



**Acoustics'08
Paris**
June 29-July 4, 2008

www.acoustics08-paris.org

Problems in sound quality evaluation in Brazil: general or cultural ones?

Stephan Paul^a and Samir Gerges^b

^aLab. of Vibrations and Acoustics, Fed. Univ. of Santa Catarina, CTC, EMC, Campus Trindade, 88040 Florianópolis, Brazil

^bFederal University of Santa Catarina (UFSC), Campus Universitario - Trindade, 88040-900 Florianópolis, Brazil
stephan.paul.acoustic@gmail.com

Within modern sound engineering the individual perception of sound events has to be considered. In the area of sound and vibration research procedures and instruments have been developed worldwide, but the latter ones are subject to several restrictions, especially in relation to their language. In order to carry out sound evaluations in Brazil adequate evaluation procedures and especially instruments are required. This article is intended to discuss some of the problems encountered by the authors when developing assessment procedures and especially instruments for evaluation of sound quality with Brazilian subjects. In particular, the work undertaken to study descriptors for sound and the associated problems will be outlined and discussed.

1 Introduction

Listening experiences with subjects sampled from a population are indispensable tools in sound quality assessment. The vast majority of approaches to assess sound quality by means of listening experiences use methods borrowed from psychology and sociology, usually psychometric ones. Acoustic community assessment tools including rating scales, semantic differentials, paired comparisons and others are commonly used. Nevertheless, these tools are highly dependent on linguistic, methodological and cultural issues and they can not be readily translated and used cross-culturally.

In the course of several projects related to sound quality issues in Brazil the author, as well as his colleagues in Brazil, have faced several problems. Some of them are of a more general nature but others seem to be particularly related to Brazilian language and culture.

These problems will be discussed in this article, which is intended mainly to report difficulties found, to open a discussion on these issues, to solicit feedback and to advise other researchers.

2 Adaptation of psychometric instruments from other studies

The adaptation of psychometric instruments from studies in other languages seems to be a suitable way to make use of the large number of psychometric instruments available, to take advantage of the knowledge related to them and to facilitate the comparison of results with the results from other studies. One important adaptation required to turn any psychometric instrument into one suitable for the assessment of sound and sound quality, considering the target population, concerns the linguistic and semantic elements which all such instruments rely on to communicate different aspects of the procedure to the individual. Within sound related investigations psychometric instruments are sometimes simply translated from one language to another and used cross-culturally.

On several occasions the possibility of such an adaptation procedure for all the different elements of a psychometric instrument has been analyzed. Extremely careful translation and adaptation of the concept names (e.g. comfort to *conforto* in Brazilian Portuguese) used in a psychometric instrument developed in other languages and with other subjects seemed to be reasonably adequate [21], but problems also appeared at the very beginning.

As an example we take Quehl's semantic differential for

aircraft interior noise and vibration [27]. According to Quehl [27], aircraft interior noise and vibration is an attribute of comfort within German speaking subjects. In a pilot study with Brazilian subjects we confirmed this point of view [6]. Thus, we studied the adaptation of Quehl's semantic differential to Brazilian Portuguese, but already in the initial stages problems arose. Quehl studied comfort in aircraft as understood by German speaking subjects. Her semantic differential was developed in German¹ and requires the evaluation of "*Flugzeuginnengeräusche*" (aircraft interior noise). It is generally accepted that the meaning of comfort differs, often significantly between cultural groups. Also, there is no true equivalent of the German root term "*Geräusch*" to English or to Brazilian Portuguese². This term describes any sound that is audible, is not used necessarily to transmit information and has neither a positive nor negative association [23, 7]. Thus, we were likely to investigate a slightly different concept, at least in Brazil, by asking the subject to evaluate *ruído dentro da aeronave* (noise in the aircraft) or *som dentro da aeronave* (sound in the aircraft), inducing a positive or negative rating tendency, as the connotation of *som* is mainly positive and the connotation of *ruído* is usually negative. This rating tendency was confirmed when asking for the adequacy of descriptors for aircraft interior noise using a rating scale. Either due to the meaning of the term *ruído* (\approx noise) or due to the general relation to aircraft interior noise subjects rated negative descriptors much more suitable than positive. The term *som* is also very likely to be misinterpreted, because naive subjects, but also some texts related to Brazilian noise policies, relate it to music and stereo systems [24] and less to the physical phenomena.

