Audio Engineering Society

Convention Paper

Presented at the 126th Convention
2009 May 7–10 Munich, Germany

Comparison of different methods for the subjective sound quality evaluation of compression drivers

José Martínez[1], Joan Cruañes[2], Jorge Francés Monllor[3] and Jaime Ramis[4]

[1] Acústica Beyma SL, Moncada, Valencia, 46113, Spain
j.martinez@beyma.com

[2] Escola politècnica Superior de Gandia, Gandia, Valencia, 46730, Spain
joacrca@upvnet.upv.es

[3] Universidad de Alicante, Alicante, 03000, Spain
jfmonllor@ua.es

[4] Universidad de Alicante, Alicante, 03000, Spain
jramis@ua.es

ABSTRACT

In this work an approach to the problem of sound quality evaluation of radiating systems is considered, applying a perceptual model. One of the objectives is to use the parameter proposed by Moore in [1] to test if it provides satisfactory results when it is applied to the quality evaluation of indirect radiation loudspeakers. Three compression drives have been used for these proposals. Recordings with different test signals at different input voltages have been done. Using this experimental base, an approach to the problem from different points of view is done: a) Taking in consideration classic sound quality parameters such as roughness, sharpness and tonality. c) Applying the parameter suggested by Moore obtained from the application of a perceptual model. Moreover, a psychoacoustic experiment has been made on a population of 25 people. The results, although preliminary and strongly dependant on the reference signal used to obtain Rnonlin, show a good correlation with the Rnonlin values.

1. INTRODUCTION

Nonlinear distortion of compression drives has been object of study in many occasions from different perspectives. Due to its mechanical complexity, the multiple sources of nonlinearities that characterize this type of systems have been studied in order to establish each source contribution to the global nonlinear distortion. In addition, these devices work attached to a waveguide or horn that is itself a source of nonlinear distortion. Some authors consider that the distortion introduced by the horn is high enough that considering improvements in the mechanical aspects of the driver in order to reduce its distortion is not necessary, since these improvements will not be appreciated by the end user due to the horn's distortion, that would mask these improvements [2]. Other, however consider that nonlinear distortion and its sources are an essential factor to take in account in the design process of compression drivers, since it is not only produced by the waveguide or the horn.

Historically can be said that there have been three main approaches to assessment of nonlinearity: identification models, measurement models and perceptual models [3]. The identification methods try to obtain as much data as it is possible about the system with the purpose of being able to predict the system behavior with an arbitrary signal. This first approaching way to the problem is outside the focus of the present work.

The second approach could be denominated measurement methods. The objective is to obtain symptoms of nonlinearity through measurement protocols such as Total Harmonic Distortion, intermodulation methods, weighting high order harmonics or the use of the coherence function.

Thirdly would be the perceptual methods. These methods are based on the simulation of psychoacoustic effects responsible of sound quality perception.

Moreover, there are other parameters that historically have been used to evaluate acoustic quality and have been of great relevance in the psychoacoustics field. Even though these parameters are not related to the distortion evaluation or measurement, they have been used to describe sound characteristics that, obviously affects sound perception. With an appropriate test signal, some of these parameters could being indicating particularities of each drive that could have any correlation with the perceived quality of these sources.

In the compression drives context there is a problem that has not yet been solved satisfactorily: does the distortion caused by the waveguide masks the intrinsic drive distortion?

This work tries to be an approximation to the sound quality evaluation of compression drives. This approach is realized from different perspectives: on one hand, an evaluation of the quality of these systems from the point of view of the psychoacoustic classic parameters described by Zwicker [4] will be done. On the other hand, a perceptual model described by Moore and based on the human auditory system will be applied. Both models will be applied in signals registered on different commercial models of compression drives. The work also includes the accomplishment of a psychoacoustic experiment in which a hearing will realize a valuation of perceived quality. Finally, one will try to establish a correlation with the information obtained in the three mentioned approaches.

