# Comparison of Different Loudness Models with Subjective Evaluations of Real Sounds: Vacuum Cleaner and Shaver Examples

Serkan Atamer, M. Ercan Altinsoy
*Chair of Acoustic and Haptic Engineering, TU Dresden, 01062 Dresden, Deutschland*
*E-Mail: serkan.atamer@tu-dresden.de*

## Introduction

Loudness is a fundamental psychoacoustical unit, representing the perception of intensity of the stimuli [1]. Available standardized loudness models are proved to be efficient to characterize the intensity perception of synthetic sounds, understanding the underlying structure, especially in the tests where subjects need to change the level of a stimuli to equalize the perceived loudness with a reference stimuli, as well as defining the threshold in quiet and equal loudness contours [2-3].

However, as a term, “loud” is rather complex, and it might be not so easy to have a common understanding when the complex real signals are involved. As a listener, what do we understand from the term “loud” is at that point important, if it is expected from the calculated loudness values to represent a perceptual attribute which is accepted amongst all of the listeners, at least the statistical majority.

The main aim of this study is to understand the perception of the loudness by using real signals as stimuli in listening tests and compare the evaluations with available standardized loudness models.

## Methodology

In order to understand the loudness perception of real, complex sound sources, vacuum cleaners and shavers are selected as example, since their almost stationary characteristics limit the effect of time variant characteristics by keeping the sound colors different for the stimuli.

Selected stimuli are presented to the subjects with a basic research question of “How loud is that sound?” Before the listening test, it is explained to the subjects, that test includes vacuum cleaner and shaver sounds, in order to clarify the context of the study. It is particularly asked from the subjects to concentrate only on “loudness” of the sounds, not the other attributes like having a high pitch, low pitch or any rough characteristics. However, it is not clearly defined to subjects what “loud” is and it is left to subjects to make estimations on their own perception of term “loudness”

At the end, statistically averaged estimated loudness values are compared with calculated loudness values, based on three available standards, ISO 532, DIN 45631 and ANSI S3.4 for two equipment under test [4-6].

## Sound Samples

For the vacuum cleaners and shavers, some particular stimuli are selected from a broad stimuli pool, such that, the calculated loudness values are arranged with a proper order, i.e. not having too much differences between the successive stimuli.

For both units, sounds are recorded in anechoic chamber. Vacuum cleaners are standing on top of a reflecting surface, with and without a carpet, in order to increase the diversity of obtained stimuli. Shavers are recorded in idle running mode, without any contact with skin surface, 15 cm away from the ear position. Shavers are always connected to external power, in order to make sure that they are always working in the high efficiency mode, since a low battery might change the speed and noise characteristics. At the end, for the vacuum cleaners 15 stimuli and for the shavers 23 stimuli are selected.

Figure 1 and 2 shows three example spectrograms of the selected stimuli for vacuum cleaners and shavers. The three spectrograms, for both cases, selected such that, the left spectrogram is for the stimuli having the lowest loudness value, middle spectrogram having the value in the middle and the right spectrogram with highest loudness values, calculated according to DIN 45631.

Vacuum cleaners usually show a broad band characteristics with some particular tonal components added to broadband characteristics, while the shavers are basically sound sources of added multiple tonal components on top of each other. At some frequency range, tonal components approach to each other so closely, giving the noise as a rough characteristics.

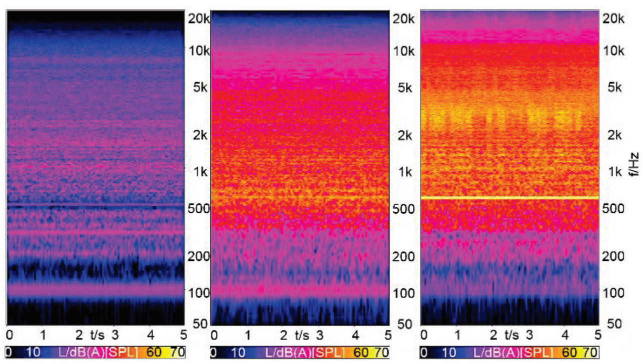


**Figure 1:** Spectrograms of three vacuum cleaners, having the minimum, middle and maximum loudness values, from left to right.

