Environmental sound description : comparison and generalization of 4 timbre studies

A. Minard, P. Susini, N. Misdariis, G. Lemaitre
STMS-IRCAM-CNRS
1 place Igor Stravinsky,
75004 Paris, France.
antoine.minard@ircam.fr

S. McAdams CIRMMT,
Schulich School of Music,
McGill University
555 Sherbrooke St. W.,
Montreal, QC,
Canada H3A 1E3
smc@music.mcgill.ca

E. Parizet LVA, Insa Lyon
25bis Avenue Jean Capelle,
69621 Villeurbanne Cedex,
France.
etienne.parizet@insa-lyon.fr

ABSTRACT
The aim of this study is to adapt the principles of sound timbre description, originally used for musical sounds, to environmental sounds. In order to reach this goal, we inventoried 4 timbre studies of diverse environmental sounds in terms of stimuli, experiments and perceptual results. Then, we tried to systematize sound description by comparing the different perceptual spaces in order to identify the main environmental sound classes and their associated timbre spaces. Thus, we identified three main environmental sound classes within our sound dataset: impact sounds, motor sounds and instrument-like sounds. These three classes each have their own timbre space. We finally used perceptually relevant acoustic features to explain these timbre spaces, according to the main acoustic specificities that define each sound class. We found that the *brightness* feature is used to discriminate sounds in several classes, while other particular features are used within each class.

Author Keywords
Environmental sounds, timbre, acoustic features, perceptual space.

ACM Classification Keywords
H.5 Information interfaces and presentation.

INTRODUCTION
The purpose of this study[1] is to inventory and compare the timbre spaces of representative environmental sound classes studied in recent years in order to adapt the principles of timbre description, originally used for musical sounds, to environmental sounds, considered, by nature, as non-musical. According to VanDerveer, an environmental sound "will include any potentially audible acoustic event which is caused by motions in the ordinary human environment" (see [18] pp. 16-17 for more detail). The principles of timbre description state that the perceptual distance between two sounds can be modeled by an Euclidean or pseudo-Euclidean distance estimated according to features calculated on the signals. In the medium term, the aim is to make the indexing and classification processes of this kind of sound automatic, which is actually essential for sound quality measurements, as well as for further sound content-based searching and browsing methods using perception models of environmental sounds. Both applications are very important in product sound design, which requires a quite exhaustive knowledge of sound database descriptions. The present study starts from 4 previously published timbre studies on environmental sounds [7, 10, 15, 17] and the timbre representations obtained in them. The aim of those studies was to apply musical timbre principles [4, 8, 11] to a given category of environmental sounds in order to reveal the qualities or characteristics that underlie the preference judgements [10, 16] of a class or category of sound objects. The assumption of this approach rests on the model suggested by McAdams [9] which postulates that the recognition of the sound sources rises from a process of analysis, calculation and extraction of a certain number of auditory features related to the acoustic parameters of the signals. The standard methodology consists, as a first step, in a dissimilarity-rating experiment, in which listeners are asked to rate the dissimilarity of a pair of sounds. Then, the obtained scores are processed using a MultiDimensional Scaling (MDS) analysis in order to obtain the main perceptual dimensions that constitute the derived timbre space. This approach presumes that the sounds can be compared along a limited number of continuous acoustical dimensions, which is valid only when the sound corpus is homogeneous, i.e. it is associated with the same type of objects (for example the class of cars). In the opposite case, when the corpus is heterogeneous (for example a car class mixed with a plane class), the comparison will be carried out on a semantic or a causal level in terms of distinctive discrete characteristics of the sound categories [5, 16]. For musical sounds, several psychoacoustic studies using MDS analysis of dissimilarities have clearly shown that timbre is a multidimensional set of features. Grey's study [4] revealed a 3-dimensional perceptual space shared by all the sounds tested. Krumhansl [6] also found a space with three shared dimensions with another technique of analysis. In the first section of this paper, we will introduce and describe the data that constitute the starting point of this review, in terms of sound datasets[2] and the timbre spaces that resulted from the MDS analyses performed during the studies. Then, in the second section, we will present the acoustic features that

[1]Part of these results comes from the SampleOrchestrator project funded by the "Agence Nationale de la Recherche" (ANR): http://www.ircam.fr/pds.html?tx_ircam_pi1[showUid]=36&ext=1

[2]http://recherche.ircam.fr/equipes/pds/projects/timbre/index.html

have proven to be successful in correlating the dimensions of the timbre spaces. Finally, in the third section, according to the comparison of the different timbre spaces and in order to go further with a description of environmental sounds, we will present this overall sound dataset organization in terms of its main environmental sound classes and intra-class perceptually relevant acoustic features.