2. CONCEPTS

The following is a brief description of the different parameters and measurements that we have applied in order to approach to the problem exposed in the above section.

2.1. Psychoacoustic parameters

The following parameters were used as they are described in [4].

2.1.1. Sharpness

Sharpness is a measure that quantifies the high frequency content of a sound. In this sense, a 'sharper' sound means that it has a great proportion of high frequencies. The measurement unit is the acum. Zwicker and Fastl define a sound of sharpness 1 acum as a narrow band noise one critical band wide at a centre frequency of 1kHz having a level of 60dB. Using Zwicker and Fastl's approach sharpness can be calculated as the weighted first moment of the specific loudness (N'). The calculation of a partial first moment at z is N'.z.dz. This partial first moment is then weighted by the function g'(z) to give g'(z).N'.z.dz. The sum of these weighted partial moments is calculated and divided by the total loudness:

$$S = c\frac{\int_0^{24bark} N' \cdot g'(z) \cdot z \cdot dz}{\int_0^{24bark} N' \cdot dz} \tag{1}$$

Where c is a proportionality constant (c=0.11)

2.1.2. Roughness

Roughness quantifies the subjective perception of rapid amplitude modulation of a sound. The unit of measure is the asper. One asper is defined as the roughness produced by a 1000Hz tone of 60dB which is 100% amplitude modulated at 70Hz. For a tone with a frequency of 1000Hz or above, the maximal roughness of a tone is found to be at a modulating frequency of

70Hz. Historically roughness has been used in the calculations of annoyance metrics and also to quantify sound quality in a large number of noise evaluation applications.

$$R = cal \cdot \int_{0}^{24bark} f_{\mathrm{mod}} \cdot \Delta L \cdot dz \qquad (2)$$

Where *cal* is a calibration factor, f_{mod} is the frequency of modulation and ΔL is the perceived masking depth. The main difficulty in roughness calculation is to obtain an accurate quantification of ΔL Because of this there are several proposed methods of calculation.

2.2. Perceptual models

As mentioned above, these methods are based on the simulation of psychoacoustic effects responsible of sound quality perception. Perceptual methods were developed to assist assessment of sound quality in speech and music compression systems since the classical subjective test are usually more expensive and time-consuming. These methods have been developed in two ways: explicit simulation of masking processes and simulation of physiological and psychoacoustical effects in hearing system.

2.2.1. Rnonlin

The Rnonlin metric was developed as an extension of the DS metric developed by Moore et. al. in [5]. The Rnonlin metric analyzes the difference between the input test signal and its distorted output. A coherence analysis is performed by taking the cross-correlation between the input and distorted output waveforms. Moreover, the metric algorithm uses a model of the frequency analysis performed in the peripheral auditory system including the filtering produced by the outer and middle ear. Figure 1 shows a block diagram of the steps involved in the human auditory periphery model applied in order to obtain the velocity signals for each frequency band.
Figure 2 shows the entire block diagram that summarizes the calculation process of the Rnonlin value. Both, the input (reference) and the distorted waveforms are time aligned to remove delays caused by the nonlinear system. These waveforms are filtered to simulate the response of the outer and middle ear using 4097 FIR filters, as described by Glasberg and Moore [6].

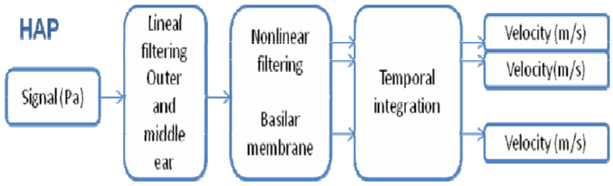

Figure 1 Block diagram of the human auditory periphery model.

Both waveforms are then filtered by an array of 40 gammatone filters with a bandwidth of 1-ERBN. This filtering provides a modeling of the auditory filtering mechanism [7].

