this parameter can give more detailed information about the train noise from the sites with a slope and/or rail curvature. Comparison between the theoretical computations and listening test results for 4 psychoacoustical parameters showed that there is the best correlation for loudness and sharpness and the worst for roughness. Thus, software calculations have good agreement with the listening test results only for sharpness and loudness. One-way ANOVA test has shown once again, that sharpness and stress have the highest effect size and therefore are the best parameters for metro train noise evaluation. It is quite intuitive that results for sharpness are so impressive since the railway noise is often associated with high-pitched sounds. Roughness and fluctuation strength are the worst. Loudness and pleasantness are good enough to be used for train noise evaluation. Correlation analysis revealed that fluctuation strength is the only parameter that does not correlate well with other parameters, and pleasantness has the best correlation with the rest of the parameters. The fact that loudness, sharpness and fluctuation strength have good mutual correlation strengthens the thought that people seem to evaluate the same signals in a similar way for most of the parameters.

The possible sources of failure and ideas concerning the future research should be mentioned as well. Generally, more signals could be used from each site and more sites for better precision. Measurements at the night time or in different season of the year could be interesting to look at. It could also be useful to add additional measurement positions to the measurement setup in order to collect more data. Even though the exact speed measurements were not crucial for this work, the more precise speed measurements could be applied instead of the speed measurements using stopwatch. The effect of different railway noise sources could be analyzed separately to find out exactly which sources should be eliminated in order to decrease annoyance caused by railway noise. As for the listening tests, the number of test participants and the attempts could be increased. The equalization of headphones to get as close to the flat frequency response as possible would also be desired in order to play the pure signal through the headphones. Also, the psycho-acoustical test using loudspeakers in a sound-proof room would be the other alternative to the listening test setup that would be interesting to have a look at. The signals that were picked up for the analysis were selected by taking into account the loudness level, which should not be harmful for the hearing of the test participants. It might lead to the obvious deviation from the real-life situation, however, the comparison between the different signals was what really mattered, and not the evaluation of each of the signals separately, and therefore this compromise was accepted. Also, there was no intention to analyze the highest possible SPL of train noise from each site and the signals that were selected were all different from each other in terms of SPL. There appeared some questions concerning the meaning of some psycho-acoustical parameters like roughness, not everyone can intuitively understand what do these parameters mean and therefore the results of the test can be somewhat subjective as different people can differently interpret different words. Some of the test participants could recognize sounds from some measurement places since they are acousticians themselves and have made measurements on the same sites. The total number of psycho-acoustical and emotional parameters could be increased and it can be curious to have a look at some other parameters like tonality or activation. Theoretical models that are the basis for software calculations are not the only existing ones, and some other models could be used to compute the psycho-acoustical parameters. The model for calculating annoyance applied to the railway noise could be proposed.

Chapter 6

Conclusion

The psychoacoustical evaluation of noise from the metro trains in Oslo was studied at four different sites. Four psychoacoustical and two emotional parameter values, based on the data obtained from the measurement campaigns, were estimated using software computations and listening test results. There was found to be good agreement between the theoretical computations using ArtemiS software and psychoacoustical listening tests in a controlled lab environment for sharpness and loudness, however, the correlation results for roughness and fluctuation strength revealed that theoretical models for these parameters do not agree with what actual test participants perceive. Thorough statistical analysis, along with the more basic one, was realized using one-way ANOVA test. Out of the four psychoacoustical parameters, sharpness seems to be a good representative for the special sound characteristics of metro noise, having a strong coupling with stress and inverse pleasantness. Loudness can also be used to describe metro train noise, though being a bit less efficient than sharpness. Roughness and fluctuation strength showed the worst results. Both emotional parameters, stress and pleasantness, although not computable and measurable, can be quite useful in describing metro train noise. Analysis of the two signals, that were selected from the same site (signals 4 and 5), has shown that sharpness values can give some additional information about the train noise at the sites, which have slopes and/or differentiate between the noise coming from the trains which ride uphill and downhill. Statistical analysis, especially the correlation between the different parameters, revealed that the nature of the signals is crucial and should be considered when talking about train noise evaluation since some of the signals show similar results despite the parameter that is being used.

Bibliography

[1] Hoffmann, Alice, *Auralization, Perception And Detection Of Tyre-Road Noise*, Thesis For The Degree Of Doctor Of Philosophy, Gothenburg, Sweden, 2016.

[2] Sørgjerd, Christian, *Her er byrådets plan: Går for T-banestasjoner på Bislett og Grünerløkka*, `https://www.aftenposten.no/osloby/i/A2nQbE/Her-er-byradets-plan-Gar-for-T-banestasjoner-pa-Bislett-og-Grnerlokka`, June 16, 2019

[3] Fastl, Hugo, *Psychoacoustics and Sound Quality*, Technical-Acoustics Group, Department of Human-Machine-Communication, Technical University of Munich, Munich.

