How many psycho-acoustic attributes are needed?

T.H. Pedersen and N. Zacharov

Delta Acoustics & SenseLab, Venlighedsvej 4, 2970 Hørsholm, Denmark
thp@delta.dk

Sounds may be characterized by objective perceptual attributes (for which there may exist physical metrics) or by subjective (affective or connotative) attributes. This paper will deal with the perceptual attributes.

Within product sound quality the metrics for classical the psycho-acoustic attributes (loudness, sharpness, roughness and fluctuation strength and maybe supplemented with tone and impulse prominence) are often used as the only attributes to characterize the sounds. But are these 4-6 attributes or dimensions sufficient to characterize a sound? Within room acoustics and reproduced sound many other attributes are used and in the language around 100 direct sound describing words may be found.

This paper will give an overview over attributes used within different acoustic areas. The latter part of the paper will discuss the role of sensory evaluation methods as a means to systematically developing attributes for the objective qualification and quantification of sound characteristics.

1 Introduction

For many legal purposes the A-weighted sound pressure level is the only metric that is used to characterize the sound/noise. Within the field of audio, the concepts of Basic Audio Quality (BAQ) or Mean Opinion Score (MOS) are considered the principle measure for audio quality. In other cases a more detailed information is needed and in such cases metrics for the classical psycho-acoustic attributes (loudness, sharpness, roughness and fluctuation strength [8] and maybe supplemented with tone prominence [13] and impulse prominence [25]) are often used as the only characteristics of the sound. The reason is obvious, because these metrics are easy to measure with commercial available instruments and software. In the fields of product sound, room acoustics and transmitted and reproduced sound many other predefined metrics and attributes are used both on a common basis and for special applications. But how are the most appropriate attributes for a given application fond and how are the relations to already defined metrics – if any? Techniques from sensory evaluation may be used to solve this problem.

2 Definitions

The definitions below are intended to give to a common understanding of the terms used in this paper. The definitions are found in or based on the references [1, 3, 4, 8, 13, 14, 24 - 26]. The original terms are not in full agreement and so some refinements have been made

Character of sound: The overall concept of a set of characteristics that portrays the sound. The "sound character" can be specified by a number of attributes and/or metrics. The sound character is not to be confused with the sound quality which involves a comparison with some desired features.

Characteristic: A metric or an attribute of noticeable prominence

Feature: A property of a sound – in some situations an attractive characteristic

Profile: A set of parameter values (e.g. sensory descriptors or/and metrics) that describes the (character of) the sound.

Attribute: A property that can be perceived (perceptual, affective or connotative)

Metric: a measure of, physical or perceptual properties

Acoustic metrics: Examples: Sound pressure level, frequency-weighted sound pressure level

Psycho-acoustic metrics: Algorithms that have been developed to substitute perceptual measurements of attributes by instrumental procedures that evaluate the waveform of the sounds. Such algorithms provide only estimates of the attributes, and their range of validity is limited. (Examples: Loudness, sharpness, roughness, roughness, fluctuation strength, tone prominence impulse prominence)

Descriptor or descriptive term: A word or phrase that describes identifies or labels an attribute or a characteristic.

Quality of sound: Assessment of quality involves a (conscious or unconscious) comparison with some desired features (a personal "reference"). For that reason quality is a subjective characteristic. The better the characteristics of the sound match the desired features, the higher the quality will be rated. The sound quality is not to be confused with the sound character.

Product sound quality: Product-sound quality is a descriptor of the adequacy of the sound attached to a product, the sound of product quality.

Sound quality: The sound quality is the quality (e.g. the fidelity of music, the intelligibility and quality of speech) of reproduced sound or generated sounds (e.g. warning signals). The term Sound Quality indicates that we are concerned with the quality of the sound itself.

Measurement: Assigning numbers to objects in a relational way, - e.g., by comparison with a standardized quantity of the same dimension (a so called unit). Specific instruments and/or a panel of expert listeners are needed.