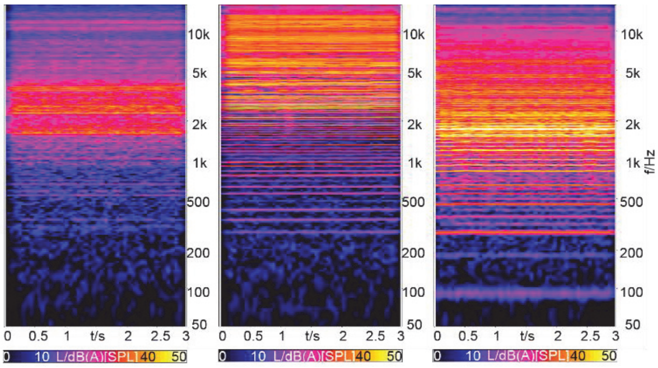


**Figure 2:** Spectrograms of three shavers, having the minimum, middle and maximum loudness values, from left to right.

## Calculated Loudness Values

Figure 3 and 4 show the loudness values for vacuum cleaner and shaver stimuli, respectively. Calculated loudness values according to DIN 45631, for the vacuum cleaners, are between 7.8 and 38.5 sone; and for the shavers between 7.4 and 15.1 sone. ISO and DIN standards are giving almost the same results, while ANSI method having the results higher than the other two methods.

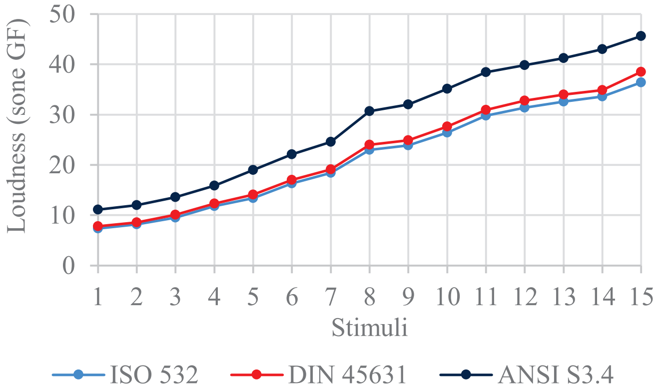


**Figure 3:** Calculated loudness values of vacuum cleaner stimuli

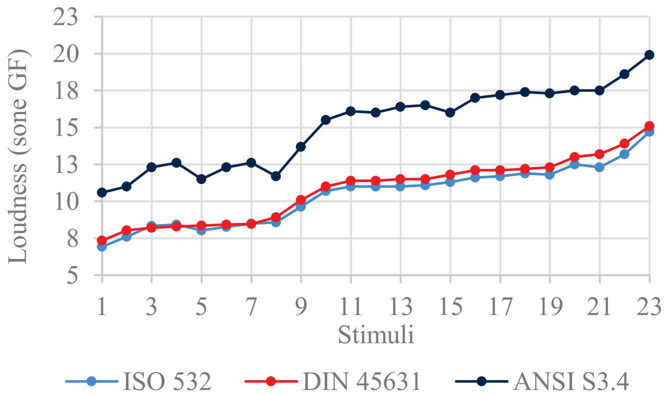


**Figure 4:** Calculated loudness values of shaver stimuli

## Listening Tests: Category Scaling

Sound samples are presented to the 20 participants, 7 women and 13 men aged between 22 and 66, through Sennheiser HD600 headphones. Experiments are conducted in a sound attenuating room. Stimuli are presented in random order and 5 random stimuli are presented before the test as sample stimuli. Every stimulus is presented three times, to check inter-individual validity. The subjects are then asked to evaluate the loudness of the sounds on a quasi-continuous scale (from 0 to 100) with equidistance neighboring categories (not at all, slightly, moderately, very, extremely).

Results of the listening tests are given in the following sections, with the comparison of calculated values for vacuum cleaners and shavers.

## Estimations vs. Calculations

Figure 5 and 6 shows the calculated (based on DIN 45631) and estimated values for vacuum cleaner and shaver stimuli, respectively. Error bars representing the standard deviation amongst the subjects. For the shavers, stimuli 1, 21 and 22 are found out that the estimated values are systematically lower than calculated loudness values according to DIN 45631.

