

THESE de DOCTORAT de l'UNIVERSITE PARIS 6

Spécialité :

**Acoustique**

Présentée

par M. Corsin VOGEL

pour obtenir le grade de DOCTEUR de l'UNIVERSITE PARIS 6

Sujet de la thèse :

**Etude sémiotique et acoustique de l'identification  
des signaux sonores d'avertissement  
en contexte urbain**

soutenue le 05 octobre 1999

devant le jury composé de :

Mme Anne BARDOT	Examineur
M. Durand BEGAULT	Rapporteur
Mme Danièle DUBOIS	Examineur
M. Jean HARDY	Rapporteur
M. Jean-Pierre PENEAU	Rapporteur et Président du jury
M. Jean-Dominique POLACK	Examineur



# Avertissement

Ce mémoire en contient déjà bien suffisamment pour que nous renoncions ici au moindre avertissement. Tout au plus pouvons-nous suggérer au lecteur un néologisme qui caractériserait ce nouveau champ de recherche pluridisciplinaire : la *sémioacoustique* (terme proposé par Monsieur Jean-Pierre Péneau à l'issue de la soutenance de cette thèse).

## Avant-propos

L'approche cognitive des sons de l'environnement, adoptée dans ce travail, est issue d'une étroite collaboration entre le Laboratoire d'Acoustique Musicale et le laboratoire Langages, Cognitions, Pratiques, Ergonomie. Les contributions associées à cette thématique de caractérisation, d'identification et de signification liées aux sons ont été initiées par le mémoire de Guyot (1996) : "*Etude de la perception sonore en termes de reconnaissance et d'appréciation qualitative : une approche par la catégorisation*". Cette recherche est aujourd'hui suivie par deux mémoires réalisés en parallèle, à savoir celui de Maffiolo (1999) : "*Contribution à la caractérisation sémantique et acoustique de la qualité de l'environnement sonore urbain*" et celui ci-présent.

Cette recherche a été menée dans le cadre de l'action thématique "Ville sensible" du Programme Interdisciplinaire de Recherches sur les Villes (PIR-Villes).

Certains exemplaires de ce mémoire sont accompagnés d'un CD-R contenant les échantillons sonores utilisés dans les expériences d'écoute.



# Remerciements

Les nombreuses personnes m'ayant aidé au cours de ce travail, de près ou de loin, intellectuellement, moralement ou matériellement, ne sauraient toutes être citées ici. Qu'elles y trouvent néanmoins toute ma reconnaissance pour tous nos échanges et nos partages.

Je voudrais tout d'abord exprimer mes meilleurs sentiments à Michèle Castellengo, non seulement pour m'avoir accueilli au sein du Laboratoire d'Acoustique Musicale, mais aussi et surtout pour m'avoir formé à de nombreux concepts de l'acoustique musicale et de la perception auditive, par l'intermédiaire de discussions passionnantes et de ses cours au C.N.S.M.D.P. Je me sens redevable envers elle de mon épanouissement au cours de ces quatre dernières années.

Je tiens également à témoigner toute ma gratitude à Jean-Dominique Polack pour avoir accepté de co-diriger cette thèse, pour ses innombrables conseils avisés et pour m'avoir formé à l'acoustique des salles. Son éclectisme est, par ailleurs, d'un enrichissement inouï.

Danièle Dubois du laboratoire Langages, Cognitions, Pratiques, Ergonomie a assuré la seconde co-direction de mes recherches. Sans elle, ce travail n'aurait pu voir le jour. Pour ses "électrochocs" conceptuels, sa patience à m'enseigner la psychologie cognitive - plusieurs "couches" auront été nécessaires, et d'autres le sont encore ! - et son enthousiasme, je veux vivement et affectueusement lui en savoir gré.

Je remercie chaleureusement messieurs Durand Begault, Jean Hardy et Jean-Pierre Péneau, pour avoir accepté d'être rapporteurs de mon travail, ainsi que Anne Bardot pour avoir bien voulu participer au jury de soutenance. L'intérêt qu'ils portent à ces recherches m'honore.