Perceptual measurement: An objective quantification of the sensory strength of individual sensory descriptors of a perceived stimulus. Perceptual tests are measurements where humans (expert assessors) are used as "measuring instruments".

Affective measurement: Subjective measurements of preference, annoyance or of connotative attributes.

Stimuli: Stimuli maybe anything that evokes a response from an assessor when presented with the stimuli. Such stimuli may stimulate one or many of the senses e.g. hearing, vision, touch, olfaction or taste.

Sensory descriptor [10]: Are used within a closed domain (e.g. a product category) and are defined to ensure their monosemy for both the panelists and persons using the results.

Monosemy [11]**:** Relation between designation and concept in which the former designates only one concept.

3 Attribute spaces

From a philosophical viewpoint all perceived sounds can be located in a space formed by a set of fundamental and mutually independent or orthogonal attributes. If it is possible to hear a difference between two sounds, then the perceived magnitude of one or more fundamental attributes are different for the two sounds.

3.1 The semantic space of sounds

Although language is not a very precise tool for characterizing sounds there are many words for describing sounds. Their meaning may not be precisely the same from person to person and there may be unambiguous relations between words and attributes. We may or may not have words for all attributes and often the words we have (descriptors or labels) are not one dimensional. Anyway it may be worthwhile to create a lexicon of sound describing words. The purpose is

- to contribute to a common and more precise language about sound characteristics,
- to be a bank of words to choose from in connection with word elicitation for listening tests, descriptive analysis of sounds for focus group discussions, etc.
- to give a list of possible words for profile diagrams
- to give inspiration for finding and defining new perceptive attributes of sound

For a start a collections of 450 words (in English and in Danish) for a lexicon have been made [26]. The sources of words were literature about psycho-acoustics, sound quality and product sound quality supplemented with findings in dictionaries and books of synonyms and thesaurus.

The words are organized in a database including definitions of the words. For easier "navigation" among the words two tools are provided. One tool is a subdivision into the following groups:

1) Direct sound descriptors (e.g. loud, bassy, shrill)
2) Words relating to perceptions from other senses than hearing (e.g. bright, dark, colorless..)
3) References to events and sound sources (e.g. howling, roaring, rattling)
4) Changes or differences in perceptions (e.g. colored, compressed, muffled)
5) Affective responses to sounds (e.g. pleasant, annoying, boring)
6) Connotative associations (e.g. sporty, luxurious, powerful)
7) Onomatopoeia (e.g. woof-woof, yap-yap)

Classes 1-4 are words related to perception, classes 5-6 are affective and connotative words and class 7 are sound imitating words. Although the distribution of the words may be discussed it is seen that the majority of the words (62%) relates to perceptual attributes.

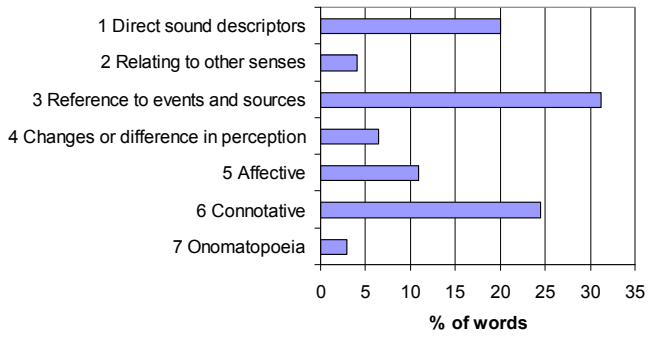

Figure 1 The relative distribution in the different word classes of the 450 words

Another means of establishing association is to find distances in a semantic space formed by assigned coordinates for 17 so called primary descriptors (Loudness, Amplitude variation, Impulse prominence, Duration, Decay, Tempo, Regularity, Roughness, Sharpness, Presence, Pitch strength, Pitch, Tone prominence, Polyphony, Harmony, Frequency variation, Localized in space).

For each of these primary descriptors a 0-10 scale with anchor points is defined. For demonstration of the principle, the words are rated by the author (it would of course be more relevant if the ratings were averages from a group of persons).