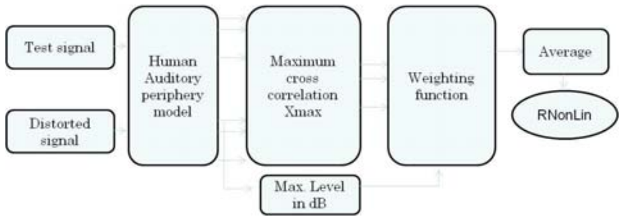

Figure 2 Block diagram for the Rnonlin calculation process.

After the human auditory periphery model, the input and output signals are split into 30ms. The next step consists in calculate the maximum value of the normalized cross-correlation between the input and output signals, Xmax. For each 30 ms frame, the Xmax values are summed across all filters. Finally, the Xmax values are averaged over all the frames resulting in the single valued, Rnonlin .

3. PROCEDURE

3.1. Test signals

The following five test signals have been used in order to make the different approaches: one single 1 kHz tone used to obtain the total harmonic distortion of each system; a 1-4 kHz multitone signal to measure intermodulation; one 1kHz signal AM modulated with with fm=70 Hz and a modulation depth of 100%. This signal is used for roughness, tonality and sharpness calculations. Finally, two music passages have also been employed for Rnonlin calculations. Both music passages were presented to the listening test subjects.

3.2. Recording setup

Three commercial drives of different manufacturers and similar characteristics and quality have been compared. In those recordings done without a horn attached to the compression drive, a plane wave tube as described in [8] was attached to it. Signal register was done at different voltage levels (from 5 to 30V) at the drive input. All signals mentioned in section 3.1 were registered. In these cases the output was recorded using a B&K high pressure microphone as shown in figure 3. 01dB Symphonie hardware and software was employed in the recording tasks.

Figure 3 Recording setup of a drive without horn.

When attaching the horn to the drives, recordings were done using a B&K measuring microphone in an anechoic chamber. The same horn was attached to all drives under test.

3.3. Listening experiment

Simultaneously with these recordings analysis, a listening test was realized to an audience of 25 people using the music passages mentioned above. One passage was a classic guitar recording of 15 seconds of duration. The other one was a 45 seconds duration percussion and saxophone music passage with a wide dynamic range. The first one became more useful to realize the test.
To the characteristic difficulties of a listening experiment [9] we must add in this case those ones introduced by our study objects, i.e. the compression drives: a standard audience is not used to hear music as it is emitted by a compression drive. In fact, these devices are used to reproduce frequency ranges above 800 Hz. Music signals were also high pass filtered with fc=800 Hz. When a listener hears a music passage emitted by this kind of sources usually qualifies it like 'bright' or 'shrill'. This unfamiliarity with the reproduced sound makes more difficult to compare or evaluate sound quality or distortion perception.

4. RESULTS

In this section the main results obtained will be presented in graphic form. We will name the three transducers as A, B and C with horn (WH) or without horn (WOH).

4.1. Psychoacoustic parameters

The following results were obtained applying the 01dB psychoacoustic parameters module to the 1 kHz modulated recorded signal.
Table one and two show the obtained values of roughness and tonality for each transducer at different input voltages.

	Vin	10	20	30
A WOH	T	0,9	0,87	0,86
	R	0,25	0,26	0,23
B WOH	T	0,89	0,89	0,87
	R	0,23	0,2	0,19
C WOH	T	0,89	0,89	0,85
	R	0,23	0,2	0,2

Table 1 Roughness and Tonality with no attached horn at different input voltages.

	Vin	10	20	30
A WH	T	1,23	1,07	1,02
	R	0,22	0,26	0,29
B WH	T	0,78	1,56	1,35
	R	0,28	0,28	0,29
C WH	T	1,12	1,12	0,88
	R	0,27	0,28	0,28

Table 2 Roughness and Tonality with attached horn at different input voltages.

Figures 4 and 5 plots the obtained sharpness and roughness.