Further problems are likely to arise with the instructions and items when trying the translation approach. Sometimes the items, before or even after translation, do not apply to the cultural circumstances of the target population (e.g. [5]) or can not, or can only broadly, be translated. The latter occurs often with items used in English or German psychometric instruments for sound and sound quality evaluation, because these languages offer more simple descriptors for sound than Romance languages, for instance Brazilian Portuguese. In several studies on Brazilian Portuguese descriptors for sound in general and vehicle and aircraft sounds in particular it has been observed that these differences are even greater considering spoken Brazilian Portuguese (e.g. [24, 28]). In some cases, such as in the example given in the next

¹The original semantic differential was obtained by personal communication with Mrs. Julia Quehl.

²The same is true for other languages like French, Spanish or Italian.

section, the meaning of a term may change completely, even when the closest translation equivalent is given from the point of view of a translator.

It should be noted that translations of items must be carried out (e.g. [19, 4, 15]), especially to English, in the case of scientific publications. Nevertheless, several authors (e.g. [30, 21]), point out that translation, even to languages with similar cultural background, is sensitive to errors and the semantic stability and validity of the instrument is very likely to be lost. This situation is even worse in the case of successive translation, translating the psychometric instrument already translated to English again to another language for use. Therefore, an approach based on translation and adaptation seems not to be appropriate when cultural questions play an important role. For example, evaluations of tone composition in music, when connotational, are likely to include cultural aspects and therefore to differ. Also, large semantic differences can exist, as is the case between Romance languages (e.g. French, Spanish, and Portuguese) and Germanic languages (e.g. English and German).

3 Problems with descriptors

The items, e.g. adjectives and adjective pairs, used in rating scales as verbal endpoints or in semantic differentials, pose another difficulty. They are usually derived from attributes related to the concept, for instance aircraft interior noise and vibration, and are often essentially characteristics of the concept that one may use to describe it. Brazilian Portuguese poses then the difficulty that (1) in general only a small number of precise descriptors for acoustical phenomena is available³, especially in spoken Brazilian Portuguese [17, 24], (2) many terms are ambiguous, (3) some terms change their significance when being translated, and (4) many descriptors are not generally understood by naive people.

Ambiguous terms include *alto-baixo* (high-low) or *abafado-não abafado* (\approx muffled-not muffled) that apply to both volume and frequency. A term that is very likely to be misinterpreted on translation is the term 'rough', which according to Helmholtz [31] refers to a hearing sensation evoked by modulated sounds, for instance by the German spoken letter "R"⁴ which is strongly modulated. Nevertheless, Brazilian subjects do not relate modulated sounds with *áspero* at all, this being the translation of rough, but agree that totally unmodulated noises such as white noise are *áspero*. This is contrary to the concept idealized by Helmholtz and considered by others

³A similar observation about the limitations of the verbal space related to sound come from French scientists concerning French speaking subjects and urban sounds in France. Guastavino *et al.* [13, 12] and Dubois [8] report that adequate description of sounds using lexical terms in French, apart from those adopted by experts, are very hard to find. They state that it is much easier to find verbal descriptors for visual items and observed the use of complex syntactic structures in description of source events. Also, for British subjects Osborne&Clarke [20] report that the proportion of simple one word descriptions for complex phrases was found to be low.

⁴That is why Helmholtz called this hearing sensation "rauigkeit" (roughness).

such as Aures [1]. Accordingly, an attempt to measure *aspereza* subjectively would assess a rather different aspect than that qualified as "Rauigkeit" (\approx roughness) by German speaking subjects. Other commonly used terms, such as *tonal* (tonal), which is of interest to the sound quality engineer, are not used or understood by the general public. In a large elicitation session with more than 250 subjects listening to different aircraft interior noises the term *tonal* was not mentioned once. When introduced by the experimenter the significance remained unclear to naive subjects. Thus, it is difficult to assess the tonal qualities of sounds with Brazilian speaking subjects. The same occurs with many other terms that might be of interest from the point of view of the engineer or are used by experts, such as *rouco* ou *ronco* for automotive sounds, often found in Brazilian automotive magazines.

