Measures to Address the Dissonance Perception of Multiple Tonal Components in Sounds

Anna Rieger[1], Christian Mühlbauer[1], Rainer Weber[2], Hans-Peter Rabl[1]
[1] *OTH Regensburg, CEEC Laboratory, 93053 Regensburg, Germany, Email: anna.rieger@st.oth-regensburg.de*
[2] *Continental Automotive GmbH, 93055 Regensburg, Germany, Email: rainer.2.weber@continental-corporation.com*

Motivation and Introduction

In the fight against global warming, local emission-free electric vehicles are expected to play a key role. But what is said to solve one problem raises another one. Namely that the electrification is about to cause novel issues concerning NVH (Noise Vibration Harshness). The familiar combustion engine is no longer available as masking element for unwanted noise of other aggregates. This means that other sound sources now elicit an annoying or disruptive hearing sensation. Due to its clearly audible whistling noises whilst driving or charging, the electric powertrain itself is one important factor. Whistling noises, in other words pure tones, make a certain sound more annoying by their presence [1].
In turn, an opposite problem is generated by the absence of noise for pedestrians. To counteract the greater risk of accidents, it is legally required to equip the vehicles with a so-called Acoustic Vehicle Alerting System (AVAS) which should be designed to emit noise up to a speed of $20\,\mathrm{km\,h^{-1}}$. Out of this, the opposite desires of euphonious and simultaneously safe mobility arise. [2]
The current situation obviously demands suitable characterization and evaluation possibilities for tonal components in technical sounds and a fundamental understanding about which features of such sounds attract attention without annoying the environment to a great extent.

Audible tones in technical sounds are already considered by the metric tonality T in tu (tonality unit) [3]. Tonality is also used for the determination of psychoacoustic annoyance PA [4, 5] and sensory euphony W [3] which are indicators for how disturbing or rather how pleasant a certain sound is perceived.
The determination of tonality includes several processing steps which were investigated for possible improvements and enlargements in a master's degree study. These are the processing of the signal in frequency and time domain followed by the application of tone extraction methods (for example Terhardt's method [6] and the procedure according to the standard DIN 45681 [1, 7]) to gain level excesses of the particular tones. When the tonal components are extracted, the determination of tonality takes place (for example according to the model of Aures [3]). It can be pointed out that the extracted tonal components are weighted by several criteria and are summed up for a particular time step to achieve tonality $T(t)$. The criteria for weighting are level excesses ΔL, the components' frequencies f and their bandwidths Δf. It has to be mentioned that every component is weighted separately, whereby the issue arises if different combinations of tonal components also elicit different auditory perceptions. Obviously a major second (the musical tone interval from the notes C to D as example) elicits an unpleasant hearing impression, whereas a triad (C, E and G together) is perceived pleasant [8, 9] in a musical point of view. This might give rise to question the appropriateness of the characterization of multiple tonal components by common models. Therefore, the research issues concerning a measure to evaluate multiple tonal components in technical sounds and concerning the appropriatness of psychoacoustic annoyance for such sounds open up.

The project of the master's program consisted of a literature research, the implementation of program code of models and measures in Matlab to create a routine as well as the organisation, setup and execution of a hearing test on the perception of dissonance and annoyance for generated test sounds. In the following, the framework of the hearing test as well as the correponding results and an outlook on the data is illustrated.

Setup

Firstly, the generated stimuli are presented. After that, it is shortly explained which listeners had participated and in which test environment. Thereupon, instructions and the evaluation method of the hearing test are illustrated.

Synthesized sounds, generated in Matlab, were used. The stimuli contained two or three frequencies, each with a maximum SPL of 60 dB. The duration of the sounds was 10 s.
It was decided to keep the focus of the experiment on sounds containing two frequencies, due to the fact that in music theory two simultaneous sounding tones build the basis of dissonance perception. Also sounds containing three frequencies were generated, as it is assumed that a certain dissonance of a three oder multi-tone complex is composed of the dissonances of the included dyads.
In the following the different sets are described in detail: The fundamental frequency of all dyads and triads in the experiment was $f_0 = 400\,\mathrm{Hz}$.

Set 1: Dyads Within the Octave

This set consisted of 7 stimuli, which have a variation in their second tone frequency. Thereby, most of the intervals are inspired by music theory. Also, intervals were chosen where a perceived roughness is present. Therefore $f_1 : f_0$ is 1.025:1, 16:15 (semitone), 5:4 (major third), 45:32 (augmented forth), 3:2 (perfect fifth), 12:7 (minor seventh) and 2:1 (octave).