[4] Lercher, Peter, *Psychoacoustic assessment of railway noise in sensitive areas and times: is a rail bonus still appropriate?*, `https://www.researchgate.net/publication/260276062_Psychoacoustic_assessment_of_railway_noise_in_sensitive_areas_and_times_is_a_rail_bonus_still_appropriate`, June 16, 2019

[5] Christian, H. Kasess, *The relation between psychoacoustical factors and annoyance under different noise reduction conditions for railway noise*, `https://asa.scitation.org/doi/10.1121/1.4982878`, June 16, 2019

[6] Patsouras, Christine, *Psychoacoustic evaluation of tonal components in view of sound quality design for high-speed train interior noise*, Acoustical Letter, September 4, 2001.

[7] Fastl, Hugo, Zwicker, Eberhard, *Psychoacoustics*, Third Edition, Springer, 2007.

[8] Wikipedia contributors, *Noise pollution*, Wikipedia, The Free Encyclopedia, `https://en.wikipedia.org/wiki/Noise_pollution`, June 16, 2019

[9] De Vos, *Paul, Railway Noise*, Chapter 5, Licitra, 2013.

[10] De Vos, Paul, *Railway noise in Europe*, State of the art report, `https://uic.org/IMG/pdf/railway_noise_in_europe_2016_final.pdf`, June 16, 2019

[11] European Environment Agency (EEA), *Noise in Europe 2014*, https://www.eea.europa.eu/publications/noise-in-europe-2014, June 16, 2019

[12] Clausen, Uwe, *Reducing Railway Noise Pollution*, Directorate General For Internal Policies, Transport and Tourism, European Parliament, http://www.europarl.europa.eu/RegData/etudes/etudes/join/2012/474533/IPOL-TRAN_ET(2012)474533_EN.pdf, June 16, 2019

[13] Purves, Dale, *Neuroscience* Chapter 12, The auditory system, 3rd edition, Sunderland, MA:Sinauer Associates Inc., 2004.

[14] ISO 3095:2013(E), *Acoustics - Railway applications - Measurement of noise emitted by railbound vehicles*

[15] Timeanddate, *Weather in Oslo, Norway*, https://www.timeanddate.com/weather/, June 16, 2019

[16] Merriam–Webster's Dictionary, *Psychoacoustics*, https://www.merriam-webster.com/dictionary/psychoacoustics, June 16, 2019

[17] Garrett, Steven L., *Understanding Acoustics*, An Experimentalist's View of Acoustics and Vibration, Springer, 2017.

[18] HEAD acoustics, *Loudness and Sharpness Calculations*, Application note, https://www.head-acoustics.com/downloads/eng/application_notes/Psychoacoustic_Analyses_I_e.pdf, June 16, 2019

[19] Daniel, Peter, *Psychoacoustical Roughness*, Bruel & Kjær, GmbH, Bremen, Germany, https://link.springer.com/chapter/10.1007/978-0-387-30441-0_19, June 16, 2019

[20] Merriam–Webster's Dictionary, *Stress*, https://www.merriam-webster.com/dictionary/stress, June 16, 2019

[21] Nordtest, *Nordtest Method*, Acoustics: Human Sound Perception – Guidelines For Listening Tests, May, 2002.

[22] Merriam–Webster's Dictionary, *Semantic*, https://www.merriam-webster.com/dictionary/semantic, June 16, 2019

[23] Susini, Patrick, *Psychological Measurement for Sound Description and Evaluation*, https://www.researchgate.net/publication/281985123_Psychological_Measurement_for_Sound_Description_and_Evaluation, June 16, 2019

[24] Haldsrud, Stian, *Fremtidens vogner*, Sporveien, Sporveisforlaget, 2016.

[25] Wikipedia contributors, *History of the Oslo Tramway and Metro*, Wikipedia, The Free Encyclopedia, https://en.wikipedia.org/wiki/History_of_the_Oslo_Tramway_and_Metro, June 16, 2019

[26] J Mackenzie, Ruairi, *One-Way vs Two-Way ANOVA: Differences, Assumptions and Hypotheses*, https://www.technologynetworks.com/informatics/articles/one-way-vs-two-way-anova-definition-differences-assumptions-and-hypotheses-306553, June 16, 2019

[27] Unknown author, *What is a 1-Way ANOVA?*, http://statistics-help-for-students.com/What_is_a_1_Way_ANOVA.htm#.XPWj69MzbfZ, June 16, 2019

[28] Unknown author, *ANOVA – Simple Introduction*, https://www.spss-tutorials.com/anova-what-is-it/, June 16, 2019

[29] Unknown author, *ANOVA Test: Definition, Types, Examples*, https://www.statisticshowto.datasciencecentral.com/probability-and-statistics/hypothesis-testing/anova/, June 16, 2019

[30] Mathworks, *anova1*, https://se.mathworks.com/help/stats/anova1.html, June 16, 2019

[31] Embleton, Tony, *Tutorial on sound propagation outdoors*, The Journal of the Acoustical Society of America 100, 31 (1996); doi: 10.1121/1.415879

[32] HEAD acoustics, *Binaural Measurement, Analysis and Playback*, Application note, https://www.head-acoustics.com/downloads/eng/application_notes/BinauralMeasurement_e.pdf, June 16, 2019

[33] Wikipedia contributors, *Box plot*, Wikipedia, The Free Encyclopedia, https://en.wikipedia.org/wiki/Box_plot, June 16, 2019