SOUND DATASET AND PERCEPTUAL DATA

This section presents the standard methodology used during the previous studies of the perception of environmental sounds, and the materials resulting from these studies as sound datasets and timbre spaces. This methodology derives from the field of musical timbre, where MultiDimensional Scaling (MDS) has proven to be quite successful.

The main paradigm that allows one to obtain a timbre space is a dissimilarity-rating experiment. Participants are asked to rate the dissimilarity of a pair of sounds on a continuous scale. This has to be done for every pair of sounds. However, the studies usually start from a high number of sounds. Consequently, the first step of every timbre study is to build a sound dataset with a reasonable number of elements. This is performed by a classification experiment, where the participants are asked to gather the sounds in the main sound categories. According to the results of this experiment, a reasonable number of sounds are selected so as to be representative of the different categories. In the previous studies, this method allowed the following sound datasets to be generated:

Study A / Car interiors [10]: Two datasets corresponding to two different engine modes:

- A1: 3^{rd} gear and 4000 RPM. 16 loudness-equalized, stereophonic sounds recorded in 16 different cars
- A2: 5^{th} gear and 3500 RPM. 16 loudness-equalized, stereophonic sounds recorded in 16 different cars

Study B / Interior air-conditioning units [17]: 19 monophonic sounds, among which 4 were synthesized sounds.

Study C / Car horns [7]: 22 loudness-equalized, monophonic sounds.

Study D / Car door closing [15]: 12 stereophonic sounds.

The dissimilarity-rating experiment performed during each study yielded a matrix with the dissimilarity score of each pair of sounds of the corresponding dataset. Each matrix was processed using an MDS analysis that resulted in the main perceptual dimensions. For the MDS analysis of study D, the INDSCAL [2] technique was used, whereas for the other studies the CLASCAL [19] technique was used. The main difference between the two methods is that the latter identifies specificities and latent classes of listeners, which the former does not. In the CLASCAL model, specificities are specific characteristics of the sounds that cannot be compared on a common scale, and latent classes of listeners do not apply the same weights on the dimensions (see [19] for more detail). With INDSCAL, each listener has a set of weights for the different dimensions. The resulting timbre spaces are respectively formed with:

Study A:

- Dataset A1: 3 dimensions with specificities and only 1 latent class.
- Dataset A2: 2 dimensions with specificities and only 1 latent class.

Study B: 3 dimensions with specificities and 5 latent classes.

Study C: 3 dimensions with specificities and 6 latent classes.

Study D: 3 dimensions.

ACOUSTIC FEATURES

In order to explain the timbre spaces of the different environmental sound classes, we used perceptually relevant acoustic features that were extracted or calculated on the basis of different toolboxes of audio features. This section presents the acoustic features that have proven to be useful for this purpose. All of them are either based on the toolbox of the CUIDADO project[3] or using the "Auditory Toolbox[4]".

Energy-related features

RMS value

The estimation of the RMS (Root-Mean-Square) value of the signal is frame-based and it is calculated every 60 ms with a Blackman windowing. The feature is then the mean value over time.

Loudness

Loudness is the intensive attribute of human hearing. It thus describes the subjective aspect of the intensity of a signal by considering masking effects that occur over the whole spectrum and the filtering steps of the hearing path. The loudness model used is the ISO 532-B model from Zwicker and Fastl [20].