It would be interesting to investigate why Latin languages differ from Germanic ones in terms of their descriptors for sounds and noises. Maybe for cultural reasons less attention has been paid to (technical) sounds in countries where Latin languages are spoken than in cultures that use Germanic languages. But one can also see that terms are adopted from other languages and cultures. This occurs for instance in the Brazilian audio engineering community where several English terms have been adopted, although they do not necessarily share a common meaning even within the small circle of audio engineers.

When interested in descriptors for sound the experimenter might choose from a list readily at hand, taking care to choose descriptors that are likely to be understood, or identify descriptors that describe the concept under investigation by elicitation. Definition must be subsequently carried out and referred to the analysis of the descriptors found, along with the determination of the attributes to be assessed and their respective items to be used in the instrument. The methods used to arrive at descriptors, and therefore at attributes and items, depend on the type of set one might be interested in and are thoroughly described in [2, 29, 25]. From the techniques described, triadic comparison [2, 18] was found to be particularly suitable for elicitation of adjectives, whereas the other techniques are likely to provide more complex syntactic structures if no special request for adjectives is made. Regardless of the elicitation techniques, subjects must be orientated properly in order to understand the aim of the elicitation session. This was best achieved by giving an example like the description of the breakfast the subject ate that morning. The experimenter together with the subject will come up with descriptors like: sweet (marmalade), hot (coffee), and so on. It is important to use an example that is unlikely to produce descriptors related to the concept or attribute of interest. For example, when the investigator is looking for attributes and descriptors related to sound they should avoid the use of music as an example, because it is very likely that subjects will only rely on the descriptors from the example, and those elicited together with the experimenter, and not give new information.

It can be noted that a focus on adjectives in the elicitation task may limit the output of the session, when the

language does not have many adjectives for sound description or these are not naturally used. According to our experience, this is the case for Brazilian Portuguese, and according to Guastavino *et al.* [13, 12] and Dubois [8] it is also the case for French.

In the elicitation sessions it was found that the verbal discrimination abilities of acousticians not working with sound perception are not necessarily better than those of naive subjects, and even the acousticians reported difficulties involved in finding adequate descriptors for sound.

After the elicitation of descriptors, which usually results in a larger list, the experimenter must define the items to be used in the preliminary instrument. One approach to defining the descriptors for a certain type of sound is their selection from a list of previously found or defined descriptors [3]. This approach is well suited for the construction of the preliminary instrument, but the suitability rating depends on the experience of the subject with the concept under investigation and the presence or not of sound. In a study on aircraft interior noise, subjects sampled from different cities in Brazil and from the naive population as well as acoustics students, were asked to evaluate the adequacy of unipolar descriptors in this way, and considered negative descriptors, such as noisy, unpleasant, etc., to be much better suited than positive descriptors, such as silent, pleasant etc. This poses difficulties when one attempts to find descriptors that could form pairs to be integrated into a semantic differential, because the output of the evaluation task did not provide the information expected. When asking for the adequacy of previously defined pairs the experimenter must consider that it will remain unclear how the rating of a pair is obtained when one pole is considered to be adequate but the other pole is not. In both cases, evaluating single descriptors or pairs, the output is likely to be biased by several factors, e.g. the social desirability bias. It can also be noted that presentation of the respective sounds in an appropriate environment is of fundamental importance and can affect the output to an important degree.

4 Problems with rating scales

The construction of rating scales with discrete verbal labels for Brazilian Portuguese is difficult because the metric properties of verbal labels in Brazilian Portuguese are still unknown, especially regarding rating scales for sound phenomena. Günther *et al.* [11] announced an explorative study on verbal qualifiers in Brazilian Portuguese for the Brazilian Version of the ICBEN noise annoyance scale, but at present these qualifiers are still not validated. Additionally the validation procedure will consider a noise annoyance scale, and the output will be, at least strictly speaking, valid only for a noise annoyance scale, but not for other scales like a preference scale. In cases where the metric properties are unknown several authors recommend that arbitrary labels are not used, but left blank (e.g. [14, 10]).

