CHALMERS

Measurement, Judgment and Prediction of the Annoyance of the Noise Produced by Seat Belt Retractors

Master's Thesis in Master Programme Sound and Vibration

ZELIN ZONG

Department of Architecture and Civil Engineering
Division of Applied Acoustics
Vibroacoustics Group
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2019
Master's Thesis: ACEX30-19-19

MASTER'S THESIS: ACEX30-19-19

Measurement, Judgment and Prediction of the Annoyance of the Noise Produced by Seat Belt Retractors

Zelin Zong

Department of Architecture and Civil Engineering

Division of Applied Acoustics

Vibroacoustics Group

CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2019

Measurement, Judgment and Prediction of the Annoyance of the Noise Produced by Seat Belt Retractors

Master's Thesis: ACEX30-19-19

Department of Architecture and Civil Engineering

Division of Applied Acoustics

Vibroacoustics Group

Chalmers University of Technology

SE-41296 Göteborg

Sweden

Tel. +46-(0)31 772 1000

Measurement, Judgment and Prediction of the Annoyance of the Noise Produced by Seat Belt Retractors

ZELIN ZONG

Department of Civil and Environmental Engineering

Division of Applied Acoustics

Chalmers University of Technology

Abstract

The automobile industry has grown sustainedly over the past decade. To enhance competitiveness, user experience and product details increasingly receive attention. Squeak and Rattle (S&R) are noises caused by the relative motion between installed components, in terms of friction and impact. The perceived annoyance of squeak and rattle noise from seat belt retractors, as might be expected, is nowhere near the sense of annoyance from other more significant sources, for instance, tires and engine. However, such a sort of objective noise is still audible and might be more predominant in future cars which have much better sound isolation.

In this thesis, three types of seat belt retractors are studied. They are both practically used inside Volvo cars. The primary objective is to design a sound quality metric for the squeak and rattle noise from seat belt retractors.

While recording the sounds, different recording environments, including laboratories and actual in-car environments, are seriatim tested. In order to obtain the perceived annoyance level of objective noises, a series of listening tests are performed in the form of MATLAB graphical user interface (GUI), with participants from different backgrounds involved. The listening test results are then combined with calculated psychoacoustic characters of sounds, thereby designing the corresponding sound quality metric using multiple linear regression.

Keywords: Seat Belt Retractors, Squeak and Rattle, Psychoacoustics, Sound Quality Metrics, Listening Tests, Graphical User Interface, Multiple Linear Regression

Acknowledgments

I would first like to thank my supervisor and examiner Jens Ahrens at the division of applied acoustics of Chalmers University of Technology. His door is always open to me for questions and requests. He also enlightened me with the application of MATLAB GUI, which broadened my programming skills to a great extent.

Besides, I feel really grateful to Anneli Rosell, my supervisor at Volvo Car Corporation. She helped me quite a lot throughout the whole process of prototype design, practical implementation and scheme improvement.

In addition, I would like to show great appreciation to Mohsen Bayani, who is also my supervisor at Volvo Car Corporation. His patient guidance can always deepen my understanding with respect to psychoacoustics.

Moreover, I would specially thank all the participants of my listening tests, including my colleagues, classmates and friends. The user study would never be possible to accomplish without their participation.

And finally, I must express my gratitude to my parents and my family. They have been supporting my decision all the time.

Zelin Zong

Gothenburg, April 2019

Contents

1. Introduction

The global automobile industry has been maintaining growth in the past decade. The worldwide automobile production increased by 1.5 times, from 61.8 to 91.5 million vehicles, since the 2008 financial crisis. Meanwhile, passenger car sales are also expected to continue to grow in 2019 [1]. With such a positive market expectation as well as the impacts of new technologies, for instance, the new energy automobiles and the self-driving technology, the established automobile companies spontaneously focus on new technology research and development, to better face the volatile market challenges and enhance their product competitiveness. Nevertheless, it is equally important to review the issued vehicles and try to improve the user experience, since no technology is entirely isolated. Any tiny improvements may receive recognition in the future.

Volvo cars has introduced a system named Squeak & Rattle (S&R) to both acoustically and mechanically evaluate and verify the passengers' experience. Absence of squeak and rattle is of particular importance to premium cars. Squeak and rattle are caused by the relative motion between installed parts, which make them interact, sometimes only at certain environmental conditions. Basically, S&R issues are identified very late in the production cycle, some even after the vehicle is launched. [2]

Due to the design concept of the mechanical mechanism of a seat belt retractor, it can generate rattle noise when the car is running at relatively high speed or on an uneven road surface. The rattle is essentially caused by the impact between a metallic ball and its rigid plastic cover. Even if the sound levels from seat belt retractors are low, the rattle noise is usually audible since it is generated intermittently and has a frequency content where the sound masking is relatively low [3].

In this study, three different types of rear-seat belt retractors are measured and evaluated. For ease of description, they are called seat belt retractor type A, B and C, respectively, hereinafter.

1.1. Aims

The aim of this study are listed as follows:

- To record the noise from various types of seat belt retractors under different recording environments.

- To evaluate the records according to the methodology from psychoacoustics (calculating psychoacoustic parameters).

- To design and conduct a series of corresponding listening tests to obtain subjective evaluations on perceived annoyance.

- To associate the perceived annoyance with psychoacoustic parameters, thereby forming a sound quality metric.

Meanwhile, there are some certain challenges:

- The quality of audio recordings should be highly regarded since it can dramatically change the judgments in the listening test.

- Determinants of perceived annoyance are various and complex.

- The difference in subjective evaluations of perceived annoyance is tremendous among individuals.

1.2. Limitations

The records used in the listening test are recorded under laboratory environments, which makes the rattle noise sound different from the real experience. Since this study concentrates on the evaluation of a specific component, which is the seat belt retractor in this case, the interferences from other noises are deliberately avoided.

The equipment for audio playback will more or less affect the fidelity, thus leading to differences in perceived characteristics of the sound. Meanwhile, the metric is derived from the calculated psychoacoustic parameters. In addition, the site where the listening tests take place is different from the actual environment inside a car. During a listening test, the sense of space can also have an influence on the judgment.

2. Theory

In this section, the relevant theories about the calculation of psychoacoustic parameters are presented. It should be noted that significant differences can appear when different methods are applied. In this study, the methods from DIN 45631/A1 and DIN 45692 are used to calculate loudness and sharpness, respectively. While roughness is calculated by Zwicker's method in 1 Bark resolution.

2.1. Calculation of loudness

Loudness is the subjective perception of sound pressure. It is dependent on the frequency content of a sound. Because of the effect of the human ear, the relation between sound pressure level and loudness is nonlinear. Loudness is usually referred to as N, and a widely-used unit of it is Sone. One sone is the loudness of a sound whose loudness level is 40 phon [4].

In 2010, the DIN 45631 has been extended with the DIN 45631/A1 to cope with the calculation of loudness of time-variant noises. Figure 2.1 shows the general procedures for loudness calculation according to DIN 45631/A1.

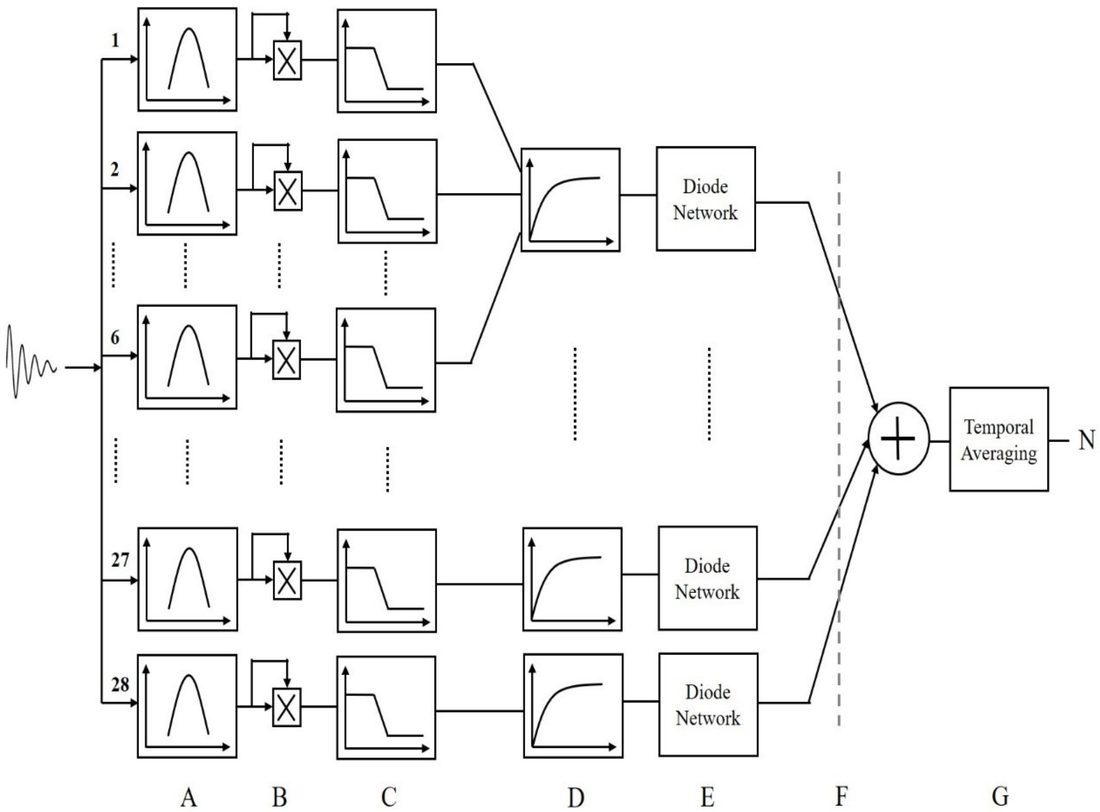


Figure 2.1: Flow chart for loudness calculation according to DIN 45631/A1

The individual steps are explained as follows [5]:

A. Calculation of the time-dependent 1/3 octave levels: 28 individual Chebyshev filters of 6th order are used.

B. Calculation of intensities: the third-octave band signals are transformed to time-dependent intensities by squaring.

C. Time-related averaging: the intensities is smoothened by feeding through frequency-dependent low-pass filters.

D. Calculation of core loudness values: the calculation of core loudness values is done according to the DIN. Thereby, intensities of the signals of lowpass filters 1 to 6, 7 to 9, as well as 10 and 11 are added. Those of the lowpass filters 12 to 28 are processed individually.

E. Generation of a fade-out time depending on duration by means of a diode network: a diode network described by Zwicker is utilized [6].

F. Calculation of the loudness summation: taking 20 core loudness values, the specific loudness distribution is calculated initially. After that, the specific partial loudness values are summed.

G. Temporal averaging of the loudness summation: The loudness summation is first filtered with two lowpass filters of 1st order (time constant 3.5 and 70 ms). Then, the total loudness is obtained from the weighted addition of these signals.

2.2. Calculation of sharpness

Sharpness is a measure of the weight of the high-frequency components of a sound. The greater the proportion of high-frequency components a sound has, the sharper it sounds. Sharpness is referred to as S, and the unit of it is acum. The narrow band of a noise, which centered at 1 kHz with a level of 60 dB and a bandwidth of 160 Hz, has an agreed sharpness value of 1 acum [4].

In DIN 45692, the method for the sharpness calculation is similar to the one developed by von Bismarck. Besides, the underlying loudness values are calculated according to DIN 45631/A. [7]

As DIN 45692 describes, the sharpness of a sound can be calculated from the following equation:

$$S = 0.11 \cdot \frac{\int_0^{24\,Bark} n'(z) \cdot z \cdot g(z) \cdot dz}{N},$$

$$g(z) = \begin{cases} 1 & , \quad z \leq 15.8\,Bark \\ 0.85 + 0.15 \cdot e^{0.42 \cdot (z-15.8)} & , \quad z > 15.8\,Bark \end{cases},$$

where $n'(z)$ is the specific loudness, z is the critical band rate and N is the total Loudness.