Harmonic emergence

This feature is a Harmonic-to-Noise ratio, designed to convey the relative amounts of harmonic (or pseudo-harmonic) energy and noise energy in the signal. It is based on the *Pm2* partial extraction method [1]. Once both harmonic and noise parts of the signal are extracted, the feature simply consists of the ratio of their respective loudnesses N_h and N_n:

$$HNR = N_h/N_n$$

Spectral features

Spectral centroid

The spectral centroid is a weighted mean frequency of the spectrum of the signal. The calculation of this feature can be more or less complex. Its definition is quite similar to Zwicker and Fastl's *sharpness* feature [20]. It uses a gammatone filterbank (from Auditory Toolbox) that is based on the ERB-rate scale z (see [8] for more detail). The resulting feature is the Perceptual Spectral Centroid:

$$PSC = \sum_z f_z \cdot N_z \Big/ \sum_z N_z$$

where N_z is the specific loudness in each channel (obtained by each gammatone filter) and f_z is the corresponding center frequency.

[3] http://www.ircam.fr/anasyn/peeters/ARTICLES/Peeters_2003_cuidadoaudiofeatures.pdf

[4] http://cobweb.ecn.purdue.edu/ malcolm/interval/1998-010/

Spectral Spread
The spectral spread describes how the spectrum is spread around its mean value, i.e. the spectral centroid defined above. The associated perceptual feature uses the same perceptual modeling as the PSC feature, thus giving the Perceptual Spectral Spread PSS:

$$PSS = \sum_z (f_z - PSC)^2 \cdot N_z \Big/ \sum_z N_z$$

Complex brightness
This feature estimates the *brightness* sensation of a sound that combines a noisy and a harmonic part. It simply corresponds to the linear combination of the PSC values of both noisy and harmonic parts (respectively PSC_h and PSC_n) and the PSS value of the whole signal:

$$Complex\,brightness = \alpha PSC_h + \beta PSC_n + \gamma PSS$$

where α, β and γ are linear coefficients.

Cleanness indicator
This feature represents the short-term variations of the loudness of the signal. These variations, which usually occur between 20 and 100 Hz, are slow enough to be heard as a temporal phenomenon, but they are too fast to be heard as separate sound events (*e.g.* bounces, rattles...). The feature corresponds to the amplitude of the spectrum of the instantaneous loudness $N(t)$, which is estimated every 3.3 ms, within this frequency band.

$$Cleanness\,indicator = \sum_{20-100Hz} \left| FFT_{256}(N(t)) \right|$$

where FFT_{256} is the 256-point Fast Fourier Transform.

Roughness
Roughness is a feature that quantifies the perceived modulation or graininess of a sound. When inharmonicity is strong, amplitude modulations can generate beating in some cases. When the beating becomes fast enough so that the modulations are no longer discriminated by the human ear, they seem to give a *rough* aspect to the sound. This roughness feature [3] mainly consists of estimating a modulation index at the output of every auditory filter, which is called the *partial roughness*. The overall roughness is the sum of all the partial roughnesses. From each auditory filter output, the modulation frequency f_{mod_i} and the modulation depth m_i are estimated with a temporal envelope calculation. The partial roughness is proportional to the product of both modulation frequency and depth $f_{mod_i} m_i$. The roughness R is then calculated as the sum of the R_i:

$$R_i = K f_{mod_i} m_i \quad \text{and} \quad R = \sum_i R_i$$

where K is the proportionality coefficient.

ENVIRONMENTAL SOUND DESCRIPTION
Each of the studies dealt with in this review was conducted independently of the others, as they were meant to fulfill different industrial needs. Thus, what is still lacking for a timbre generalization is the comparison of the description of the different kinds of environmental sounds. In other words, it is necessary to investigate the relations between our different datasets and to wonder whether their respective sounds are perceived on the basis of the same features. As a matter of fact, some of the datasets are quite similar in terms of sound production mechanism and acoustical nature, which makes them perceptually different from the others. On the other hand, the similar datasets seem to have in common the same kind of discriminating acoustic features. Consequently, one can represent this overall sound dataset structure as a 2-level organization. The upper level corresponds to a categorical structure emphasizing some of the main environmental sound classes. The lower one, which intervenes within a given class, is a continuous description of the sounds that corresponds to the timbre space. Firstly, we will present the main environmental sound categories, and then, we will generalize the timbre space of each category by defining a new set of describing features.

