Perception of the timbre of car horns

Pacs reference: 43.66.Jh

Guillaume Lemaitre[1,2], Patrick Susini[1], Suzanne Winsberg[1], Stephen McAdams[1]

[1]Ircam-CNRS,
Musical Perception and Cognition Team
1, Place Igor Stravinsky,
F-75004 Paris, France

[2]Laboratoire d'Acoustique de l'Université du Maine
Avenue Olivier Messiaen
F-72000 Le Mans, France

Correspondance should be adressed to: lemaitre@ircam.fr or susini@ircam.fr

ABSTRACT
Due to the technologies used, there exist only a few different kinds of car horn sounds.Prior to the design of new sounds, our study aims to define these sounds from a perceptual standpoint, using a two-step procedure: The first step was a free classification experiment. We found a structure of 9 families of sounds. The second step was a dissimilarity judgment experiment based on these families. The results are represented by a timbre space with three dimensions. Three psychoacoustic descriptors match the perceptual dimensions, allowing us to characterize the perception of the different horn technologies.

INTRODUCTION

Sound design is an important issue nowadays for the automotive industry. A lot of attention is paid to all of the sounds produced by cars. And car horns are no exception.
Designing car horn sounds raises two main issues for car builders. Firstly, it allows them to make their cars sound different from those of their competitors. Secondly, it is important to match the sounds of a car to its price category.
However, and before all these marketing aspects, the function of a car horn is to alert people to potential danger related to the car. Designing the sound of the car horns thus involves a compromise between the need to customize the sound and the necessity of providing efficient warning signals.
This paper deals with the timbre of contemporary car horn sounds. Prior to the design of new sounds, our goal was to characterize people's perception of the timbre of car horns. There exist nowadays only a few different kinds of car horn sounds, but they are well known by people and identified without any ambiguity. People do have a sort of "prototypical" representation of car horn sounds [1]. The study proceeded in two steps, following the method described in [2]: in the first step, a free classification task was used, providing us with the main families of sounds perceived by people. The second part aimed to characterize the perception of these families. A dissimilarity rating experiment, with a multidimensional scaling analysis revealed psychophysical descriptors that potentially explain the perceptual factors underlying the dissimilarity ratings.

THE FIXATION PROBLEM: A PRELIMINARY STUDY

When car horns are tested, they are usually fixed on a heavy metal bar, and recorded in an anechoic chamber (laboratory recording). However, they sound different when they are recorded attached to the body of a car (car body recording). Hence the first task of our study was to determine the validity of a study based on car body recordings. The question was then: "Is the perception of the sounds of a set of car horns different when they are fixed on a metal bar compared to when they are attached to a car body?".
To answer this question, we performed a preliminary study. Space does not allow a detailed description of this study. So only the main results

will be presented. We recorded 43 car horn sounds either fixed on a metal bar in an anechoic chamber or attached to the bodies of two different cars. Then, to characterize the perception of the three sets of sounds, we carried out a dissimilarity-rating experiment. We then compared the perception (i.e. the dissimilarity ratings) of the three sets with an analysis of variance (Anova). The results showed that when people were asked to compare directly a car horn recorded in the laboratory and on a car body, they were able to hear a small difference (far smaller than the differences between two different horns). However, the global relationships between sounds were not modified significantly by the way the horns were fixed. In other words, two horns that are judged different when they are fixed in the laboratory are still judged as different when attached to a car body. Our interpretation was that the global perception of a set of sounds of car horns is not influenced by the way they are fixed. Hence we concluded that the results obtained in a study based on sounds recorded in the laboratory could be generalized to the situation in which they are attached to a car.

THE SOUND CORPUS

Car Horn Technologies

Before going further, let us take a close look at car horns. There exist three kinds of devices. The first, that we will call the "horn-like" device, is based on a electrodynamic driver loaded by a horn. The second kind is also made of an electrodynamic driver, but there is a metal plate attached to the membrane. We will call this one the "plate-like" device. The third kind is a pneumatic driver loaded by a horn, which we will call "pneumatic-driven" device.

Recordings

The horns were recorded in the anechoic chamber at Ircam. They were all 550 ms in duration, and were equalized in loudness in a preliminary level-adjustment experiment. The corpus included sounds of horn-like, plate-like, and pneumatic-driven devices. They were recorded individually (monophonic case), and in some cases in twos or threes to obtain chords (multiphonic case). This made a total of 43 sounds.