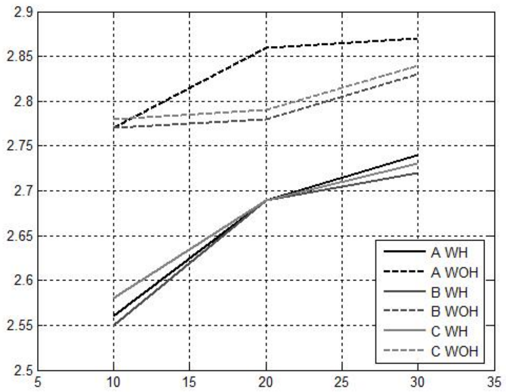

Figure 4 Sharpness vs input voltage obtained for the three devices with and without attached horn.

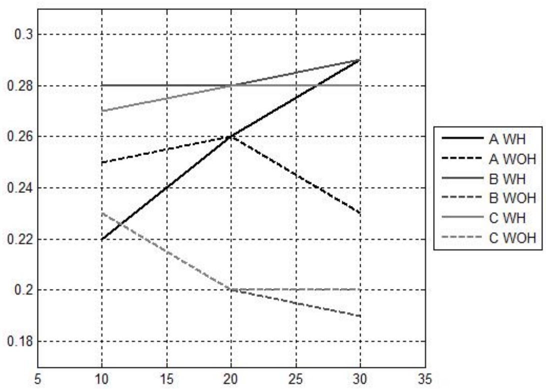

Figure 5 Roughness vs input voltage obtained for the three devices with and without horn.

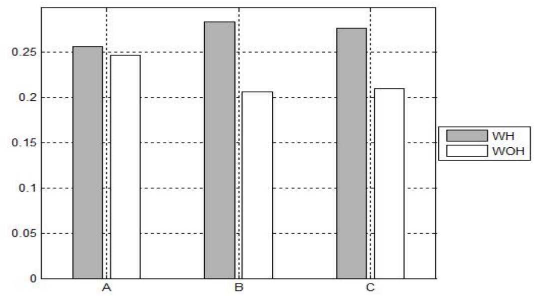

Figure 6 Averaged roughness value for each drive.

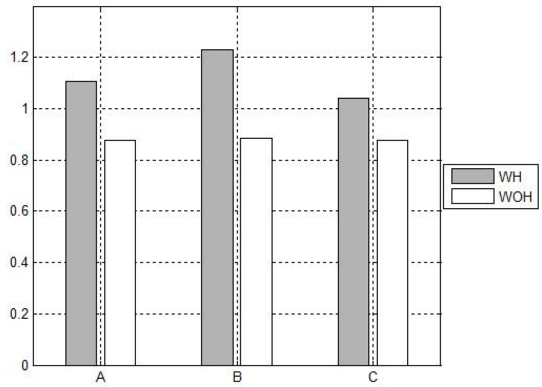

Figure 7 Averaged tonality value for each drive.

4.2. Rnonlin

Taking as the reference signal the one registered in device B with an input voltage of 5 V (as this is the recording with less THD) the following values of Rnonlin are obtained:

.

WOH	A	B	C
10V	0,3913	0,5835	0,3989
20V	0,3942	0,5811	0,4017
30V	0,4111	0,4280	0,3743

Table 3 Rnonlin values with no attached horn at different input voltages.

WH	A	B	C
10V	0,4009	0,3965	0,3904
20V	0,4446	0,4447	0,4318
30V	0,4081	0,4815	0,4563

Table 4 Rnonlin values with attached horn at different input voltages.

Figure 7 shows the averaged Rnonlin values for each device.

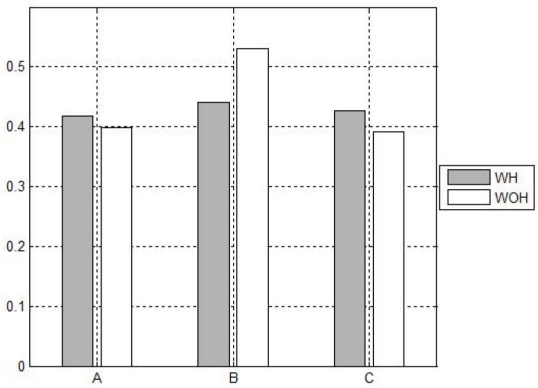

Figure 7 Averaged Rnonlin value for each drive.

