PERCEPTION OF DISSONANCE IN TIME-VARYING TECHNICAL SOUNDS

Benedikt Bugl[1], Hans-Peter Rabl[1], Sabrina Freisleben[1], Stephan Töpken[2], Eike Claaßen[2], Steven van de Par[2], Anna Rieger[1]

[1] *OTH Regensburg, Regensburg, Germany*
[2] *Carl von Ossietzky Universität Oldenburg, Department of Medical Physics and Acoustics,* Acoustics Group and Cluster of Excellence "Hearing4all", Oldenburg, Germany

This paper is a revised version of the authors' submitted manuscript that was peer-reviewed by at least two anonymous reviewers.

Abstract

While dissonance is traditionally studied in musical contexts using rather stationary tonal structures, technical sounds require a perceptual perspective in which dissonance may also be assessed for time-varying spectral content.

This study investigates the perception of dissonance and unpleasantness in synthetic acceleration and deceleration sounds typical for electric vehicle interior noise. A laboratory listening experiment was conducted with 12 participants across two appointments. Long (8 s) instationary sounds and short (1 s) stationary and time-varying excerpts were rated. The stimuli consisted of broadband background noise combined with tonal orders, presented at two nominal tone levels. Participants rated unpleasantness and perceived dissonance on a categorical scale.

Results show that perceived dissonance is primarily determined by interval structure, as long as tones are sufficiently audible. Sounds are perceived more unpleasant, when higher frequencies are present and for more dissonant interval combinations, particularly for the higher tone level. Deceleration sounds are rated slightly less unpleasant than acceleration sounds, while dissonance ratings show no systematic difference between the two conditions. Comparisons between stationary and time-varying short stimuli reveal only minor differences.

Keywords: *psychoacoustics, dissonance, time-varying sounds, listening experiment*

1. Introduction

With the increasing market penetration of electric vehicles (EVs), interior vehicle noise is dominated by new key factors influencing perceived sound quality and user acceptance. Unlike combustion engines, electric powertrains lack broadband masking noise, which can lead to more salient tonal components originating from the electric motor and gearbox [1]. Especially in this case, multiple tonal components can occur simultaneously at different, sometimes neighboring frequencies.

In music perception, the simultaneous occurrence of tones is linked to a sensation of consonance and dissonance, depending solely on the frequency interval between the tones [2]. While the concept of dissonance has been extensively studied in musical contexts, less research has addressed how the presence of multiple tones influences the perception of complex, non-musical sounds such as vehicle or machinery noise. From a psychoacoustical point of view the sensation of dissonance can be influenced by many factors, such as the magnitude of tonal content, loudness or variations of a sound over time [3].

EV-related sounds are inherently time-varying and tonal frequencies and levels of noise sources change with rotational speed during acceleration and deceleration. For a psychoacoustic model aiming at a prediction of perceived dissonance this raises the question how short parts of a sound with different properties affect the perceived dissonance of a longer sound as a whole. It is also of methodological interest whether short stationary sound excerpts can adequately represent short time-varying signals in such a model but also in perceptual experiments regarding dissonance.

The present study investigates the perception of dissonance and unpleasantness in synthetic acceleration and deceleration sounds typical for electric vehicle interior noise. By systematically varying tonal interval structures, tone levels, temporal characteristics, and driving conditions, this work aims to clarify the role of tonal interactions in dissonance perception and to assess their contribution to perceived unpleasantness

Corresponding author: *benedikt.bugl@oth-regensburg.de.*

in time-varying technical sounds.

2. Methods

2.1. Procedure

The study was conducted as a laboratory listening experiment with two appointments. Participants were randomly assigned to one of two groups. Participants were assigned to a counterbalanced acceleration/deceleration order across two appointments. Each appointment consisted of four parts. In Part 1, 31 8-s run-up sounds were rated with respect to unpleasantness. In Part 2, 176 1-s excerpts of these sounds were rated for unpleasantness. In Parts 3 and 4, the same procedures were repeated, with participants rating the sounds for perceived dissonance instead of unpleasantness. Unpleasantness ratings were obtained first to minimize potential bias from a prior consideration of dissonance and to facilitate a clearer separation of unpleasantness and dissonance judgments.