EXPERIMENT 1: FREE CLASSIFICATION OF THE SOUNDS

The aim of this experiment is to determine families of sounds with similar timbres. Since 43 sounds are far too many to use in a dissimilarity-rating experiment, this classification would allow us to choose sounds that are representative of the perceived families.

Method

Subjects: 28 subjects (15 men and 13 women) volunteered as listeners and were paid for their participation. They were aged from 18 to 34 years. All reported normal hearing.

Stimuli: The 43 sounds (described above) were played at 83 dB SPL.

Apparatus: The test took place in the IAC sound-attenuated rooms at Ircam. The experiment was run on a Personal Computer under Linux, and the graphical interface was implemented under Matlab. The sounds were amplified through a Yamaha amplifier and sent to Sennheiser HD 520 II headphones.

Procedure: The subjects all were given written instructions explaining the free classification task. Emphasis was placed on what timbre is (neither pitch nor perceived duration nor loudness). The subjects saw a white screen on which stars labeled from 1 to 43 were drawn, each star corresponding to a sound. The labelling was different for each subject. They could hear the sound by double-clicking on a star. Subjects were asked to move the stars in order to group together the sounds they heard as having the same timbre. They were allowed to form as many groups as they wished, and to put as many sounds in each group as they desired. The data for each subject consisted of a incidence matrix i.e. a matrix in which a one indicates that the two sounds have been classed together, and a zero that they have been classed in different groups.

Results

By averaging the individual incidence matrices, we obtained a co-occurrence matrix, which can be interpreted as a Euclidian proximity matrix [3].

Correlation between the subjets: An analysis by principal components showed that the responses of the subjects were consistent.

Hierarchical tree representation analysis: We derived a hierarchical tree representation of the data using an unweighted arithmetic average clustering (UPGMA) analysis procedure. In such a representation, the distance (according to the co-occurence matrix) between two sounds is represented by the height of the node which links them [3].

According to this classification, sounds of car horns of the same technology are classed together (horn-like devices with horn-like devices , multiphonic with multiphonic, etc.: typological classification). However, it was not easy to divide the tree into clusters. As the tree was rather homogenous, there was no particular reason at this stage to cluster at any given level.

« Bootstrap » analysis: In order to evaluate the most stable level of clustering across listeners' responses, we used a bootstrap algorithm [3]. The hierarchical tree obtained from the averaged co-occurence matrix reflects only the relationships of proximity inside the co-occurence matrix. It shows the hierarchical relationships perceived between the sounds, but it doesn't give any reason to favour one level of clustering over another. The upper part of the tree (close to the root) is a *super-ordinal* level of classification, which is much too ordered to reflect the subjects' responses. The lower part (the leaves) is a *sub-ordinal* level of classification which contains very little information.
What we are looking for is rather an optimal (or principal) partition of the sounds, which corresponds to perceptual categories.
The algorithm developed in [3] is based on the principle of bootstrap. It generates a great number of subsets of subjects on the basis of random sampling with replacement as the data are sampled. New hierarchical tree representations are then computed based on the new subsets of subjects. It then compares all the representations to find what part of the trees is stable between all the new trees. As a result we found a stable tree with 9 classes as illustrated in Figure 1.

<u>Discussion</u>
The relationships between this timbre-based classification and the typological classification of the car horns should be considered. First of all, three large clusters appear. The first one (classes 1 to 4) groups together all the monophonic devices, including the pneumatic horn. The second one (classes 5 et 6) is made of the multiphonic sounds from both plate-like and horn-like devices. The last one (classes 7 to 9) groups together the monophonic plate-like devices.
The distinction between monophonic and multiphonic sounds is quite clear. At this step, it can be however concluded, that there are two main families of timbres for car horn sounds: horn-like and plate-like devices. If we go deeper into the details of the classification tree, some horns are distinguished from the others. One new sound (class 1: horn-like), which was at the time of the study only a prototype, was heard as a horn, but somehow set apart from the standard ones. The ship horn (class 3) was also separated, which confirms the validity of this classification: subjects didn't confuse it with car horns. A very low horn was also set apart and was curiously grouped together with the plates. This classification gives us an idea of the main families of sounds perceived by the subjects.
We used this classification in order to choose a subset of sounds, both representative of the variety of the car horns and not too large in number.