2.3. Calculation of roughness

Roughness is a complex measure which quantifies the subjective perception of rapid (from 15 to 300 Hz) amplitude modulation of a sound. Roughness is referred to as R, and its unit is asper. One asper is defined as the roughness produced by a 1 kHz tone of 60 dB which is 100% amplitude modulated at 70 Hz. [4]

The roughness of a sound can be evaluated from the following equation [8]:

$$R = cal \cdot \int_0^{24\,Bark} f_{mod} \cdot \Delta L \cdot dz\,,$$

where cal is a calibration factor, f_{mod} is the frequency of modulation and ΔL is the perceived masking depth [4]. In this study, the calibration factor cal is selected to be equal to 1.

3. Implementation

This chapter gives an overall view of the rattle noise recording and the listening test operation.

The recording sessions take place in three different experimental environments, which are the workshop at Volvo PV, the Volvo's test track and the anechoic chamber at Division of Applied Acoustics. The quality and similarity of the records from different environments can be tested and compared. In addition to this, in order to enhance the comparability and replicability of this study, the same established vibration generator system is applied.

The recorded objective noises are subsequently screened and processed to be more representative. Besides, two forms of listening tests are conducted sequentially. The similarities and differences between them will also be discussed hereinafter.

3.1. Objective sounds recording

3.1.1. Vibration excitation system

3.1.1.1. Vibration exciter

To generate vibration to the seat belt retractors, exciter "Energizer RED" is used. Detailed engineering data can be found in Appendix A "Energizer RED Exciter Specifications". This sort of generator system is only applied under laboratory environments which include Volvo workshop and the division's anechoic chamber. Figure 3.1 below is the sketch of the laboratory setup including names of main components, in both side and top view. Three bold-marked axes will also be referred to in the following sections.

3.1.1.2. Vibration excitation signal

In order to simulate the vibration, an excitation signal is fed through a piece of software, called "MB-Dynamics Random Vibration Control System", to the shaker. This signal is summarized from a series of the precedent studies and presented in the form of a power spectral density (PSD) function, which reflects the general car body vibration along the vertical direction. For the relevant documents are classified, figure 3.2 shows merely a diagrammatic drawing of the PSD. According to the figure, the PSD has a small rise from 10 Hz to 30 Hz. However, since the figure illustrates the density of power (acceleration), which can be seen that the power is positively related to the frequency in this frequency range with a coefficient which is slightly higher than 1. A relatively rapid increase occurs after 30 Hz. The power density, as a result, reaches the peak value of approximately 0.6 m/(s^2·Hz) and decays afterward. A feedback accelerometer is mounted on top of the shaker to adjust the excitation.

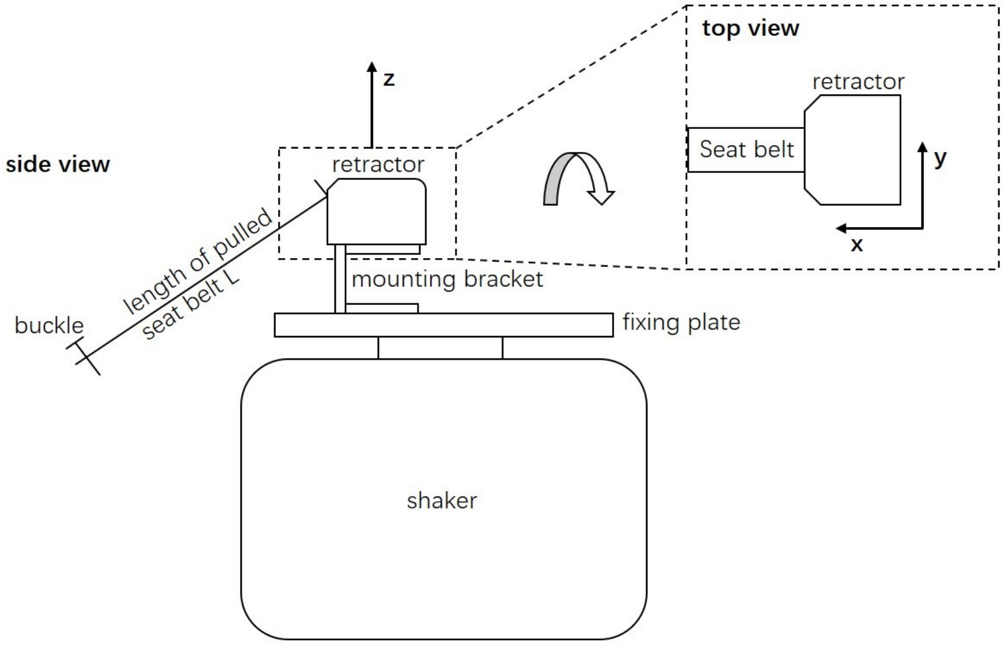


Figure 3.1: Sketch of the laboratory erection

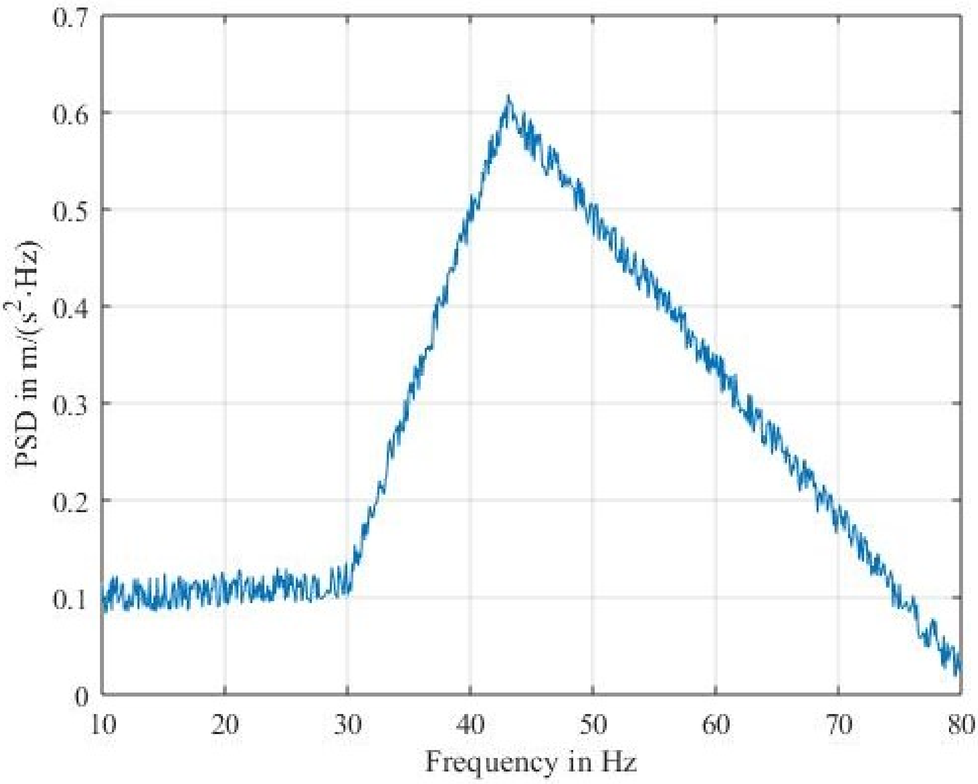


Figure 3.2: Diagrammatic drawing of the power spectral density (PSD) of the excitation signal

3.1.2. Recording sessions

3.1.2.1. Recording in Volvo's workshop

The recording work is initially taken in Volvo's workshop. Due to there was another study concerning the seat belt retractors' noise recorded on the same site, to make the results comparable, the recording setup is restored as much as possible. The slot surface, where the seat belt can be pulled out, is seen as the front. As figure 3.3 shows, two microphones are placed in the front (along the x-axis) and one side(along the y-axis). Each of them has a distance of 500 mm to the slot and a 45-degree angle to the ground.

Figure 3.3: Recording layout in Volvo's workshop

Both of the three types of seat belt retractors are tested in this recording environment, successively with a seat belt pulled-out length of 1,300 mm and 400 mm. These two pulled-out length data are concluded from precedent measurements, where the longer length is to simulate the situation when a passenger is belted in the car and the shorter length is the pull-out length when no one is belted. However, only the sounds captured by the front microphone with 1,300 mm seat belt pulled out were included in the subsequent analysis, which is due that is the condition closest to the most-affected passenger's experience in a real car. Simultaneously, the vertical accelerations on top of the retractors and the mounting bracket, respectively, are also recorded to populate the database. Hence, for instance, the acceleration transfer function along the vertical direction through these types of seat belt retractors can be conducted in possible future studies.

As recording work goes on, a severe problem is exposed that the shaker itself generates noise at a considerable level in comparison with the objective sounds from seat belt retractors. While measuring the shaker's noise, a bunch of metal cubes is stuck on the top of the mounting bracket to simulate the weight of the retractors, thus simulating the mass of the retractor with its noise avoided. As a result, the shaker‘s noise reaches an SPL of 27 dBA under workshop condition (including background noise). Since the overall SPL of the quietest retractor is about 36 dBA and the shaker's noise can be easily distinguished in by listeners, it is realized that some signal processing work, in terms of filtering, is required in order to minimize the interference to the judgments from the succeeding listening test. Moreover, the spectrum of shaker's noise, unsurprisingly, would change when the weight on the load-bearing changes, which means different filters should be applied to each type of seat belt retractors.

Meanwhile, relatively-high-level background noise is sometimes observed in the workshop. Though the workshop has foam boards on walls and locates at a relatively independent area, the background noise level is still considerable. Therefore, the records from this workshop are not qualified for the further user study. Nevertheless, this recording session is not good for nothing, whereby the significant interfering noise is noticed. The idea of seeking for a better recording environment is spawned.

In Volvo's workshop, the noises from all three types of seat belt retractors are recorded. For retractor type A and B, ten samples are tested. Besides, due to the limit of samples for retractor type C, only one sample involved in this recording session.

3.1.2.2. Recording in Volvo's test track

The following recording work takes place in Volvo's test track, located at Hällered. As it is shown in figure 3.4, two microphones are placed at the position where human ears are likely located with the support of a customized microphone stand. Because all the prepared seat belt retractors are designed for the left rear seat, only the headrest of the left rear seat is equipped with microphones.

Figure 3.4: In-car recording layout in Volvo's test track

At Hällered Proving Ground, nine road surface types are tested. According to an established driving test specification for S&R performance, the names of road surface types and their respective driving speed requirements (including initial speed and steady speed), are illustrated in table 3.1. Based on the site settings, the length of each type might differ.

Table 3.1: Name of test-road surface types and driving speed requirements

Road surface type	Initial speed in km/h	Steady speed in km/h
Comfort track 1	10	40
Comfort track 2	10	40
Man hole covers	10	50
Painted lines	10	50
Vienna stone	20	40
Country Road		
Concrete high way (LA freeway)	0	90
Bardfield road	0	70
Ucklum	50	70
Route Bonde	30	50

As expected, other interfering noise, especially the tire noise, is considerably louder, thus affecting the judgment in a way. However, owing to the practical experience, it can be asserted that the rear-seat seat belt retractors can generate audible noise, which is not entirely masked by other noises. Once the listener is used to the seat belt retractor rattling, it can be easily distinguished.

In Volvo's test track, the noises from retractor type A and B, with three samples for each, are recorded. This is due to the demand for compressing test duration and the lack of test vehicles suitable for retractor type C. The three samples are selected to be the retractors with the highest, the second-highest and eighth-highest average loudness values among each's ten samples. Besides, the pulled-out seat belt length is 1,300 mm, invariably.

3.1.2.3. Recording in the division's anechoic chamber

In order to acquire high-quality records, in which the objective noise dominates, the recording session is ultimately implemented in the anechoic chamber in the division of applied acoustics of Chalmers. Owing to the exceedingly low background noise level (17 dBA) [9] and few reverberations, the sounds captured in this laboratory can be nearly seen as pure and direct objective noise. Since these records are used in the following listening tests and analysis, the evaluation is turned to another focus, the sound quality of retractor noise in a quasi-free field.

