1997 - 102nd CONVENTION of A.E.S. Munich

Sound Characterization Of Urban Environments : an Approach Based on Ecological Validity

Corsin VOGEL, Valérie MAFFIOLO, Jean-Dominique POLACK and Michèle CASTELLENGO

Laboratoire d'Acoustique Musicale, Université Paris VI, CNRS,

4 place Jussieu, 75252 Paris Cedex 05, France

Abstract

In order to characterize typical soundscapes of urban environments, a sound corpus is selected by methods akin to sociological surveys. In a second step, recording and reproduction techniques best adapted to the soundscapes are selected by psychoacoustic tests. The results indicate that specific recording arrangements are not appropriate outdoors for moving sources (work supported by PIR-Villes, CNRS, France).

1. INTRODUCTION

The perceptive and acoustical characterization of sound environments is nowadays an important direction of research among those concerning the quality of life. Many works are used to reduce, in term of sound level, the effect of various annoying sources, but are insufficient to treat the problems of sound quality. Indeed the qualitative appreciation by an audience listening to a sound environment depends also on others components of human perception [1]. Consequently we take the point of view of the human person to study the perception of the sound space.

In order to characterize urban sound ambiences, that is, scene-like fragments of the sound environment bearing meaning and identity, it is first necessary to establish a corpus of recorded soundscapes. The work consists in setting a method for selecting soundscapes that can be considered as representative of the studied city. In a second step, it aims at selecting the recording and reproduction techniques of sound sequences closest to human perception, that is, the one creating the best illusion so that subjects react as if they were in the real environment. The research has been applied to Paris in its exploratory phase.

2. SELECTION OF TYPICAL SOUNDSCAPES

Sociology has identified techniques that can be used in the selection of soundscapes from surveys with users of a city, that are completed here by contributions from cognitive

psychology and linguistic. In order to select the sound extracts to be recorded, an investigation technique developed by Amphoux [2] was used which takes into account the acoustical, topological and sociocultural dimensions of cities.

2.1. Methodology

The aim of this methodology is to define criteria for selecting soundscapes considered as representative of the sound identity of a city, and to obtain qualification criteria based on direct perception of recorded sound spaces. We have intentionally limited our study to sound memorization and particularly to the realisation of what Amphoux calls "mental sound maps" ("cartes mentales sonores"). In a first step the process consists in asking subjects to draw "sound sketches" ("dessiner du sonore"). This forces the subjects to change their analysis by executing a cartographical representation (that is visual) of sounds. In a second step, the subject verbally describes sound experiences which he/she has been through (and so memorized). This method makes it possible to establish a corpus of typical soundscapes of a city.

2.2. Interviews

In order to establish a representative corpus of locations which would have particular acoustic richness, we carried out an enquiry in Paris. It consisted in asking different subjects living in the city :

1 - to establish a graphical representation of the sound space in Paris by drawing sketches which answered to the following question : "What is the sound of Paris according to you?", and then to comment the obtained drawing,

2 - to enumerate locations or routes having, to their ears, particular acoustic features.

During the interviews, we used the expression "sound space representation" ("représentation de l'espace sonore") rather than "mental sound map" ("carte mentale sonore") used by Amphoux, in order to avoid inciting the subject to use a specific analysis of the city, such as a cartographical analysis. Indeed the word "map" has a very topographical connotation. Actually the aim was to leave the subjects free to approach the drawing as they liked and to represent the sounds or sound locations as they wish.

30 persons (17 men and 13 women, 20 to 58 years old) have been interviewed by two researchers. The duration of the interview is expected to be brief and to give rise to spontaneous representations. In most cases, it lasted about 30 minutes but in some cases it lasted longer giving more valuable information.