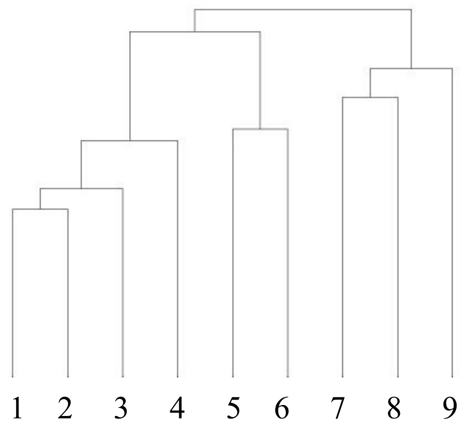

Figure 1. Shematic version of the hierarchical tree computed from the data of the Experiment 1, after bootstrapping. The labels of the leaves refer to Table 1.

Class	Label	Number
1	Special low monophonic horn-like	1
2	Standard monophonic horn-like	11
3	Ship horn-like	1
4	Special very high-pitched horn-like, both mono- and multiphonic	6
5	Multiphonic standard plate-like	5
6	Multiphonic standard horn-like	6
7	Special plate-like, both mono- and multiphonic	6
8	Standard monophonic plate-like	6
9	Special very low horn-like	1

Table 1. Horns within the nine classes.

EXPERIMENT 2: DISSIMILARITY-RATING EXPERIMENT

The strategy followed here is based on [4]. The main idea is first to represent the proximity between the sounds heard by the subjects using a spatial (Euclidian) representation. This spatial model is considered to represent the perceptual dimensions that underlie the dissimilarity ratings. The issue is thus to find the psychoacoustical correlates that match the perceptual dimensions. As this study does not aim to explore timbre, we didn't seek to find new descriptors but rather used descriptors discovered by previous studies on timbre ([5-9] for instance).

Method

Subjects: 41 subjects (20 men and 21 women) volunteered as listeners and were paid for their participation. They were aged from 18 to 34 years. All reported normal hearing.

Stimuli: 22 sounds were chosen from the 9 classes obtained from the classification task. They were played at the same level as the previous experiment.

Apparatus: The test took place in the IAC sound-attenuatation rooms at Ircam. The experiment was run on a Personal Computer under Linux, and the graphical interface was implemented under Matlab. The sounds were amplified through a Yamaha amplifier and sent to Sennheiser HD 520 II headphones.

Procedure: Subjects all received written instructions explaining the task. They were told that they were to make judgments on the timbre, and the meaning of word timbre (neither pitch, nor perceived duration nor loudness) was explained to them. Particular emphasis was placed one ignoring pitch [6]. The experiment was performed in a single session, but subjects were allowed to take a break during the session. All 241 different pairs among the 22 sounds were presented. At the beginning of the session, the subject listened to all of the samples in a random order to get a sense of the range of variation possible. Next, 5 training trials were presented to familiarize the subject with the rating task. On each trial, a pair of sounds was presented, separated by a 500-ms silence.

The subject saw a horizontal slider on the computer screen with a cursor that could be moved with the computer mouse. The scale was labelled "Very Similar" at the left end and "Very Dissimilar" at the right end. A rating was made by moving the cursor to the desired position along the scale and clicking on a button to record it in the computer.

Results of the MDS Analysis

Coherence of the responses: An analysis of the correlations between the responses of the subjects revealed that one subject was correlated negatively with the others. This subject was removed from subsequent analyses.

Multidimensional scaling analysis with Clascal: The Clascal analysis is described in more detail in [10]. We give here only a short description. In this analysis, dissimilarities are modelled as distances in an extended Euclidian space of R dimensions. In the spatial representation of the N stimuli a large dissimilarity is represented by a large distance. The Clascal model for the distance between stimuli i and j postulates common dimensions shared by all stimuli, specific attributes, or "specificities", particular to each stimulus, and latent classes of subjects. These classes have different saliences or weights for each of the common dimensions and the set of specificities. The class structure is latent: there is no a priori assumption concerning the latent class to which a given subject belongs. The Clascal analysis yields a spatial representation of the N stimuli on the R dimensions, the specificity of each stimulus, the probability that each subject belongs to each latent class and the weights or saliences of each perceptual dimension for each class.

Results and discussion:

We found a spatial model of 3 dimensions with specificities and 6 latent classes. The figure shows the projections of the sounds in the space on the D1-D2 and D2-D3 planes. The distances computed in the spatial model were correlated with the raw dissimilarity data and accounted for 87% of the variance ($r(20)=0.93$, $p<0.01$).

These results are coherent with the perceptual and typological classifications: horn-like sounds are separated from plate-like sounds along dimension 2, whereas dimension 1 separates multiphonic from monophonic sounds.

If we go deeper into details on the high number of latent classes, there is no clear evidence for 9 of the subjects belonging to a given latent class. The huge number of latent classes added to this seems to indicate that the latent classes should not be interpreted here as different strategies of response shared by the subjects but rather as an indication that the subjects each have their own way to respond.