Les nombreux échanges avec les doctorants et stagiaires du L.A.M., ainsi que leur soutien inconditionnel auront été déterminants tout au long de cette thèse. En particulier, j'ai partagé toute la première année de thèse avec Valérie Maffiolo, sur la partie de validité écologique des échantillons sonores : nous avons réalisé ensemble la majorité des enregistrements sonores dans Paris, les entretiens et les tests d'écoute relatifs aux ambiances sonores urbaines représentatives de Paris. La dernière année, Chysanthie Nathanail m'a considérablement aidé grâce au soutien mutuel de rédaction que nous nous sommes dévoués. Je leur exprime ma profonde gratitude et n'oublie évidemment pas Claire qui a dû me supporter, Bertrand, Cécile, Myriam, Jean et Laurent.

Merci à Catherine Fourcin pour avoir géré avec brio tous mes documents administratifs.

Mes remerciements s'adressent également aux doctorants du PIR-Villes, ainsi qu'à Michèle Ansidéi et à Jean-Marc Rennes, chargés de mission au PIR-Villes, pour toutes les opportunités d'ouverture sur diverses thématiques de la ville. Ces échanges et expériences auront été particulièrement enrichissants.

Une attention toute particulière à Benoît Fabre qui m'a conseillé, soutenu et aidé dans les moments difficiles, et avec qui j'ai appris tant de notions sur la prise de son et l'écoute critique. Saint-Cloud !

Une mention "spéciale" à Frédérique Guyot pour ses conseils et ses corrections, ainsi qu'à mes parents, Anja et Zeno pour leurs remarques et leur soutien sans faille.

Merci enfin à tous les sujets qui ont accepté de participer aux tests d'écoute.

Au risque d'oublier certaines personnes (qu'elles ne m'en veuillent pas, par pitié !), je ne puis m'empêcher de manifester ici toute mon affection à Bolélé, Bim, Céline, Diane, Jean-Christophe, Marianne, Nat, Philippe, Régis et Véronique.



# Table des matières

<b>Avertissement</b>	3
<b>Avant-propos</b>	3
<b>Remerciements</b>	5
<b>Introduction</b>	11
1. Pourquoi étudier les signaux sonores d'avertissement ?	11
2. Comment étudier les signaux d'avertissement ?	12
3. Organisation de ce mémoire	12
<b>Chapitre 1 : Les signaux sonores, leur perception et leur reconnaissance</b>	15
1. Introduction	15
2. Représentations acoustiques et unités de mesure	16
2.1 Trois représentations à deux dimensions	16
2.2 Une représentation à trois dimensions : le sonagramme	17
2.3 Grandeurs physiques issues de la psychoacoustique	18
2.4 La localisation des sons	21
3. Psychoacoustique des signaux d'avertissement	22
3.1 Une approche behavioriste	22
3.2 Une approche par la détection de signal	22
3.3 Une approche par la notion d'urgence	24
3.4 Les temps de réponse	26
4. Reconnaissance et identification de formes sonores	28
4.1 Les apports de la Gestalttheorie	28
4.2 Événements sonores, reconnaissance et identification	30
4.3 La catégorisation prototypique	32
5. Vers des signaux qui font sens...	34
5.1 Les apports de la psychologie écologique	34
5.2 Des études centrées sur les sujets	35
5.3 Signaux, langage et communication	38
6. Discussion	43
6.1 Les approches physicalistes	43
6.2 L'homme "écoutant"	45
6.3 La cognition et la psycholinguistique	46
7. Conclusion	47
<b>Chapitre 2 : Approche méthodologique</b>	49
1. Introduction	49
2. Les échantillons sonores	50
2.1 Des échantillons sonores écologiquement valides	50
2.2 Importance du contexte d'écoute	51
3. Les procédures des situations d'écoute	54
3.1 La validité écologique par rapport aux sujets	54
3.2 Le choix des sujets	54
3.3 Tâches demandées aux sujets	55
3.4 Les réponses : la verbalisation	56
4. Le traitement des données	57
4.1 Respect des formes linguistiques produites par le sujet	57
4.2 Les organisations cognitives	57
4.3 L'acoustique en aval	57
5. Conclusion	58