[34] Mathworks, *boxplot*, https://se.mathworks.com/help/stats/boxplot.html, June 16, 2019

[35] Mathworks, *lsline*, https://se.mathworks.com/help/stats/lsline.html, June 16, 2019

[36] Minitab Blog Editor, *Regression Analysis: How Do I Interpret R-squared and Assess the Goodness-of-Fit?*, The Minitab Blog, https://blog.minitab.com/blog/adventures-in-statistics-2/regression-analysis-how-do-i-interpret-r-squared-and-assess-the-goodness-of-fit, June 16, 2019

[37] Frost, Jim, *Standard Error of the Regression vs. R-squared*, Statistics By Jim, https://statisticsbyjim.com/regression/standard-error-regression-vs-r-squared/, June 16, 2019

Appendix

6.1 Resistivity

TABLE I. Characterization of various ground surfaces. The values of effective flow resistivity were obtained by matching transmission spectra measured *in situ* with spectra predicted theoretically [from Ref. 24].

Description of surface	Effective flow resistivity kPa s/m^2
0.1-m new fallen snow, over older snow	10–30
sugar snow	25–50
floor of evergreen forest	20–80
airport grass or old pasture	150–300
roadside dirt, ill-defined, small rocks up to 0.01-m mesh	300–800
sandy silt, hard packed by vehicles	800–2500
thick layer of clean limestone chips, 0.01- to 0.025-m mesh	1500–4000
old dirt roadway, small stones with interstices filled by dust	2000–4000
earth, exposed and rain-packed	4000–8000
very fine quarry dust, hard packed by vehicles	5000–20 000
asphalt, sealed by dust and use	~30 000
upper limit, set by thermal conduction and viscosity	2×10^5 to 1×10^6

6.2 Introduction to the listening test

Velkommen til lyttetest!

Hovedmålet med testen er å finne ut hvordan folk oppfatter støy fra T-banen i Oslo ved å vurdere målte lyder basert på ulike psykoakustiske parametere. Oppsettet for testen består av en bærbar PC og hodetelefoner. Dette er en digital test programmert ved hjelp av Matlab.
Oppgaven for deltakeren er å vurdere seks psykoakustiske parametere for fem ulike lydsignaler, via en lydspiller i Matlab. Testen består av to forsøk, og har en varighet på ca. 30 minutter. Det finnes ingen riktige svar, bare dine spontane avgjørelser.

Følgende uttalelser blir brukt under testen for hver av de seks parameterne:

Psychoacoustic parameter	Statement	Scale
Loudness	Lydnivået er 'høyt'	uenig/enig
Sharpness	Lyden er 'skarp'	uenig/enig
Roughness	Lyden har 'ruhet'i seg	uenig/enig
Fluctuation Strength	Lyden er 'fluktuerende'	uenig/enig
Annoyance/Pleasantness	Lyden er 'trivelig'	uenig/enig
Stress	Lyden er 'stressende'	uenig/enig

Lydspilleren brukes på følgende måte:
Trykk på «SPILL» for å starte avspilling.
Trykk på «STOPP» for å stoppe avspilling.
Gi vurdering til parameteren ved å velge ett tall fra 1 til 7.
Du kan skrive kommentarer i kommentarfeltet hvis det er ønskelig.
Trykk på «NESTE» for å gå videre til neste lydfil.
Obs! Du kan spille et signal flere ganger hvis det er ønskelig, men hele signalet må spilles av minst én gang for å kunne fortsette til neste lydfil.

Eksempel på lyttetesten, screenshot fra eksperimentet

Dine rettigheter som deltaker: Du kan stoppe testen når som helst for å stille tilleggsspørsmål eller slappe av, og avslutte testen uten begrunnelser. Lydtrykksnivået er justert slik at testen ikke skal føre til hørselskader. Testen er anonym. Data som blir lagret ila testen kan brukes for videre forskning.

Jeg har lest og aksepterer vilkårene .... (Signatur)

Figure 6.1: Introduction to the listening test in Norwegian

6.3 Table of statements in English

Table 6.1: Table of statements

Psychoacoustic parameter	Statement	Scale
Loudness	Sound level is 'high'	agree/disagree
Sharpness	Sound is 'sharp'	agree/disagree
Roughness	Sound is 'rough'	agree/disagree
Fluctuation Strength	Sound is 'fluctuating'	agree/disagree
Stress	Sound is 'stressful'	agree/disagree
Annoyance/Pleasantness	Sound is 'pleasant'	agree/disagree