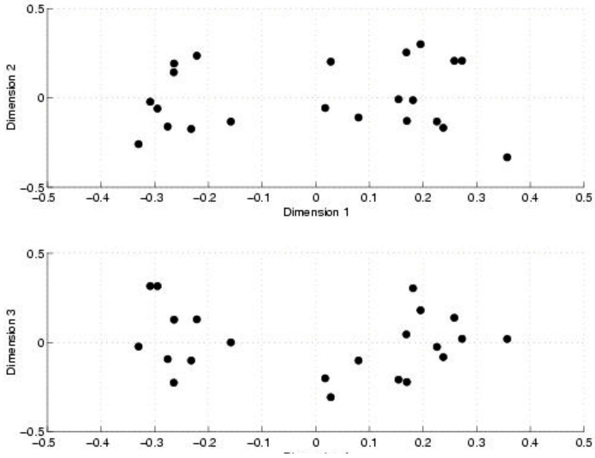

Figure 2. Timbre space of the car horns.
Specifities are not represented

One further illustration of this is given by performing an Exscal analysis. This analysis is similar to a Clascal analysis, except that there is no latent class: each subject has individual weights for the three dimensions and the set of specificities. The figure shows the weights over the three dimensions for the 40 subjects. It seems very difficult to interpret these weights in terms of response strategies. For instance, we cannot clearly separate subjects who made their ratings based on a subset of dimensions. Rather, each subject seems to have an individual rating strategy. Finally, we decided to retain the solution having 3 dimensions with specificity and 6 latent classes as the best compromise.

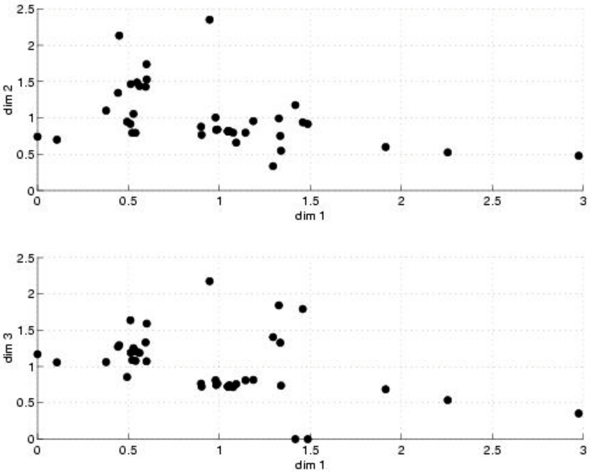

Figure 3. Individual weights over the three dimensions of the timbre space for the Exscal analysis.

PHYSICAL PARAMETERS UNDERLYING THE PERCEPTUAL SPACE

Pyschoacoustic Correlates

At this stage of the analysis we have obtained a spatial model representing the perceptual structure of of the car horn sounds. It is then important to give a physical interpretation of the perceptual dimensions revealed by the MDS analysis. According to previous studies on timbre [5,6,11], we perform an acoustical analysis of our sounds, in which the psychoacoustic descriptors found to be relevant to timbre perception by previous studies are computed. These descriptors are based on physiological models of the auditory system and computed directly on the signal.

	d1	d2	D3
Descr. 1	0.015	0.97**	0.080
Descr. 2	-0.86**	-0.058	-0.051
Descr. 3	0.34	-0.38	-0.82**

Table 2 : Correlations between psychoacoustic descriptors and perceptual dimensions. (There are 20 degrees of freedom in each case.)

Three descriptors have been found to match the perceptual dimensions. Table shows the correlation between those descriptors and the dimensions.
The first descriptor is the central spectroid, as computed in [6]. It has often been associated with the semantic attribute of "brightness".
The second descriptor is roughness, as described in [9].
The third descriptor is spectral deviation [5;6]. It is related to fine structure of the spectral envelope.

Specifities

The specificity values were very weak, except for two sounds. One had a fundamental frequency much lower than the other sounds. The other was the sound of a new prototype of horn. The specifity value indicates that even if it shares perceptual dimensions with the "traditional" sounds', subjects still judged it has dissimilar to the others.