The setup in the anechoic chamber is shown in Figure 3.5. Due to the limit of available microphone standing positions and the height of standing height, the distance from the retractor to the microphone is slightly increased to 515 mm, while the angle to the ground is 24 degree. Since the primary noise source is in the flank of the seat belt retractor, to record with two microphones could in a way provide a stereo feeling during the listening tests.

In the anechoic chamber, the noises from all three types of seat belt retractors are recorded. For the retractor type A and B, the same samples as in the road tests are selected. While for retractor type C, still, the only sample is used. All samples are tested under operating state (with 1,300 mm seat belt pulled out).

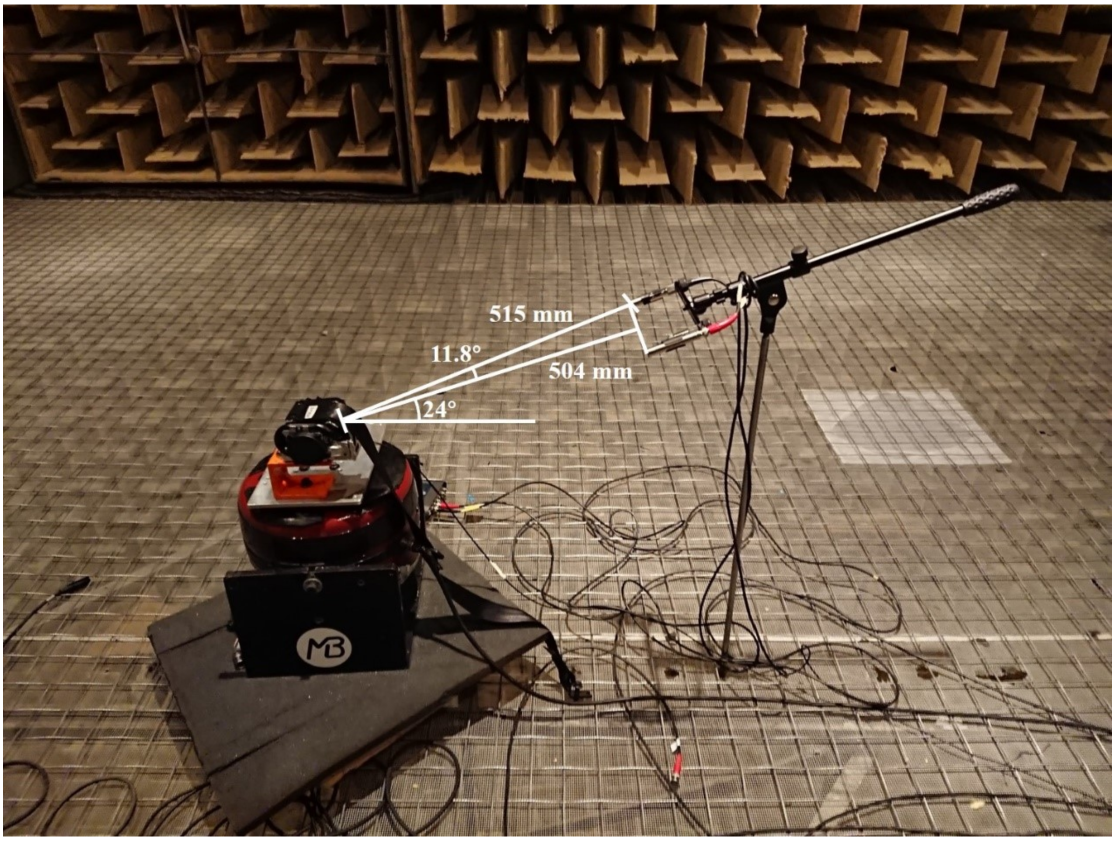


Figure 3.5: Recording layout in Chalmers' anechoic chamber

3.2. Listening test

3.2.1. Test sound preparation

3.2.1.1. Pretreatment to records

The quality of records is crucial for the listening test results and the representativeness of sound quality metrics. Therefore, the sounds recorded from the anechoic chamber are undoubtedly preferable. As mentioned previously, three samples respectively of retractor type A and B, as well as the only sample of retractor type C are tested in this recording environment. In order to shorten the length of the listening test in a proper range, the records from two samples of retractor type A and B are chosen to represent. They are both selected to be the second-loudest and the eighth-loudest samples out of ten. Then, including retractor type C, five stimuli would be evaluated in the listening test. All stimuli are clipped to five-second audio files.

3.2.1.2. Shaker noise removal processing

As mentioned above, since the shaker is not an ideal silent shaker, it generates interfering noise which also varies with different load mass. This sort of interfering noise would have an impact on the judgment during listening tests for sure, thus increasing the error in sound quality metrics. It becomes essential to examine the feasibility and validity of shaker noise removal.

Figure 3.6 shows the spectrum of the shaker noise with different types of seat belt retractors mounted, where the mass of retractor type A and C are very close (1308 g and 1357 g, respectively) while type B weighs 1155 g. The spectrums are obtained from the left channel of anechoic chamber records. Meanwhile, in order to avoid retractor noise, some metallic cubes are applied to simulate the load mass. As the figure shows, the shaker noise is more significant in lower frequencies, which makes it feasible and convenient to filter out the interfering shaker noise from records.

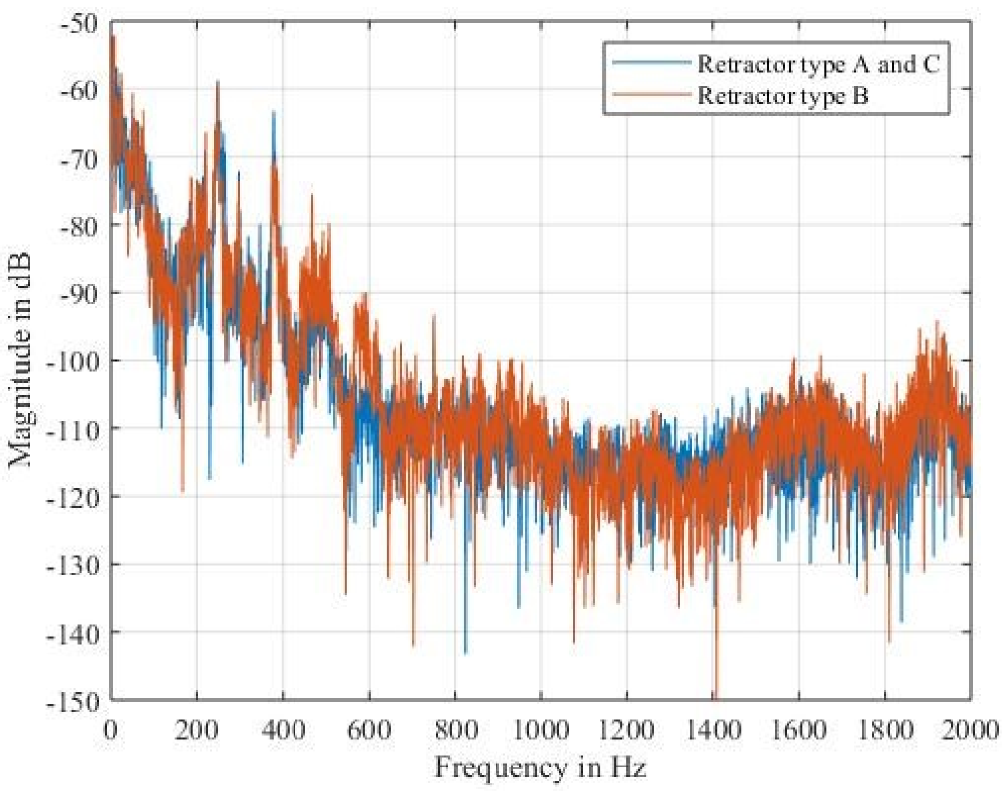


Figure 3.6: Spectrum of shaker noise with different types of retractors mounted

Figure 3.7, 3.8 and 3.9 illustrates the level difference between shaker noise and the overall noise (including both shaker noise and retractor noise), for three types of seat belt retractors, respectively. It can be figured out that, within the frequency range of human hearing, the shaker noise is relatively comparable with the overall noise below

650 Hz. In the frequency range above 650 Hz, the level difference maintains around 20 dB. Though it barely fulfills the masked threshold [10], the objective retractor noise is dominating.

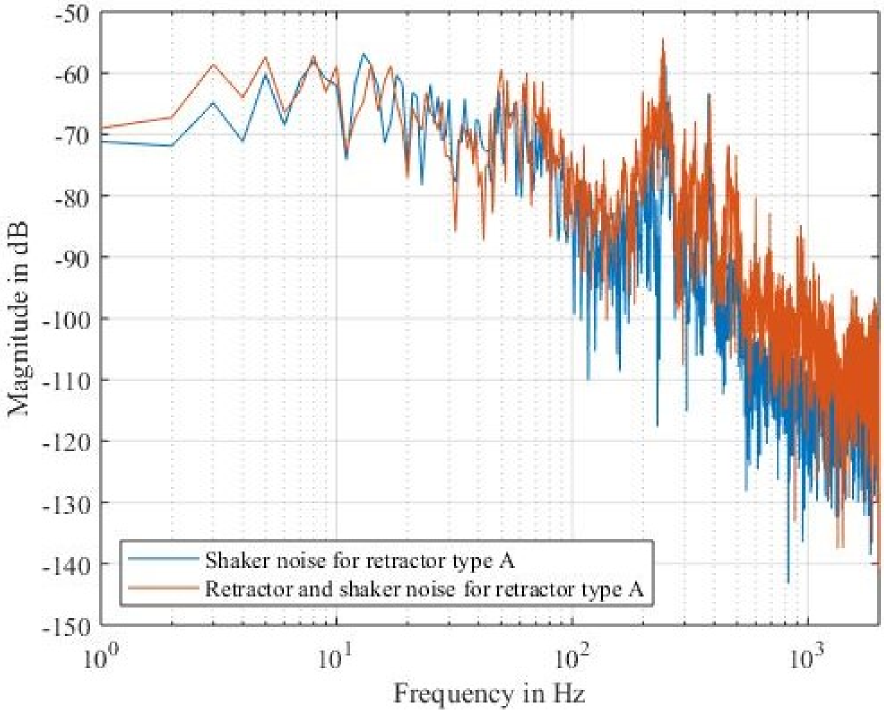


Figure 3.7: The level difference between shaker and overall noise, retractor type A

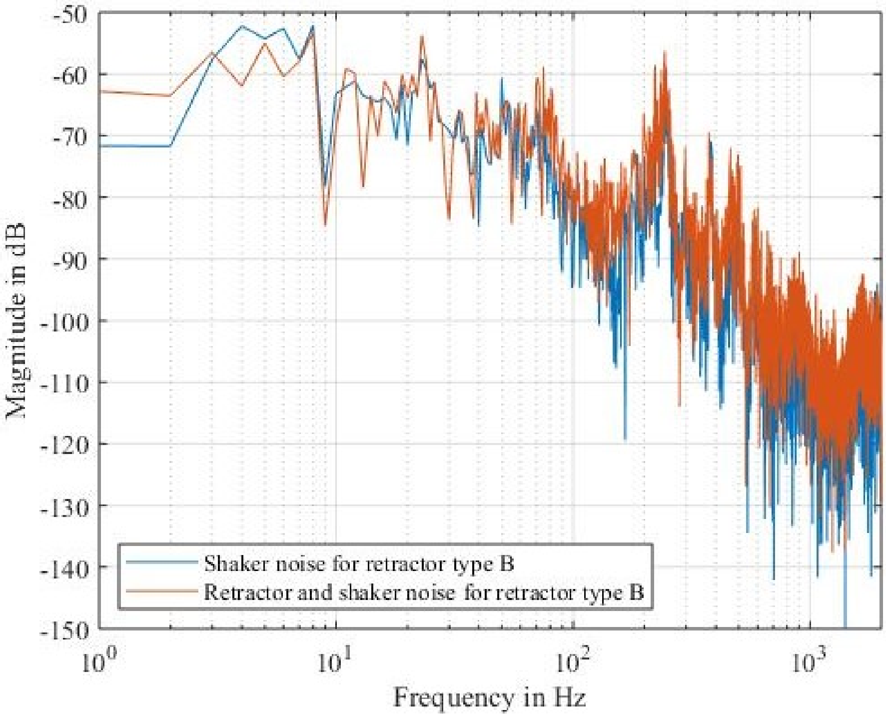


Figure 3.8: The level difference between shaker and overall noise, retractor type B