It has been proven in explorative interviews that rating scales, particularly numbered ones, are difficult for the

inexperienced, untrained subjects to use successfully, although they provide relative ease of data processing for the experimenter. This was also confirmed in other studies, e.g. [22, 23, 16, 26], as they do not allow the subjects to express their impressions in an easy and natural way.

It is said that subjects would rather describe the object using linguistic expressions. Unfortunately the possibilities for the description of sound and vibration by simple one word descriptions in Brazilian Portuguese, as well as other Latin languages (see e.g. [9, 13, 12]), seem to be relatively limited, as experienced in several studies in which the author participated, as described earlier in this article.

5 Conclusions

This article discusses some problems encountered when working with sound quality evaluation in Brazil. Some of the problems related in this article are of a more general nature whilst others are particularly related to Brazilian Portuguese, such as problems of limited semantic space regarding the description of sound phenomena. It must be concluded that acoustics-related research in Brazil still has a long way to go in order to provide reliable tools for the assessment of sound, sound quality and sound related annoyance. The main aim of this article is to describe the problems encountered and give some orientation for future research.

Acknowledgements

I am very grateful to my colleagues Ricardo P. Leite and Raquel Bitencourt for many helpful discussion on the topic.

References

- [1] W. Aures. Ein Berechnungsverfahren der Rauhigkeit. *Acustica*, 58:268–281, 1985.
- [2] Jan Berg and Francis Rumsey. Spatial attribute identification and scaling by repertory grid technique and other methods. In *Proc. of 16th International Conference of Audio Engineering Society*, pages 51–66. AES, 1999.
- [3] David L. Bowen and Richard H. Lyon. Mapping perceptual attributes of sound to product design choices. *Noise Control Engineering Journal*, 51(4):271–279, 2003.
- [4] S. Buss, N. Chouard, and B. Schulte-Fortkamp. Semantic differential tests show intercultural differences and similarities in perception of car-sounds. In *Fortschritte der Akustik - DAGA, 2000*. Dt. Gesell. f. Akustik DEGA, 2000.
- [5] D. de A. da Cruz, C.M. Fontes, C.G. Braga, M.P. Volpato, and S.L. de Azevedo. Adaptation to the portuguese language and validation of the lunney scoring method for rating accuracy of nursing diag-

