

# Validation subjective de la prise de son en extérieur

C. VOGEL, V. MAFFIOLO, J.-D. POLACK, M. CASTELLENGO

Laboratoire d'Acoustique Musicale

UMR 9945 CNRS, Ministère de la Culture

Université Pierre & Marie Curie Paris 6

Tour 66, Case 161, 4, place Jussieu, F-75252 Paris cedex 05

**Summary :** In order to study the characterisation of soundscapes, it is necessary to record and reproduce them. The aim is to select the recording and reproduction techniques nearest to our perception, that is, the one creating the best illusion so that subjects react as if they were in the real environment. Test sequences were recorded at arbitrary locations in Paris with traditional recording methods, in order to determine which recording arrangement is best suited for specific ambiances. Playback was done on loudspeakers in fairly anechoic surrounding. Subjects were presented 14 pairs of sequences, each pair recorded simultaneously. In order to avoid any aesthetic judgement, the real goal of the experiment, that is, the comparison of the sound recordings, was hidden. For each pair, subjects were asked to choose in which sequence it was easier for them to feel inside the recorded soundscape. Results indicate that configurations used indoors for fixed sources cannot adequately reproduce sources in motion.

## 1. INTRODUCTION

Cette étude s'inscrit dans le cadre d'une recherche sur la caractérisation des ambiances sonores urbaines et fait suite à la communication "Ambiances sonores représentatives d'une ville : le cas de Paris". Son objectif est de sélectionner des techniques d'enregistrement de séquences sonores qui permettent de rendre le mieux compte de notre environnement quotidien. Notre souci premier lors d'enregistrements des ambiances sonores urbaines est donc de conserver le caractère écologiquement valide des échantillons sonores. Cela signifie qu'il faut essayer de donner à l'auditeur *"l'illusion crédible qu'il existe entre et au-delà des enceintes, un environnement acoustique dans lequel des sources sonores recréées peuvent être localisées et dissociées dans l'espace"* (Hugonnet et Walder, 1995). Cet environnement perçu à l'écoute de l'enregistrement doit être le plus proche possible de celui que l'on perçoit dans la réalité. Par conséquent, nous avons testé un certain nombre de configurations de prise de son en milieu urbain afin de sélectionner les plus crédibles.

## 2. CHOIX PRELIMINAIRES DE PRISE ET DE RESTITUTION DU SON

Pour des questions de validité écologique des échantillons sonores et de contraintes matérielles, nous avons choisi :

- une prise de son stéréophonique afin d'obtenir une image sonore proche de notre perception qui rende compte à la fois de la localisation et de l'effet d'espace.
- des microphones électrostatiques (Schoeps MK5, MK6 et MK21) qui ont l'avantage d'avoir une bonne réponse en fréquence pour reproduire une ambiance telle que les bruits de circulation, très chargés en basses fréquences (20-100 Hz notamment).
- les enregistrements sont effectués à l'aide d'un enregistreur R-DAT portable (hhb Portadat), échantillonnés à 48 kHz.
- une reproduction en chambre semi-anéchoïque sur haut-parleurs (Studer A723), celle au casque donnant une impression d'ambiance confinée dans la tête ce qui a pour effet de réduire considérablement l'illusion d'être en situation réelle. Le niveau de restitution est fixé 10 dBA en dessous du niveau réel ambiant, ce dernier ayant été estimé insupportable dans la chambre semi-anéchoïque. Les sujets sont placés à environ 1m 50 des enceintes.

## 3. PREMIER TEST

### 3.1 Méthode

Une première expérience a été réalisée en vue de déterminer les écartements optimaux entre microphones pour des angles physiques d'ouverture donnés. Trois directivités avec quatre écartements ont été testées :

- **O60** (omnidirectionnel, 60 cm d'écartement, 0° d'ouverture), **O70**, **O80**, **O90**,
- **I20** (infracardioïde, 20 cm d'écartement, 0° d'ouverture), **I40**, **I60**, **I70**,
- **C20** (cardioïde, 20 cm d'écartement, 120° d'ouverture), **C30**, **C40**, **C60**.