<b>Chapitre 3 : Expériences préliminaires relatives à la validité écologique</b>	59
<b>1. Introduction</b>	59
<b>2. Choix des lieux d'écoute</b>	60
2.1 Méthode	60
2.2 Procédure et sujets	60
2.3 Résultats	61
2.4 Des ambiances à enregistrer	63
<b>3. Enregistrement des échantillons sonores</b>	64
3.1 Matériel utilisé	64
3.2 La prise de son	65
<b>4. La restitution des échantillons sonores</b>	68
4.1 Choix du mode de diffusion des échantillons sonores	68
4.2 Le choix des haut-parleurs : un test d'écoute	68
4.3 Sélection d'un local d'écoute	70
<b>5. Premier test : écartements entre microphones</b>	71
5.1 Méthode	71
5.2 Résultats	73
5.3 Défauts de l'expérience	74
<b>6. Second test : configurations de prise de son</b>	75
6.1 Méthode	75
6.2 Résultats	76
6.3 Interprétation	78
<b>7. Conclusion</b>	79
<b>Chapitre 4 : Seuils d'identification en contexte urbain et identification hors contexte</b>	81
<b>1. Introduction</b>	81
<b>2. Méthode</b>	82
2.1 Corpus sonore	82
2.2 Procédures et sujets	84
2.3 Audiogrammes	85
2.4 Une petite précaution...	85
<b>3. Résultats de l'épreuve verbale</b>	86
3.1 Désignations les plus fréquentes	86
3.2 Désignations les plus disponibles	87
3.3 Interprétation	88
<b>4. Résultats des seuils de perception</b>	89
4.1 Itérations et seuils	89
4.2 Analyse quantitative des résultats	91
4.3 Variation des seuils selon le contexte	93
<b>5. Résultats de la verbalisation des sujets</b>	95
5.1 Répertoire lexical disponible sur l'ensemble des données	95
5.2 Description des 15 signaux dans les différentes situations d'écoute	96
<b>6. Signaux univoques : les prototypes</b>	105
6.1 Signaux prototypes des désignations les plus fréquentes	105
6.2 Interprétation	110
<b>7. En résumé</b>	111
<b>8. Conclusion</b>	112

<b>Chapitre 5 : Identification en contexte</b>	113
1. Introduction	113
2. Méthode	114
2.1 Corpus sonore	114
2.2 Procédure et sujets	115
3. Résultats de l'identification verbale	116
3.1 Répertoire lexical disponible sur l'ensemble des données	116
3.2 Description des 15 signaux dans les différentes situations d'écoute	117
4. Prototypes issus de l'identification verbale	124
4.1 Signaux prototypes des désignations les plus fréquentes	124
4.2 Interprétation	129
5. Les signaux "semblables"	130
5.1 Résultats	130
5.2 Interprétation	132
6. Conclusion	133
<b>Chapitre 6 : Le sens et le son ; analyses acoustiques de propriétés pertinentes</b>	135
1. Introduction	135
2. Le sens issu des expériences...	136
2.1 Répertoire lexical disponible sur l'ensemble des données	136
2.2 Description des 15 signaux dans les différentes situations d'écoute	138
2.3 Signaux prototypes des désignations les plus fréquentes	139
3. ...et le son des expériences	145
3.1 Acoustique des contextes urbains	145
3.2 Acoustique des 15 signaux d'avertissement	148
3.3 Sémiotique associée aux signaux acoustiques	164
3.4 Propriétés acoustiques et signification des signaux	166
4. Conclusion	168
<b>Conclusion générale</b>	169
1. Les résultats	169
2. Les perspectives	170
<b>Références bibliographiques</b>	173
<b>Bibliographie complémentaire</b>	177
<b>Annexes</b>	179
Annexe 1 : les questionnaires	181
Annexe 2 : résultats bruts du test d'identification à niveau seuil et hors contexte (chapitre 4)	195
Annexe 3 : résultats bruts du test d'identification et de comparaison en contexte (chapitre 5)	217