Set 2: Dyads Beyond the Octave

Set 2 also contained 7 stimuli, but with frequency ratios greater than an octave. Also the behavior beyond the octave was studied, due to the fact that in technical sounds various components in different frequency regions are likely to appear. $f_1 : f_0$ is 2.3, 3, 3.6, 4, 4.5, 9 and 10. It is assumed that also beyond the octave certain intervals show specific patterns.

Set 3: Exemplary Level Variations

This set contained 8 sounds with level variations on previously described dyads. The aim was to observe possible changes in dissonance perception due to the variations. The intervals 12:7 and 2:1 have been chosen exemplarily. Firstly, due to the fact that 12:7 is assumed to be more dissonant than 2:1 and secondly because the influence of roughness perception is neglectable on these intervals. The following level variations (SPL's in dB) were chosen: $\mathrm{SPL}_0 : \mathrm{SPL}_1$ is 60:55, 60:50, 55:60 and 50:60.

Set 4: Triads

Set 4 did only contain sound examples with three simultaneous sounding tones and included a total of 11 sounds. The sounds are followed the condition: $f_0 : var \cdot f_0 : 3 \cdot f_0$, with $var = \{1.05, 1.15, 1.25, 1.5, 1.75, 2, 2.25, 2.5, 2.75, 2.85, 2.95\}$. The aim was to see in which way the dissonance of dyads is influencing the dissonance perception of triads.

A total of 22 listeners participated in the test. Most of them were students in bachelor or master degree studies with no special musical knowledge. The investigations were done at a university room with low background noise. A consistent test environment for every subject was ensured.

Prior to listening to the test sounds, the subjects received several instructions. First of all, they were informed about the research project and what the sense of the investigations should be. This was done in order to give the tests a meaning for the subjects and to strengthen their attention. Then, explanations to the sensations dissonance and annoyance were introduced. They were taught to rate dissonance and annoyance for every of the presented test sounds. Which one they should rate first was not dictated. As in the literature, there are various explanations on dissonance, couple of them were presented to the subjects in order not to precondition them too strong.
Annoyance was described as a sensation which describes the perceived annoyance of the sound as a whole and that it might be easier to do an annoyance rating on the sounds if the subjects would imagine to be exposed to the sounds for a longer time. Dissonance was explained to be only one part of sensations which the subjects are able to perceive while listening. They were taught that sounds which are dissonant sound *blurred* or are rather inharmonious or the particular tones seem to be *a mismatch*. The latter one was the kind of explanation the subjects liked the most, as it was most general. Consonant sounds in contrast were defined to be sounds that are clear, euphonious, harmonious or the particular tones seem to be *a good match*. Additionally, they were told if they found a certain dissonance would have an influence on annoyance they should consider this in their evaluation.
To highlight the concepts practically, three exemplary sounds with different dissonance were also presented to the subjects. If they were satisfied with the explanations and instructions and seemed to understand the concepts, the presentation of the test sounds started.
For every subject a random playback sequence was chosen and they could hear the test sounds as often as they wanted or needed. This was achieved by generating a graphical user interface (GUI) in Matlab. They were instructed to place their answers on two nine categories scales ranging from *not dissonant* (1) to *very dissonant* (9) and *not annoying* (1) to *very annoying* (9). Furthermore, the subjects had to do the ratings on the test sounds twice. The second run had taken place one week after the first one.

Results

Subsequently, the results of the previously described hearing test are drawn. The graphs depict the means of all listeners after statistical evaluation. In addition to the representation of the results an assessment on observable tendencies regarding psychoacoustics and music theory is done.

Set 1: Dyads Within the Octave

Figure 1 shows that dissonance for the first two stimuli (from left to right) is high. This is the case when the sounds are perceived rough to a certain extent. As the stimuli of set 1 cover musical intervals it can be pointed out that the third and the fifth sound present local minima. In music theory those intervals are considered consonant. Also the fifth stimuli has a lower dissonance as the fourth, what is also depicted by music theory. The last sound, an octave, reaches the global minimum of the curve, what is musically also expected. That these results, or rather the visible trends, are reproducable can be observed in investigations of Fingerhuth in [10].
Although the octave is the most consonant interval in this test, the lowest mean value of dissonance only reached the category *little* dissonant.
Principally, the form of the curves of dissonance and annoyance is not the same. The maximum and minimum values of dissonance for example are spread over more than 6 categories, while the mean of the annoyance curve is only spread over approximately 2 categories.