2.3. Results

The verbal descriptions allowed us to verify that for almost the totality of subjects the sound of Paris is essentially made of traffic noises. The locations and activities most often mentioned in the conversations, are in particular : the traffic of the ring road, of Saint-Germain and Hôpital boulevards, of the place de la Bastille, the birds and the child of the jardin des Plantes, of the Buttes Chaumont and of the bois de Vincennes, the markets of Aligre and of the Mouffetard street, the bustles in the streets of Montmartre, the Halles and

the parvis de Beaubourg, the walks in the Marais and the Coulée verte. These locations display soundscapes that can be considered as typical of Paris.

Actually it is possible to further develop the analysis of the results. Although this analysis is not directly relevant to the present topic, we decided to present it here in order to show how graphical and verbal descriptions can be used in psychology under the paradigm of "free categorization" [1].

The main difficulty of the subject task is the translation of an auditory percept into a graphical representation. Indeed we obtained four classes of sound space representations (Figure 1) : cartographies of Paris (20%), specific locations with generic or proper names (23%), unorganized sources in a location (50%) and abstract schemata (13%) (some people taking several strategies). This differenciation between classes of representation is difficult to compare with Amphoux's observation where he distinguishes three general forms of representations (general map, itineraries and discontinuous and fragmentary unities as a whole) and three specific modes of representation (codified symbolization of the location, expression of a type of space and icon representing the type of sound sources).

Through the analysis of the verbalisation used by subjects when they were asked to comment their sketches or to enumerate locations having particular features, we noticed that the subjects talk about locations in different ways : in terms of topological description, in terms of qualitative appreciation, but also according to the type of location, their positions in the location and their activities. Indeed the linguistic analysis [3] shows that the qualifiers used to talk about the underground, for example, allow one to identify two subject positions : in the underground train or on the platform. In general we can know how the subjects project themselves into the location by asking the questions : Are they inside or outside the location? If they are inside, are they active or not? If they are active, do they produce sound or not? It appears then that the subjects structure and describe the location in relation to themselves.

When we take into account all the ways that can be used to talk about a location, it is possible, on one hand, to understand what a location is from the point of view of the speakers and how it appears to them, and on the other hand, to analyse how the space is identified in general.

In the present case, only the first step of the results was kept for the second part of the research. The list of locations was used to select typical recording settings in the city of Paris in order to constitute a corpus of soundscape extracts.

3. RECORDING AND REPRODUCTION OF SOUNDSCAPES

Our second aim is to select recording techniques that account best for our everyday environment. Retaining the ecological validity of the sound patterns is vital. This environment perceived when listening to the recording should be as near as possible to the environment perceived in reality. Thus, we have tested some sound recording configurations in urban context in order to select the most credibles.

3.1. Preliminary Selection of the Recording and Reproducing Techniques

First, we had to make some choices based on ecological validity of the sound patterns and of material constraints, such as ruling out transaural reproduction technique :

- We chosed sterephonic sound recording in order to obtain sound images near to our perception, that is, which permits localisation and spatial effects [4].
- We used electrostatic microphones (Schoeps MK5, MK6 and MK21) with wide, flat frequency responses and wide dynamic ranges in order to reproduce ambiences such as traffic noise which is very rich in low frequencies, particularly from 20 to 100 Hz.
- We recorded the sounds with a portable R-DAT (HHb Portadat), sampled at 48 kHz.
- Playback was done on loudspeakers (Studer A723) in fairly anechoic surrounding, headphones giving the impression of ambiences confined in the head (for example a car passing in front of the microphones is perceived passing through the head). This reduces significantly the illusion of being in real situation and thus, the ecological validity of the reproduced sounds. The sound restitution level was fixed at 10 dBA below the effective sound level, the later being judged as too loud in the fairly anechoic room. Subjects are placed 1.5 m in average from the loudspeakers.

3.2. First Test

3.2.1. Method

A first test was realized in order to find the optimal spacing between microphones for given physical spread angles. Three directivities with four spacings have been tested :

- O60 (omnidirectional, 60 cm spacing, 0° spread), O70, O80, O90,
- W20 (wide cardioid, 20 cm spacing, 0° spread), W40, W60, W70,
- C20 (cardioid, 20 cm spacing, 120° spread), C30, C40, C60.

