Audio Engineering Society

Conference Paper

Presented at the 2022 International Conference on
Automotive Audio
2022 June 8–10, Dearborn, MI, USA

Enhanced Perceptual Rub & Buzz Measurement for Testing Automotive Loudspeakers

Steve Temme

Listen, Inc., Boston, MA, 02118, USA

Correspondence should be addressed to Steve (stemme@listeninc.com)

ABSTRACT

Loudspeaker Rub & Buzz faults are a problem for automotive manufacturers as they sound harsh and immediately give the perception of poor quality. There are two places such faults can occur - during speaker manufacturing and installation of the speaker in the car. A buzzing loudspeaker in a car is disappointing to a customer and is costly to replace. It is also challenging for a service center to determine exactly where the buzzing is coming from and whether it is caused by a faulty loudspeaker or bad installation. Perceptual distortion measurements are often considered the holy grail of end-of-line testing because rejecting speakers with only audible faults increases yield. Although such measurements have been around since 2011, production line adoption has been slow because until now, sensitivity to background noise has made limit-setting challenging. In this paper, a new algorithm is introduced that uses advanced technology to reduce the impact of background noise on the measurement and offer more repeatable results. This facilitates limit setting on the production line and makes it a truly viable production line metric for increasing yield. This same metric may also be used for end-of-line automotive quality control tests. Results from various algorithms will be shown, and their correlation to subjective and other non-perceptual distortion metrics explained.

1 Introduction

The automotive industry's stringent quality expectations make end-of-line quality testing on automotive speakers and drivers absolutely critical. End-of-line tests typically measure a range of parameters including frequency response, THD, and polarity. Manufacturing-introduced defects such as Rub & Buzz and Loose Particles [1] are also measured. Reliable, automated testing has been available for decades now, and most large manufacturers rely on these software-based systems for identification and rejection of defective products. While these tests do an excellent job of identifying defective units, there is always a certain level of false rejection where units with some distortion fail even though it is completely inaudible to the human ear.

From a manufacturing perspective, higher yields and therefore greater profitability is always desirable. This has driven the development of perceptual distortion measurements - automated measurements that replicate the human hearing to detect only audible distortion defects. Such metrics increase production line yield by passing products with inaudible distortion, as the product will still sound exactly as the manufacturer intended.

Perceptual methods are very simple to configure for production line use. Since they return a result in Phons, an absolute measurement that can be easily correlated to the listener’s threshold of hearing, the operator can set a fixed limit across the board, regardless of product. Naturally, the price point and quality expectations for the product may influence the level of distortion that is deemed acceptable.

Our algorithm [2], introduced in 2011, was the first commercial perceptual distortion metric, although in the past couple of years, other test system manufacturers have also started to offer perceptual distortion tests. It offers excellent correlation with human hearing and performs well in laboratory tests. However, like the human ear, repeatability decreases in the presence of background noise. This is not a failure of the algorithm as such, but an indication that the algorithm performs just like a human listener; when background noise is high, audible distortion is masked. This limitation restricts the value of such algorithms on the production line, as with today's high-volume manufacturing, there is only time for one fast test sweep. If this sweep gets a different result under changing background noise conditions, limit setting becomes challenging, and repeatability and reliability is decreased. Similar algorithms from other test system manufacturers also suffer from the same problems.

This paper details efforts to create an algorithm that hears like a human in quiet conditions, e.g. in a living room or passenger automotive cabin, under the less-than-perfect conditions of a manufacturing environment where considerable and varying background noise may be present. In other words, a perceptual model that is more independent and reliable than the human ear when it comes to noisy environments.

The resulting new algorithm overcomes these limitations to offer repeatable end-of-line test results, even in noisy environments. It incorporates noise reduction techniques and enhanced perceptual filters to overcome the reliability and high frequency masking issues of earlier versions. In short, the algorithm offers the performance of an 'enhanced' human ear - it detects distortion like an ear in a quiet environment, even when there is background noise. This makes it a viable solution for production line use.