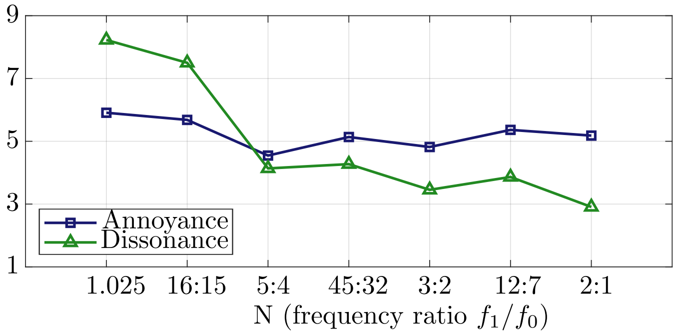

Figure 1: Dissonance and Annoyance ratings - Set 1 (Sound no. 1-7).

Set 2: Dyads Beyond the Octave

Beyond the interval of an octave several points were chosen to see what the understanding of dissonance is if intervals between tones get greater, see figure 2.

Dissonance as well as annoyance values are roughly spread over 3 categories. Also the form of the curves of both, annoyance and dissonance, is quite similar.

It seems like, the greater the intervals between tones are getting the harder it is for subjects to recognize a *harmonious* relation. This is striking, due to the fact that the 3rd and 4th octave are indeed local minima in the depicted graphs, but rated higher than the octave of set 1. Also, the 9th and 10th octave which are basically also in a harmonious relation are rated high in dissonance. It is also again slightly visible, that like predicted in music theory, integer intervals are perceived more consonant than odd ones. This trend can be observed by comparing the interval of 2.3 to 3 and the interval 3.6 to 4.

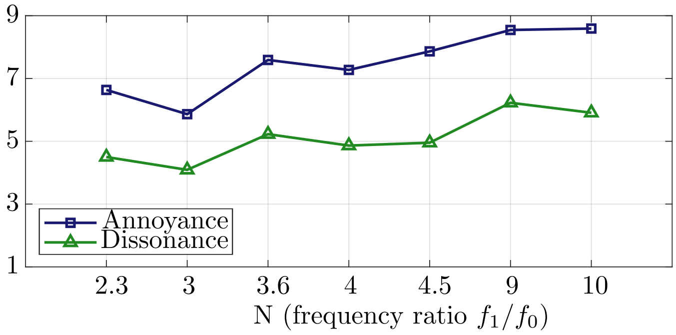

Figure 2: Dissonance and Annoyance ratings - Set 2 (Sound no. 8-14).

Set 3: Exemplary Level Variations

In this set some changes in SPL were applied to particular test sounds, to see if this has a remarkable influence on dissonance perception, compare figure 3. The additional information, given as annotations in the graph, depicts the relation of the levels SPL_0/SPL_1 in dB respectively.

Generally, the means of dissonance and annoyance resemble one another.

As in set 1 it becomes obvious that an interval of 12:7 is rated higher in dissonance than 2:1. No apparent differences of dissonance perception (at least for those examples) due to level changes can be observed. In annoyance ratings, the dyads where the second tone was reduced in SPL are rated a little lower annoying than those where the first tone is reduced.

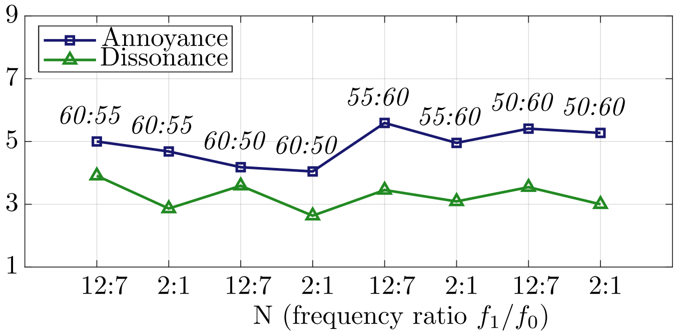

Figure 3: Dissonance and Annoyance ratings - Set 3 (Sound no. 15-22).

Set 4: Triads

The factor, which was changed in this set was the frequency of the second tone (f_1), f_0 as well as f_2 were kept constant. Figure 4 shows the results of set 4.

Perceived dissonance and annoyance for these sounds also resemble one another.

Dissonance and annoyance are on a maximum when two tones are separated by 20 Hz. The global minimum of the dissonance curve is shown by the interval of the octave. In this case the three tones form a harmonic complex tone. Again, local minima can be found for more simple intervals, like $N = 1.5$ or $N = 2.5$.