The test consisted in listening to 12 traffic sequences during about one minute each, devided in three series according to the three directivities. The sequences were recorded near a crossroad, on the footpath (1.6 m from ground). For each spacing series, subjects were asked to select the nearest to reality. The 12 subjects (9 men and 3 women, 24 to 58 years old), all acousticians, could listen again to the sequences and change the order of listening.

3.2.2. Results

The responses indicate a preference for the recording configurations with spacing of 60 and 70 cm. Indeed for narrower spacing, the stereophonic image is narrow. For larger spacing, the stereophonic image presents a "hole" in the center.

Besides, it proved essential to realize simultaneous recordings of sound sequences in order to facilitate their comparison. In the first test, the sound content of each sequence could drastically change from one recording to another, influencing the subject judgements.

Moreover, the duration of each sequence was to long and subjects could relax their attention. The sound level of the samples was also perceived unequal.

Last, the notion of realism wasn't appropriate for some subjects in a simulation context.

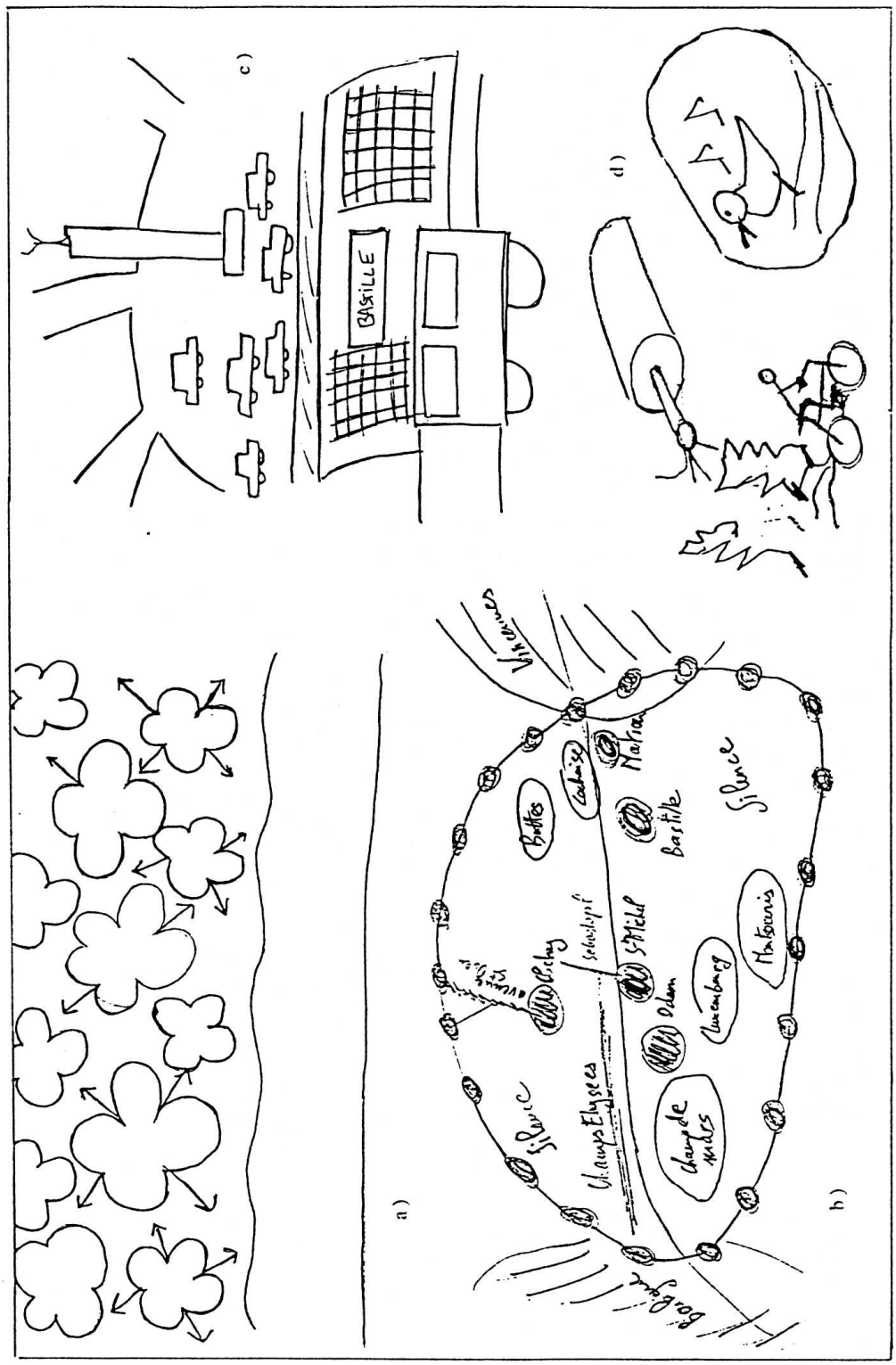

Figure 1: 4 types of sound space representation : a) abstract schemata, b) cartography, c) specific locations and d) unorganized sources in a location.

3.3. Second Test

3.3.1. Method

A second test followed in order to determine configurations most appropriate to urban sound ambiences. The three configurations kept from the first test, O70, C60 (100°) and W70 (100°), were compared to two well tried indoor configurations : ORTF (cardioid, 17 cm, 110°, French broadcasting) and C35 (90°, approximate the Dutch broadcasting system NOS). All the critics from the first test were taken into account. The experiment consisted in pair comparisons, a method already used in our previous works [5]. 14 pairs of ambiences (8 traffic ambiences and 6 park ambiences), each recorded simultaneously, were presented to 48 subjects (32 men and 16 women, 21 to 58 years old) at equal A-weighted sound pressure level. In order to avoid any aesthetic judgement from the subjects, they were asked to choose the sequence which they could more easily feel within. They had then to estimate the dissimilarity between both sequences on a continuous scale, graduated from 1 to 7 (1 : weakly dissimilar, 7 : strongly dissimilar).