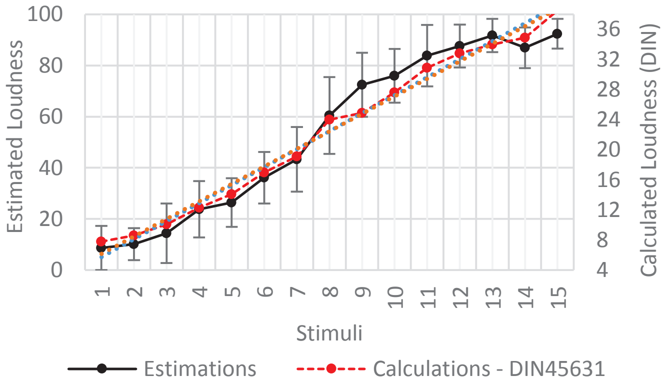


**Figure 5:** Estimated vs. calculated loudness for vacuum cleaners, error bars representing standard deviation

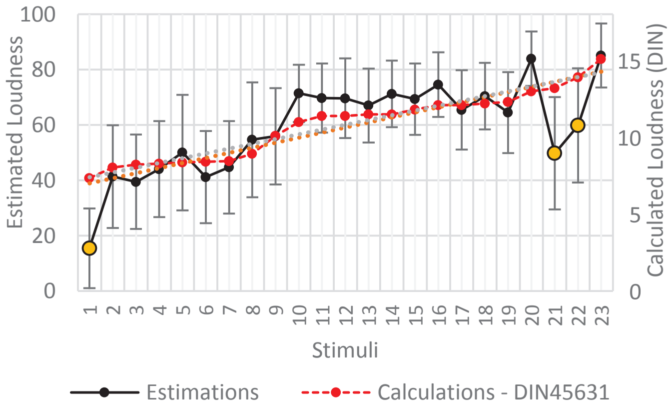


**Figure 6:** Estimated vs. calculated loudness for shavers. Error bars representing standard deviation. Note that stimuli 1, 21 and 22 have lower estimation values

## Results of Listening Tests

For both cases, listening test results show normal distribution. Inter-individual differences are all in the acceptable range, and the results are given in mean values amongst the subjects and repetitions.

For the vacuum cleaners, calculated loudness values and estimated loudness values are showing good correlation, however, for the shavers, stimuli 1, 21 and 22 have less perceived loudness values than calculated loudness values.

Figure 6 is given in comparison with estimations vs. DIN 45631 norm, although for the other calculation standards, it is possible to see the same particular trend for stimuli 1, 21 and 22.

## Psychoacoustical Parameters

In order to understand the results for shavers in detail, for the 23 shaver stimuli, A-weighted sound levels, roughness, sharpness and tonality values are calculated and the results are given in Figure 7-10. Sharpness calculations are based on DIN 45692 standard, roughness calculations are based on Aures model and tonality calculations are based on the publications by Terhardt and Aures.