In this paper we explain how the algorithm works, demonstrate how the results compare with earlier perceptual algorithms and show its correlation with human hearing and conventional distortion algorithms. We also compare its performance in the presence of background noise to other perceptual algorithms by adding recorded factory background noise to the signal before passing it through the algorithms.

2 Algorithm Development

The new algorithm was based on the author's earlier algorithm [2]. Although this was well received by the industry, repeatability issues were identified when used on the production line; small variations in results when there is only one fast measurement made during the end-of-line test makes it challenging to set limits and increases the possibility of false rejects or acceptances.

This algorithm primarily set out to resolve these issues, but also used the company's original research to improve the perceptual filtering algorithms for greater correlation with human hearing, especially at higher frequencies.

The construction of the algorithm is explained in figure 1. It begins with a stepped-sine wave stimulus. This is the same test signal commonly used for other end-of-line tests. This means that perceptual distortion analysis can be measured in parallel with other metrics such as frequency response, conventional distortion, etc. with no increase in test time - an important consideration for high-speed production lines.

A proprietary noise reduction algorithm minimizes the effect of background noise, before two separate analyses are then carried out on the response signal, and then combined into a numeric perceptual distortion measurement in Phons. In one analysis, the frequency response of the ear is modelled, following the sound principles outlined in the ITU PEAQ standard for CODEC sound quality evaluation [3]. First, auditory filter bands convert the FFT spectrum (constant bandwidth) to a Bark scale to replicate human ear filtering. Next, an ear weighting filter compensates for the transfer function of the outer to inner ear, and finally, the internal noise of the ear (noise floor due to blood flow) is added. A frequency spreading function – a simplified mathematical representation of auditory masking curves - is then applied to mimic the psychoacoustic filters of the ear in hearing Rub & Buzz defects. These masking curves change with frequency and level [4, 5]. The fundamental and its masking effects are then removed to give the distortion of the speaker plus noise. This is summed over the frequency range to give the perceptual partial loudness for a single tone of the input signal [6, 7].

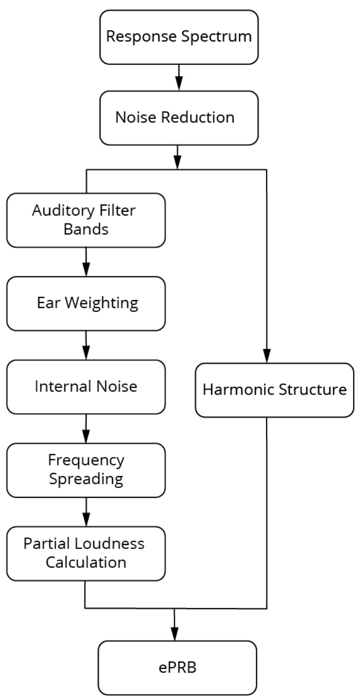

Figure 1. Schematic of the ePRB algorithm

On the other branch of the algorithm the harmonic structure of the response is quantified using the power cepstrum (a cepstrum is a spectrum of a log spectrum) [8]. A strong and extended harmonic structure is a signature of Rub & Buzz.

The most significant difference between this and the earlier algorithm is the addition of a unique noise reduction algorithm in the initial step to mitigate the effects of background noise. Additionally, changes to the masking and weighting in the perceptual filtering part of the algorithm effectively amplify the relative magnitude of the distortion in the highly sensitive 1 to 5 kHz part of the ear's hearing range, resulting in improved correlation to human hearing.

The precise methodologies used in these two enhancements form part of a pending patent application and therefore cannot be described here in detail. Nevertheless, the results of these enhancements are demonstrated fully below.

3 Comparison with Earlier Algorithms

3.1 Repeatability

To evaluate repeatability, three speakers - good, bad, and borderline sounding loudspeakers, were examined using the new algorithm (Algorithm 2), and compared to both our original 2011 algorithm (Algorithm 1) and another recently launched perceptual Rub & Buzz algorithm (Algorithm 3). Three speakers of the same type - a typical small loudspeaker - were analyzed.