If, for example, we seek words with small distances from "thundering" we get the words (ordered according to Euclidean distance): Bumpy, blasting, rumbling, resonant, dynamic, thud, banging, gnashing, grinding, roaring. Words with large distances are: Twitter, cheep, howling, whine, jangling, chirping, pinging, squealing, wailing, shrill, beeping, and whistling.

The large number of sound describing words implies that there is a need for a detailed characterization of sounds. The words used to characterize sound may to some extent depend on the situation, but anyway it seems as the many facets of the sound cannot be uncovered by a one digit number of attributes.

This semantic space of sound forms a large database of terms that may be referred to during more elaborate sensory descriptor development, as discussed in section 4.

3.2 Metrics in different domains

A number of metrics are defined in different domains of sound, based upon different methods, practices and needs in each respective domain. The metrics are often algorithms that have been developed to represent perceptual attributes by instrumental procedures. In this section several examples domain are presented with regards their approach to the definition of attributes and sensory descriptors. Formerly, the complexity of the metrics was limited by the technical limitations of the measuring instruments, now the limitations may have other reasons. The metric may be have been defined by watchful and skilled specialists or they may have been the result of systematic listening tests or sensory evaluations.

It is characteristic that each domain has its own metrics and attributes.

General psychoacoustics

The general psycho-acoustic attributes are Loudness (metric: sone), Loudness level (metric: phon), Sharpness (metric: acum), Fluctuation strength (metric: vacil), Roughness (metric: asper) and Prominence of audible tones and impulses. These are supported by instrumental procedures giving estimates of the attributes. Also attributes for pitch and pitch strength are defined.

Product sound

This domain has adopted the traditional psycho-acoustic metrics supplemented with metrics for tonal prominence. The characteristics of the product sounds and the target sound may be displayed in a profile.

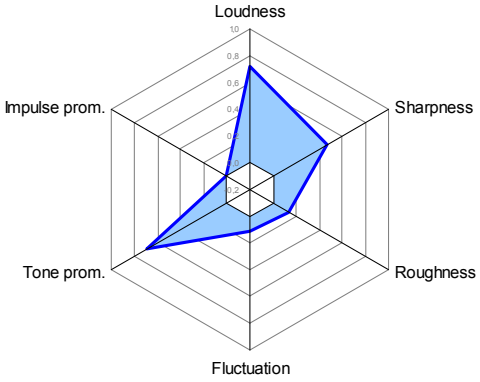

Figure 2 Example of a normalized instrumental profile for ventilation noise

Especially in this domain many connotative attributes (sporty, durable, luxurious…) are used.

Reproduced sound

A number of domains related to transmitted or reproduced sound, exist: Hearing aids, PA-systems, room acoustics, communication speech etc. Each has their own attributes. In the areas of sound reproduction, a number of studies have been performed to define the sensory descriptors for particular domains of application (e.g. multichannel, headphone, hearing aid, etc.). Many of the studies have been formed using sensory evaluation techniques and some of the sensory descriptor sets are illustrated in **Table 1**. Whilst all studies where performed with different assessor panel and stimuli, using different sensory evaluation techniques, many common terms and concepts appear between studies. This illustrates two key issues

- The potential to define a core set of sensory descriptors,
- That there is a lot more to the perceptual characterization of sound, which is often very domain specific.

4 Overview of sensory evaluation

Sensory evaluation (SE) forms a collection of techniques used to study the complex multidimensional perceptual characteristics of stimuli. Sensory evaluation considers both the effective (perceptual / objective) and affective (subjective / hedonic) responses of assessors to the stimuli. Whilst it is typical in the domain of sound to consider one sense alone, e.g. sound, these methods are equally well suited to multi-sensory stimuli [16]. The process of sensory evaluation is described by Lawless and Heymann [19] to comprise of the following steps (see Figure 3):

- **Evoke.** Comprising of stimulus generation, preparation and presentation to assessors in order to evoke their response,
- **Measure.** Associated with ***quantitative*** rating of attributes[1] by assessors,
- **Analyse.** The robust, and often statistical, analysis of assessor ratings,
- **Interpret.** The complex interpretation of data based upon knowledge of the application/limitation of methods employed. Possible extrapolation of the implications of the data to large populations.