# Introduction

Le monde ne peut être dit "humain"  
que dans la mesure où il signifie quelque chose  
Greimas (1970)

## 1. Pourquoi étudier les signaux sonores d'avertissement ?

Le développement urbain durant ces dernières décennies a provoqué une multiplication dramatique des signaux visuels et sonores dans les villes, encombrant ainsi le paysage urbain. Les citoyens sont aujourd'hui non seulement quotidiennement exposés à ces signaux codifiés et chargés de sens, mais sont également censés y répondre par une modification de leur conduite, ce qui présuppose la compréhension de ces signaux. Les publicitaires excellent d'ailleurs dans cet art de chercher à faire réagir le "grand public". Mais d'autres types de signaux doivent communiquer des informations plus importantes, parfois urgentes, voire vitales. Dans ces situations de danger, les signaux sonores prédominent car le danger peut provenir de n'importe quelle direction et doit être compris immédiatement par les personnes concernées, sans que ces dernières aient à se retourner pour interpréter visuellement le risque.

Or, il s'avère que de nombreuses situations de la vie courante ne permettent pas d'identifier et/ou de localiser rapidement des signaux d'avertissement. Le remède habituel consiste à rehausser le niveau sonore des signaux, uniquement dans un but de détection plus précoce, c'est-à-dire à une plus grande distance de l'individu percevant le signal (Eldred et Sharp, 1972). Toutefois, cette démarche ne garantit ni une meilleure identification de la source, ni une meilleure localisation de celle-ci. Si la détection d'un signal est évidemment nécessaire à la capacité d'un individu à réagir, elle n'est certainement pas suffisante. Discutant de l'utilité et de l'efficacité des signaux d'avertissement, Eldred et Sharp (1972) en arrivent à la conclusion que "la solution est clairement de trouver une meilleure approche [que celle de l'intensité sonore] pour signaler aux conducteurs de véhicules qu'un danger provenant d'un autre véhicule est imminent". Nous verrons que, pour être efficace, un signal doit être identifié par l'auditeur - ou tout au moins reconnu comme un avertisseur - et doit donc avoir une signification précise et univoque dans le contexte d'écoute donné. Ainsi, une solution au problème consisterait plutôt à déterminer des signaux *facilement identifiables* et par conséquent, chargés de sens, dont les caractéristiques acoustiques permettent une *localisation aisée* et qui ne soient *pas gênants* pour les personnes non concernées par l'avertissement.

Mais qu'est-ce qu'un signal d'avertissement pour les usagers de la ville ? Nous avons réalisé un questionnaire auquel 28 personnes (essentiellement des étudiants en mécanique) ont répondu. 48% des réponses concernent des signaux auditifs, 32% des signaux visuels et 20% les autres sens. Par ailleurs 80% des sujets considèrent un signal d'avertissement comme un moyen de prévenir ou d'avertir d'un danger. Les autres le décrivent plutôt comme un moyen d'informer, d'attirer l'attention ou de communiquer. Ainsi, *un signal d'avertissement est-il avant tout considéré comme sonore et prévenant d'une situation dangereuse*. Si l'on demande maintenant aux sujets de citer trois signaux d'avertissement, 59% des signaux cités sont

sonores ("klaxon", "sirène"...), 37% visuels ("feu tricolore de signalisation", "clignotant" de voiture...) et 4% sont olfactifs ("odeur de gaz"). Les exemples produits par les sujets confirment donc la prédominance de la dimension sonore des signaux d'avertissement.