Instructions were given in written form to reduce experimenter bias. Before the respective dissonance-tasks, participants received a detailed explanation of the concept of dissonance using auditory examples. Musical intervals ordered by increasing dissonance according to music theory were presented as synthetic piano tones as well as pure sinusoidal tones. In addition, pink noise (as a non-tonal sound) and combinations of pink noise with musical intervals were demonstrated. For all four tasks, an orientation phase preceded the ratings, during which participants were familiarized with the range of the sound material and the rating scale.

Subjects rated the unpleasantness and the dissonance of the sounds on an equidistant 9-point scale (0-8) with 5-point Rohrmann labels at 0, 2, 4, 6 and 8 [4], ranging from 0 ("not") to 8 ("very").

At the end of the psychoacoustic experiment, participants were asked for feedback on the experiment using an open question.

2.2. Stimuli

The basis of the experiment was a set of 31 time-varying acceleration sounds with 8 s duration, created with a custom synthesizer [5] and resembling the interior noise of an electric vehicle between 250 and 2500 rpm, with tonal components representing motor and gear orders that varied with rotational speed, as well as background noise. Inverter-related tones were explicitly not included in the present study.

Out of the 31 sounds, 11 were selected for further processing. Figure 1 illustrates how each selected 8-s sound (a) was segmented into eight consecutive 1-s time-varying snippets (b). The black lines indicate an exemplary order run-up during acceleration.

For each time-varying snippet, a corresponding stationary 1-s sound (c) was generated. These stationary sounds were designed to match the spectral content and values of the sound pressure level and frequencies at the temporal midpoint of the respective time-varying snippet.

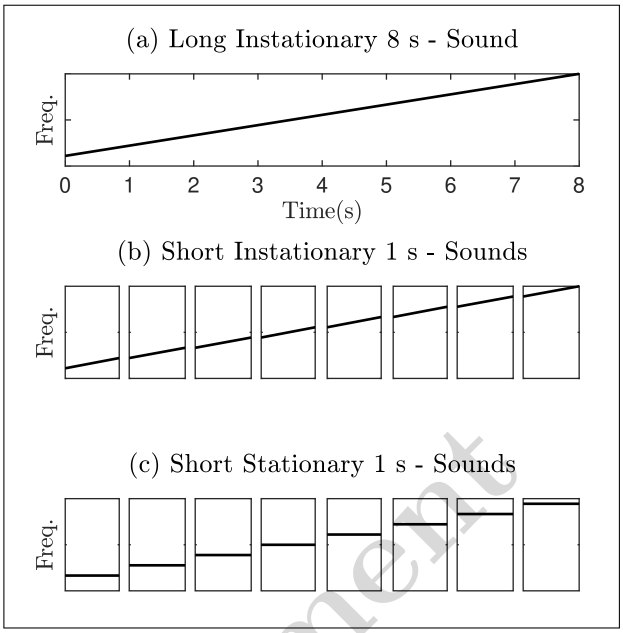

Figure 1: Visualization of the sounds used in the experiment. The plotted lines represent one single exemplary motor order of a sound in acceleration.

A previous categorical rating experiment with systematically varied stimulus durations, showed that the differentiation between stimuli with more and less dissonant intervals is markedly reduced for durations shorter than approximately 0.5 s. In contrast, durations of 0.5 s and longer produce clearly more distinguished and stable categorical judgments, thereby supporting the use of 1-s stimuli in the present study [6].

In total, each short-duration stimulus set comprised 88 time-varying and 88 stationary 1-s sounds.

The 8-s sound stimuli consist of background noise and synthesized tonal components.

The background noise is low-cut at 20 Hz and has a flat spectrum up to 40 Hz (to prevent low-frequency clipping), followed by a 12 dB/octave decay towards higher frequencies. Its sound pressure increases linearly from 56 dB to 74 dB SPL during the 8 seconds. The tones, representing typical motor or gearbox orders, are synthesized with a fixed sound pressure level, independent of motor RPM, constant over time.

The sounds contained up to four orders. Two nominal tone levels were used, defined with respect to the 24th order: 25 dB SPL (low) and 39 dB SPL (high). The levels of higher orders decrease with 6 dB per octave. This means that at the 25 dB nominal level, the 24th order sweeps from 100 to 1000 Hz at 25 dB, while the 48th order (one octave higher) sweeps from 200 to 2000 Hz at 19 dB (6 dB less).

Tone order combinations are selected to form intervals known from music theory: a Fourth, regarded as consonant in music theory, a Minor Second, a very dissonant interval and also conditions with only harmonic tones.