6.4 Headphone frequency response

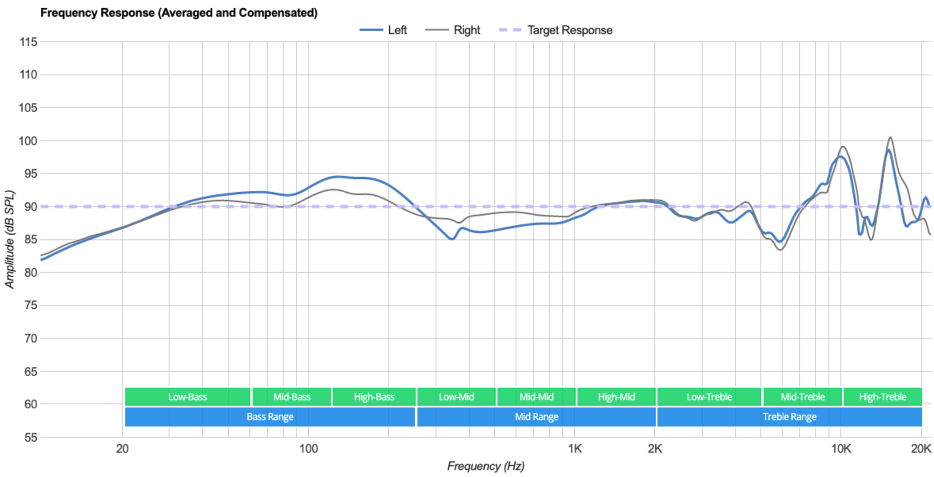


Figure 6.2: Headphone ATH-M40x frequency response(left and right channel), retrieved from `www.rtings.com`

6.5 Raw data from the measurement sites

11.02.19

Voksenlia

IN/OUT	TRAIN #	FILE	SPEED	# OF CARS	COMMENTS
IN	70	338	-	3	Error on channels
OUT	55	339	-	3	Error on channels
-	-	341	-	-	Background noise
OUT/IN	98/55	342	-	3/3	Double passing
IN	08	343	7.47	3	
OUT	80	344	7.22	3	
IN	98	345	-	3	No speed measurement
OUT	69	346	7.21	3	
IN	80	347	-	3	No speed measurement
OUT	40	348	7.47	3	
IN	69	349	6.94	3	
OUT	42	350	7.13	3	
IN/OUT	40/70	351	7.28/7.13	3	Double passing, no signal in Left
OUT	42	353	7.03	3	
IN	75	354	7.00	3	Car pass
-	-	356	-	-	Background noise

12.02.19

Dalbakkveien

IN/OUT	TRAIN #	FILE	SPEED	# OF CARS	COMMENTS
OUT	09	358	5.47	6	
IN	36	359	5.19	3	
-	-	360	-	-	Background noise
OUT	06	361	3.04	3	
IN	11	362	-	6	No speed measurement
-	-	364	-	-	Background noise
OUT	52	365	6.00	6	
IN	78	366	5.87	6	
OUT	06	367	6.00	6	
IN	06	368	3.06	3	
OUT	14	369	2.93	3	
IN	82	370	5.87	6	
OUT	22	371	6.78	6	
OUT	03	372	5.87	6	
IN	02	373	7.09	6	
IN	36	374	3.00	3	
OUT	06	375	2.57	3	
IN	09	376	3.71	3	
OUT	98	377	5.67	6	

25.02.19

Voksenlia, second measurement

IN/OUT	TRAIN #	FILE	SPEED	# OF CARS	COMMENTS
IN	30	357	6.81	3	Car passing close
OUT	76	358	7.07	3	Passed very close to previous
-	-	359	-	-	Background noise
IN	67	360	6.97	3	Birds chirping
OUT	31	361	7.00	3	Person with dog passing
-	-	362	-	-	Background noise
OUT	26	363	6.94	3	
IN	76	364	7.16	3	
-	-	365	-	-	Recording disrupted
OUT	47	366	7.28	3	
IN	26	367	6.97	3	
OUT	04	368	7.00	3	
IN	95	369	7.00	3	

25.02.19

Borgen

IN/OUT	TRAIN #	FILE	SPEED	# OF CARS	COMMENTS
IN	33	371	8.38	6	
OUT	71/52	372	7.88	6	Passed very close to previous
-	-	373	-	-	Background noise
OUT	72/58	374	8.60	6	
IN	62/85	375	7.69	6	
IN	79/40	378	7.85	6	
OUT	28/98	380	6.91	6	
-	-	381	-	-	Background noise
IN	25/100	382	7.78	6	
OUT	44	383	3.84	3	
IN	02/65	384	6.81	6	
IN	68/56	385	8.28	6	
IN	55/77	386	8.34	6	
OUT	54/09	388	7.84	6	
-	-	389	-	-	Background noise
IN	94/04	390	8.25	6	
OUT	78	391	3.84	3	
OUT	21/82	392	8.13	6	
IN	57/71	393	8.31	6	

28.02.19

Tjensrud

IN/OUT	TRAIN #	FILE	SPEED	# OF CARS	COMMENTS
-	-	357	-	-	Background noise
OUT	78/07	358	8.25	6	
IN	06/14	359	8.10	6	
OUT	19/58	360	8.22	6	Children playing, airplane
IN	79/38	361	8.16	6	
OUT	01/39	363	8.69	6	
OUT	02/61	364	9.72	6	
IN	23/14	365	8.00	6	
OUT	05/27	366	8.22	6	
IN	39/01	367	8.12	6	
OUT	83/64	368	8.12	6	
IN	61/02	369	8.10	6	