Subjective examination of the speakers identified one "good" speaker, one "bad" speaker with high levels of Rub & Buzz distortion, and one "borderline" speaker which exhibited some distortion, but at a level considerably less audible than the speaker with high distortion. These three particular devices were chosen from a much larger batch as examples of good, mildly distorted and severely distorted output in order to clearly show contrast in the results. The three speakers were subjectively assessed by a panel of 7 engineers experienced with listening to distortion artifacts, and the results are shown in table 1.

Speaker #	Determination	Listener comments
97	Good	No audible distortion
76	Borderline	Some audible Rub & Buzz at low frequencies only
57	Bad	Very audible Rub & Buzz over most of the frequency range

Table 1. Subjective analysis of measured loudspeakers

The test configuration shown in figure 2 was used to evaluate the speakers. This is a basic loudspeaker test setup consisting of analysis software (SoundCheck), connected to the speaker under test and the measurement microphone via an audio interface.

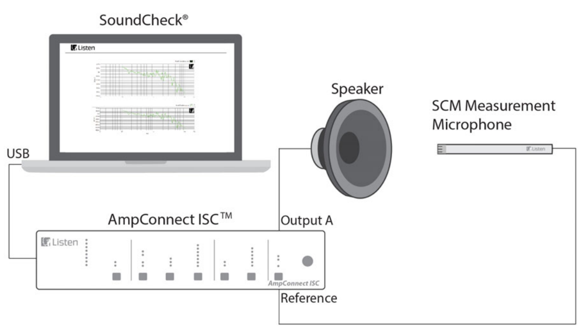

Figure 2. Test configuration for measuring speakers

Ten measurements were made for each speaker (90dBA@1kHz), in a quiet office (~50dBA) with no acoustic treatment or isolation, and the captured response waveforms saved. The same waveforms were then input into all three algorithms to ensure consistency in input data.

Figure 3 shows the results of three algorithms with the same input data.

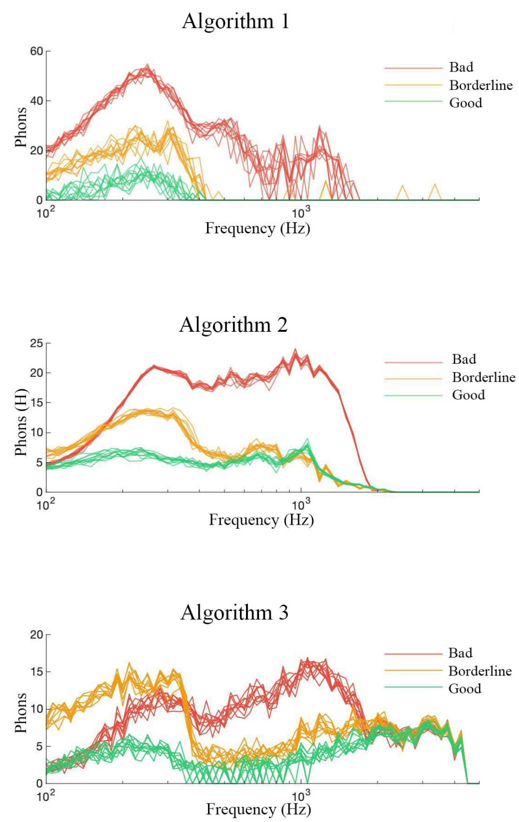

Figure 3. Comparison of three different perceptual Rub & Buzz algorithms

In the top graph, Algorithm 1 clearly shows correlation to subjective listening results. The bad speaker shows the highest level of audible distortion, the good speaker has minimal measured distortion, and the borderline speaker is somewhere in the middle. However, small variations are observed between the 10 measurements for each speaker. This is primarily due to background noise, and in a production environment where there is only one test sweep on each device, this variation makes limit-setting challenging.