A subset of SE techniques is referred to as descriptive analysis (see [22][19] for an overview) which are employed to qualify and quantify the character of the stimuli in an objective manner, using an expert assessor panel [15] of > 12 assessors. Such panels are used as a means of objectively and repeatably assessing the character of the stimuli, resulting in a so-called sensory profile. Descriptive (quantitative analysis) is defined as "the use of descriptive terms in evaluating the sensory *attributes* of a sample and the intensity of each *attribute*"[1] according to [14] and allows for a detailed quantitative comparison of the perceptual characteristics of products/stimuli. Several different techniques for the development of sensory descriptors exist, which will be described briefly in section 4.1.

In order for the sensory evaluation to be objective, the performance of the panel is paramount. Assessors and the panel as a whole need to be constantly monitored to ensure they are kept in *calibration* and provide objective and reliable/repeatable data. A number of techniques have been developed for assessor performance evaluation e.g. [5][27] and most recently the development of open source software for this purpose [23], all providing powerful means to calibrate assessors and panels.

4.1 Development of sensory descriptors

One of the first steps to in descriptive analysis is to define sensory descriptors. This can be performed two main manners, briefly described herein.
Consensus vocabulary development procedures include methods such as quantitative descriptive analysis (QDA) [28] for example. The main idea of these methods is to employ an expert assessor panel to develop a set of common perceptual attributes to describe the sensory properties of the stimuli under investigation. Standard methodologies have been developed for this type of consensus vocabulary development process [12] and this method has proved to be successful in the food industry and is starting to be applied in the field of sound and vision. Individual vocabulary development procedures include methods such as free choice profiling (FCP) [30], repertory grid technique (RGT) [17][29], flash profile [7] and

[1] In this context *attribute* should be replaced with current day term "Sensory Descriptor" [10].

