How different sources fill the soundscape. A perceptive and acoustical approach.

DEFREVILLE Boris[1], LAVANDIER Catherine[2], LANIRAY Marc[3]
[1]LASA, 236 bis rue de Tolbiac 75013 Paris, France, defreville@lasa.fr
Tel : + 33 143 133 400, Fax : + 33 143 133 409
[2]LMRTE Univ. Cergy-Pontoise, Dept Génie Civil - Rue d'Eragny, Neuville/Oise, 95031 Cergy-Pontoise , France
[3]01dB Acoustics&Vibration, 200, Chemin des Ormeaux, 69578 Limonest Cedex, France

Introduction

The objective characterization of an urban environment is generally assessed by the sound level evolution, where equivalent level L_{Aeq} or percentile levels as L_{A10} or L_{A5} can be calculated. If these measures can characterize a thorough fare where there's a continuous flow of vehicles, it is not enough to characterize a calm street where vehicles are perceived independently from the others [1].

This paper deals with the perception of urban environment sounds and focusing particularly on the more encountered vehicles in an urban environment (cars, mopeds, buses and motorcycles).

A sample of 20 stimuli of 15 seconds each, is extracted from a set of stereo recordings in urban street. The type and number of vehicles by stimulus is presented in table 1. The same stimuli are recorded with a measurement system. Criteria are calculated on the equivalent continuous sound pressure level $L_{Aeq.}$. A hearing test is next carried out and subjects are asked to evaluate 3 aspects of source perception: "prominence", "presence and "proximity".

Method

Recordings: Two kinds of recording were used simultaneously: one for measurement, the other for listening tests

For the listening tests, the sound environment samples were recorded with a pair of microphones in an ORTF-AB stereo configuration. This technique is known to give good results on spatial localization of the sound sources during the restitution.

For measurements, the sound acquisition set (*01dB symphony*® system) consists of one transducer linked to a small computer unit (a single channel omni directional *Bruel&Kjaer* class1 microphone), which transfers data in real-time to a notebook computer. Stimuli were recorded, in 2 different streets with the same U shape, quite the same speed of vehicles (about 50 km/h), and with the same type of road coating (asphalt) with variable granulometry. The standards for environment measurements were respected: 3 meters from the source and 1.5 meters above ground level.

Semantic description: General perception of a sensitive world can be structured, thanks to multidimensional approaches. In the book "Ambiances Architecturales et Urbaines", P. Woloszyn and D. Siret [2] try to elaborate a perceptual model of ambient objects such as sounds, smells or sight. They propose a perceptual spatial representation built on three notions: the "prominence", the "presence" and the "proximity" of the object. Contrary to the complexity of these notions in their paper, they have been simplified for the auditory tests. The "prominence" has been presented as the importance of a sound source in its environment, the "presence" as the duration of appearance, and the "proximity" as the distance to the source.

	Car	Moped	Motorcycle	Bus
Stimulus 1	4			
Stimulus 2	3	2		
Stimulus 3	3		1	1
Stimulus 4	5			
Stimulus 5	1	1		
Stimulus 6		1		1
Stimulus 7		2		
Stimulus 8	2	1	1	
Stimulus 9	1	1		
Stimulus 10				1
Stimulus 11	3	1		
Stimulus 12	1			
Stimulus 13	3			
Stimulus 14	3	1	1	
Stimulus 15	4			
Stimulus 16	1			
Stimulus 17	3	2		
Stimulus 18	3			1
Stimulus 19				2
Stimulus 20			1	

Table 1: Number and type of vehicles in the 20 stimuli

Hearing test: An experiment was conducted with a corpus of 20 subjects (15 men and 5 women), most of them being acoustician or sound engineers. For each stimulus presented randomly and through a *Seinheiser* HD 600 headphone, they had to rate the "prominence", the "presence" and the "proximity", for each type of source, on a 1-10 dynamic scale: 1 was "few prominent" (respectively few present and very far) and 10 was "very prominent" (respectively very present and very close). Training stimuli were presented at the beginning of the test to help subjects to respect the imposed scale and the sounds were presented as the same level as recorded. All signals were filtered by the inverse transfer function of the headphone. The test lasted around 30 minutes.

The answers, for the names such as "prominence", "presence" and "proximity" and for each type of vehicle, are transformed into a matrix of histograms. To reduce dimensions of this matrix, answers "0" are counted in a class named "not", answers from "1" to "3" are gathered into a class named "a few", answers from "4" to "6" are gathered into a class named "enough" and answers from "7" to "10" are gathered into a class named "a lot". For example: 6 subjects gave answers between "1" and "3" for the presence of the cars in the stimulus number 5. It means 6 subjects answer “a few cars” for the stimulus number 5.