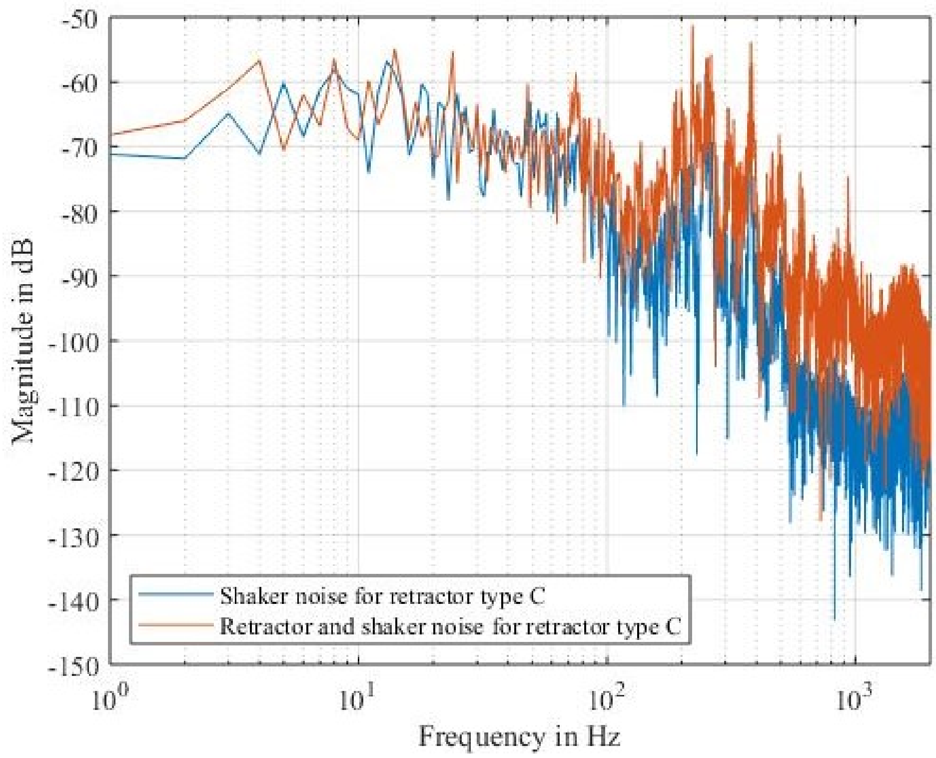


Figure 3.9: The level difference between shaker and overall noise, retractor type C

After a brief series of listening tests, which aims to ensure the objective retractor noise is not distorted, the high-pass filters are chosen to be at 500 Hz with 6 dB roll-off for seat belt retractor type A/C, and at 650 Hz with 24 dB roll-off for the seat belt retractor type B.

The main benefit of the shaker noise filtering is, it can effectively remove the interfering shaker noise. The participants' judgment, therefore, would be more targeted to the objective retractor noise during the following listening tests. However, which should not be overlooked is, the objective noise is also influenced by the filtering process to a certain extent. Because of in the psychoacoustic category, the actual listening experience is crucial. The processed audios are compared with the originals in the form of a brief listening test. The feedbacks are both positive.

3.2.2. Paper-based listening test

The listening test takes place first in the "Sound Quality & Sound Design" room at Volvo, where it allows up to four participants to do the test at the same time. The evaluation is anonymous. The stimuli were played from the computer through an audio interface to Sennheiser HD 600 headphones. The full process is well-calibrated to reproduce the perceived sounds as realistically as possible.

In this paper-based listening test, two test types, full-paired comparison and category judgment, are adopted. The questionnaire can be found in Appendix B1. Three records from the workshop also involve in the test, to verify the applicability for the designed metrics. The first section of the listening test is the full-paired comparison among five anechoic-chamber-recorded stimuli, which is to say, it includes 20 pairs of stimuli. The order is disrupted. The second section includes six stimuli pairs which means the full-paired comparison of workshop records. While the third and the last section is the category judgment, where eight stimuli are evaluated separately. For the soundtracks are edited ahead, excluding instruction reading time, each test has a fixed duration of eleven minutes. After every test, the results and feedbacks are collected.

During the listening test, two significant issues are uncovered. The first one is due to the soundtrack is prepared with set intervals between sounds, for some less-experienced participants, the relatively intense test rhythm may increase their sense of anxiety. The test results, therefore, would also be influenced to a large extent [11]. The second one is caused by the prepared soundtrack as well. It provides no possibility to playback, thus increasing the difficulty of selection especially for those pairs contain similar stimulus, for instance, the records from different samples of the same retractor type. For these reasons, a better test form is demanded. Till this test stops, twelve participants have involved.

3.2.3. Listening test via MATLAB GUI

3.2.3.1. Audio interface calibration

While using the MATLAB graphical user interface, the perceived sound quality can either be affected by the computer output volume, the audio interface adjustment or the headphone. In order to enhance the sound fidelity in the listening test, a pre-calibration is indispensable. Figure 3.10 shows the setup for devices calibration in Volvo's workshop, where MOTU audio interface Ultra Lite mk4 and HEAD acoustics HMS IV dummy head are used.

To minimize the number of variables, the output volume of the computer is kept at 100%. The noise from each type of retractors is played via GUI and then recorded by the dummy head sequentially. Two types of headphones, Sennheiser HD 280 and HD 600, are tested. Since the loudness of a sound generally plays the most critical role in perceived annoyance according to [12] and [13], the primary goal is to obtain the closest loudness value in comparison with the records played through HEAD acoustics SQuadriga II (including the matched headphone), which has a relatively high audio reproduction. After trying with different attenuation levels in the audio interface, HD 280 gives closer sharpness values while the loudness is as close as possible, obtaining which, the attenuation is set to -19 dB in the MOTU interface. The calibration factors, hence, are determined to be 100% computer output volume, headphone Sennheiser HD 280 and -19 dB in MOTU audio interface, which shall be fixed in the following GUI listening tests.

Figure 3.10: Layout of audio devices calibration in Volvo's workshop

3.2.3.2. Test implementation

Since the devices are both portable, the GUI listening tests take place at different places. However, the test sites must fulfill two main requirements, to be quiet enough and with few moving objects in the ambiance.

Same as the paper-based, the two test types, full-paired comparison and category judgment, are utilized, with a pretty much similar section layout. Apart from the conspicuous benefits, in terms of the feasibility of playback and full mastery of test rhythm, one remarkable change is the third section, which means the category judgment, becomes available for relative comparison (See Appendix B2).

Twenty participants take part in the GUI listening time. All user data are collected for the subsequent analysis. For the GUI listening test can only operate by one participant at the same time, though the single test duration is slightly shortened (to 10 min averagely), the total test hours are longer than the paper-based test.

3.3. List of equipment

- HEAD acoustics SQuadriga II
- Two omnidirectional microphones, G.R.A.S. 40AE
- Two preamplifiers, G.R.A.S. 26CA
- Calibrator, Brüel & Kjær 4231
- Monodirectional accelerometer, PCB JM352 C65
- Two monodirectional accelerometers, Brüel & Kjær 4507B
- Shaker, Energizer RED
- Audio interface, MOTU Ultra Lite mk4
- Four semi-open headphones, Sennheiser HD 600
- Closed headphone, Sennheiser HD 280
- Dummy head, HEAD acoustics HMS IV
- Cables
- Microphone standings
- Metallic cubes
- Computers

4. Result and Discussion

4.1. Recording environments comparison

Though in the end only the records from the anechoic chamber are used during the listening test, it is meaningful to compare the records from different recording environments. While comparing the coherence between two audio signals from the same type of seat belt retractor under different environments, some representative environment interactions might be discovered, thereby enriching the operability for any possible similar studies.

One sample from the seat belt retractor type A is taken as an example, where only the left channel involves the calculation. As figure 4.1 shows, the coherence values between the sound recorded in the anechoic chamber and the nine sorts of road surface types, are mostly below 0.5, which suggests the two signals are dissimilar in the frequency domain. This is predictable because of in the test tracks, the objective noise is mostly masked by the noises from other more conspicuous sources, for instance, the tire noise and the wind noise. Moreover, the excitation signal applied in the anechoic chamber, as well as in the workshop, is summarized from historical data and present randomly, which leads to a considerable difference in the excitations. Nevertheless, some relatively distinct peaks can be found in the second graph. A dogmatic conclusion could be, among these nine types of road surfaces, the behavior of “Comfort track 2” is somewhat closer to the excitation signal at some specific frequencies.

The coherence between signals recorded in the anechoic chamber and the workshop is considered as well. Thus, the room effect on the recording quality can be known. However, it can not be neglected that the retractor is not an ideal omnidirectional source, the angle difference can have an impact on the record’s spectrum, which is also considered as a part of the room effect herein.

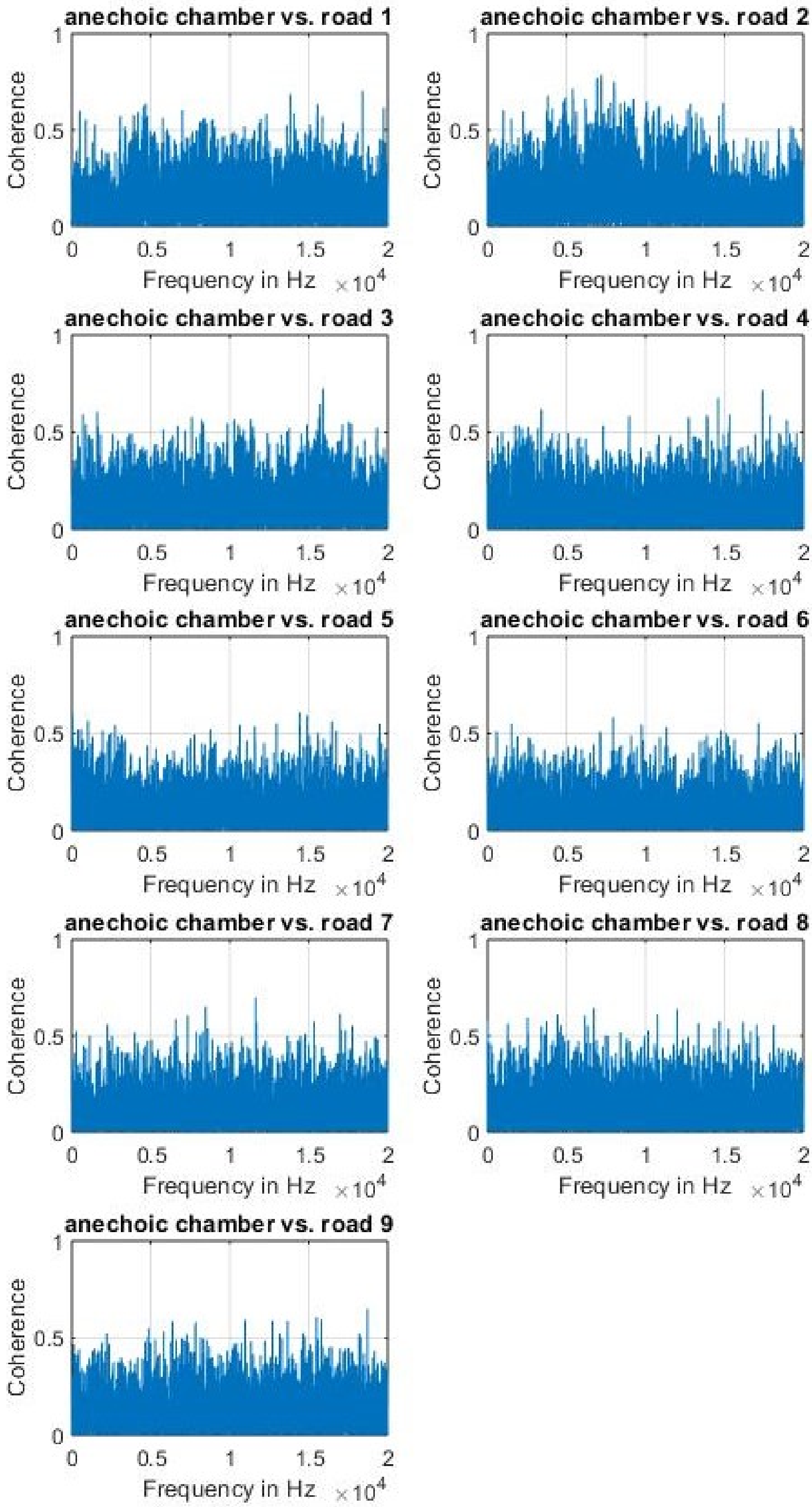


Figure 4.1: Coherence between mono audio signals recorded in the anechoic chamber and different test roads

Figure 4.2 illustrates the aforementioned coherence. It can be seen that the coherence value is mostly below 0.5, which means these two signals have an extremely weak correlation. Hence it can be said that the records acquired from the anechoic chamber and the Volvo’s workshop are incomparable due to the strong room effect in the workshop.