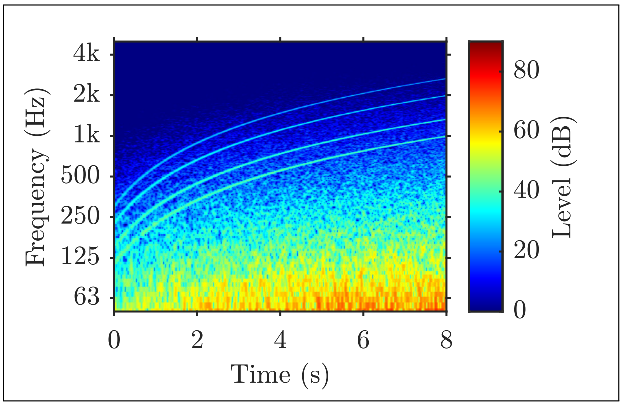

Figure 2: Exemplary spectrogram of an 8s-sound with four tone orders, forming a Fourth with a second harmonic, at 39 dB tone level.

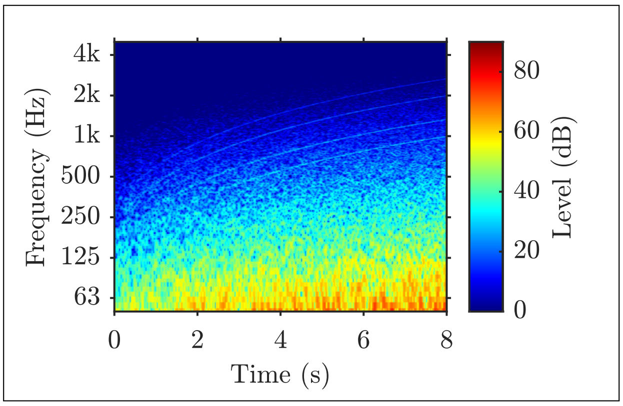

Figure 3: Exemplary spectrogram of an 8s-sound with four tone orders, forming a Fourth with a second harmonic, at 25 dB tone level.

All deceleration sounds were generated by reversing the acceleration sounds.
Figures 2 and 3 show the spectrograms of sounds with a 24th, a 32nd, a 48th and a 64th order at the higher (Figure 2) and lower (Figure 3) tone level. Especially at the lower tone level the high relative level of background noise is masking the orders at low frequencies at the beginning of the run-up. Towards higher RPM, the frequency increases and the local tone-to noise ratio gets higher due to the steep spectral decay of the background noise. The higher the order, the more it stands out, because of the lower decay over order compared to the decay over frequency of the background noise. So an order that is one octave higher has a 6 dB higher local tone-to-noise ratio.

2.3. Listening Setup

The listening experiment was carried out in a sound attenuated listening booth. The sounds were played from a computer running a Matlab-App connected to Sennheiser HD650 open headphones via a RME Fireface UFX III audio interface. The listening setup was calibrated using a GRAS 43AA-S2 CCP Ear Simulator and a 1 kHz sine with a defined level. All sounds were presented diotically with a 20 ms fade-in and fade-out using half-Hann (raised-cosine) envelopes.

2.4. Participants

Overall, 12 volunteers participated in the listening experiments. Subjects with self-reported hearing loss did not participate in the study. The Joint Ethics Committee of Bavarian Higher Education Institutions (GEHBa) had no objections regarding the study (GEHBa-202405-V-186).

3. Results

Figure 4 shows unpleasantness and dissonance ratings for three sounds presented at the higher tone level (39 dB). The blue data points correspond to a sound containing only the 24th and 48th orders, resembling an octave or respectively a sound with a second harmonic. The green data points represent a sound with additional 32nd and 64th orders, which form a Fourth relative to the base frequencies. The red sound includes 25.4th and 50.8th orders, corresponding to a Minor Second interval relative to the 24th and 48th. Ratings are shown for successive time segments of the short sounds. The rightmost data point represents the rating of the corresponding long sound.
The unpleasantness ratings increase towards the later sections of the sound in all three order combinations. The difference between the curves also seems to get smaller towards the later sections, especially between the Minor Second (red) and the Fourth (green). The general increase in unpleasantness towards the later sections of the sound might be explained by the increasing loudness, as the tone orders move to a frequency range, where the ear is more sensitive [7] and the background noise level increases, as well as the higher tone to noise ratio causing an increase in tonality, which has been shown to correlate to higher unpleasantness ratings in a similar experiment [8]. The lower number of tones in the sounds can therefore also partially explain the lower ratings for the sound with no interval tones (blue). For most datapoints, the more dissonant tone combination is also perceived as more unpleasant. The curves getting closer towards the later sections could be explained by a perceptual dominance of sharpness over the dissonance sensation, as the tones in the stimuli are reaching frequency ranges where sharpness has a strongly increasing perceptual effect [7]. The ratings of the corresponding 8-s sounds (grey shaded area) are close to those of the most unpleasant sections for the 39 dB tone level.