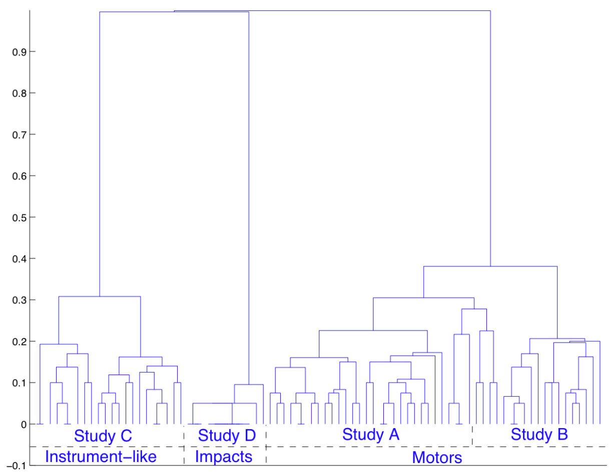

Figure 1. Hierarchical tree representation of the 83-sound dataset

Main environmental sound categories
In order to identify the main environmental sound categories that correspond to our overall 83-sound dataset, we conduted a categorization experiment with this dataset, in which 20 participants with normal hearing took part. The task was to classify the 83 sounds in as many categories as they wished according to their own criteria. The test took place in the IAC sound booths at Ircam. The sounds were played through a RME Fireface 400 sound card and Sennheiser HD 520 II headphones. We derived a hierarchical tree representation from the resulting data using an unweighted arithmetic average clustering (UPGMA) analysis procedure. In such a representation, the distance between two sounds is represented by the height of the node which links them. This hierarchical tree representation is shown on figure 1. It can easily be seen that 3 categories constitute the sound dataset. Consistently, some of the studies' datasets are quite similar, in terms of either sound features or sound production mechanism. More precisely, sound datasets of both study A and study B correspond to sounds having two discriminable parts: a harmonic part with a quite low fundamental frequency produced by a

"motor", and a noisy part produced by air turbulence. This particularity of these three sound datasets makes them different from the sounds of studies C and D, which do not contain several discriminable simultaneous parts, although they present other specific features. As a consequence, from those five sound datasets, one can consider these 3 main categories:

- "Impact" sounds from study D (closing car doors). Actually, one can easily discriminate these sounds from the others because of their temporal structure. This idea is consistent with discrimination of percussive and sustained sounds among musical sounds. Indeed, impact sounds of the environment are quite close to musical percussive sounds in terms of sound production.
- The "motor" sounds, which include the sounds from both datasets in study A (car interiors) and those from the study B dataset (air-conditioning units). These studies sounds contain a low-frequency harmonic part due to the motor itself and a noisy part due to air turbulence that occurs because of this motor.
- The "instrument-like" sounds from study C (car horns), which include only one or several higher tones, closer to those produced by music instruments than those generated by motors.

This categorization seems to be consistent with the product sound classification of Özcan et al. [14], who defined 6 sound categories: *air, alarm, cyclic, impact, liquid* and *mechanical*. Even though, unlike in our sound dataset, these product sounds were mainly from a domestic context, they found an *impact* sounds category as well, and their *alarm* category seems to correspond to our instrument-like class, while one could relate our motor sounds either to their *air* sounds or to their *mechanical* sounds.

Timbre space dimensions

In this review, we tried to go further in the description of the different timbre dimensions corresponding to each of the formerly identified sound classes. More precisely, we tried to generalize the timbre description principles of the environmental sounds, in other words to systematize the correlated acoustic features for every sound class or at least for the sounds within the same class. In order to provide an optimized set of describing features, we then looked for those that best fit the timbre dimensions of each class. Those features are presented in the following sections.

Impact sounds

This sound category corresponds to the sound dataset from study D. Its MDS analysis resulted in a 3D timbre space. According to the correlation scores in Tab. 1, three different acoustic features, presented below, were found to correlate with those dimensions.