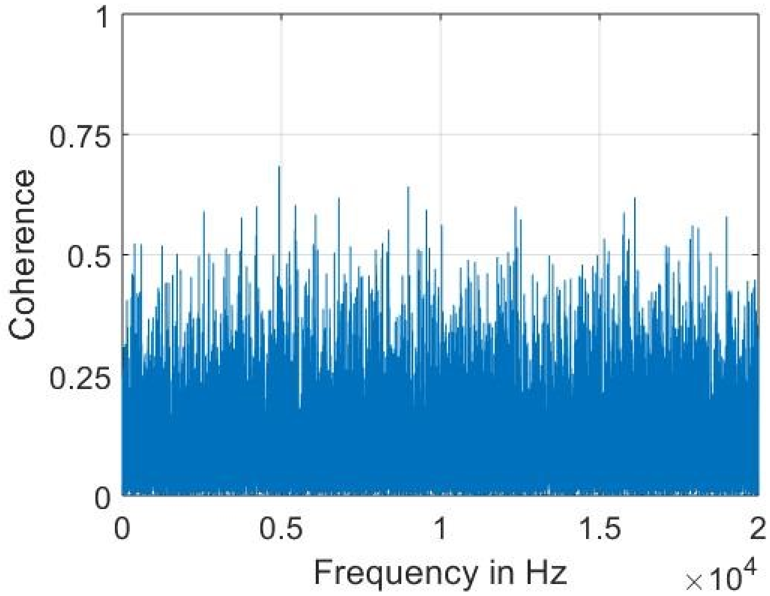


Figure 4.2: Coherence between mono audio signals recorded in the anechoic chamber and the workshop

However, it is not neglectable, that two records from different environments can still sound similar though the coherence between them is considerably low. Since the coherence reflects the correlation between two signals in the frequency domain, low coherence demonstrates that significant numerical differences can exist in their psychoacoustic parameters. Therefore, it is apparently risky to apply the sound quality metric derived from the anechoic chamber records to the records from other recording environments.

4.2. Listening test result analysis and statistics

The listening test result analysis includes two main steps. The individual consistency of judgment in the full-paired comparison is first checked, which aims to filter out some potential poor-quality data, due to the significant deviations can become an interference factor in the metric design. Whereafter in the remaining results, the data distribution and some statistical values of perceived annoyance are enumerated and discussed.

4.2.1. Consistency of judgment

The consistency of judgment is calculated by the results from the first section, where the full-paired comparison of five anechoic-chamber-recorded stimuli is tested. In this section, the same pair of stimuli is played twice with the stimulus order swapped. The consistency of the judgment on the more annoying stimulus becomes the indicator.

For a full-paired comparison with five objects, the total of pairs is 20. The selecting result for each pair is compared with its corresponding swapped pair in the sequence. After a statistical summary, three out of twenty test results show a relatively low consistency of judgment, coincidently all equal to 40%. The judgment consistencies of all the rest results are above 60%. However, after further observation, it is noticed that one single test result shows almost opposite data. Since it is opinionated to assume the participant oppositely understands the test questions, this result is excluded as well. As for the rest sixteen results, the overall consistency of judgment is 74%.

4.2.2. The degree of perceived annoyance

Figure 4.3 shows the 95% confidence interval of the perceived annoyance from the category judgment section, where the range of perceived annoyance degree is from 0 to 10. The circle and line inside the boxes are the average value and median value, respectively. The number of stimuli, in turn, represents the two samples of seat belt retractor type A, the two samples of retractor type B, and the only sample of retractor type C. In the box figure, the shorter boxes means the distribution of selection is more concentrated, which suggests the corresponding stimulus is relatively more characteristic. According to the figure, the marks of stimulus 1, 2 and 4 are distributed in a smaller range, while the marks of the last three stimuli are more likely to fulfill normal distribution since the average and median values appear in the middle.

Table 4.1 presents the mean, the median and the standard deviation for each stimulus from the valid test results. Except for the fifth stimulus, which has the most significant standard deviation equals to 2.8, the standard deviation values for the rest stimuli are all around 2.0. However, considering the permissible range is from 0 to 10, the results can be regarded as rational. The mean values are used in the following sound quality metric design.

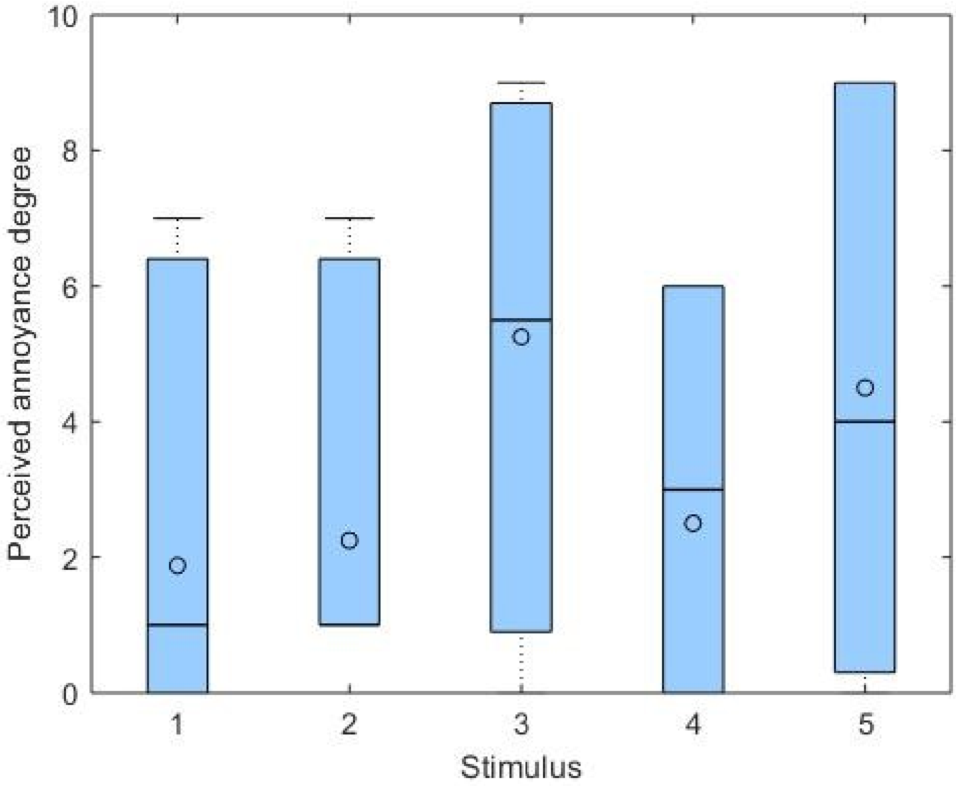


Figure 4.3: 95% confidence interval of perceived annoyance degree

Table 4.1: Mean values, median values and the standard deviation of perceived annoyance degree for each stimulus

Stimulus	Mean value	Median value	Standard deviation
1	1.875	1.0	2.0
2	2.250	1.0	1.8
3	5.250	5.5	2.1
4	2.500	3.0	1.9
5	4.500	4.0	2.8

4.3. Characteristics of objective sounds (psychoacoustic parameters)

In order to obtain a more accurate metric model, four different statistical values of the sound characteristics are calculated and compared with the annoyance sequence. They are the arithmetical average, the maximum, 95th percentile and 90th percentile, all calculated in HEAD acoustics ArtemiS using the calculation methods described in the Theory chapter. Table 4.2 shows the statistical values and their respective correlation coefficient to the annoyance marks. For ease of analysis, the listed values are averaged from respective psychoacoustic parameters of two sound channels. The correlation coefficient value in bold for each psychoacoustic parameter represents the most correlated statistical value among the four.

Table 4.2: Four statistical values for sound characteristics and their each correlation coefficient against the perceived annoyance degree

Loudness N in Sone					
Stimulus	Annoyance	N_{avg}	N_{max}	N_5	N_{10}
1	1.875	2.760	5.975	4.625	4.060
2	2.250	3.450	7.965	6.055	5.300
3	5.250	3.465	8.460	5.685	5.115
4	2.500	2.870	6.330	4.410	3.950
5	4.500	3.470	7.185	5.475	4.680
Correlation coefficient against annoyance	1	**0.69**	0.65	0.43	0.45
Sharpness S in acum					
Stimulus	Annoyance	S_{avg}	S_{max}	S_5	S_{10}
1	1.875	3.045	3.605	3.415	3.340
2	2.250	3.110	3.785	3.535	3.415
3	5.250	2.300	2.770	2.585	2.500
4	2.500	2.300	2.830	2.605	2.545
5	4.500	2.300	2.930	2.605	2.545
Correlation coefficient against annoyance	1	-0.73	-0.72	-0.73	**-0.75**
Roughness R in asper					
Stimulus	Annoyance	R_{avg}	R_{max}	R_5	R_{10}
1	1.875	0.0228	0.0375	0.0331	0.0306
2	2.250	0.0274	0.0503	0.0422	0.0388
3	5.250	0.1125	0.3060	0.2565	0.2050
4	2.500	0.0377	0.0824	0.0659	0.0581
5	4.500	0.0262	0.0678	0.0461	0.0387
Correlation coefficient against annoyance	1	0.72	**0.78**	0.74	0.73

It can be seen in the table, each parameter has at least two statistical values which show relatively high correlation coefficients (with the absolute value above 0.6). Both characteristics, therefore, will be taken into consideration during the metric design, using the most correlative statistical value.

Nevertheless, it is noteworthy that sharpness shows a negative correlation with perceived annoyance, which is very likely counterintuitive. A dogmatic explanation could be, that within such a sharpness range (from 2 to 4 acum), the effect of sharpness on the perceived annoyance of seat belt retractor noise happens to be like this. However, more relevant studies are required to certify this conclusion.

4.4. Sound quality metric design

4.4.1. Metric for perceived annoyance

As stated above, the arithmetical average of loudness N_{avg}, the 90th percentile of Sharpness S_{10} and the maximum of roughness R_{max} are selected to form the annoyance metric.

Using multiple linear regression, the metric become:

$$\text{Perceived annoyance, } PA = 1.87 \cdot N_{avg} - 1.50 \cdot S_{10} + 4.42 \cdot R_{\text{max}} + 1.10,$$

where the coefficients of each term are supposed to include the unit which can offset the unit of each parameter. To make the metric concise, they are not reflected.

Table 4.3 shows the calculated results that the metric gives and the subjective judgments from the listening test.

Table 4.3: Calculated and questionary degrees for the perceived annoyance of the anechoic chamber records

Stimulus	1	2	3	4	5
Calculated degree	1.42	2.66	5.19	3.02	4.08
Questionary degree	1.85	2.25	5.25	2.50	4.50

The calculated results and the questionary perceived annoyance for each stimulus are shown in figure 4.4. As the figure shows, apart from the third stimulus, apparent deviations can be found. The maximum error appears at the fourth stimulus, which reaches 0.52. However, considering the scale is from 0 to 10, with an RMS error of 0.41, the error rate is 4.1%. This metric can be considered as a fairly reliable prediction method for the seat belt retractor noise.