Le problème auquel nous sommes confrontés est donc de caractériser les signaux sonores d'avertissement du point de vue de leurs propriétés acoustiques afin de pouvoir les modifier, en prenant comme point de départ la signification qui leur est attribuée par les usagers de la ville. De toute évidence, la problématique repose sur différentes disciplines scientifiques, dont notamment l'acoustique et la sémiotique, théorie qui étudie les signes au sein de la vie sociale.

## **2. Comment étudier les signaux d'avertissement ?**

De nombreux chercheurs, issus de divers domaines de recherche, se sont penchés sur la question de la perception des sons. Outre les travaux en physiologie de l'oreille, qui traitent du fonctionnement du système auditif, différentes approches ont été développées en acoustique, psychoacoustique, sociologie, psychologie... Toutefois, très peu d'études ont été réalisées sur la perception des signaux d'avertissement en contexte urbain. Nous intéressant avant tout à l'individu et à sa perception du monde sonore, il nous est apparu essentiel de comprendre comment chercheurs et protagonistes du milieu sonore, tous domaines confondus, appréhendent les signaux sonores, c'est-à-dire les sons chargés de sens, censés susciter l'attention d'un auditeur. Nous étendrons par conséquent notre champ d'investigation à la perception des sons, des signaux sonores et, du fait de leur nature indissociable de notre objet d'étude, à leur identification et à leur signification. Ainsi, les différentes approches d'élaboration de paramètres acoustiques liés à des évaluations subjectives telles que la reconnaissance, l'identification ou l'appréciation qualitative d'objets sonores seront décrites dans ce mémoire. Nous nous positionnerons alors par rapport à ces démarches et justifierons celle adoptée dans notre recherche, inspirée de la linguistique et de la psychologie cognitive. Nous proposerons ensuite une méthodologie basée sur ces considérations théoriques et l'appliquerons à plusieurs tests perceptifs. La méthodologie adoptée doit minimiser les restrictions ou contraintes qui pourraient être introduites par la méthode d'analyse, et clairement identifier celles qui sont inévitables. Sans quoi, il ne sera pas possible d'extrapoler les résultats à des situations réelles de l'environnement urbain.

## **3. Organisation de ce mémoire**

Dans le premier chapitre, nous ferons le point sur l'état des connaissances concernant la perception des signaux sonores, mais également sur les différentes approches d'étude et d'analyse du son et de sa perception. Il s'agit là avant tout d'éclairer le lecteur sur les différents fondements théoriques appliqués au sonore, depuis la physique jusqu'à la psychologie, son "complémentaire".

Partant de ce constat, nous développerons au chapitre 2 notre propre démarche, résolument pluridisciplinaire, qui consiste à jeter des ponts entre l'acoustique, la psychoacoustique, la psychologie cognitive et la linguistique. Cette approche, centrée sur le

sujet, la représentation mentale du monde qu'il en a, et la façon dont il exprime cette expérience vécue, représente selon nous le seul moyen de mettre en évidence les paramètres acoustiques liés à la signification du milieu sonore.

Les trois chapitres suivants sont consacrés à la mise en application de notre méthodologie. Les expériences décrites concernent la validité écologique des échantillons sonores à travers l'écoute des sujets (chapitre 3), les seuils de perception de quinze signaux d'avertissement "usuels" pour deux fonds sonores urbains représentatifs de la ville (chapitre 4), ainsi que leur identification au niveau seuil ou hors contexte (chapitre 4) et à deux niveaux d'émergence du bruit de fond (chapitre 5).

Enfin, nous tenterons de mettre en relation les résultats verbaux issus de la sémiotique des signaux d'avertissement en contexte urbain avec les propriétés acoustiques de ces signaux (chapitre 6), l'objectif étant de déterminer des propriétés acoustiques signifiantes des avertisseurs.