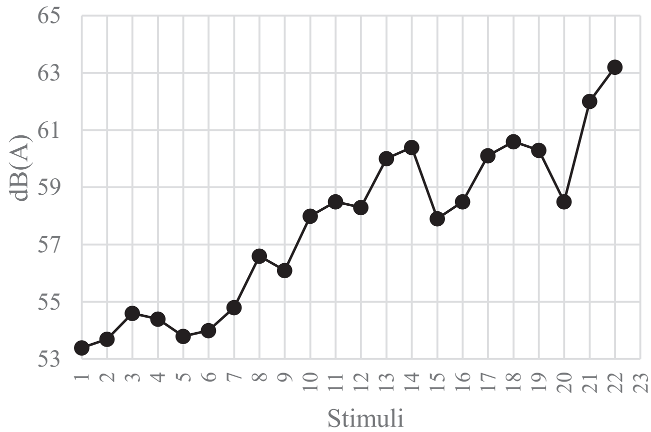


**Figure 7:** A weighted sound levels for the shaver stimuli

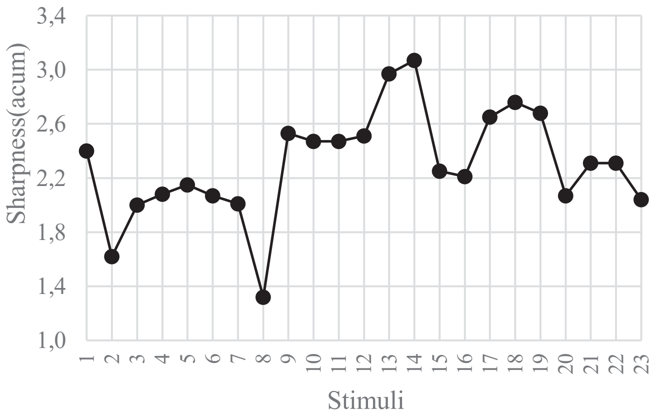


**Figure 8:** Sharpness of the shaver stimuli

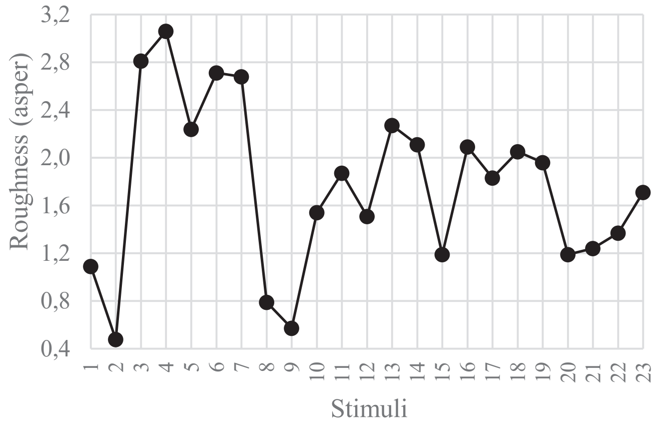


**Figure 9:** Roughness of the shaver stimuli

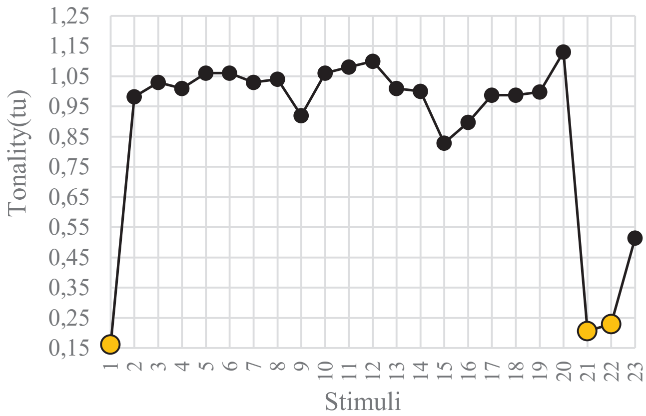


**Figure 10:** Tonality of the shaver stimuli, having 1, 21 and 22 has lower tonality values than most of the other stimuli

## Results and Discussions

In this study, vacuum cleaner and shaver sounds are used to understand the perceived loudness and its comparison with the calculated loudness values based on available loudness models, particularly DIN 45631. 15 stimuli for vacuum cleaners and 23 stimuli for shavers are used for the listening tests. Subjects are asked to evaluate the loudness of a sound sample, consisting all real recordings without any synthetic stimuli.

For the listening tests, category scaling method is used with a quasi-continuous scale and for both cases, results show normal distribution and the results are given in mean values with standard deviations.

For the vacuum cleaners, loudness estimations show great correlations with calculated loudness based on DIN 45631, however, for the shavers, although the overall trend is quite similar between estimated and calculated values, three particular stimuli, named within the study as stimuli 1, 21 and 22, shows systematically lower estimated values than calculated values.

In order to understand the possible reason behind this fact, different features of the shaver stimuli is calculated, such as A-weighted sound levels, sharpness, roughness and tonality values.

No correlations are observed, especially for the stimuli 1, 21 and 22 between A-weighted sound levels, roughness and sharpness values. Relation between those three dimensions seems to be not correlated with the estimated loudness values.