The second graph shows the results for Algorithm 2, also with the exact same 10 repeated measurements for each speaker. It is immediately apparent that the results are more consistent, as is expected with less sensitivity to noise. This improved repeatability makes it considerably more reliable than the Algorithm 1 and easier to set limits, therefore making it more suitable for production line use. Additionally, the enhancements to the algorithm's masking curves reveal increased detail in the ear's highly sensitive 500Hz – 2kHz range, which more accurately mirrors human perception.

The third graph, Algorithm 3, demonstrates similar repeatability issues to Algorithm 1.

3.2 Correlation with other methods and human perception

The results in figure 3 for Algorithms 1 and 2 demonstrate a clear correlation between the measured results and subjective measurements. They clearly show a higher level for the "bad" speaker, significantly lower, but still elevated level for the "borderline" speaker and low level across the spectrum for the "good" speaker. Algorithm 3 features an interesting anomaly with the borderline speaker. Although the curves are the same overall shape as algorithm 2, it rates the distortion level of the "borderline" speaker approximately equal to the "bad' speaker, in fact even higher than the "bad" speaker towards the lower end of the frequency range.

The perceptual results can be validated by examining the harmonic distortion using a 'conventional' FFT analysis (Figure 4), as well as correlating them to listener experience. To simplify both the harmonic spectrum and the listening experience, since it is very challenging for even a trained listener to hear distortion in a fast sweep, a fixed tone of 256 Hz, near the loudspeaker's resonance and maximum cone excursion, was used to evaluate the same three speakers. This frequency was selected as it reflected the frequency at which both the bad and borderline speakers exhibit a peak in perceptual distortion measurement and therefore would be the easiest to hear.

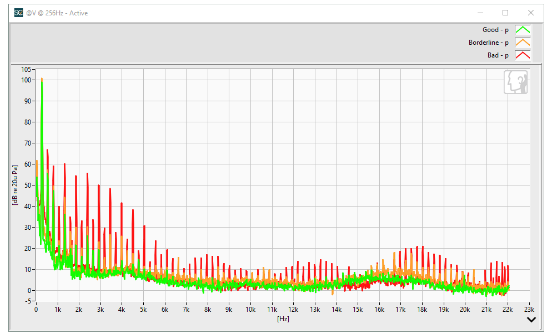

Figure 4. Conventional (HOHD analysis) of 3 test speakers

The distortion spectra clearly show minimal high order harmonic distortion for the good loudspeaker (green line), which correlates well with the perceptual measurement, as one would expect. Both the bad and borderline speakers show elevated high order harmonic distortion, again, in line with algorithms 1 and 2. The bad speaker (red) clearly has more distortion than the borderline speaker (orange) throughout the spectrum. At no point does the borderline speaker exhibit more distortion. As one might expect, the perceptual distortion algorithm amplifies the relative magnitude distortion in the highly sensitive part of the ear's range.

It is interesting that algorithm 3 demonstrated little difference in the maximum distortion audibility of the two speakers that exhibit Rub & Buzz, although human listeners perceived a significant difference in distortion level. Furthermore, at the 256 Hz frequency measured above, it actually measured greater distortion with the borderline speaker than the bad speaker, which did not correlate with either the higher order harmonic distortion measurement or listener experience.

3.3 Noise immunity

An important goal in the creation of this algorithm was to reproduce the listening performance of a human in a quieter environment, such as a quiet car, by removing the variability of factory noise, which changes constantly due to the environment, for example, air compressors, forklifts, machine noise, items dropping etc. In order to evaluate the noise-immunity of the new algorithm, the performance of Algorithms 1 and 2 were compared in the presence of factory background noise. Factory background noise was recorded inside the end-of-line loudspeaker test chamber at a customer site.