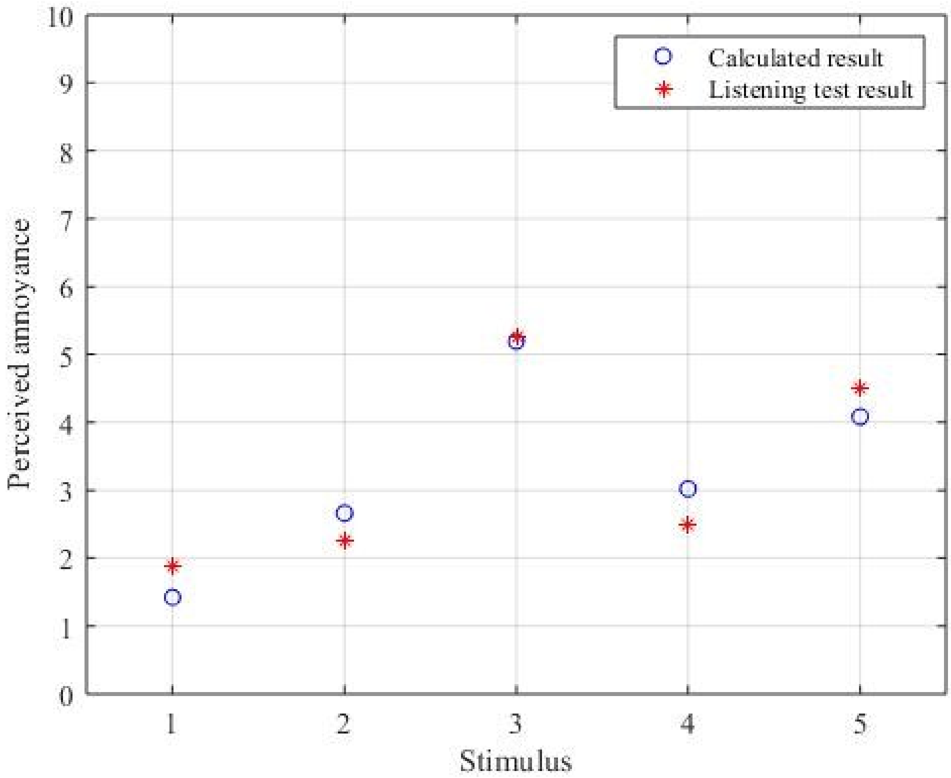


Figure 4.4: Comparison of the calculated results and the listening test results

Figure 4.5 shows the tendency for all considered sound characteristics when the perceived annoyance increases. The stimuli are sorted from the least annoying to the most annoying along the x-axis. Their corresponding sound characteristics are normalized between zero and one along the y-axis. The normalizing factors are the maximum values of each sound characteristic. For ease of description, the intervals along the x-axis are numbered sequentially.

According to interval 1 and 4, loudness affects the perceived annoyance degree significantly when sharpness and roughness are almost equivalent. Though the increases of loudness are similar, a huge difference in perceived annoyance increases happens. This is due to the change of perceived annoyance depends on the synthetical functions of various sound characteristics. An obvious discrepancy can be seen in the sharpness and the roughness of these two corresponding stimuli pairs, thus leading to the observed growth difference. While sharpness, as mentioned above, has a negative trend against perceived noise. It is particularly evident in interval 1, where the two stimuli have the top two sharpness values, but the lowest two perceived annoyance degrees instead. Interval 5 shows that when loudness and sharpness are at the same level, roughness plays a considerable role in perceived annoyance judgment. For these five stimuli, roughness shows the biggest deviation in percentage among the three selected psychoacoustic parameters.

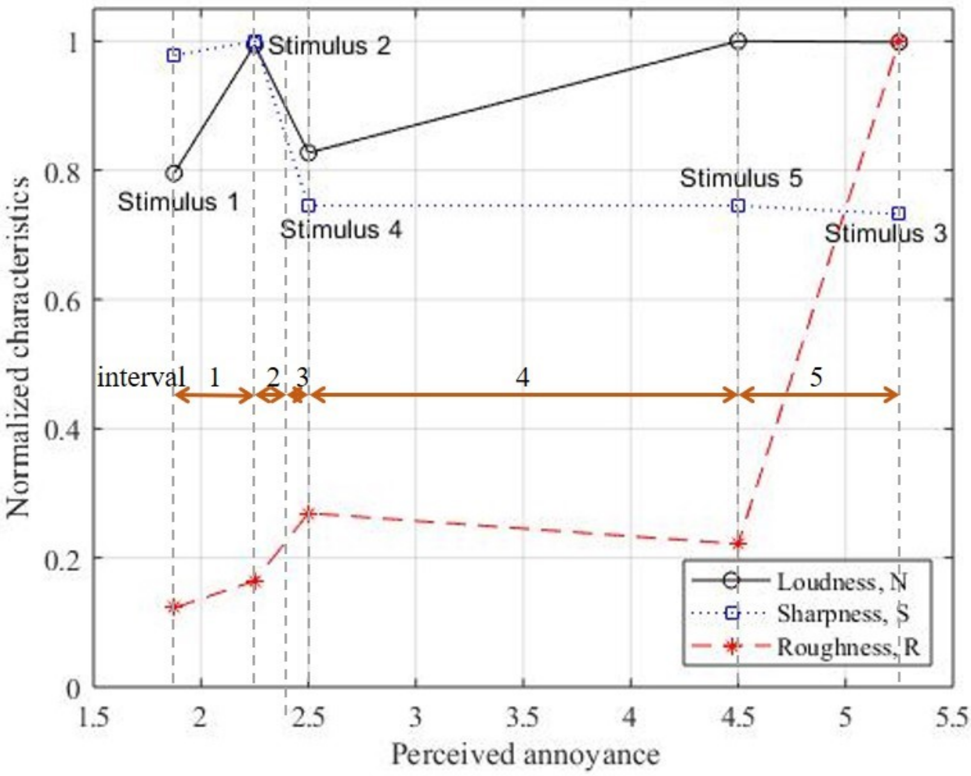


Figure 4.5: Tendency chart of perceived annoyance with respect to sound characteristics

4.4.2. Application of the perceived annoyance metric to the sounds recorded in the workshop environment

Although, as previously discussed in section 4.1, the metric is very likely not applicable to the records from different environments, it is still significative to ascertain the degree of deviation. Since the annoyance degrees of the workshop records are also asked during the listening tests, it is possible to verify the applicability of the perceived annoyance metric over the workshop recording environment.

Table 4.4 shows the calculated results and the subjective judgments of the workshop records, where stimulus 6 to 8 represent the workshop records of seat belt retractor type A, B and C, respectively.

Table 4.4: Calculated and questionary degrees for the perceived annoyance of the workshop records

Stimulus	6	7	8
Calculated degree	10.00	6.38	5.43
Questionary degree	8.94	5.69	5.13

Figure 4.6 illustrates the calculated results and the questionary perceived annoyance for each workshop stimulus. It can be seen that the metric always gives higher annoyance degrees than subjective judgments. However, with an RMS error of 0.75, the performance of the metric can be considered acceptable. It becomes meaningful to verify the general applicability of this metric by using the records from different recording environments or applying to the noises of other types of seat belt retractors recorded in the anechoic chamber.

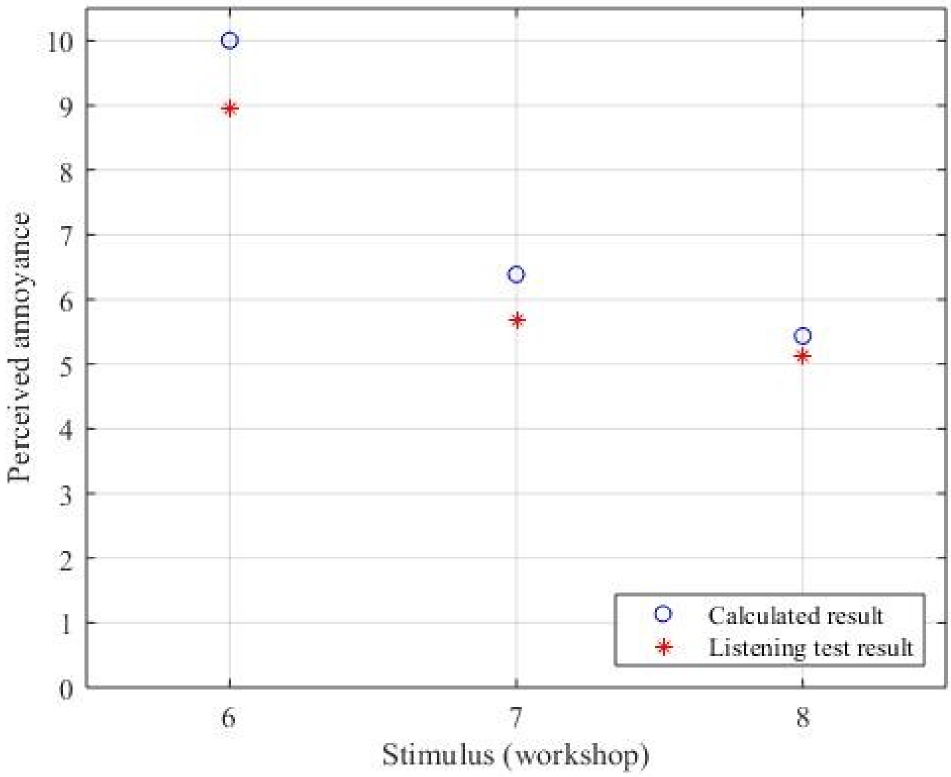


Figure 4.6: Comparison of the calculated results and the listening test results for workshop stimuli

4.4.3. Metric for Volvo Score

Since there is another subjective evaluation system widely used at Volvo, the details of which can be seen in figure 4.7, the perceived annoyance metric is then transformed into this scale to make the metric possible to be used directly.

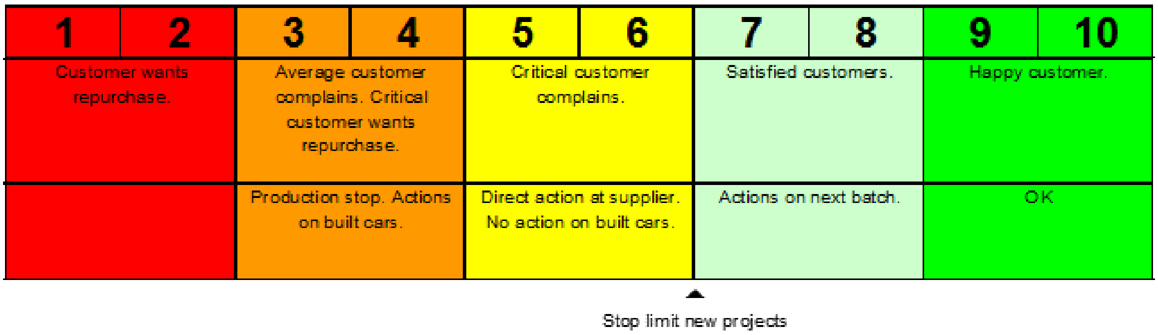


Figure 4.7: The Volvo scale of subjective evaluation

Due to the results from the listening test reflect the perceived annoyance while this scoring system assesses satisfaction, an assumption becomes crucial to the metric transformation. The assumption is the perceived annoyance and the satisfactory always sum to unity. For instance, as figure 4.8 shows, if the distance from slider to the left end records the perceived annoyance degree, the distance to the right end is assumed to represent the satisfactory degree, which can be imprecise since the judgment would very likely show a different degree of satisfaction on the same stimulus when the question was reversed due to there is supposed to be a middle ground between an annoying sound and a satisfactory sound.