3.3.2. Results

Subjects were inclined to choose the second sequence. To rule out artifacts, we made a second version of the test with 24 out of the original 48 subjects, where the sequences were inverted in each pair. As a matter of fact, an order effect appears in the results and introduces a significant shift (Figures 2 and 3). This order effect is predominant for pairs 1, 2, 3, 6, 7, 11 and 14 and doesn't permit any other conclusion. When listening to the first sequence, the subjects trie to recognize and identify the ambience ; whereas, when listening to the second sequence, they know its content and can become integrated in the ambience more easily.

Only pairs 10 and 13 give significant results : the order effect is not important and preferences are pronounced. So, for the park, W70 and C60 are prefered to O60. For pairs 4, 5, 9 and 12, the order effect is not too predominant and we can conclude that ORTF is prefered to O70 for the park, W70 is prefered to O70 and C60 to ORTF for traffic noise. In our listening configuration, subjects consider that the low frequency in omnidirectional recordings is excessive.

A dissimilarity scale analyse shows that the pairs considered as weakly dissimilar are the wide cardioid / cardioid pairs and the ones considered as strongly dissimilar are the omnidirectional / cardioid pairs. This confirms the fact that the omnidirectional configuration is very different from the other configurations. Every comparison to omnidirectional configuration is significant. On the contrary, the comparison between wide cardioid and cardioid is the most influenced by the order effect. Subjects can't choose easily, they mark the pair weakly dissimilar.

The sequences duration, reduced to 30 seconds, was still considered too long, subjects complained that the sequences were too rich, with a risk for forgetting past events.