Figure 5 shows the calibrated noise spectrum (blue curve), which is pretty typical of production line noise.[9] Alongside this, the fundamental frequency response (green curve) of a buzzing loudspeaker in the presence of the factory noise is plotted. The test chamber attenuates the high frequencies well but is less effective at frequencies below 300 Hz where Rub & Buzz is typically more dominant. In fact, the background noise is higher than the loudspeaker's output below 60Hz, resulting in a negative signal-to-noise ratio at 50Hz.

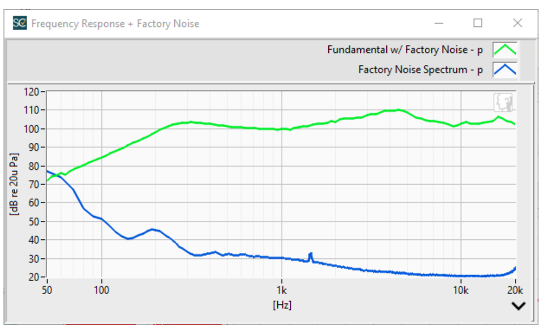

Figure 5. Frequency response of a buzzing loudspeaker in the presence of factory background noise

To accurately compare the noise immunity of the algorithms, the waveform of a buzzing loudspeaker was first recorded in a quiet room, using a stepped-sine sweep (figure 6). This was then mixed with the first 4 seconds of the factory background noise (figure 7), to create a composite recording (figure 8), as if the loudspeaker had been measured in the customer's actual test box. This allows a direct comparison of the two algorithms with and without background noise. The schematic diagram for this is shown in figure 9.

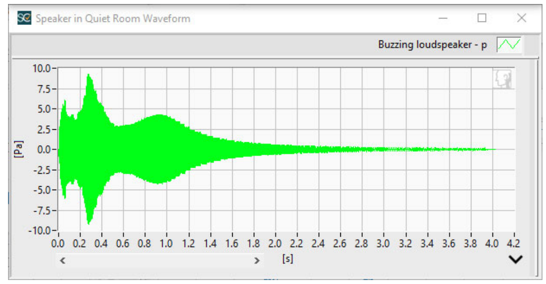

Figure 6. Recorded Time Waveform of buzzing loudspeaker in a quiet room

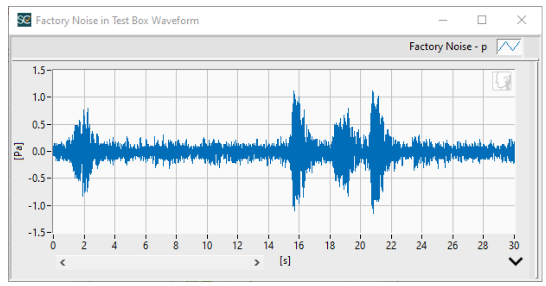

Figure 7. Recorded Time Waveform inside loudspeaker test box in noisy factory

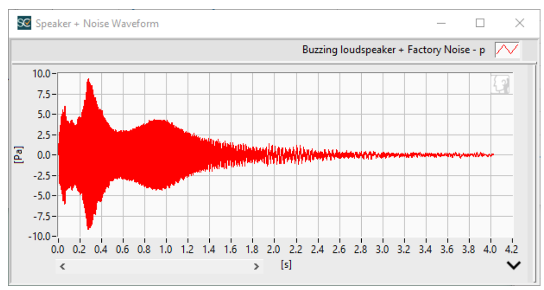

Figure 8. Resulting Time Waveform of factory background noise mixed with buzzing loudspeaker recording

Figure 10 shows the results of Algorithm 1. The green curve indicates the perceptual Rub & Buzz without background noise, and the lower-level red curve is the measurement with the background noise. This clearly indicates the algorithm's lower sensitivity to distortion due to background noise, as would be expected in an algorithm designed to mimic human perception.