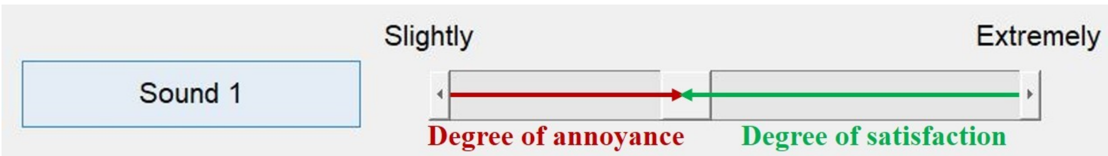


Figure 4.8: An example of the relation between perceived annoyance and satisfaction under the assumption

A brief listening test involves six participants, who are familiar to the Volvo evaluation system, is then conducted. The two stimuli from seat belt retractor type B are played during this listening test. The participants are asked to assess them with respect to the Volvo scale. The Volvo scores of the two stimuli are obtained by averaging the judgments of all participants. By comparing the intervals between two stimuli, a scaling factor of 0.56 is figured out.

The transformation formula is as follows:

$$\text{Volvo score, } VS = ((10 - PA) - (10 - 5.25)) \cdot 0.56 + 4,$$

where *10* is the upper limit of perceived annoyance in the listening test, *5.25* is the perceived annoyance degree of the most annoying stimulus (stimulus 4). Therefore, the term *(10 – PA)* represents the calculated satisfaction for each stimulus, while the term *(10 – 5.25)* represents the subjective satisfaction of stimulus 4 from the listening tests. Moreover, *0.56* is the scaling factor, and *4* is the corresponding Volvo score of stimulus 4.

After collating, the metric for Volvo score becomes:

$$VS = -1.02 \cdot N_{avg} + 0.82 \cdot S_{10} - 2.41 \cdot R_{max} + 6.26.$$

Table 4.5 presents the calculated Volvo scores of each stimulus.

Table 4.5: Calculated Volvo scores for the stimuli recorded in the anechoic chamber

Stimulus	**1**	**2**	**3**	**4**	**5**
Volvo score	6.1	5.4	4.0	5.2	4.6

5. Conclusion

This thesis provides a straightforward method to evaluate and predict the perceived annoyance of the rattle noise produced by seat belt retractors. As long as the rattle noise is recorded in an anechoic chamber of similar condition, the sound quality metric can be directly applied. For the five retractor samples used in this thesis study, the selected statistical values of the psychoacoustic parameters show a relatively strong correlation to the perceived annoyance obtained from the listening test. The metric, therefore, can give a fairly accurate result, though a great deviation appears in the judgments of perceived annoyance among individuals. It is constructive to apply this method to other types of seat belt retractors thereby verifying its applicability. Moreover, according to the preliminary exploration in 4.4.2, the metric is likely also applicable to the retractor's noise recorded from other lab environments.

The rattle noise of seat belt retractors, which is the object of study in this research, is so unimpressive that seldom considered important in comparison with other vehicle noises. However, its particularity is this sort of noise is produced inside the cab. This study may have an increasing significance while the sound insulation of the bodywork progresses.

6. Future research

This topic can be extended in various ways. Some suggestions are provided as follow:

- To have more stimuli and participants involved, thereby enhancing the applicability of the sound quality metric.

- To conduct the listening test in a car, better in a running car, thus reproducing a more authentic experience.

- To take more psychoacoustic parameters into account, for example, tonality and fluctuation strength.

- To try to form the metric with the parameters with different exponentiations, or in other forms. One possible form is the unbiased annoyance (UBA) metric [14].

While doing a comparative study, it should be noted that the psychoacoustic parameters must be calculated according to the same standards. And the listening test is supposed to designed and conducted in a similar way.

Bibliography

[1] OICA, "Estimated worldwide automobile production from 2000 to 2018 (in million vehicles)", 2019.

[2] GRAS Sound & Vibration, "Buzz, Squeak and Rattle Noise Testing".

[3] G. Cerrato-Jay, J. Gabiniewicz, J. Gatt and D. J. Pickering, "Automatic Detection of Buzz, Squeak and Rattle Events". In: *JOURNAL OF PASSENGER CAR: MECHANICAL SYSTEMS JOURNAL,* 2001.

[4] Zwicker E., Fastl H. "Psychoacoustics: Facts and Models", 1990.

[5] Deutsches Institut Fur Normung E.V. "DIN 45631/A1: Calculation of loudness level and loudness from the sound spectrum - Zwicker method - Amendment 1: Calculation of the loudness of time-variant sound", 2010.

[6] Zwicker, E. "Dependence of post-masking on duration". In: *J. Acoust. Soc. Am., Vol 75, No. 1,* 1984.

[7 Deutsches Institut Fur Normung E.V. "DIN 45692: Measurement technique for the simulation of the auditory sensation of sharpness", 2009.

[8] Widmann U, Fastl H. "Calculating Roughness Using Time-Varying Specific Loudness Spectra". In: *Sound Quality Symposium*, 1998.

[9] Facilities: Anechoic chamber. URL: http://www.ta.chalmers.se/ (visited in 2019).

[10] Gelfand, S.A. "Hearing – An Introduction to Psychological and Physiological Acoustics 4th Ed.". In: *Marcel Dekker, New York,* 2004.

[11] Yo In'nami. "The effects of test anxiety on listening test performance". In: *System,* 2006.

[12] W. Ellermeier, A. Zeitler and H. Fastl. "Predicting annoyance judgments from psychoacoustic metrics: Identifiable versus neutralized sounds". In: *the 33rd International Congress and Exposition on Noise Control Engineering, Prague, Czech Republic,* 2004.

[13] Andrew M. Willemsen and Mohan D. Rao. "Characterization of sound quality of impulsive sounds using loudness based metric". In: *Proceedings of 20th International Congress on Acoustics*, *Sydney, Australia,* 2010.

[14] Fastl H. "Psychoacoustics and Sound Quality Metrics". In: *Sound Quality Symposium,* 1998.

A
Appendix - Equipment

Energizer RED Exciter Specifications

MB Dynamics — TECHNOLOGY, TEST PROCESSES & ENGINEERING SERVICES

Vibration Exciter System
MB Energizer RED

Features:

- Quiet package: no exciter noise, no cooling required, optimized design for Squeak & Rattle (S&R) testing
- S&R background noise: running typical S&R random vibration profile of 0.36 g RMS, 8 to 100 Hz, microphone positioned 10" (250 mm) above mounting table; Sound Pressure Level (25 dBA); N10 time varying loudness: <0.1 sone
- Force: 90 lbf pk (400 N pk) forced-air cooled; 40 lbf pk (200 N pk) convection cooled
- Force: 75 lbf RMS (250 N RMS) forced-air cooled; 30 lbf RMS (125 N RMS) convection cooled
- Maximum stroke: 1.5" (38 mm) p-p
- Maximum acceleration: 30 g's pk, bare mounting table
- Bandwidth, full performance: DC to 6000 Hz
- Bandwidth, usable: DC to 9000 Hz
- Armature / mounting table weight: 3.4 lbs (1.6 kg)
- Large mounting table: 4" (100 mm) diameter, with 2" x 2" grid for fixtures (50 mm x 50 mm); ¼"-28 (M6) inserts
- Maximum payload: 13 lbs. (6 kg) vertical or horizontal (without optional air spring)
- Achieves full force with 10 lbs (4.5 kg) payload, vertical or horizontal
- Stiff load support with patented stiff, quiet flexures resist overturning moments due to offset or "high CG" payloads; horizontal CG offset ≤150 mm; safely supports non-centered unbalanced payloads
- Maximum overturning moment: 100 in-lbs. (11 Nm)
- Suspension driven-axis stiffness: 50 lbf/in. (8.8 kN/m)
- Suspension transverse stiffness: 2500 lbf/in. (425 kN/m)
- Lightweight shaker design with trunnion base: 80 lbs (36 kg); 18.5"L x 19"W x 18.5"H (470 mm x 483 mm x 470 mm) footprint when mounted inside trunnion base
- Energizer RED dimensions: 14"L x 12"W x 12.4"H (355 mm x 305 mm x 315 mm)
- CE marked
- Optional shaker cooling: auxiliary fan cooling package or factory air
- Optional portable, bolt-on inertial masses for heavy payload or high-g combinations

Vibration Exciter System
MB Energizer RED

Applications:

- Quiet operation for S&R evaluations and other external/ambient noise sensitive testing
- Robust construction for durability testing
- NVH testing
- Sensor manufacturing quality assurance testing
- In-laboratory R&D instrumentation verifications
- Recommended pairing with with MB500VI amplifier

MB Dynamics, Inc.
25865 Richmond Road, Cleveland, OH 44146 USA
Tel: +1 216 292 5850 · Fax: +1 216 292 5614
For Sales: sales@mbdynamics.com · For Service: service@mbdynamics.com
www.mbdynamics.com

MB-ENRGZRRED-1018

MB Dynamics
TECHNOLOGY, TEST PROCESSES & ENGINEERING SERVICES

Vibration Exciter System
MB Energizer RED

Parameter	
Maximum excitation force	Requires cooling, uncooled operation up to 50% of max. excitation force
Sine	400N pk
Random	250N RMS
Peak force (instantaneous)	1000N pk
Max. payload, excitation (both vertical and horizontal)	6 kg (12.5 bs.)
Max. payload with air spring, vertical	1.6 kg (3.4 lbs.)
Weight-Armature plus mounting table	N/A
Acceleration, 5 kg payload, uncooled	N/A
Acceleration, 15 kg payload, uncooled	N/A
Acceleration, 30 kg payload, uncooled	N/A
Accelerations with forced air cooling	2X uncooled
Mounting table diameter (standard)	100 mm (3.9")
Surface mounting pattern	50 mm x 50 mm (M6)
Inserts and bolt-hole pattern	M6 (1/4"-28), 50 mm x 50 mm
Frequency range (less with air spring)	DC to 6000 Hz
Operating background noise	Using a typical S&R test profile
Noise rating curve (NR)	NR16
Sound Pressure Level [dB(A)]*	<25 dB(A)
N10 Time Varying Loudness**	<0.1 sone
Stroke pk-pk	38 mm
Suspension driven-axis stiffness	8.8 kN/m (50 lb/in)
Suspension transverse-axis stiffness	425 kN/m (2500 lb/in)
Max. displacement at test item, pk-pk	38 mm (1.5")
Max. velocity, pk	1.3 m/s (51 in/sec)
Weight including trunnion and HMT mounting base	36 kg (79 lbs.)
Energizer mounting base footprint	470 mm x 483 mm (18.5" x 19")
Height to top of mounting table	315 mm (12.4")
CE mark	Yes
Power requirements	1000 watts, max.
Recommended amplifier	MB500VI
Amplifier input power	120 V or 220 V
Amplifier dimensions, 19" rack; taller cabinets availabile for PC & instruments	N/A
Weight of power amplifier	15 kg (33 lbs.)
Temperature and overtravel control	Optional

* A-weighted, FAST (125 ms) Sound Pressure level over the complete AUDIO frequency range from 20 Hz to 20 kHz
** N10 Time Varying Loudness in accordance with DIN45631/A1, measured in accordance with GMW14011

MB Dynamics, Inc.
25865 Richmond Road, Cleveland, OH 44146 USA
Tel: +1 216 292 5850 · Fax: +1 216 292 5614
For Sales: sales@mbdynamics.com · For Service: service@mbdynamics.com
www.mbdynamics.com

MB-ENRGZRRED-1018

B1
Appendix – Paper-based Listening test questionnaire

Annoyance of squeak and rattle noise

from seat belt retractors

Listening test questionnaire

Test date: ___________________

Interview

How old are you? __________

What is your gender?

male   female   other

What is your knowledge and experience in psychoacoustics?

expert   very good   good   limited   none

What is your experience in doing Squeak & Rattle subjective testing?

expert   very good   good   limited   none

How many listening tests have you attended?

many (>5 times)   several (3-5 times)   few (1-2 times)   none (0)

Do you have normal hearing?

yes   no (Please mention what kind of hearing difficulties you have)

____

____

____

____

Page 1/8

Section one

Instructions:

1. You will hear 20 pairs of noise samples separated by a one-second silence, followed by a four-second silence where you can rate. A serial number will be played before each pair.