1^{st} **dimension: Spectral centroid /** The best-fitting feature for this dimension is the Perceptual Spectral Centroid (PSC). Indeed, this dimension describes the *brightness* of the sounds.

Dimension	1	2	3
PSC	**-0.89****	0.05	0.08
Cleanness indicator	-0.18	**0.90****	0.24
RMS value	-0.28	0.27	**0.86****

Table 1. Correlation scores between acoustic features and the dimensions of timbre space D ($df = 10$, ** $p < 0.01$).

2^{nd} **dimension: Cleanness indicator /** It seems, when listening to the sounds along this scale, that this dimension is linked to the *cleanness* of the sounds. More precisely, it discriminates sounds containing only one impulse from those where one or more impulses follow the main one (rattle, bounce...).

3^{rd} **dimension: RMS value /** The RMS value is correlated with this dimension. Indeed, the dimension seems to be somehow related to pulse amplitude.

Motor sounds

One of the main characteristics of this kind of sound is that it contains two different simultaneous parts. The first one corresponds to a harmonic pattern that we can easily model by a sum of sinusoids, and the second one corresponds to the noise resulting from the air turbulence. Perceptually, those two parts are highly discriminable. Consequently, unlike in both other sound classes, both parts need to be taken into account when estimating the acoustic features, and it is thus essential to separate both parts from the original sound. This is the reason why we tested several harmonic separation methods and eventually used them in order to describe both parts, as well as their mutual interaction. This sound category regroups sound datasets A1, A2 and B. The MDS analyses of those datasets resulted in 3D timbre spaces, except for A2 dataset, which gave a 2D timbre space. Because of the attributes that define this type of sound, and which these three datasets have in common, two dimensions of these three timbre spaces were explained with similar acoustic features. Tab. 2, 3 and 4 show the correlation scores for the three datasets, respectively.

Dimension	1	2	3
HNR	**-0.93****	0.12	-0.22
Complex brightness	0.09	**0.86****	-0.17
PSS	0.07	-0.15	**0.83****

Table 2. Correlations between acoustic features and the dimensions of timbre space A1 ($df = 14$, ** $p < 0.01$).

Dimension	1	2
HNR	**0.83****	-0.17
Complex brightness	-0.34	**0.90****

Table 3. Correlations between acoustic features and the dimensions of timbre space A2 ($df = 12$, ** $p < 0.01$).

1^{st} ***shared* dimension: Harmonic emergence /** For all three datasets, several acoustic features seemed to correlate highly with this shared dimension, but none of them were significant. Furthermore, only one feature correlated well

Dimension	1	2	3
HNR	**0.91****	-0.07	0.47
Complex brightness	-0.52	**0.81****	0.00
Loudness	0.42	-0.07	**0.84****

Table 4. Correlations between acoustic features and the dimensions of timbre space B ($df = 17$, ** $p < 0.01$).

with the dimension for the three datasets: the Harmonic-to-Noise Ratio (HNR). As a matter of fact, it seems, when listening to the sounds along this dimension, that it is related to the amount of harmonic (or pseudo-harmonic) energy in the signal. Consistently, the HNR feature correlates with this dimension. The other features that correlated highly with this dimension were usually spectral envelope features. Actually, those high correlation scores are consequences of the HNR correlation. Indeed, both parts of the sounds have quite different spectral envelope behaviors, and when the proportion of both parts is modified, the overall spectral aspect of the sound is also modified.

2^{nd} ***shared* dimension: Complex brightness /** For the three datasets, when listening to the sounds along this scale, *brightness* features, such as spectral centroid or sharpness, seemed to explain the dimension. However, for the two datasets where the harmonic part is the most prevailing, i.e. studies A1 and B, this perception of *brightness* seemed to depend on the harmonic proportion. Indeed, the *brightness* perception of a predominantly noisy sound is not the same as that of a predominantly harmonic sound, all the more because both parts have quite different spectral behaviors. It is thus essential to take into account both the harmonic part and the noise part in the *brightness* estimation. For each dataset, this feature with specific linear coefficients has been found to correlate significantly with this dimension. However, we did not find a unique set of linear coefficients that resulted, for the three datasets, in a significant correlation score with this dimension. Indeed, each one depends on the relative proportion of the harmonic and the noisy part.