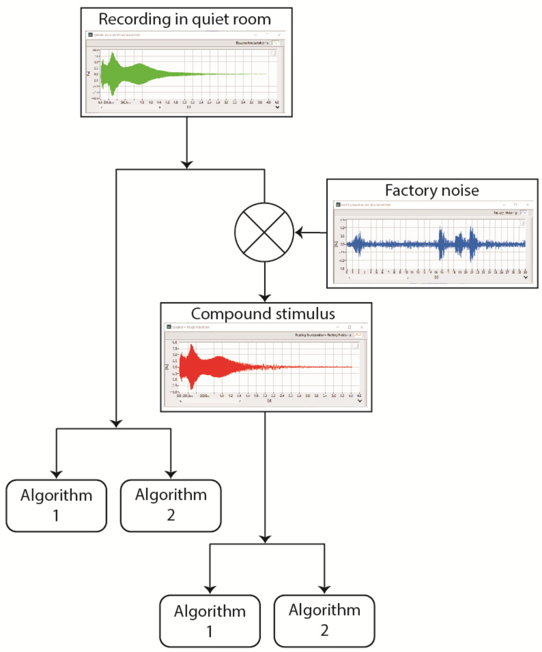

Figure 9. Schematic of methodology used to compare algorithms with and without background noise

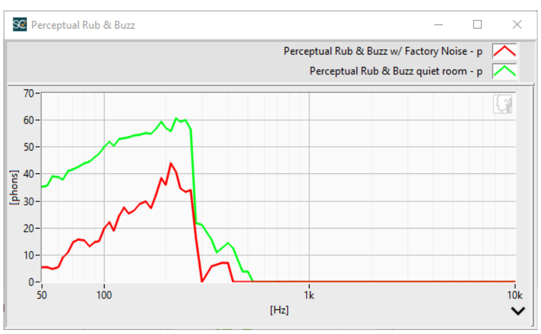

Figure 10. Comparison of results with and without background noise for Algorithm 1

Figure 11 shows the performance of the new algorithm, Algorithm 2 with the same input signals. It clearly shows that there is very little difference in the measured perceptual Rub & Buzz with and without factory background noise, demonstrating the algorithm's superior performance under noisy conditions. The only slight discrepancy is at the very lowest frequencies where the factory background noise is highest, about 80dB SPL, and the loudspeaker under test has very low output, about 70dB as can be seen in figure 5. At these frequencies, we have a negative signal-to-noise ratio, and although Algorithm 2 effectively attenuates the background noise, no amount of filtering and signal processing can easily overcome a negative signal-to-noise ratio!

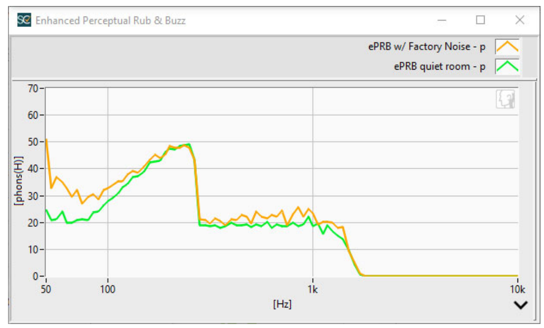

Figure 11. Comparison of results with and without background noise for algorithm 2

4 Discussion

The new algorithm, Algorithm 2 offers significantly improved performance over existing algorithms (Algorithms 1 and 3) for perceptual end-of-line detection of Rub & Buzz defects. The results shown above clearly demonstrate greater repeatability which makes it more reliable and easier to set limits. It offers excellent correlation to both subjective listening tests and other automated Rub & Buzz detection methods. Most significantly, its immunity to background noise means that its performance in the presence of typical factory background noise correlates well to the human hearing in a quiet environment.