Le test consistait à faire entendre 12 séquences de bruits de circulation d'environ une minute chacune à 12 sujets, tous acousticiens. Les séquences sonores ont été enregistrées près d'un carrefour, en milieu piétonnier (sur le trottoir, à 1m 60 du sol). Nous demandions de sélectionner parmi les 4 premières séquences celle qui semblait au sujet la plus proche de la réalité, de même pour les 8 séquences suivantes. Le sujet était libre de réécouter les séquences et de varier l'ordre d'écoute.

### 3.2 Résultats

Les réponses des sujets ont mis en évidence une préférence pour les configurations dont les écartements sont compris entre 60 et 70 cm. En effet, en-deçà, l'image stéréophonique reste étriquée. Au-dessus, il apparaît un "trou" au centre dans l'image stéréophonique, le décalage temporel des signaux arrivant sur les deux microphones en incidence latérale devenant trop important.

Par ailleurs, il s'est avéré primordial de réaliser des enregistrements simultanés des séquences sonores afin de faciliter leur comparaison. En effet, le contenu sonore de chaque séquence peut considérablement varier d'une prise à l'autre et influencer le jugement des sujets.

De plus, il est apparu que la durée de chaque extrait sonore était trop longue, les sujets ayant tout le temps de se déconcentrer, et que le niveau sonore des échantillons était perceptiblement inégal.

Enfin, l'utilisation de la notion de réalisme n'était pas appropriée car elle introduit un concept que certains sujets n'acceptent pas dans un contexte de simulation.

## 4. DEUXIEME TEST

### 4.1 Méthode

Une seconde expérience de prise de son a été effectuée afin de déterminer les configurations les mieux adaptées aux ambiances sonores de la ville. Les trois configurations retenues à l'issue du premier test, **O70**, **C60** (100°) et **I70** (100°), ont été comparées à deux configurations éprouvées en intérieur, **ORTF** (cardioïde, 17 cm, 110°, radiodiffusion française) et **C35** (90°, proche du système NOS de la radiodiffusion hollandaise). Tous les problèmes dégagés des résultats du premier test ont été pris en considération. L'expérience a consisté en une comparaison par paires, méthode déjà utilisée dans des travaux antérieurs (Vogel, 1995). 14 paires de séquences (8 de circulation et 6 de jardin public) enregistrées simultanément ont été présentées au même niveau sonore (dBA) à 48 sujets. Pour éviter tout jugement esthétique de la part des sujets, il leur était demandé de choisir la séquence dans laquelle ils s'intégraient le plus facilement. Ils devaient ensuite estimer sur une échelle continue, graduée de 1 à 7, la dissemblance entre les deux extraits sonores (1 : peu dissemblable, 7 : très dissemblable).

### 4.2 Résultats

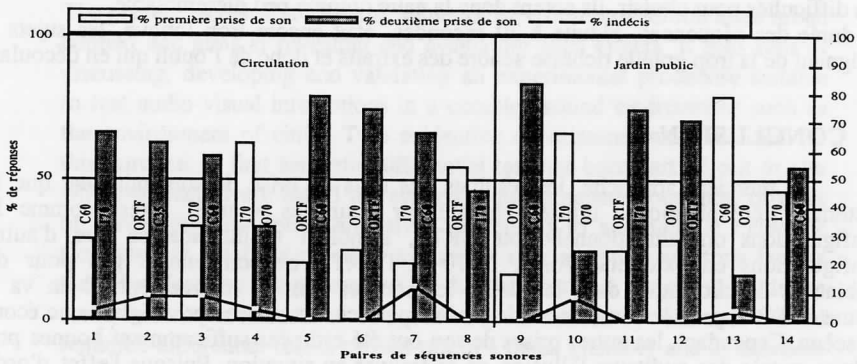


Fig. 1 : résultats du second test, première série (24 sujets)