Calculation of different criteria:

On the L_{Aeq} curve of each stimulus, descriptors related to the levels (average L_{Aeq}, percentile levels L_{A10}, L_{A5}, maximum level L_{Amax}), descriptors related to the same criteria over each type of sources (coded sources), and to the time (duration ratio of presence (%D), number of vehicles) are calculated.

Results

Correspondence Analysis (CA): this analysis allows us to extract parameters that describe the variability of the answers. The reduced structure proposes 4 factors that explain 84% of the data variance.

The first dimension: 35% seems to explain the identification of buses, from “not” (no bus) to “a lot” (a lot of buses). This first dimension is correlated to criteria reflecting the levels as well as the temporal ones (table 2). The consensus of the answers here is probably due to the maximum level of buses which is more important than the other vehicles. The first dimension is significantly correlated to the prominence, the presence and the proximity ($r>0.8$, $p<0.001$).

	Dim 1		Dim 2		Dim 3
	Bus	Car	Moped	Car	Motorbike
Prominence	-0.86*	0.64*	0.77*	-0.65*	-0.88*
Presence	-0.87*	0.59*	0.76*	-0.55	-0.84*
Proximity	-0.84*	0.66*	0.77*	-0.56*	-0.85*
L_{Amax}	-0,87*	0.55	0.79*	0.58*	-0.80*
%D	-0.88*	0.40	0.60*	-0.76*	-0.86*
Number	-0.87*	0.41	0.67*	-0.56	-0.82*

Table 2: correlation coefficients between stimuli coordinates and acoustic criteria; (*) indicates probability $p<0.01$

The second dimension, explaining 23% of the answer variability, is related to the identification of moped. As for the buses, the three aspects of perception seem to be mixed up. Nevertheless, the level criterion is better correlated with this dimension than the temporal one.

The identification of cars is distributed over the two first dimensions (table 2). One explanation could come from the construction of the test: the more numerous the cars are, the less numerous the mopeds or the buses are (table 1). It can be noticed that the temporal criterion (%D) explains the discrimination between stimuli over the second dimension.

The third dimension explains the identification of the motorbikes, especially for the stimulus number 20 where the motorbike is the only vehicle. No aspect of perception seems to be better mixed up than another and it is the same for the criteria calculated on the L_{Aeq} curve (table 2).

The fourth dimension explains the errors of motorbike and car identifications for the stimuli “3” and “8” which are opposed, on this fourth dimension, to the stimuli “20” and “14” (fig 1). There are no errors for these two last stimuli. Of course no acoustic criterion is significantly correlated to this dimension.

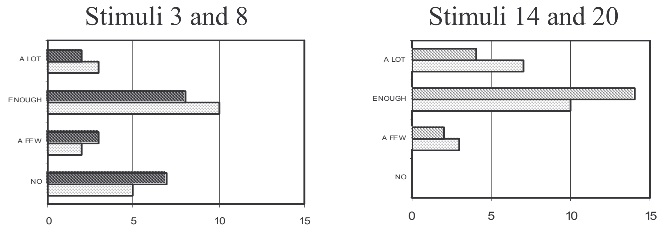

Fig 1: Histogram description of marks concerning the “presence” of the motorbikes for four different stimuli. We oppose the errors in the stimuli 3 and 8 to the correct marks in the stimuli 14 and 20.

Discussion and conclusion

Identification of vehicle pass-bys has been expressed through “prominence”, “presence” and “proximity”. The answers concerning four types of vehicles are analyzed with a multidimensional correspondence analysis. This reveals four main dimensions. Then, correlations allowed to find a selection of acoustic criteria that could explain these dimensions.

Generally, “prominence”, “presence” and “proximity” have been interpreted in the same way (they have the same distribution over the perceptual space) except for cars.

For buses it seems that subjects express a consensus which can be explained by level indicators as well as temporal indicators.

For the cars, “prominence” is not distributed in the same way as “presence” and “proximity”. The variability of car numbers in the different stimuli helps to better discriminate the three notions.

To sum-up, some acoustic criteria, extracted from the temporal evolution of the coded source levels, can explain consensus over sound identification of different types of vehicles in urban streets. This codification poses two problems. First, the coding doesn’t allow to discriminate two different sources occurring at the same time. Second, to be manually coded, the stimuli need to be heard; therefore they need to be recorded. This process takes time and space on computers. It should be interesting to find a process to automatically separate the L_{Aeq} temporal evolutions of each type of sources.

References

[1] Raimbault M., Lavandier C., Berangier M., «Ambient sound assessment of urban environments: field studies in two French cities». Applied Acoustics, Vol. 64 (12), pp.1241-1256, December 2003.

[2] Woloszyn P., Siret D., « Du complexe au simplexe, le modèle des objets ambiants » in Cahiers de la Recherche Architecturale, Paris, Octobre 1998.