Discussion

We have found psychoacoustic descriptors that explain a significant portion of the variance in the perception of the sounds of the auto horns. The descriptor (roughness) matching the first dimension characterizes the classification between monophonic and multiphonic sounds. It should be noted that the sounds were distributed continuously along this dimension. Subjects did not categorize in binary fashion the sounds between monophonic and multiphonic, but rather performed a continuous rating from pure periodic sounds (one single harmonic series) to sounds made of the addition of two periodic sounds. The spectral analysis of the intermediate sounds reveals that they are made of harmonic series based on the fundamental frequency, added to a second attenuated sub-harmonic series, which progressivly increases the perceived roughness.
The descriptor matching the second dimension (brightness) distinguishes perceptually between the acoustic signals of horn-like and plate-like devices, allowing us to describe what are, from a perceptual standpoint, those two main families of car horn sounds. The third dimension (spectral deviation) appears related to a descriptor characterizing fine-grained spectral aspects. This dimension is the only one that distinguishes sounds belonging to the predefined families of car horns related to the first two dimensions. Whereas these two dimensions match the typology (number of notes, technology of the device), this third dimensions explains the differences between the different clusters, within the main categories.

CONCLUSION

We have described the results of two experiments that sought to analyse the perception of contemporary car horns. The first experiment revealed a classification of our database of sounds into nine families. From this classification, we chose a sample of twenty two sounds representative of the variety of the original set of 43 sounds. The second experiment was a dissimilarity-rating experiment. We found three common perceptual

dimensions underlying the perception of the sounds, and three psychoacoustic descriptors correlated with these dimensions.
These two experiments provide us with a description of the different classes of sounds perceived by the subjects based on the acoustical signal. With these descriptors, we are now able to define the relevant acoustic features that are characteristic of contemporary car horns.
The next step will be the evaluation of the perceived urgency of the horn sounds. This problem deals not only with the acoustically-based perceived properties of the sounds, but also with the meaning carried by them. Once we have assessed the efficiency of these sounds to warn people about danger, we will also be able to relate it to the acoustical features described above. This would then allow us to create sounds that are different from those available today, but which are still perceived as car horns and good warning signals.

ACKNOLEDGMENT

This work is part of a PhD thesis financed by SCE Klaxon, part of FIAMM group, and is a collaboration between Ircam and the Univerisity of Le Mans. We would like to thank Boris Letinturier and Gilbert Hurst at Klaxon for their help, Jeremy Marozeau, Alain de Cheveigné and Olivier Houix at Ircam for successful discussion, and Vincent Rioux at Ircam for the Matlab graphical interface.

REFERENCES

[1] James A. Ballas. Common Factors in the Identification of an Assortment of Brief Everyday Sounds. *Journal of Experimental Psychology: Human Perception and Performance*, 19(2), 1993, pp. 250, 267.

[2] Patrick Susini, Ivan Perry, Xavier Rodet, Stephen McAdams and Sandrine Vieillard. Sensory Evaluation of Air-conditionning Noise: Sound Design and Psychoacoustic Evaluation. *International Conference on Acoustics*, Rome, 2001 (ICA 01).

[3] Olivier Houix. Structures Vibrantes et Catégorisation Auditive. *PhD Dissertation*, Université du Maine. In preparation.

[4] Patrick Susini, Stephen McAdams, and Suzanne Winsberg. A multidimensional Technique for Sound Quality Assessment. *Acta Acustica*, 85, 1999, pp. 659-656.

[5] Jochen Krimphoff. *Analyse Acoustique et Perception du Timbre.* DEA thesis, Université du Maine, 1994.

[6] Jeremy Marozeau, Alain de Cheveigné, Stephen McAdams and Suzanne Winsberg. The Dependency of Timbre on Fundamental Frequency. In preparation.

[7] Daniel Pressnitzer. Perception de Rugosité Psychoacoustique: D'un Attribut Elementaire de l'Audition à l'Ecoute Musicale. *PhD Dissertation*, Université de Paris 6, 1998.

[8] Geoffroy Peeters. Timbre Toolbox Documentation. *Technical Report*, Ircam, Analysis and Synthesis Team, 2000.

[9] P. Daniel and R. Weber. Psychoacoustical Roughness: Implementation of an Optimized Model. *Acta Acustica* 83, 1997, pp. 113-123.

[10] Suzanne Winsberg and Gert de Soete. A Latent Class Approach to Fitting the Weighted Euclidian Model, CLASCAL. *Psychometrika*, 59(2), 1993, pp.315-330

[11] Stephen McAdams, Suzanne Winsberg, Sophie Donnadieu, Gert de Soete and Jochen Krimphoff. Perceptual Scaling of Synthesized Musical Timbre: Common Dimensions, Specificities and Latent Subject Classes. *Psychological Research* 58, 1995, pp.177-192.