The dissonance ratings increase toward the later sections of the sound only very slightly, indicating that subjects rated dissonance relatively independently of the increasing tone-to-noise ratio and frequency content at this higher tone level. The overall arrangement of the curves in ascending dissonance suggests that the tone combinations have meaningful influence on perceived unpleasantness, highlighting the importance of dissonance perception.

Figure 5 shows the results for the sounds with the

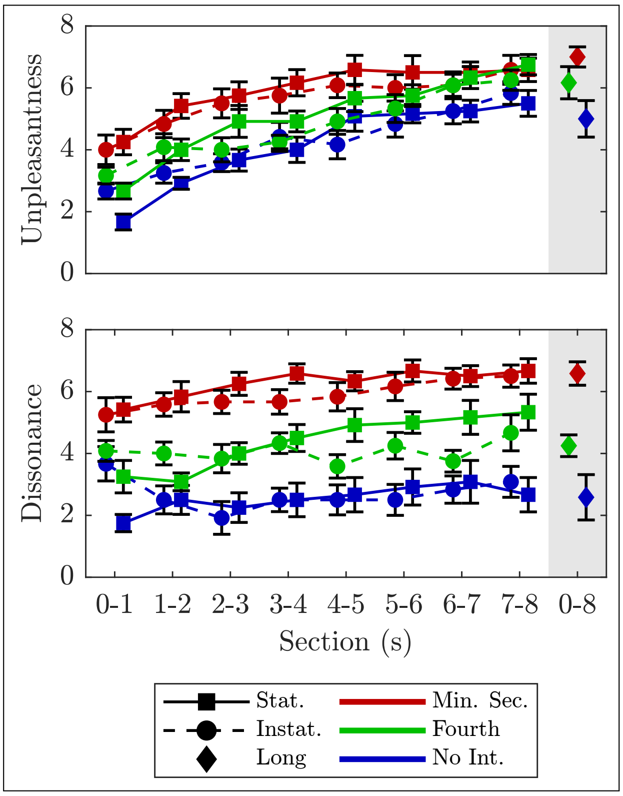

Figure 4: Unpleasantness and Dissonance ratings for short and long sounds at 39 dB tone level.

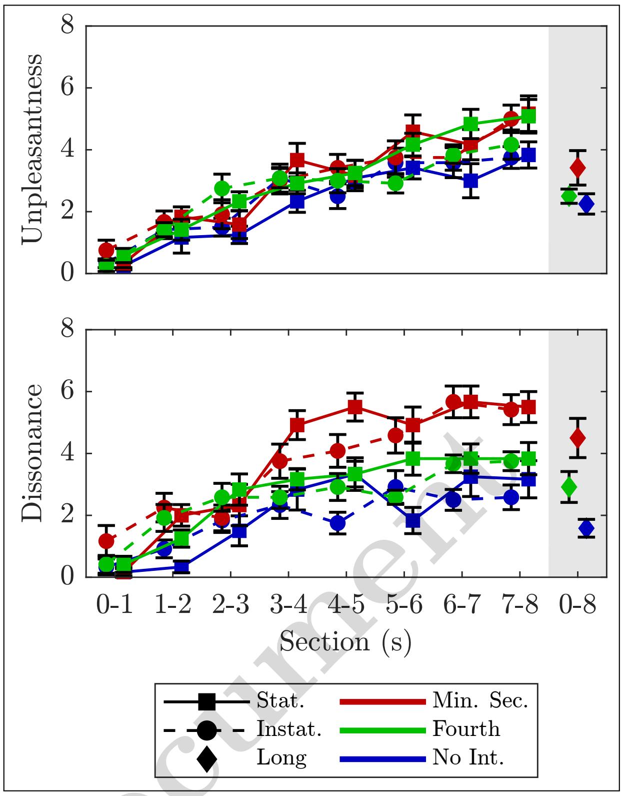

Figure 5: Unpleasantness and Dissonance ratings for short and long sounds at 25 dB tone level.