# Chapitre 1 :

## Les signaux sonores, leur perception et leur reconnaissance

the mind is  
not an object  
but a picture  
of the mind  
is an object  
& pictures  
in the  
mind  
are objects  
*Robert Lax (1993)*

### 1. Introduction

Les signaux acoustiques ont pour effet d'attirer l'attention des individus. Lorsqu'ils sont délibérément associés à la notion d'avertissement, ils s'accompagnent d'une signification et deviennent des signaux sonores qui ont pour fonction première d'avertir : les signaux d'avertissement. Mais avertir de quoi ? Selon le domaine d'application, ils peuvent informer de l'état d'un équipement ou d'un processus dynamique, du début ou de la fin d'événements particuliers (Ballas, 1993) ; ou bien, ils constituent une aide aux malvoyants pour traverser les rues en toute sécurité (Szeto *et al.*, 1991a) ; ou encore, ils doivent permettre de réduire les accidents de travail et d'améliorer les performances (Edworthy *et al.*, 1991). Dans un environnement urbain, le but recherché est sensiblement le même, puisqu'il s'agit de permettre aux usagers de la ville de réagir rapidement et de façon sécuritaire à des situations de danger imminent. Pour ce faire, un signal sonore doit évidemment être détecté dans une ambiance donnée, mais aussi être identifié et faire sens pour les personnes percevant ce signal.

Cette recherche bibliographique s'organise selon une démarche progressive, en partant de l'acoustique physique et la psychoacoustique, s'orientant ensuite vers une perception plus "naturelle", initiée par les théories de la psychologie de la forme et de la psychologie écologique, pour aboutir à des approches centrées sur l'individu, son interprétation du monde perçu et sa manière de communiquer. Il s'agit alors d'accéder aux *représentations mentales* des individus, représentations en mémoire d'événements signifiants. Ainsi, nous aborderons le signal sonore en tant que phénomène acoustique, en termes d'effets psychophysiques sur l'individu exposé, mais également par rapport à une description événementielle ou encore psychologique. Une fois ces différentes recherches décrites, nous discuterons des méthodes et des résultats en nous efforçant de soulever les points importants, utiles à notre démarche.

## 2. Représentations acoustiques et unités de mesure

Les théories de la physique s'appuient sur un formalisme mathématique, une signification physique et une validité expérimentale. Elles reposent donc avant tout sur des représentations des états et des propriétés observables des systèmes étudiés. L'acoustique physique de l'audible cherche à quantifier des objets et des phénomènes observables du monde sonore, c'est-à-dire à établir des *représentations* visuelles du signal acoustique selon ses dimensions fondamentales et à déterminer les *grandeurs mesurables* liées aux phénomènes observables de la perception humaine. Dans cette approche psychoacousticienne (cf. § 2.3), la notion de perception est très proche de celle de sensation, au sens où l'on étudie des grandeurs "ressenties" par les individus.

### 2.1 Trois représentations à deux dimensions

L'approche physicaliste classique du phénomène sonore consiste à rechercher les propriétés acoustiques des signaux à partir de grandeurs physiques mesurables. Pour ce faire, le signal est décomposé selon trois axes fondamentaux : l'intensité, la fréquence et le temps. Il en découle trois formes de représentations du signal selon deux dimensions : le plan *dynamique* désigne l'intensité en fonction du temps (cf. Figure 1-1), le plan *spectral* correspond à l'intensité en fonction de la fréquence (cf. Figure 1-2), et le plan *mélodique* visualise le temps en fonction de la fréquence (Leipp, 1971).