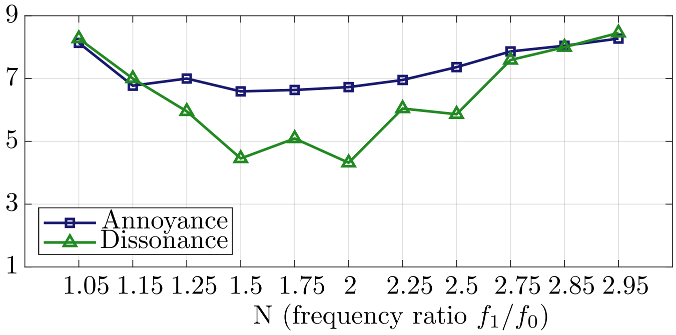

Figure 4: Dissonance and Annoyance ratings - Set4 (Sound no. 23-33).

Outlook

Based on the data collected, correlation analyses on existing literature models will be done. The aim is to investigate if significant correlations can be achieved by already existing models. The models to be investigated are the dissonance model according to Plomp and Levelt [11], Harmonic Entropy [12] and psychoacoustic annoy-

ance [5] (enlarged version of Zwicker's original annoyance model [4] including tonality.)
It will be interesting to see if any of the models concerning dissonance are able to illustrate the subjectively gained evualtion scores properly. In addition to that it is investigated in which way the psychoacoustic annoyance model fits the subjective evaluation scores of annoyance. If a suitable illustration of dissonance can be achieved, it might also be of interest if including dissonance in the psychoacoustic annoyance model would also increase the correlation. Nevertheless, it has to be stated that further hearing tests on dissonance and annoyance perception should be done. Inter alia to see if the visible trends are also reproducable in other frequency regions for example. Also other variations in level as well as increasing the number of simultaneous sounding tones would be appropriate in order to find valid relations for various technical sounds.

References

[1] DIN 45681: *Akustik - Bestimmung der Tonhaltigkeit von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschimmissionen*. Berlin: Beuth Verlag, Mar. 2005.

[2] SPIEGEL ONLINE: *Erfolg für Blindenverbände: Elektroautos müssen künftig wie Benziner und Diesel klingen*. Copyright: SPIEGEL ONLINE, Hamburg, Germany. 30.10.2018. URL: `http://www.spiegel.de/auto/fahrkultur/elektroautos-muessen-kuenftig-wie-verbrenner-klingen-a-1224344.html` (sighted: 11/20/2018).

[3] W. Aures: "Berechnungsverfahren fuer den sensorischen Wohlklang beliebiger Schallsignale". In: *Acustica* 59 (1985), pp. 130–141.

[4] Hugo Fastl and Eberhard Zwicker, eds.: *Psychoacoustics. Facts and Models*. 3rd edition. Springer-Verlag Berlin Heidelberg, 2007.

[5] Guo-Qing Di et al.: "Improvement of Zwicker's psychoacoustic annoyance model aiming at tonal noises". In: *Applied Acoustics* 105 (2016), pp. 164–170.

[6] Ernst Terhardt, Gerhard Stoll, and Manfred Seewann: "Algorithm for extraction of pitch and pitch salience from complex tonal signals". In: *The Journal of the Acoustical Society of America* 71 (1982), pp. 679–688.

[7] DIN 45681: *Akustik - Bestimmung der Tonhaltigkeit von Geräuschen und Ermittlung eines Tonzuschlages für die Beurteilung von Geräuschimmissionen, Berichtigungen zu DIN 45681:2005-03*. Berlin: Beuth Verlag, Aug. 2006.

[8] Soenke Kraft: *Musiktheorie*. Copyright: Soenke Kraft, Hannover, Germany. 2001. URL: `http://www.sackpfeyffer-zu-linden.de/Musik.html#0316` (sighted: 02/08/2018).

[9] J. Gedan: *Musiktheorie Teil II. Grundlagen der Harmonielehre*.

[10] Sebastian Fingerhuth: "Tonalness and Consonance of Technical Sounds". Dissertation. RWTH Aachen, 2009.

[11] R. Plomp and W. J. M. Levelt: "Tonal Consonance and Critical Bandwidth". In: *The Journal of the Acoustical Society of America* 38 (1965), pp. 548–560.

[12] William A. Sethares: *Tuning, Timbre, Spectrum, Scale*. 2nd edition. Springer-Verlag London, 2005.