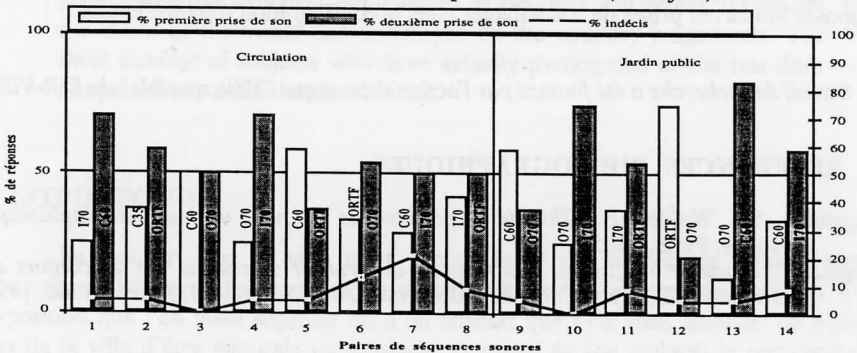


Fig. 2 : résultats du second test, deuxième série (24 sujets)

Nous avons constaté que les auditeurs avaient tendance à choisir la deuxième séquence (Fig. 1 et 2). Pour savoir si cette attitude est liée à la séquence écoutée, nous avons fait une deuxième série de test (24 des 48 sujets) où les séquences de chaque paire sont inversées. Il apparaît effectivement un effet d'ordre dans les résultats, celui-ci étant prépondérant dans les paires 1, 2, 3, 6, 7, 11 et 14. En fait, le sujet, à l'écoute de la première séquence, serait dans une phase de reconnaissance et d'identification de l'ambiance. A l'écoute de la seconde séquence, connaissant son contenu sonore, il s'y intégrerait plus facilement.

Seules les paires 10 et 13 donnent des résultats pertinents : l'effet d'ordre n'y apparaît pas et les préférences sont marquées. Ainsi, pour le jardin public, I70 et C60 sont préférés à O70. Pour les paires 4, 5, 9 et 12, l'effet d'ordre est faible et permet d'en déduire que, pour le jardin public, ORTF est préféré à O70 et que, pour la circulation, I70 est préféré à O70 et C60 à ORTF. Dans la configuration d'écoute retenue, les sujets considèrent que les basses fréquences des ambiances enregistrées en omnidirectionnel sont excessives.

En dépouillant les échelles de dissemblance, il apparaît que la paire considérée comme peu dissemblable est la paire infra/cardio et que la paire très dissemblable est la paire omni/cardio. Cela nous conforte dans les déductions faites précédemment car la configuration qui sort nettement du lot est bien celle avec les microphones omnidirectionnels. Toute comparaison avec celle-ci donne des résultats tranchés. De même, la comparaison infra/cardio est la plus influencée par l'effet d'ordre. Les sujets ont des difficultés pour choisir, ils notent donc la paire comme peu dissemblable.

La durée des séquences, réduite à 30 secondes, était encore trop longue, les sujets se plaignant de la trop grande richesse sonore des extraits et donc de l'oubli qui en découlait.

## 5. CONCLUSION

En première approche, les résultats des tests de prise de son indiquent que les systèmes stéréophoniques utilisés en intérieur pour des sources fixes, comme les configurations omnidirectionnelle ou ORTF, semblent moins adaptés que d'autres configurations en extérieur. Ainsi, l'O70 et l'ORTF ne conviennent pas pour des ambiances de circulation dans lesquelles les sources sont en mouvement. Il en va de même de l'O70 pour les ambiances de jardins publics, ambiances privilégiant une écoute de scène. Cependant, les autres prises de son ont été estimées suffisamment bonnes pour induire auprès des auditeurs l'illusion de se trouver en situation. Puisque l'effet d'ordre est prépondérant dans les résultats de comparaison de ces prises de son, nous pouvons supposer toutes ces prises de son valables dans la configuration d'écoute choisie.

*Ce travail de recherche a été financé par l'action thématique "Ville sensible" du PIR-Villes*

## 6. REFERENCES BIBLIOGRAPHIQUES

- Hugonnet, C., Walder, P., *Théorie et pratique de la prise de son stéréophonique*, Eyrolles, Paris, 1995.  
Vogel, C., *Etude de l'influence des niveaux de bruit et vibrations sur le confort des passagers ATR*, rapport de DEA, Université Aix-Marseille II, Aérospatiale, 1995.