4. CONCLUSION

The selection of urban soundscapes has allowed us to obtain much information from verbal and graphical descriptions of the perception of the sound environment. Indeed the verbal description analysis from talks has allowed us to identify locations and activities most often mentioned. We also showed that the analysis can be pushed further to extract information on the structure in the mental representation of a soundscape, not only in terms of topology and qualitative appreciation, but also according to the subjects, the position they have in the location and their activities. The four classes of sketches, ranging from maps to abstract locations, shows the interest of a pluridisciplinary process which allow the analysis of an auditory percept from its graphical and verbal expression.

We then tested the ecological validity of the simulation for some recording techniques. In a first approximation, results indicate that the stereophonic omnidirectional configuration O70 is not appropriate for outdoor recordings. Furthermore, the ORTF system doesn't seem to be appropriate for traffic noise, that is, for sources in motion. In a park, the motion of sources is not so important, the listening is more static and the ORTF system is not rejected. However, the other recording configurations were estimated to better provoke the illusion of being in real situation. The order effect is predominant for all these pairs so that we can consider all these recording techniques adequate in the chosen listening configuration. But, it could be very interesting to compare more precisely some recording systems for which the order effect was predominant, in particular the configurations C60, ORTF and W70.

We shall now record the soundscapes of typical locations of Paris, as selected above, with the configuration C 60 (cardioid microphones being the only one not rent) which seems to be well adapted for each of them. Then we intend to perform a subjective analysis by means of categorization and dissimilarity tests, including the analysis of the verbalisation used by subjects. An objective analysis by sonogram will also be performed in an attempt to extract the physical parameters characterizing each soundscape.

Acknowledgments : The authors wish to thank Danièle DUBOIS and Sophie DAVID, LCPE-Paris, for valuable discussions and inputs to the research.

5. BIBLIOGRAPHY

[1] Guyot F. (1996) *Etude de la perception sonore en termes de reconnaissance et d'appréciation qualitative : une approche par la catégorisation.* Ph.D. thesis, Université du Maine, France.

[2] Amphoux P. (1991) *Aux écoutes de la ville - La qualité sonore des espaces publics européens - Méthode d'analyse comparative - Enquête sur trois villes suisses.* Research report 94. IREC, architecture department, EPF Lausanne, Switzerland.

[3] David S., Dubois D., Rouby C., Schaal B. (1997) *L'expression des odeurs en français : analyse morphosyntaxique et représentation cognitive.* Intellectica (in press).

[4] Hugonnet C., Walder P., (1995) *Théorie et pratique de la prise de son stéréophonique.* Eyrolles, Paris, France.

[5] Vogel C., (1995) *Etude de l'influence des niveaux de bruit et vibrations sur le confort des passagers ATR.* DEA report, Université Aix-Marseille II, Aerospatiale, France.