4.3. Listening experiment.

Only the results confirmed by a 70% or more of the hearing (18/25) are remarked here:
Drive without attached horn:

- The drive A is perceived like better when the input level increases, but people refers to this fact as more clarity or definition.

- The drive B is perceived worse with an input signal of 25 V or above.

- The perception of drive C hardly changes in the dynamic range.

Drive with attached horn:

- When installing a horn the drive A, it is better qualified when the input voltage is around 20 V.

- Drives B and C improve the rating when increasing the input level with an attached horn.

In general terms, with the attached horn recordings, a 80% of the subjects consider that perceived distortion is smaller for the device B, A and C (in this order).
When installing the horn, perceived distortion is still smaller for drive B but there is not a clear difference between drives A and C .

5. CONCLUSIONS

These preliminary results can cause a deceptive optimism since there have only have been presented and commented a few results that seem to indicate that there is correlation between the distortion perceived by the hearing and the parameter based on a perceptual model. Nevertheless, there are several aspects of this work that must be improved:
The first one is a constant in the scope of the psychoacoustic experiments: the election of the reference signal with which we must compare. In this case, for the calculation of RNonLin the signal that has been used is that which presented less harmonic distortion. Further work should focus on check the detected tendencies when another reference or pattern signal is considered.
The psychoacoustic experiment must be repeated in order to get really trustworthy results. Indeed, the concept of ‘Perceived Distortion’ is not habitual for most of the hearing. Once analyzed the found difficulties, probably it is necessary to have recourse to a more well-known and suggestive vocabulary for most of the people.
However, the results are satisfactory. The great amount of data the study has contributed, together with other that has the company, lead the way to the proposal of several parameters dependant functions to evaluate the improvement degree that one will obtain when making small changes in a drive model or when installing a certain horn.

6. ACKNOWLEDGEMENTS

The authors would like to thank Mrs. Cathaysa González López for her essential help in data processing.

7. REFERENCES

[1] C. Tan, B.C.J. Moore, “Predicting the perceived quality of nonlinearly distorted music and speech signals”. Journal of the Acoustical Society of America 52 (2004) 699–711

[2] Earl R. Geddes et Al, “Subjective testing of compression driver”, J. Audio Eng. Soc., Vol. 53, No. 12, 2005 December.

[3] A. Voishvillo, “Assessment of nonlinearity in transducers and sound systems. From THD to the perceptual models”. 121st AES Convention.

[4] E. Zwicker and H. Fastl, “Psychoacoustics: Facts and Models”. Springer-Verlag, Berlin, Heidelberg, 1990.

[5] C. Tan, B. C.J. Moore, and N. Zacharov, “The effect of nonlinear distortion on the perceived quality of music and speech signals”. Journal of the Audio Engineering Society, 51(11):1012–1030, November 2003.

[6] B.R. Glasberg and Brian C.J. Moore, “A model of loudness applicable to time-varying sounds”. Journal of the Audio Engineering Society, 50:331–342, May 2002.

[7] R.D. Patterson, M.H. Allerhand, and C. Giguere, “Time-domain modeling of the peripheral auditory processing: A modular architecture and a software platform”. Journal of the Acoustical Society of America, 98:1890–1894, 1995.

[8] R. Magalotti, C. Zuccatti, and P. Pasini, “Building a Plane-Wave Tube: Experimental and Theoretical Aspects,”J. Audio Eng. Soc. (Engineering Reports), vol. 47, pp. 596–601 (1999 July/Aug.).”

[9] Jens Blauert. “Spatial Hearing. The psychophysics of human sound localization”. MIT Press. 1997.