lower tone level.
The unpleasantness ratings are generally lower compared to the higher-level conditions which can be explained by lower loudness and tonality. For the long sounds, unpleasantness ratings show the same order as for the higher tone level, although the differences between conditions are small in this figure. For the short stimuli the separation of the three interval conditions, as observed in Figure 4, is no longer clear.
The Dissonance ratings increase initially over the first four sections, followed by a plateau for the later sections, where the dissonance ratings seem to approach the interval order as seen in the plot for the higher tone level. With a look at the spectrogram of a sound in the lower tone level condition in Figure 3, the low dissonance ratings at the first sections of the sound might be explained by the masking of the tone orders by the background noise. A masking of the tones also reduces the salience of low-frequency envelope fluctuations associated with dissonance perception [2]. In conclusion these results support the hypothesis, that tones must be sufficiently audible for dissonance to be perceived.

Figure 6 shows the ratings for acceleration (red) and deceleration (grey) in a sound with the Minor Second interval at the 39 dB tone level. The segments of the deceleration sound are plotted in reverse time order, so that the sections with the identical RPM-range are at the same x-tick each.

Regarding the unpleasantness ratings, both long time-varying and short time-varying sounds are slightly lower for deceleration than for acceleration.The dissonance ratings are very close together, except for the 0-1 s acceleration section, where the tones are least audible in both conditions. This similarity may be attributed to the generally stable perception of dissonance when the frequency interval is constant throughout the sound.

To derive a single prediction value for the perceived unpleasantness or dissonance of the 8-s sounds from the ratings of the time-varying 1-s segments, several aggregation metrics were evaluated (Figure 7). Specifically, the mean, median (50th percentile, p50), and upper quantiles (60th–90th percentile, p60-p90) of the segment ratings were compared with respect to their ability to predict the ratings of the full 8-s sounds. Prediction quality was quantified by the mean squared error (MSE) between the aggregated metric from the short sections and the rating of the corresponding long stimulus.
This comparison is relevant for perceptual modeling because many time-varying sound descriptors produce temporal sequences of values. To predict a single perceptual rating for a longer sound, it is important to understand how these time-series values must be integrated to a representative single value.
For the unpleasantness ratings of the 25 dB tone level stimuli, the 60th and 70th percentiles yielded the

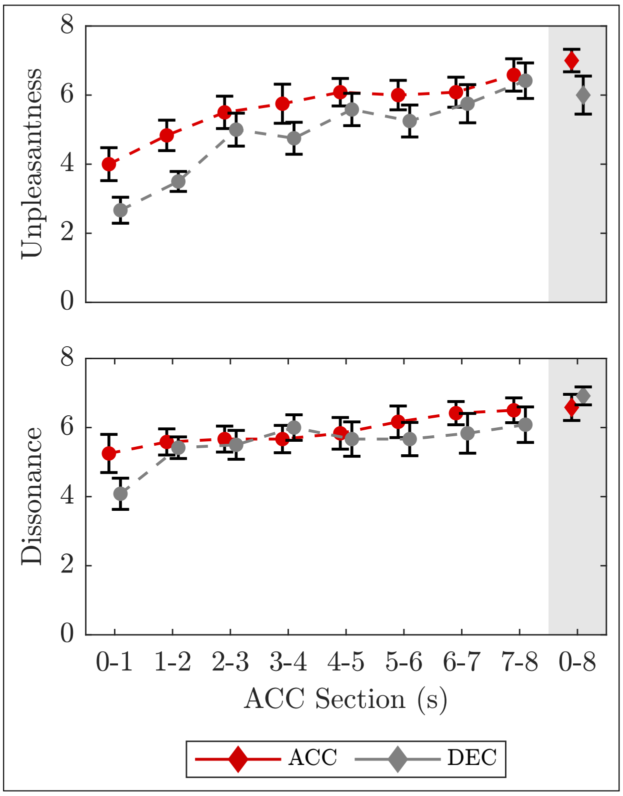

Figure 6: Unpleasantness and Dissonance ratings for acceleration and deceleration (reversed x-axis for better comparison) in the Minor Second condition at the 39 dB tone level.