However, it is observed that, tonality values for stimuli 1, 21 and 22 are quite lower than the rest of the stimuli. Most of the stimuli of shavers having tonality values around 1.05 tu, while these three particular stimuli around 0.2 tu.

It is observed that, in the absence of the tonal components, when compared with the other tonal stimuli in the stimuli pool, perceived loudness, or perceived intensity is lower, although the calculated standards give more or less the same loudness values.

Figure 11 and 12 shows the frequency content of the two neighboring stimuli in shavers. Figure 11 shows the results for stimuli 20 while Figure 12 shows the stimuli 21. In terms of DIN 45631, calculated loudness values for those two stimuli are 13 and 13.2 sone GF, while A-weighted sound levels are 58.7 dB(A) and 62.1 dB(A), respectively. However, estimated loudness values, averaged amongst the repetitions and subjects are 84 points for stimuli 20, while 50 points for stimuli 21, in a range between 0-100.

It is observed that, when more tonal components in a critical band are present, in comparison to a broad band noise falling in the same critical band, calculated loudness for the broadband noise is more than the estimated loudness results. This fact could be taken into account to modify the available loudness calculation standards, in order to have more efficient standardized loudness calculation methods for real signals, representing the stimuli perception of majority of the listeners without hearing impairments in a more efficient manner.

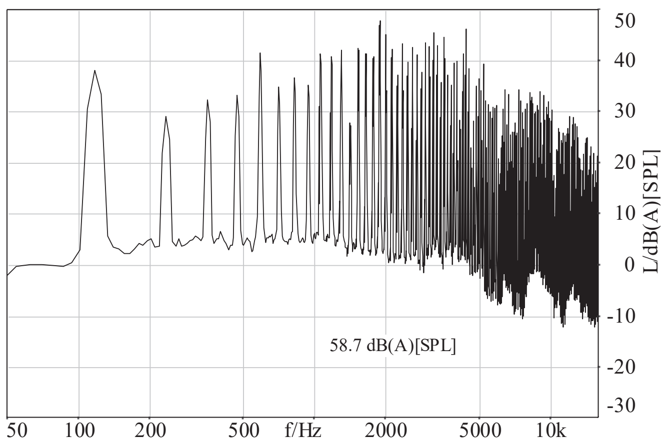


**Figure 11:** Frequency content of the stimuli 20 (spectrum size 4096)

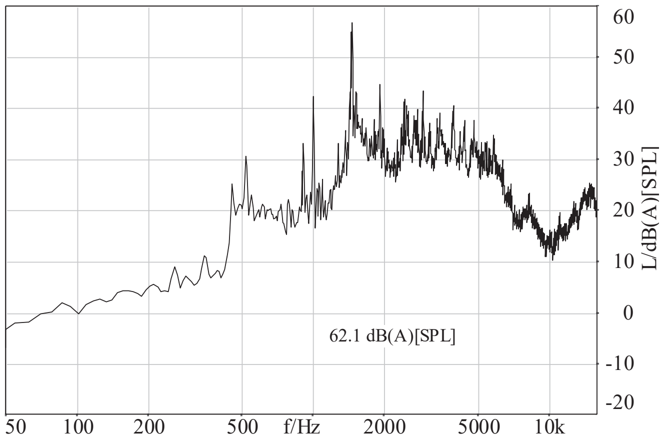


**Figure 12:** Frequency content of the stimuli 21 (spectrum size 4096)

## References

[1] Fastl, Hugo, and Eberhard Zwicker. Psychoacoustics: facts and models. Vol. 22. Springer Science & Business Media, 2006.

[2] Fastl, Hugo, Florian Völk, and Martin Straubinger. "Standards for calculating loudness of stationary or time-varying sounds." Proc. Inter-Noise 2009, Ottawa, Canada. 2009.

[3] Moore, Brian CJ. "Development and current status of the “Cambridge” loudness models." Trends in hearing 18 (2014).

[4] ISO 532 (1975), “Method for calculating loudness level”

[5] DIN 45631 (1991), “Berechnung des Lautstärkepegels und der Lautheit aus dem Geräuschspektrum – Verfahren nach E. Zwicker”

[6] ANSI S3.4-2007, “Procedure for the Computation of Loudness of Steady Sounds”