3^{rd} ***unshared* dimension /** This dimension is not the same between the datasets:

Study A1: The Perceptual Spectral Spread seems to correlate fairly well with this dimension.

Study A2: This timbre space is only 2-dimensional.

Study B: Loudness was found to correlate significantly with the dimension.

The loudness is a perceptually strong characteristic that can easily prevent light variations of other features from emerging. As the sounds of dataset B were not loudness-equalized, the fact that a dimension of the timbre space is correlated with the loudness is logical. Moreover, the fact that no third perceptual dimension was obtained for dataset A2 can be related to the predominance of the noisy part that can mask some variations of other features. For these reasons, we did not go further in the study of this dimension.

Instrument-like sounds

This sound category corresponds to the dataset of study C. Its MDS analysis resulted in a 3D timbre space. The correlation scores are shown in Tab. 5.

Dimension	1	2	3
Roughness	**-0.93****	-0.06	0.37
PSC	0.04	**0.97****	0.05
PSS	0.05	-0.11	**-0.90****

Table 5. Correlations between acoustic features and the dimensions of timbre space C ($df = 20$, ** $p < 0.01$).

1^{st} **dimension: Roughness /** This dimension seems to discriminate the monophonic from the polyphonic sounds. When listening to the sounds along this scale, one goes from perfectly harmonic tones towards successively pseudo-harmonic tones (tones with inharmonicity relationships between their partials) and polyphonic sounds (with several tones). Consistently, roughness correlates significantly with this dimension.

2^{nd} **dimension: Spectral centroid /** When listening to the sounds along this dimension, the relation to the *brightness* of the sounds is obvious. Consistently, the Perceptual Spectral Centroid gives the best correlation score.

3^{rd} **dimension: Spectral spread /** This dimension was the one whose interpretation was the most difficult just by listening to the sounds along this scale, even though it correlated fairly well with the Perceptual Spectral Spread, which usually describes the spectral richness of the sounds.

CONCLUSION

The aim of this paper was to compare the results obtained in different studies of the timbre of environmental sounds, and to apply in a more systematic manner the principles of musical timbre description towards such non-musical sounds. We first identified important categories of environmental sounds such as impact sounds, motor sounds and instrument-like sounds, according to the source of sound production. This seems to be coherent with the categories of product sounds that Özcan et al. found [14]. In a second approach, we compared the timbre spaces of the sounds within each category in order to find common perceptual dimensions of the timbre description. Thus, we identified:

- One feature that is preponderant for the description of all sound categories, i.e. the *brightness* feature, usually based on spectral envelope features. Therefore, this perceptual feature appears to describe musical sounds as well as environmental sounds.
- One or two features, in each class, that are related with a specificity of the corresponding sounds:
  - an important part of the perceptual discriminability of impact sounds is related to a temporal behavior feature, describing the sounds *cleanness*,
  - motor sound perception is largely characterized by the mixture of two highly discriminable parts, in terms of either energy or spectral content,

– instrument-like sounds present timbre features, originally derived for the description of musical sounds.

Note that, contrary to musical timbre for which attack time is an important cue of the timbre space, the studies revealed no temporal features corresponding to the two last classes, but that is mainly due to the quasi-stationary nature of those sounds. Nonetheless, a temporal parameter such as the cleanness indicator, attack time or impulsiveness is likely to discriminate two main categories of sounds from our environment such as impulsive (car door closing) and non-impulsive sounds. However, some issues still need refinement. We were faced with several constraints related to the datasets upon which we based the generalization. In fact, we based our sound classification on the sound production mechanism and on the observations on the chosen types of acoustic features of the studies, and this was coherent with other classifications we found in the literature. Nevertheless, even though the identification of these three classes with a categorization experiment seemed obvious, they have to be investigated with less specific sounds in order to define their boundaries precisely. That is actually the next step in our research: we wish to conduct a forced-choice classification experiment with an enlarged dataset including much more diverse sounds in each of the three classes. Furthermore, in order to systematize sound description in categorical as well as in continuous levels of organization, it would be very interesting to look for a way of defining the categories in terms of acoustic feature values, in order to automate sound description, for example. Finally, according to the Özcan et al. study [14], other major sound classes, such as liquid or cyclic sounds, exist and need a definition as well, and their main perceptual features must be investigated.