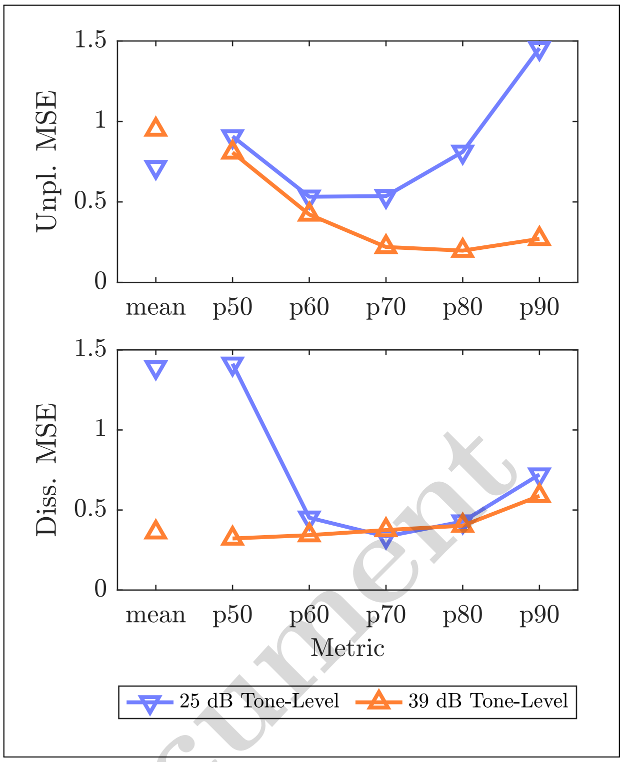

Figure 7: Mean squared errors (MSE) between the ratings of the long sounds and different metrics calculated from the ratings of the corresponding time-varying sounds.

lowest MSE, while for the 39 dB tone level stimuli the 80th percentile performed best. This suggests that segments with unpleasantness ratings above the mean value drive the overall judgment.

For dissonance, the pattern differed. For the 39 dB stimuli, the median (50th percentile) provided the best prediction, whereas for the 25 dB tone level stimuli the 70th percentile performed best and the median resulted in substantially larger prediction errors. One possible explanation for this is that only parts of the 25 dB tone level stimuli contained clearly audible orders. Consequently, the dissonance ratings of the long sounds may primarily reflect those segments in which tonal components are perceptually salient, making higher quantiles of the segment ratings more representative of the overall percept.

The statistical analysis of the data regarding drive condition, stationarity and appointment sequence did not show significant main effects. On average, the stationary short sounds were rated slightly more unpleasant and more dissonant.

4. Discussion

The findings indicate, that the interval structure plays a clear role in dissonance perception as long the tonal components are sufficiently audible. At the high tone level, the three interval conditions were consistently differentiated in terms of perceived dissonance across most sections of the sounds. The ordering of the ratings largely followed expectations based on music theory, with the Minor Second producing higher dissonance ratings than the Fourth, and the condition without interval tones producing the lowest ratings.

The results obtained at the lower tone level indicate that a certain amount of perceptual salience of the tonal components is necessary for dissonance judgments. At this level, the tonal components were partially masked by the background noise for the early sections of the acceleration sounds, resulting in reduced differentiation between the interval conditions. These findings suggest that the perception of dissonance in technical sounds depends not only on the tonal frequency relationships but also on the audibility of the interacting tones.

The observed relationship between dissonance and unpleasantness further highlights the role of tonal interactions in the judgement of sound quality. When tones were sufficiently audible, sounds containing more dissonant interval structures were also rated as more unpleasant. However, unpleasantness ratings were influenced by additional factors beyond interval structure. In particular, unpleasantness increased toward the later sections of the acceleration sounds, even when dissonance ratings remained relatively stable. This trend might be attributed to the increasing sharpness and loudness due to the tonal components as their frequency moved into regions of higher auditory sensitivity which seems to dominate the unpleasantness ratings in this case. These effects are consistent with previous research showing that loudness, tonality, and sharpness contribute to the perceived unpleasantness of technical sounds [7], [8].