6.6 F-table

		Degrees of freedom in the numerator								
Degrees of freedom in the denominator	p	1	2	3	4	5	6	7	8	9
28	.100	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87
	.050	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24
	.025	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61
	.010	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12
	.001	13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.69	4.50
29	.100	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86
	.050	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22
	.025	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59
	.010	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09
	.001	13.39	8.85	7.12	6.19	5.59	5.18	4.87	4.64	4.45
30	.100	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85
	.050	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21
	.025	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57
	.010	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07
	.001	13.29	8.77	7.05	6.12	5.53	5.12	4.82	4.58	4.39
40	.100	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79
	.050	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12
	.025	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45
	.010	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89
	.001	12.61	8.25	6.59	5.70	5.13	4.73	4.44	4.21	4.02
50	.100	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76
	.050	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07
	.025	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38
	.010	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78
	.001	12.22	7.96	6.34	5.46	4.90	4.51	4.22	4.00	3.82
60	.100	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74
	.050	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04
	.025	5.29	3.93	3.34	3.01	2.79	2.63	2.51	2.41	2.33
	.010	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72
	.001	11.97	7.77	6.17	5.31	4.76	4.37	4.09	3.86	3.69
100	.100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.69
	.050	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97
	.025	5.18	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.24
	.010	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59
	.001	11.50	7.41	5.86	5.02	4.48	4.11	3.83	3.61	3.44
200	.100	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66
	.050	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93
	.025	5.10	3.76	3.18	2.85	2.63	2.47	2.35	2.26	2.18
	.010	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50
	.001	11.15	7.15	5.63	4.81	4.29	3.92	3.65	3.43	3.26

Figure 6.3: F-table, retrieved from `http://www.stat.purdue.edu`

6.7 Technical drawings of metro cars

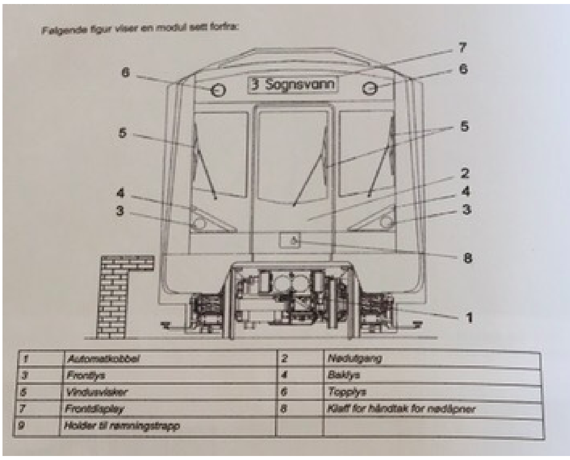


Figure 6.4: Frontal view of metro train, retrieved from Haldsrud, S. (2016). Fremtidens vogner. Sporveisforlaget

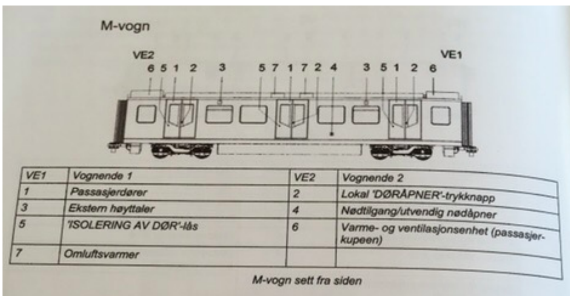
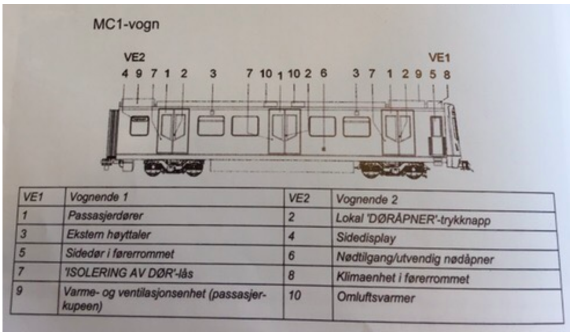
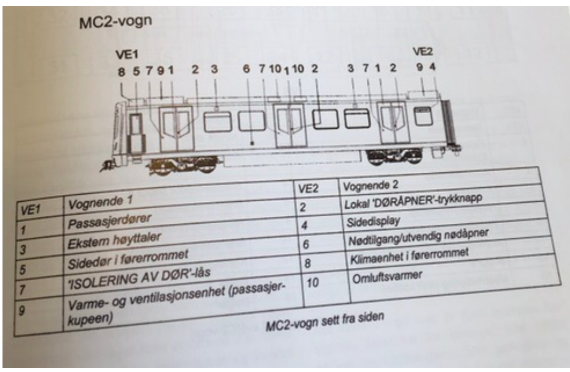


Figure 6.5: Side view of M car, retrieved from Haldsrud, S. (2016). Fremtidens vogner. Sporveisforlaget



VE1	Vognende 1	VE2	Vognende 2
1	Passasjerdører	2	Lokal 'DØRÅPNER'-trykknapp
3	Ekstern høyttaler	4	Sidedisplay
5	Sidedør i førerrommet	6	Nødtilgang/utvendig nødåpner
7	'ISOLERING AV DØR'-lås	8	Klimaenhet i førerrommet
9	Varme- og ventilasjonsenhet (passasjer-kupeen)	10	Omluftsvarmer