2. Within one pair, the first sound is named as A, and the second is named as B in this questionnaire.

3. In this section, you are only supposed to rate which sample is more annoying within the same pair.

Example:

According to the experience, sound A is considered more annoying.

Example 1	A	B
	×	

Pair 1	A	B

Pair 2	A	B

Pair 3	A	B

Pair 4	A	B

Pair 5	A	B

Pair 6	A	B

Pair 7	A	B

Pair 8	A	B

Pair 9	A	B

Pair 10	A	B

Pair 11	A	B

Pair 12	A	B

Pair 13	A	B

Pair 14	A	B

Pair 15	A	B

Pair 16	A	B

Pair 17	A	B

Pair 18	A	B

Pair 19	A	B

Pair 20	A	B

Section two

Instructions:

1. You will hear 6 pairs of noise samples separated by a one-second silence, followed by a four-second silence where you are can rate. A serial number will be played before each pair.

2. Within one pair, the first sound is named as A, and the second is named as B in this questionnaire.

3. In this section, you are only supposed to rate which sample is more annoying within the same pair.

Example:

According to the experience, sound B is considered more annoying.

Example 2	A	B
		×

Pair 1	A	B

Pair 2	A	B

Pair 3	A	B

Pair 4	A	B

Pair 5	A	B

Pair 6	A	B

Section three

Instructions:

1. You will hear 16 noise samples, separated by a six-second silence where you can rate how satisfactory the sound is. A serial number will be played before each sound.

2. For each sample, you are supposed to rate the level of annoyance within the range from 1 to 5 on the graduated bar, where 5 represents least annoying and 1 represents most annoying. Arbitrary position is accepted.

More information concerning the grading can be found in the table below,

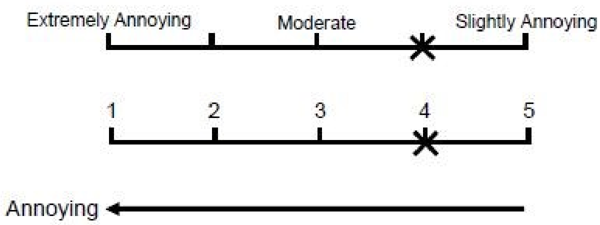
Grade	Corresponding to
1	extremely annoying (horrible)
2	very annoying
3	annoying
4	little annoying
5	slightly annoying (acceptable)

Note: within this section, two types of graduated bars are applied (See example).

3. To make the rate well-spread, the first sample is calculated to be one of “the most annoying” sounds, whilst the second sample is calculated to be one of “the least annoying” sounds. This is for you only as reference, please rate the level based on your own experience.

Example:

According to the experience, the sound is considered to have an annoyance level of 4 (little annoying).



Sound 1:

Extremely Annoying | Moderate | Slightly Annoying

Sound 2:

Extremely Annoying | Moderate | Slightly Annoying

Sound 3:

Extremely Annoying | Moderate | Slightly Annoying

Sound 4:

Extremely Annoying | Moderate | Slightly Annoying

Sound 5:

Extremely Annoying | Moderate | Slightly Annoying

Sound 6:

Extremely Annoying | Moderate | Slightly Annoying

Sound 7:

Extremely Annoying | Moderate | Slightly Annoying

Sound 8:

Extremely Annoying | Moderate | Slightly Annoying

Page 7/8

Annoying ←

Sound 9:

1 2 3 4 5

Sound 10:

1 2 3 4 5

Sound 11:

1 2 3 4 5

Sound 12:

1 2 3 4 5

Sound 13:

1 2 3 4 5

Sound 14:

1 2 3 4 5

Sound 15:

1 2 3 4 5

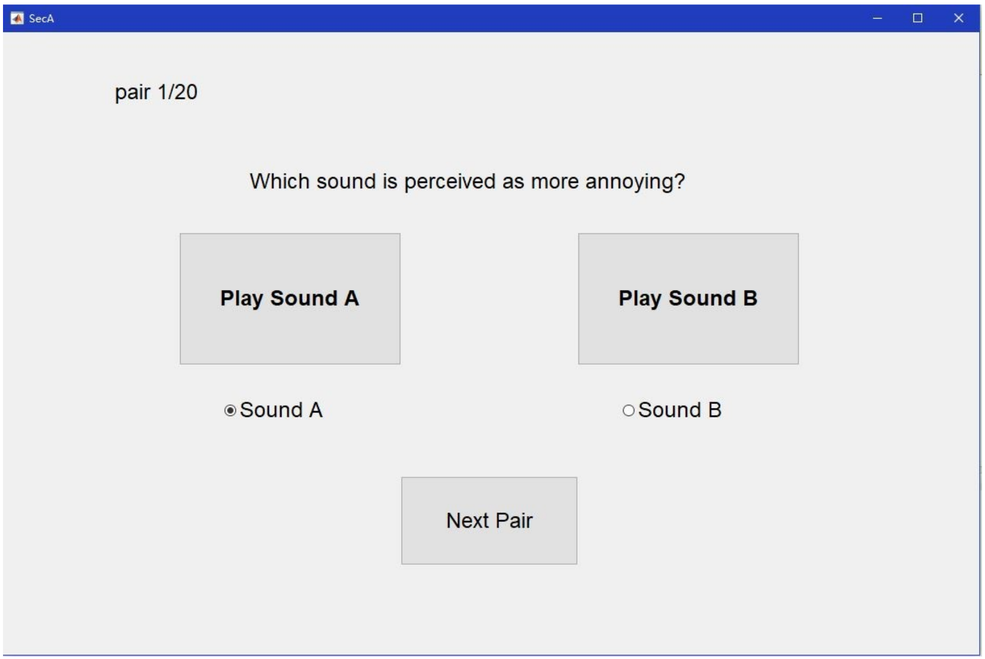
Sound 16:

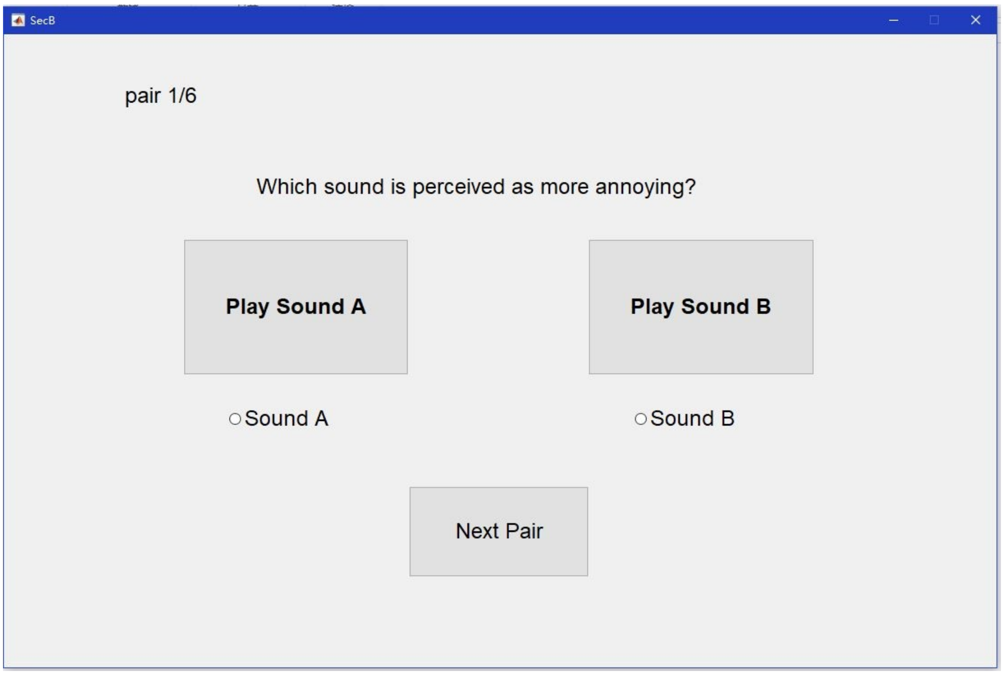
1 2 3 4 5

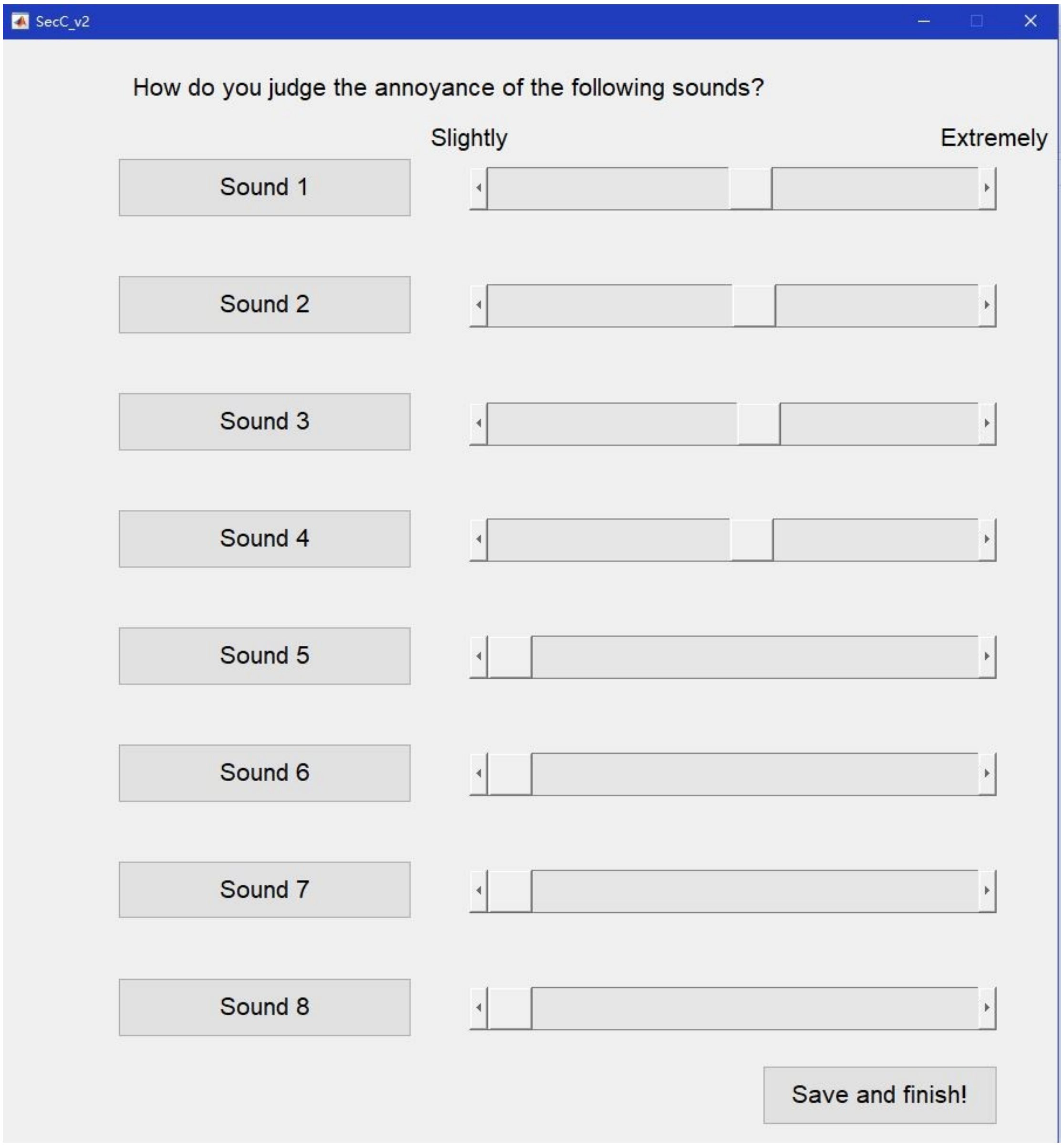
Annoying ←

Page 8/8

B2
Appendix – MATLAB GUI Listening test interface






SecC_v2
How do you judge the annoyance of the following sounds?
Slightly
Extremely
Sound 1
Sound 2
Sound 3
Sound 4
Sound 5
Sound 6
Sound 7
Sound 8
Save and finish!