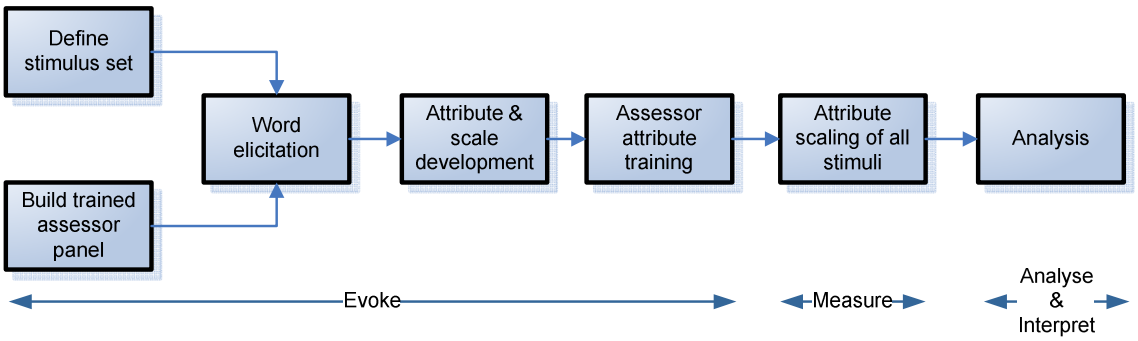

Figure 3 Overview of the descriptive analysis process

Gabrielsson & Sjögren [9]	Berg & Rumsey [2]	Koivuniemi & Zacharov [18]	Lorho [21]	Choisel & Wickelmaier [6]
Loudspeaker, headphone and hearing aid sound reproduction	*Loudspeaker spatial sound reproduction*	*Loudspeaker spatial sound reproduction*	*Headphone stereo sound enhancement*	*Multichannel sound reproduction*
Clearness / Distinctness	*Localisation*	*Sense of direction*	**Sense of distance**	Width
Sharpness/hardness – Softness	**Depth/distance**	Sense of depth	*Sense of direction*	Elevation
Brightness – Darkness	**Envelopment**	***Sense of space***	Sense of movement	***Spaciousness***
Fullness – Thinness	Width	Sense of movement	Ratio of localizability	**Envelopment**
Feeling of space	***Room perception***	Penetration	Quality of echo	**Distance**
Nearness	Externalisation	**Distance to events**	Amount of echo	*Brightness*
Disturbing sounds	Phase	Broadness	***Sense of space***	Clarity
Loudness	Source width	Naturalness	Balance of space	Naturalness
	Source depth	Richness	Broadness	
	Detection of background noise	Hardness	Separability	
	Frequency spectrum	Emphasis	*Tone Color*	
		Tone colour	Richness	
			Distortion	
			Disruption	
			Clarity	
			Balance of Sounds	

Table 1 Example sensory descriptors form a number of studies of different type of sound reproduction systems. Fonts and shading illustrate potentially common sensory descriptors across studies

individual vocabulary profiling (IVP) [20] for example. In these methods, each assessor develops his or her own set of attributes, which removes the need for construct alignment between the assessors. Such methods are powerful as a means to rapidly get started with descriptive analysis and require more advanced statistical analysis techniques to establish the common and perceptually salient characteristics of the sensory profile.

Sensory evaluation and descriptive analysis provide a number of advantages for the acoustics community to develop beyond the bounds of Basic Audio Quality, Mean Opinion Score and traditional sound quality metrics (e.g. see Figure 2), where the need is identified. Descriptive analysis allows us to

- Identify salient perceptual attributes, without prior knowledge,
- Establish a common set of sensory descriptors that allow for communication between engineers, end-users, marketers, etc., alike,
- Study uni-modal and multimodal perception / interaction,
- Measure effective perceptual characteristics in a statistically robust manner,
- Provide a basis for our community to develop metrics for salient sensory descriptors,
- Provide an opportunity to define a frame of reference, as discussed in [3], for the communication and training of the meanings of sensory descriptors,
- Progress towards a greater in-depth knowledge of the perception of complex stimuli.

Additionally, descriptive analysis provides a first step towards the greater knowledge that may be gained from the process of preference mapping.

5 Conclusion

From a review of the field of different areas of sound, speech, audio, product sound, it become apparent that there are many terms used to describe the characteristics of sound. Whilst in some fields it is common to attempt to qualify these characteristics using a small number of attributes (e.g. BAQ, MOS, loudness, pitch, etc.), it is apparent in the application of descriptive analysis that there are often a different and more elaborate set of sensory descriptors required to characterize the perceptually salient characteristics of the stimuli under test.

The application of sensory evaluation and descriptive analysis within the fields of speech, audio, product sound, etc., has been shown to provide a new depth of perceptual understanding. By developing a common set of sensory

descriptive we can evolve a common means for communication of complex perceptual concepts between consumers, engineers, marketeers, etc. and also provide a firm basis for physical metric development.

References

[1] S. Bech and N. Zacharov, "Perceptual Audio Evaluation. Theory, Method and Application", *John Wiley & Sons, Ltd*, 2006

[2] J. Berg and F. Rumsey, "In search of the spatial dimensions of reproduced sound: Verbal protocol analysis and cluster analysis of scaled verbal descriptors," in *Proceedings of the Audio Engineering Society 108th International Conference*, Audio Eng. Soc., 2000

[3] J. Blauert and U. Jekosch, "Concepts behind Sound Quality: Some Basic Considerations", *proceedings of Internoise*, 2003

[4] J. Blauert and U. Jekosch "Sound-Quality Evaluation - A Multi-Layered Problem", *ACUSTICA united with Acta Acustica,* 83, 747-753, 1996

[5] P. M. Brockhoff, "Statistical testing of individual differences in sensory profiling", *Food Quality and Preference,*14(5-6), 2003

[6] S, Choisel and F. Wickelmaier, "Evaluation of multichannel reproduced sound: Scaling auditory attributes underlying listener preference", *Journal of the Acoustical Society of America*, 121, 2007