Figure 6.6: Side view of MC1 car, retrieved from Haldsrud, S. (2016). Fremtidens vogner. Sporveisforlaget



VE1	Vognende 1	VE2	Vognende 2
1	Passasjerdører	2	Lokal 'DØRÅPNER'-trykknapp
3	Ekstern høyttaler	4	Sidedisplay
5	Sidedør i førerrommet	6	Nødtilgang/utvendig nødåpner
7	'ISOLERING AV DØR'-lås	8	Klimaenhet i førerrommet
9	Varme- og ventilasjonsenhet (passasjer-kupeen)	10	Omluftsvarmer

Figure 6.7: Side view of MC2 car, retrieved from Haldsrud, S. (2016). Fremtidens vogner. Sporveisforlaget

6.8 Matlab code

Statistical analysis

```
vector_tot = {'AB1.csv'; 'AB2.csv'; 'And1.csv'; 'And2.csv'; 'AndersB1.csv'; 'AndersB2.csv';...
    'DSH1.csv';'DSH2.csv';'GA1.csv';'GA2.csv';'GB1.csv';'GB2.csv'; ...
    'HAA1.csv';'HAA2.csv';'HJ1.csv';'HJ2.csv';'JT1.csv';'JT2.csv'; ...
    'Kar1.csv';'Kar2.csv'; 'KH1.csv'; 'KH2.csv'; 'LP1.csv'; 'LP2.csv'; 'MM1.csv'; 'MM2.csv'; ...
    'OJ1.csv'; 'OJ2.csv';'PC1.csv'; 'PC2.csv'; 'SO1.csv'; 'SO2.csv'; ...
    'Rob1.csv'; 'Rob2.csv';'TB1.csv'; 'TB2.csv'; ...
    'TFE1.csv'; 'TFE2.csv'; 'Theres1.csv';'Theres2.csv'; 'Tru1.csv'; 'Tru2.csv'};


m = 0;
M_sum = zeros;
M_tot = zeros(30,1);
for k = 1:numel(vector_tot)
    m = m + 1;

    data{k} = readtable(vector_tot{k});
    data{k} = sortrows(data{k}(:,[2 3 4]),2); %sort by Question/parameter(2nd column)
    M = table2array(data{k});
    M_sorted_5 = sortrows(M(1:5,:),1); %sort by Signal/places(1st column)
    M_sorted_10 = sortrows(M(6:10,:),1);
    M_sorted_15 = sortrows(M(11:15,:),1);
    M_sorted_20 = sortrows(M(16:20,:),1);
    M_sorted_25 = sortrows(M(21:25,:),1);
    M_sorted_30 = sortrows(M(26:30,:),1);
    M = [M_sorted_5; M_sorted_10; M_sorted_15; M_sorted_20; M_sorted_25; M_sorted_30;];
    %Avg
    M_sum = M_sum + M;
    if m == 42
        M_avg = M_sum/42;
    end

    %for Std
    M_vect{k} = M;
```

```

    %for box plot 42 signals
     M_tot = [M_tot M(:,3)];
end
%Std
summe_new = zeros(30,3);
for n = 1:numel(vector_tot)
    summe = (M_vect{n} - M_avg).^2;
    summe_new = summe_new + summe;
end
standr = sqrt(summe_new./41);

%box for 42 signals
M_tot = M_tot(:,2:43);

%box for 21 signals (participants)
M_box = zeros(30,1);
rpt = 1;
for g = 1:21
     M_average = (M_tot(:,rpt) + M_tot(:,rpt+1))/2;
     M_box = [M_box M_average];
     rpt = rpt + 2;
end
M_box = M_box(:,2:22);

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%Plot avg
M_avg(:,2) = [1.02 0.98 1 1 1 2 2 2 2 2 3 3 3 3 3 4 4 4 4 4
     5 5 5 5.05 4.95 6.02 6 5.98 6 6];
colorstring = {[1 0 0]; [0 0 1];[0 1 1]; [1 0 1]; [0 1 0]};

figure(1); cla;
hold on
grid on
z = 0;
for n = 1:6

     for i = 1:5
     plot(M_avg(i+z,2),M_avg(i+z,3),'X','linewidth',2.5,'
         MarkerSize',25,'MarkerFaceColor',colorstring{i},'
         MarkerEdgeColor',colorstring{i},'color',colorstring{
         i})
     end
     z = z + 5;
end
```

```
legend('Sig 1','Sig 2','Sig 3','Sig 4','Sig 5')
set(gca,'FontSize',20)
xlim([0 7])
ylim([0 8])
names = {'loud'; 'sharp'; 'rough'; 'pleasant'; 'fluctuating'; 'stressful'};
numbers = {1 2 3 4 5 6 7};
set(gca,'xtick',1:6,'xticklabel',names,'ytick',1:7,'YTickLabel',numbers)
xlabel('Parameters')
ylabel('1:do not agree - 7:agree')