REFERENCES

1. N. Bogaards, A. R obel and X. Rodet, *Sound Analysis and Processing with AudioSculpt 2*, Proc. Int. Computer Music Conference (ICMC), 2004.
2. J. Carroll and J. Chang, *Analysis of individual differences in multidimensional scaling via an* n*-way generalization of Eckart-Young decomposition*, Psychometrica 1970;35:283-319.
3. P. Daniel and R. Weber, *Psychoacoustical roughness: Implementation of an optimized model*, Acustica united with Acta Acustica, 83, 113123,1997.
4. J. M. Grey, *Multidimensional perceptual scaling of musical timbres*, J. Acoust. Soc. Am., Vol. 61, No. 5, May 1977.
5. J. M. Hajda, R. A. Kendall, E. C. Carterette and M. L. Harshberger, *Methodological issues in timbre research*, In Irène Deliège and John Sloboda, editors, Perception and cognition of music, chapter 12, pages 253-306, Psychology Press, 1997.
6. C. L. Krumhansl, *Why is musical timbre so hard to understand?* in S. Nielzén & O. Olsson (Eds.), Structure and Perception of Electroacoustic Sound and Music, pp. 43-53, Elsevier (Excerpta Medica 846), Amsterdam.
7. G. Lemaitre, P. Susini, S. Winsberg, B. Leinturier and S. McAdams, *The sound quality of car horns: a psychoacoustical study of timbre*, Acta Acustica, 93, 457-468, 2007.
8. J. Marozeau, A. de Cheveigné, S. McAdams, and S. Winsberg, *The dependency of timbre on fundamental frequency*, J. Acoust. Soc. Am., 114(5) :2946-2957, 2003.
9. S. McAdams, *Recognition of auditory sound sources and events*, in Thinking in sound: the cognitive psychology of human audition (S. McAdams and E. Bigand, eds.), Oxford University Press, 1993.
10. S. McAdams, P. Susini N. Misdariis and S. Winsberg, *Multidimensional characterisation of perceptual and preference judgements of vehicle and environmental noises*, Proceedings of Euro-Noise 98, Munich, Germany.
11. S. McAdams, S. Winsberg, S. Donnadieu, G. De Soete and J. Krimphoff, *Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes.*, Psychological Research, 58, 177-192, 1995.
12. B. C. J. Moore and B. R. Glasberg, *A Model of Loudness Applicable to Time-Varying Sounds*, J. Audio Eng. Soc., Vol. 50, No. 5, 2002 May.
13. B. C. J. Moore and B. R. Glasberg, *Suggested formulae for calculating auditory-bandwidths and excitation patterns*, J. Acoust. Soc. Am., 74, 750-753, 1983.
14. E. Özcan and R. van Egmond, *Memory for product sounds: The effect of sound and label type*, Acta Psychologica (2007), doi:10.1016/j.actpsy.2006.11.008.
15. E. Parizet, E. Guyader and V. Nosulenko, *Analysis of car door closing sound quality*, Appl. Acoust., 2006, doi:10.1016/j.apacoust.2006.09.004.
16. P. Susini, S. McAdams and S. Winsberg, *A multidimensional technique for sound quality assessment*, Acta Acoustica, 85, 650-656, 1999.
17. P. Susini, S. McAdams, S. Winsberg, I. Perry, S. Vieillard and X. Rodet, *Characterizing the sound quality of air-conditioning noise*, Appl. Acoust., 2004, doi:10.1016/j.apacoust.2004.02.003.
18. N. J. VanDerveer, *Ecological acoustics: human perception of environmental sounds*, PhD thesis, Cornell University, 1979.
19. S. Winsberg and G. De Soete, *A latent class approach to fitting the weighted Euclidean model, CLASCAL*, Psychometrika, 58, 315-330,1993.
20. E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*, Springer Verlag, 1990.