The development of this viable perceptual distortion measurement metric for end-of-line testing raises the question of whether such measurements should replace conventional higher order harmonic distortion (HOHD) measurements. While they certainly could, with no decrease in perceived product quality, there is value in measuring both simultaneously. Perceptual metrics offer an excellent pass/fail criterion for whether the product is acceptable to ship, alongside other metrics such as frequency response, THD etc. However, conventional HOHD and Loose Particle metrics offer greater insight into the health of the production line, particularly if trends are monitored over time. An awareness of any parameters that may be drifting, even if they would not be heard by the consumer, permits early correction of manufacturing problems before they lead to audible defects. Since both enhanced Perceptual Rub & Buzz and Normalized Rub & Buzz [9] (higher order harmonic distortion) use the same stepped sine sweep stimulus, as do most other rapid end of line tests such as loose particles, frequency response, THD and polarity, perceptual measurements can be measured simultaneously with other end-of-line tests with no increase in test time.

5 Conclusions

Our new enhanced Perceptual Rub & Buzz algorithm offers significantly better performance for production line perceptual distortion detection than other algorithms included in end-of-line test software, including our own earlier algorithm. The repeatability and noise immunity that is achieved make this the first truly viable perceptual metric for end-of-line distortion measurement, and it is expected to gain popularity for end-of-line pass/fail tests, both for production line speaker test and end-of-line in-car measurements.

6 Further Research

Although laboratory tests demonstrate excellent results, validation with large-scale production line data would be valuable. This algorithm is included in the latest version of our software and can be added to an end-of-line suite of tests with no increase in test time or interference with existing test data. Assistance is available with large-scale evaluation to further validate the algorithm.

In addition to testing automotive loudspeakers, this metric could be used for testing “Squeak and Rattle” when the loudspeakers are mounted in the car [10]. The principal is the same as loudspeaker Rub & Buzz, but the Squeak and Rattle comes from the loudspeakers vibrating and exciting resonances and loose components in the car such as door panels, dashboards, wiring, clips and other components attached to or close to the mounted loudspeaker.

References

[1] Pascal Brunet, Evan Chakroff and Steve Temme, “Loose Particle Detection”, presented at the 115th AES Convention, New York, 2003.

[2] Steve Temme, Pascal Brunet and D.B. (Don) Keele, “Practical Measurement of Loudspeaker Distortion Using a Simplified Auditory Perceptual Model”, presented at the 127th AES Convention, New York, 2009.

[3] T. Thiede, et al., “PEAQ- The ITU Standard for Objective Measurement of Perceived Audio Quality”, J. Audio. Eng. Soc., vol.48, no 1/2, pp. 3-29, (2000 Jan./Feb.).

[4] Brian C. J. Moore and Joseph I. Alcántara, “Masking patterns for sinusoidal and narrow-band noise maskers” J. Acoust. Soc. Am. 104 (2), Pt. 1, August 1998.

[5] Louis D Fielder, Dolby, “Perceptual Assessment of Headphone Distortion”, Presented at the 143rd AES Convention, New York, 2017

[6] Steve Temme, Pascal Brunet and Brian Fallon, “Practical Implementation of Perceptual Rub & Buzz Distortion and Experimental Results”, presented at the 129th AES Convention, San Francisco, 2010.

[7] S. Temme, P. Brunet, and Parastoo Qarabaqi, “Measurement of Harmonic Distortion Audibility Using a Simplified Psychoacoustic Model,” presented at the 133rd AES Convention, San Francisco, 2012.

[8] R.B. Randall, Bruel & Kjaer, "Cepstrum Analysis and Gearbox Fault Diagnosis", application Note No. 1 3—1 50, 1973.

[9] Steve Temme, "How to Graph Distortion Measurements", presented at the 94th AES Convention, Berlin, 1993.

[10] S. Temme and V. Dobos, "Evaluation of Audio Test Methods and Measurements for End-of-Line Automotive Loudspeaker Quality Control", presented at the 142nd AES Convention, Berlin, 2017.

[11] Steve Temme and Patrick Dennis, "In-Vehicle Audio System Distortion Audibility versus Level and Its Impact on Perceived Sound Quality", presented at the 141st AES Convention, Los Angeles, 2016.