[7] V. Dairou, J. M. Sieffermann, "A comparison of 14 jams characterized by conventional profile and a quick original method, the flash profile", *Journal of Food Science*, 67(2), 2002

[8] H. Fastl and E. Zwicker, "Psycho-acoustics. Facts and models", *Springer Verlag,* 1999

[9] A. Gabrielsson and H. Sjögren, "Perceived sound quality of sound-reproduction systems", *Journal of the Acoustical Society of America*, 65(4), 1979

[10] A. Giboreau, C. Dacremont, C. Egoroff, S. Guerrandd, I. Urdapilleta, D. Candel and D. Dubois "Defining sensory descriptors: Towards writing guidelines based on terminology", *Food Quality and Preference*,18 (2) March 2007

[11] ISO 1087, "Terminology - Vocabulary", *International Standards Organization*, 1990

[12] ISO 11035, "Sensory analysis - Identification and selection of descriptors for establishing a sensory profile by a multidimensional approach", *International Standards Organization*, 1994

[13] ISO 1996-2, "Acoustics – Description, assessment and measurement of environmental noise. Annex C: Objective method for assessing the audibility of tones in noise", *International Standards Organization,* 2007

[14] ISO 5492, "Sensory analysis – Vocabulary", *International Standards Organization*, 1992

[15] ISO 8586-2, "Sensory analysis - General guidance for the selection, training and monitoring of assessors - Part 2: Experts", *International Standards Organization*, 1994

[16] D. N. Johnson "Sensory testing and mechanical perceptions of quality - A novel application of quick individual vocabulary profiling", Master's thesis, *Helsinki University of Technology*, 2006

[17] G. A. Kelly, "The psychology of personal constructs", *Norton*, New York, USA, 1955

[18] K. Koivuniemi and N. Zacharov, "Unraveling the perception of spatial sound reproduction: Language development, verbal protocol analysis and listener training", *Proceedings of the Audio Engineering Society 111th Convention*, 2001

[19] H. T. Lawless and H. Heymann, "Sensory evaluation of food", *Chapman and Hall*, 1998

[20] G. Lorho "Individual Vocabulary Profiling of Spatial Enhancement Systems for Stereo Headphone Reproduction", *Proceedings of the 119th Convention of the Audio Engineering Society*, 2005

[21] G. Lorho, "Evaluation of Spatial Enhancement Systems for Stereo Headphone Reproduction by Preference and Attribute Rating", *presented at the AES 118th Convention*, Barcelona, Spain, 2005

[22] M. Meilgaard and G. V. Civille and B. T. Carr, "Sensory evaluation techniques", *CRC Press*, 1991

[23] A. Nilsen, O. Tomic and T. Næs, "Use of PanelCheck to compare panel performance from inter collaborative test", *The 7th Pangborn Sensory Science Symposium*, Minneapolis, USA, August 2007

[24] Nordtest ACOU 111, "Human Sound Perception – Guidelines for listening tests", 2002 `www.nordicinnovation.net/nordtestfiler/acou111.pdf`

[25] Nordtest ACOU 112, "Acoustics: Prominence of impulsive sounds and for adjustment of L_{Aeq}", 2002. `www.nordicinnovation.net/nordtestfiler/acou112.pdf`

[26] T. H. Pedersen, "The Semantic Space of Sounds, Lexicon of Sound-describing Words", *DELTA Tech. note*, March 2005

[27] P. Schlich, N. Pineau, D. Brajon, and E. M. Quannari, "Multivariate control of panel performances", *7th Sensometrics Meeting*, Davis, CA, USA, 2004

[28] H. Stone, J. Sidel, S. Oliver, A. Woolsey and C. Singleton, "Sensory evaluation by Quantitative Descriptive Analysis", *Food Technology*, 28(1), 1974

[29] D. M. H. Thomson and J. A. McEvan, "An application of the repertory grid method to investigate consumer perception of foods", *Appetite*, 10, 1988

[30] A. A. Williams and S. P. Langron, "The use of free choice profiling for the evaluation of commercial ports", *J. Sci. Food Agric.*, 35, 1984