The comparison between acceleration and deceleration sounds provides further insight into the temporal integration of sound quality judgments. The trend of lower unpleasantness ratings for deceleration sounds may indicate a recency effect, whereby the final portions of a sound exert a stronger influence on the overall evaluation. In the present case, acceleration sounds typically end with higher-frequency components, louder background noise but also higher tone to noise ratios, whereas deceleration sounds terminate with lower frequencies, quieter background noise and lower tone to noise ratios. Such recency effects have previously been reported for retrospective sound-quality evaluations [9]. In contrast, studies on loudness perception suggest that earlier portions of a sound with a duration of about one second can receive greater perceptual weight, indicating a more pronounced primacy effect in loudness judgments (e.g. [10]). The dissonance ratings show little difference between acceleration and deceleration which might be explained by the fixed the interval structure in both cases. Thus, in the context of electric vehicle drive train sounds, where the order ratios are usually not time-dependent, one might assume that dissonance ratings for acceleration and deceleration will yield similar results.

The comparison of aggregation metrics suggests that median or higher quantiles are more suitable than simple averages for deriving a single number prediction value from a time series of perceptual estimates. The results further suggest that the perceptual salience of tonal components should be explicitly represented in predictive models, since listeners appear to base their ratings primarily on segments in which tonal orders are clearly audible.

In summary, the method of using the 1 s sound sections in the listening experiments provided valuable insights into perceptual differences between individual sections of the sounds and their contribution to the overall rating.

Future work will build on these findings in several ways. One promising direction is the investigation of temporally varying interval structures, for example by systematically varying the progression of interval combinations over time, to better understand how dynamic spectral relationships influence perceived dissonance in technical sounds. The aspect of temporal integration of perceptual judgments needs further research to better understand how single-value evaluations of longer sounds are formed from perceptual impressions of shorter sections, and which temporal weighting mechanisms may best represent listeners' overall judgments.

Acknowledgements

The authors acknowledge the financial support for the research project *AudioOpt* by the German Federal Ministry of Research, Technology and Space as part of the funding program *VIP+* (contract number 03VP11931). The authors would also like to thank the industrial partners in the associated committee.

References

[1] V. Abbink, D. Landes, and M. E. Altinsoy, "Experimental determination of the masking threshold for tonal powertrain noise in electric vehicles," *Acoustics*, vol. 5, no. 4, pp. 882–897, 2023. DOI: 10.3390/acoustics5040051.

[2] W. A. Sethares, *Tuning, Timbre, Spectrum, Scale*, 2nd ed. London: Springer, 2005, p. 426, ISBN: 978-1-85233-797-1. DOI: 10.1007/b138848.

[3] A. Rieger et al., "Psychoacoustic dissonance theory supported in listeners' ratings of synthesized electric vehicle acceleration sounds," in *Fortschritte der Akustik - DAGA 2026*, Dresden, Germany, 2026. DOI: 10.71568/daga2026.557.

[4] B. Rohrmann, "Empirische Studien zur Entwicklung von Antwortskalen für die sozialwissenschaftliche Forschung (empirical studies of rating scales for social science research)," *Zeitschrift für Sozialpsychologie*, vol. 9, pp. 222–245, 1978.

[5] B. Bugl, H.-P. Rabl, S. Töpken, E. Claaßen, S. van de Par, and A. Rieger, "Implementation of a synthesis tool for stationary and nonstationary interior noise," in *Proc. DAS—DAGA 2025*, Copenhagen, Denmark, 2025, pp. 1499–1502.

[6] B. Bugl, H.-P. Rabl, S. Töpken, E. Claaßen, S. van de Par, and A. Rieger, "Influence of signal duration on subjective dissonance ratings," in *Fortschritte der Akustik - DAGA 2026*, Dresden, Germany, 2026. DOI: 10.71568/daga2026.260.

[7] H. Fastl and E. Zwicker, *Psychoacoustics: Facts and Models*. Berlin, Heidelberg: Springer-Verlag, 2006, ISBN: 3540231595.

[8] S. Töpken, E. Claaßen, S. van de Par, H.-P. Rabl, B. Bugl, and A. Rieger, "Tonality ratings and unpleasantness judgements for synthesized EV acceleration sounds," in *Proceedings of the Aachen Acoustics Colloquium 2025*, Aachen, 2025, pp. 1–8.

[9] D. Västfjäll, "The "end effect" in retrospective sound quality evaluation," *Acoustical Science and Technology*, vol. 25, no. 2, pp. 170–172, 2004. DOI: 10.1250/ast.25.170.

[10] E. Ponsot, P. Susini, G. Saint Pierre, and S. Meunier, "Temporal loudness weights for sounds with increasing and decreasing intensity profiles," *The Journal of the Acoustical Society of America*, vol. 134, no. 4, EL321–EL326, 2013. DOI: 10.1121/1.4819184.