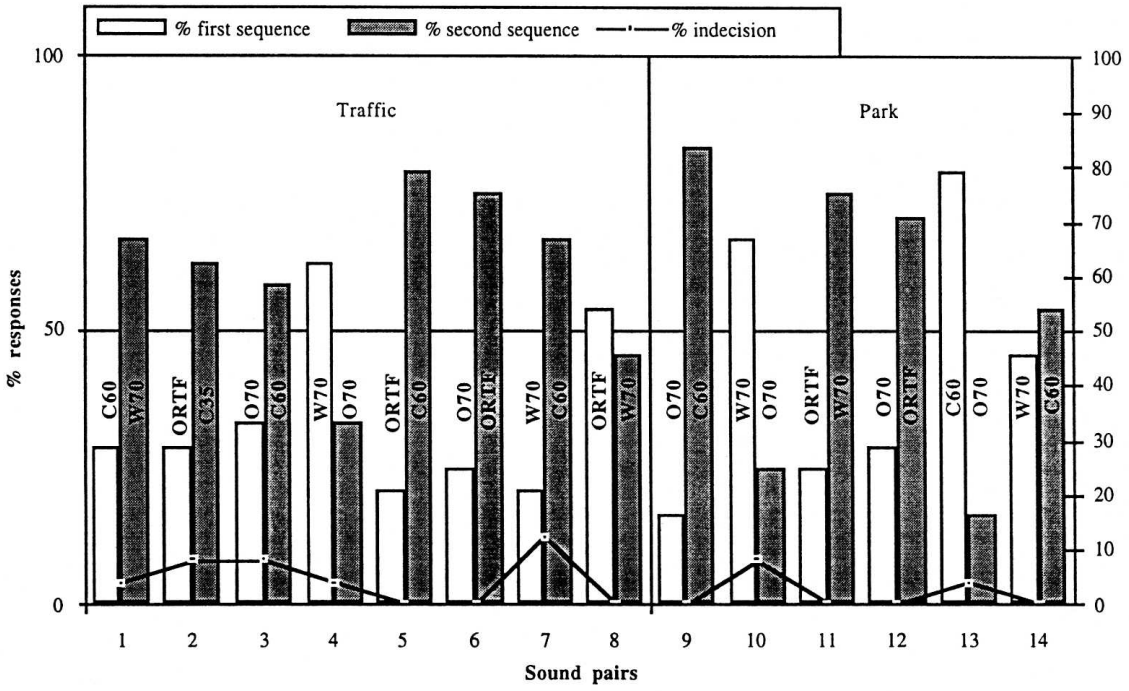

Figure 2 : results from second test, first set (24 subjects).

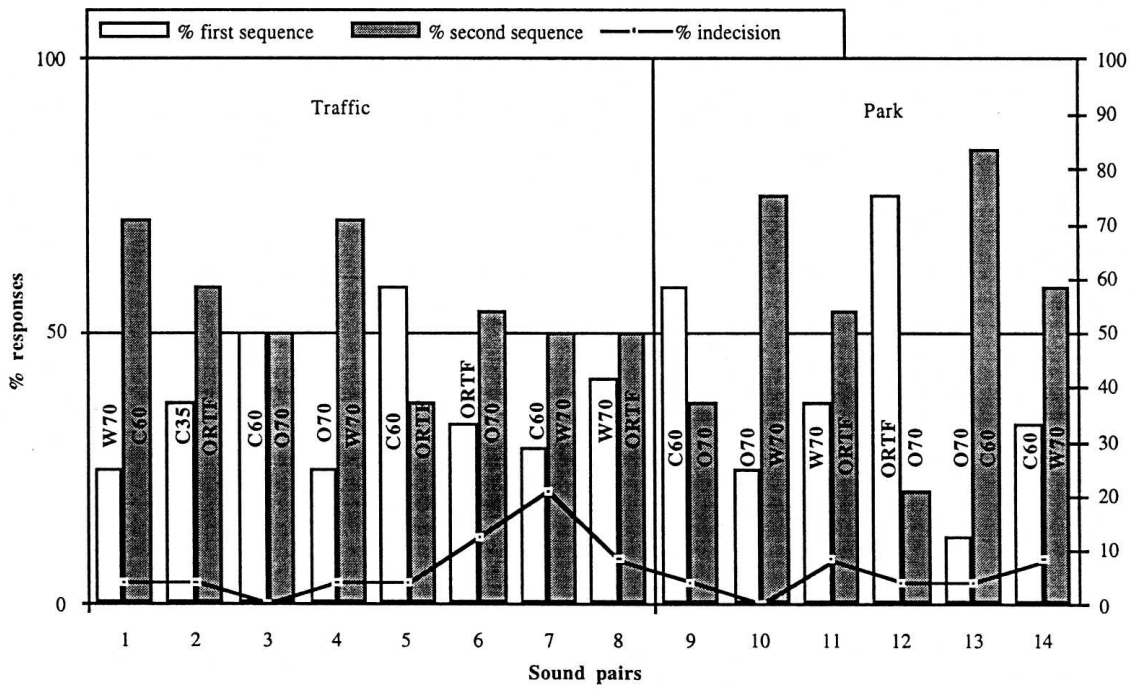

Figure 3 : results from second test, second set (24 subjects).