title('Averaged listening test results')


%Plot std
titles = {'Loudness'; 'Sharpness'; 'Roughness'; 'Pleasantness'; 'Fluctuation'; 'Stress'};
figure(2); cla;
z = 0;
for n = 1:6
subplot(2,3,n);
for i = 1:5
    errorbar(M_avg(i+z,1),M_avg(i+z,3),standr(i+z,3),'X','linewidth',2,'MarkerSize',15,'MarkerFaceColor',colorstring{i},'MarkerEdgeColor',colorstring{i},'color',colorstring{i})
    hold on
end
z = z + 5;
grid on;
set(gca,'FontSize',16)
xlim([0 6])
ylim([0 8])
names = {'Sig 1'; 'Sig 2'; 'Sig 3'; 'Sig 4'; 'Sig 5'};
numbers = {1 2 3 4 5 6 7};
set(gca,'xtick',1:5,'xticklabel',names,'ytick',1:7,'YTickLabel',numbers)
ylabel('1:do not agree - 7:agree')
title(titles{n})
end

%Plot box 42 signals
% titles = {'Loudness'; 'Sharpness'; 'Roughness'; 'Pleasantness'; 'Fluctuation'; 'Stress'};
% figure(3); cla;
```

```
% z = 0;
% for n = 1:6
% subplot(2,3,n);
%     boxplot((M_tot((1+z):(5+z),:))')
% z = z + 5;
% grid on;
% set(gca,'FontSize',16)
% xlim([0 6])
% ylim([0 8])
% names = {'Sig 1'; 'Sig 2'; 'Sig 3'; 'Sig 4'; 'Sig 5'};
% numbers = {1 2 3 4 5 6 7};
% set(gca,'xtick',1:5,'xticklabel',names,'ytick',1:7,'
    YTickLabel',numbers)
% ylabel('1:do not agree - 7:agree')
% title(titles{n})
% end

%Plot box 21 signal
titles = {'Loudness'; 'Sharpness'; 'Roughness'; '
    Pleasantness'; 'Fluctuation'; 'Stress'};
figure(4); cla;
z = 0;
for n = 1:6
subplot(2,3,n);
    boxplot((M_box((1+z):(5+z),:))')
z = z + 5;
grid on;
set(gca,'FontSize',16)
xlim([0 6])
ylim([0 8])
names = {'Sig 1'; 'Sig 2'; 'Sig 3'; 'Sig 4'; 'Sig 5'};
numbers = {1 2 3 4 5 6 7};
set(gca,'xtick',1:5,'xticklabel',names,'ytick',1:7,'
    YTickLabel',numbers)
ylabel('1:do not agree - 7:agree')
title(titles{n})
end
%
%Plot comparison
loud = M_avg(1:5,3);
sharp = M_avg(6:10,3);
rough = M_avg(11:15,3);
fluc = M_avg(21:25,3);
%Channel 3 vectors
l_3 = [24.2 26.7 23.8 37.7 29.5];
```

```
s_3 = [1.3  1.38  1.4  2.91  3.03];
r_3 = [0.0286  0.0327  0.0305  0.0293  0.0279];
f_3 = [0.0331  0.0183  0.0164  0.0253  0.0268];

labels = cellstr({'1','2','3','4','5'});

figure(5); cla;
%loudness
subplot(2,2,1);
c_l = linspace(1,10,length(l_3(1:5)));
scatter(l_3(1:5),loud(1:5,:),250,c_l,'*')
set(gca,'FontSize',20)
xlim([20  40])
ylim([1  7])
numbers = {1 2 3 4 5 6 7};
set(gca,'ytick',1:7,'YTickLabel',numbers)
xlabel('ArtemiS loudness [sone]')
ylabel('1:do not agree - 7:agree')
h_l = lsline(subplot(2,2,1));
h_l.Color = 'r';
text(l_3(1:5),loud(1:5,:),labels,'VerticalAlignment','
    bottom',...
    'HorizontalAlignment','right','FontSize',25)

%sharpness
subplot(2,2,2);
c_s = linspace(1,10,length(s_3(1:5)));
scatter(s_3(1:5),sharp(1:5,:),250,c_s,'*')
set(gca,'FontSize',20)
xlim([1  3.5])
ylim([1  7])
numbers = {1 2 3 4 5 6 7};
set(gca,'ytick',1:7,'YTickLabel',numbers)
xlabel('ArtemiS sharpness [acum]')
ylabel('1:do not agree - 7:agree')
h_s = lsline(subplot(2,2,2));
h_s.Color = 'r';
text(s_3(1:5),sharp(1:5,:),labels,'VerticalAlignment','
    bottom',...
    'HorizontalAlignment','right','FontSize',25)

%roughness
subplot(2,2,3);
c_r = linspace(1,10,length(r_3(1:5)));
scatter(r_3(1:5),rough(1:5,:),250,c_r,'*')
```

```
set(gca,'FontSize',20)
xlim([0.025 0.035])
ylim([1 7])
numbers = {1 2 3 4 5 6 7};