- noses. *Rev Esc Enferm USP*, 41(1):127–134, March 2007. [Article in Portuguese].
- [6] Raquel Fava de Bitencourt, Stephan Paul, Alexandre Luiz de Andrade, and Samir N.Y. Gerges. Relevância dos aspectos vibro-acústicos no conforto no interior de aeronaves. In *Anais do XXI Encontro da Sociedade Brasileira de Acústica - SOBRAC*, 2006.
- [7] DIN. 1320 Akustik: Begriffe. German technical standard. Deutsches Institut für Normung, 1996.
- [8] D. Dubois. Categories as Acts of Meaning: The Case of Categories in Olfaction And Audition. *Cognitive Science Quarterly*, 1:35–68, 2000.
- [9] Danièle Dubois, Catherine Guastavino, and Manon Raimbault. A cognitive approach to urban soundscapes: Using verbal data to access everyday life auditory categories. *Acustica Acta Acustica*, 92(6):865–874, 2006.
- [10] J. M. Fields, R.G. De Jong, T. Gjestland, I.H. Flindell, R.F.S. Job, S. Kurra, P. Lercher, M. Vallet, T. Yano, and etal. Standardized general-purpose noise reaction questions for community noise surveys: research and a recommendation. *Journal of Sound and Vibration*, 242(4):641–679, 2001.
- [11] Hartmut Günther, Fabio Iglesias, and Juliana Moraes de Sousa. Note on the development of a brazilian version of a noise annoyance scale. *Journal of Sound and Vibration*, 308:343–347, 2007.
- [12] C. Guastavino, B. Katz, J.D. Polack, D. Levitin, and D. Dubois. Ecological Validity of Soundscape Reproduction. *Acustica Acta Acustica*, 91:333–341, 2005.
- [13] Catherine Guastavino. *Étude sémantique et acoustique de la perception des basses fréquences dans l'environnement sonore urbain*. PhD thesis, Université Paris 6, Paris, France, 2003.
- [14] Rainer Guski. Psychological Methods For Evaluating Sound Quality And Assessing Acoustic Information. *Acustica Acta Acustica*, 83:765–773, 1997.
- [15] Thomas Hempel and Nicolas Chouard. Evaluation of interior car sound with a new specific semantic differential design. In *Proc. of ASA-EAA-DAGA-Joint Meeting 1999*, Berlin, Germany, 1999.
- [16] J.F. Huber, J. Hüsler, M.D. Zumstein, G. Ruffin, and M. Lüscher. [visual circle scale (vcs) - a patient-friendly scale to measure pain compared to vas and likert scale.]. *Z Orthop Unfall*, 145(6):795–797, 2007.
- [17] Fabiano Reikdal Lima. Capacitação técnica. Technical Report Modulo QS Volume 4, UFSC, 2005.
- [18] William L. Martens and Charith N.W. Giragama. Relating multilingual semantic scales to a common timbre space. In *Proc. of 113th Convention of the Audio Engineering Society*, Los Angeles, USA, October 2002. AES.
- [19] Uwe Müller and Martin Schütte. Sound engineering for aircraft (sefa), first results of listening examinations. In *Proc. of the 2006 Congress and Exposition on Noise Control Engineering Internoise*, 2006.
- [20] D.J. Osborne and M.J. Clarke. The development of questionnaire surveys for the investigation of passenger comfort. *Ergonomics*, 16(6):855–869, 1973.
- [21] Charles E. Osgood. The cross-cultural generality of visual-verbal synesthetic tendencies. *Behavioral Science*, 5:146–169, 1960.
- [22] Norm Otto, Scott Amman, Chris Eaton, and Scott Lake. Guidelines for Jury Evaluations of Automotive Sounds. *Journal of Sound and Vibration*, pages 1–14, April 2001.
- [23] Stephan Paul. Subjektive Beurteilung von Fahrzeugeräuschen - Qualitative Verfahren im Fahr Simulator und im realen Fahrzeug. Master's thesis, Technische Universität Berlin, Berlin, Germany, June 2004.
- [24] Stephan Paul. A first exploration of auditory descriptors for Brazilian Portuguese. In *Proc. of the 2005 Congress and Exposition on Noise Control Engineering Internoise*. Braz. Soc. of Acoustics SOBRAC, 2005.
- [25] Stephan Paul, Fernanda Higashi de Vasconcelos, and Thiago R.L. Zmijevski. Developing a computerized interface for sound quality evaluation. In *Proc. SAE Brasil 2008 NVH Congress*, 2008.
- [26] Marcelo Portela and Stephan Paul. Auralização de voz e violão – uma abordagem comparativa de diferentes métodos de simulação. *Acústica & Vibrações*, 39, 2008.
- [27] Julia Quehl. *Comfort Studies on Aircraft Interior Sound and Vibration*. PhD thesis, University of Oldenburg, Oldenburg, Germany, 2001.
- [28] Yuri Adson Ribeiro, Ricardo Penna Leite, Raquel Fava de Bitencourt, Thiago Rodrigo Lóss Zmijevski, and Samir N. Y. Gerges. Uma abordagem do ruído de janelas elétricas através dos conceitos e técnicas da qualidade sonora. In *XXI Encontro da Sociedade Brasileira de Acústica e I Simpósio de Acústica de Salas, Edificações e Escolas*, page 9, São Paulo, Brasil, Novembro 2006. Sociedade Brasileira de Acústica – SOBRAC.
- [29] Francis Rumsey. Subjective assessment of the spatial attributes of reproduced sound. In *Proc. of AES 15th International Conference*, pages 122–135. AES, 1998.
- [30] Brigitte Schulte-Fortkamp. The evaluation process of sound and vibration - integrating interdisciplinary concepts. *Noise Control Engineering Journal*, 51(4):239–243, Jul-Aug 2003.
- [31] H. von Helmholtz. *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*. F. Vieweg und Sohn, 1870.