set(gca,'ytick',1:7,'YTickLabel',numbers)
xlabel('ArtemiS roughness [asper]')
ylabel('1:do not agree - 7:agree')
h_r = lsline(subplot(2,2,3));
h_r.Color = 'r';
text(r_3(1:5),rough(1:5,:),labels,'VerticalAlignment','
    bottom',...
     'HorizontalAlignment','right','FontSize',25)

%fluctuation strength
subplot(2,2,4);
c_f = linspace(1,10,length(f_3(1:5)));
scatter(f_3(1:5),fluc(1:5,:),250,c_f,'*')
set(gca,'FontSize',20)
xlim([0.015 0.035])
ylim([1 7])
numbers = {1 2 3 4 5 6 7};
set(gca,'ytick',1:7,'YTickLabel',numbers)
xlabel('ArtemiS fluctuation strength [vacil]')
ylabel('1:do not agree - 7:agree')
h_f = lsline(subplot(2,2,4));
h_f.Color = 'r';
%legend(c_f,{'Sig 1','Sig 2','Sig 3','Sig 4','Sig 5'})
%,'FontWeight','bold'
text(f_3(1:5),fluc(1:5,:),labels,'VerticalAlignment','
    bottom',...
     'HorizontalAlignment','right','FontSize',25)


%%Plot ANOVA1
loudness
loud_a1 = (M_box(1:5,:));
[p1,t1,s1]=anova1(loud_a1')
figure
[c1,m1,h1,nms1] = multcompare(s1);
ethasqare1an= (t1{2,2})/ (t1{4,2})
partethasqare1an= (t1{2,5}*t1{2,3})/ (t1{2,5}*t1{2,3}+t1
    {3,3})

%sharpness
sharp_a1 = (M_box(6:10,:));
```

```
[p2,t2,s2]=anova1(sharp_a1')
figure
[c2,m2,h2,nms2] = multcompare(s2);
ethasqare1sc= (t2{2,2})/ (t2{4,2})
partethasqare1sc= (t2{2,5}*t2{2,3})/ (t2{2,5}*t2{2,3}+t2
    {3,3})

%roughness
rough_a1 = (M_box(11:15,:));
[p3,t3,s3]=anova1(rough_a1')
figure
[c3,m3,h3,nms3] = multcompare(s3);
ethasqare1la= (t3{2,2})/ (t3{4,2})
partethasqare1la= (t3{2,5}*t3{2,3})/ (t3{2,5}*t3{2,3}+t3
    {3,3})

%Pleasanatness
pleasant_a1 = (M_box(16:20,:));
[p4,t4,s4]=anova1(pleasant_a1')
figure
[c4,m4,h4,nms4] = multcompare(s4);
ethasqare1ra= (t4{2,2})/ (t4{4,2})
partethasqare1ra= (t4{2,5}*t4{2,3})/ (t4{2,5}*t4{2,3}+t4
    {3,3})

%F.S.
fs_a1 = (M_box(21:25,:));
[p5,t5,s5]=anova1(fs_a1')
figure
[c5,m5,h5,nms5] = multcompare(s5);
ethasqare1pi= (t5{2,2})/ (t5{4,2})
partethasqare1pi= (t5{2,5}*t5{2,3})/ (t5{2,5}*t5{2,3}+t5
    {3,3})

%Stress
stress_a1 = (M_box(26:30,:));
[p6,t6,s6]=anova1(stress_a1')
figure
[c6,m6,h6,nms6] = multcompare(s6);
ethasqare1st= (t6{2,2})/ (t6{4,2})
partethasqare1st= (t6{2,5}*t6{2,3})/ (t6{2,5}*t6{2,3}+t6
    {3,3})


%%Compute correlation
```

```
M_avg_corr = [M_avg(1:5,3) M_avg(6:10,3) M_avg(11:15,3)
    M_avg(16:20,3) M_avg(21:25,3) M_avg(26:30,3)]';
%Loudness L
 for i = 2:6
L{i}=[M_avg_corr(1,:); M_avg_corr(i,:)];
 end

%Sharpness S
 for i = 3:6
S{i}=[M_avg_corr(2,:); M_avg_corr(i,:)];
 end

%Roughness R
 for i = 4:6
R{i}=[M_avg_corr(3,:); M_avg_corr(i,:)];
 end

%Pleasantness P
 for i = 5:6
P{i}=[M_avg_corr(4,:); M_avg_corr(i,:)];
 end

%FS with Stress
FS=[M_avg_corr(5,:); M_avg_corr(6,:)];

%L corr
 [R_ls,P_ls]=corrcoef(L{2}')
 [R_lr,P_lr]=corrcoef(L{3}')
 [R_lp,P_lp]=corrcoef(L{4}')
 [R_lf,P_lf]=corrcoef(L{5}')
 [R_lst,P_lst]=corrcoef(L{6}')
%S corr
 [R_sr,P_sr]=corrcoef(S{3}')
 [R_sp,P_sp]=corrcoef(S{4}')
 [R_sf,P_sf]=corrcoef(S{5}')
 [R_sst,P_sst]=corrcoef(S{6}')
%R corr
 [R_rp,P_rp]=corrcoef(R{4}')
 [R_rf,P_rf]=corrcoef(R{5}')
 [R_rst,P_rst]=corrcoef(R{6}')
%P corr
 [R_pf,P_pf]=corrcoef(P{5}')
 [R_pst,P_pst]=corrcoef(P{6}')
%FS with Stress corr
```

```
[R_fst, P_fst]=corrcoef(FS')
```