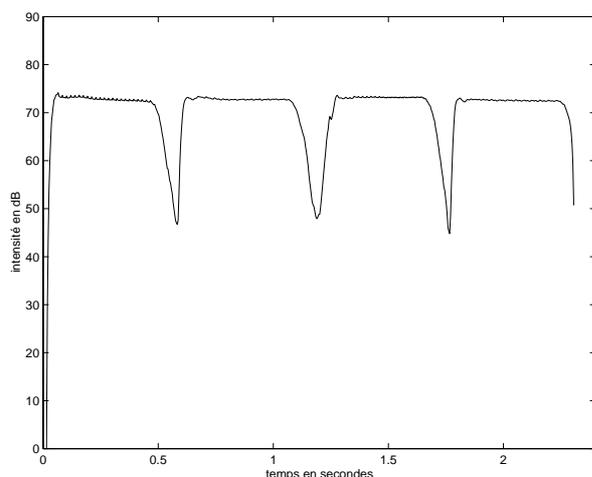


Figure 1-1 : Enveloppe dynamique (intensité-temps) de deux périodes d'une sirène de police.

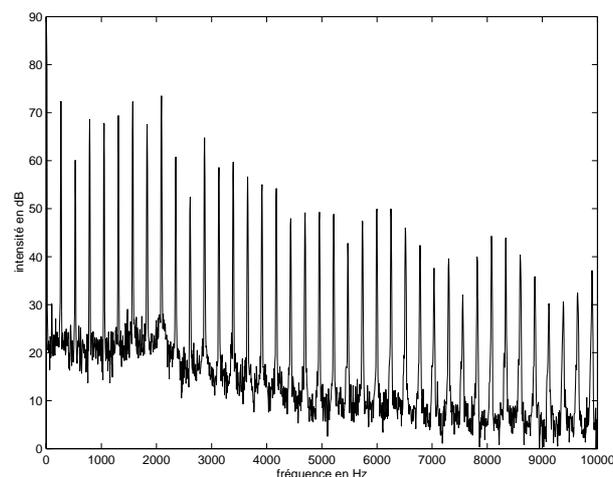


Figure 1-2 : Représentation spectrale (intensité-fréquence) de la première note de la même sirène.

En acoustique physique, un signal sera considéré comme "efficace" si certaines composantes du signal émergent du bruit de fond, en l'occurrence - concernant plus spécifiquement notre propos - un fond sonore urbain. C'est le rapport signal/bruit. La représentation spectrale permet de visualiser l'émergence de certains pics à des fréquences données et d'envisager de changer la hauteur d'un signal et de déplacer les "pics" qui le constituent dans des zones de fréquences pour lesquelles l'émergence par rapport au bruit de fond est maximale. La représentation dynamique permet, quant à elle, de visualiser l'émergence d'un signal en fonction du temps, et notamment lorsque la source sonore se déplace par rapport à l'instrument de mesure. Enfin, la représentation "mélodique" permet de suivre l'évolution temporelle des sons et notamment les changements de hauteur, les transitoires d'attaque, ainsi que les extinctions plus ou moins rapides des différentes fréquences qui composent les sons. Elle constitue le canevas de la représentation sonographique présentée ci-après.

## 2.2 Une représentation à trois dimensions : le sonagramme

Une autre forme de représentation visuelle du son a été développée au milieu du XX<sup>ème</sup> siècle, à l'origine pour étudier la parole. Elle permet de visualiser dans le plan mélodique, outre la fréquence et le temps, l'intensité - et donc les trois dimensions fondamentales du signal physique. Si l'intensité ne peut être décrite aussi précisément que dans les plans dynamique et spectral, la noirceur et l'épaisseur du trait, qui la caractérisent dans cette représentation, permettent de distinguer de faibles variations de sensation subjective d'intensité. Il s'agit de la *représentation sonographique* (cf.

Figure 1-3) dont Leipp (1965) a illustré les possibilités, mais aussi les limites. En effet, le grand avantage des sonagrammes est de permettre d'ajuster la sensibilité de contraste de manière à être proche de celle d'une perception humaine, et ainsi, de visualiser les formes acoustiques, par analogie avec les formes visuelles décrites par la Gestalttheorie (cf. § 4.1.1). Leipp (1971, p.90) insistait d'ailleurs sur le fait que "la précision des documents est comparable à celle de l'audition".

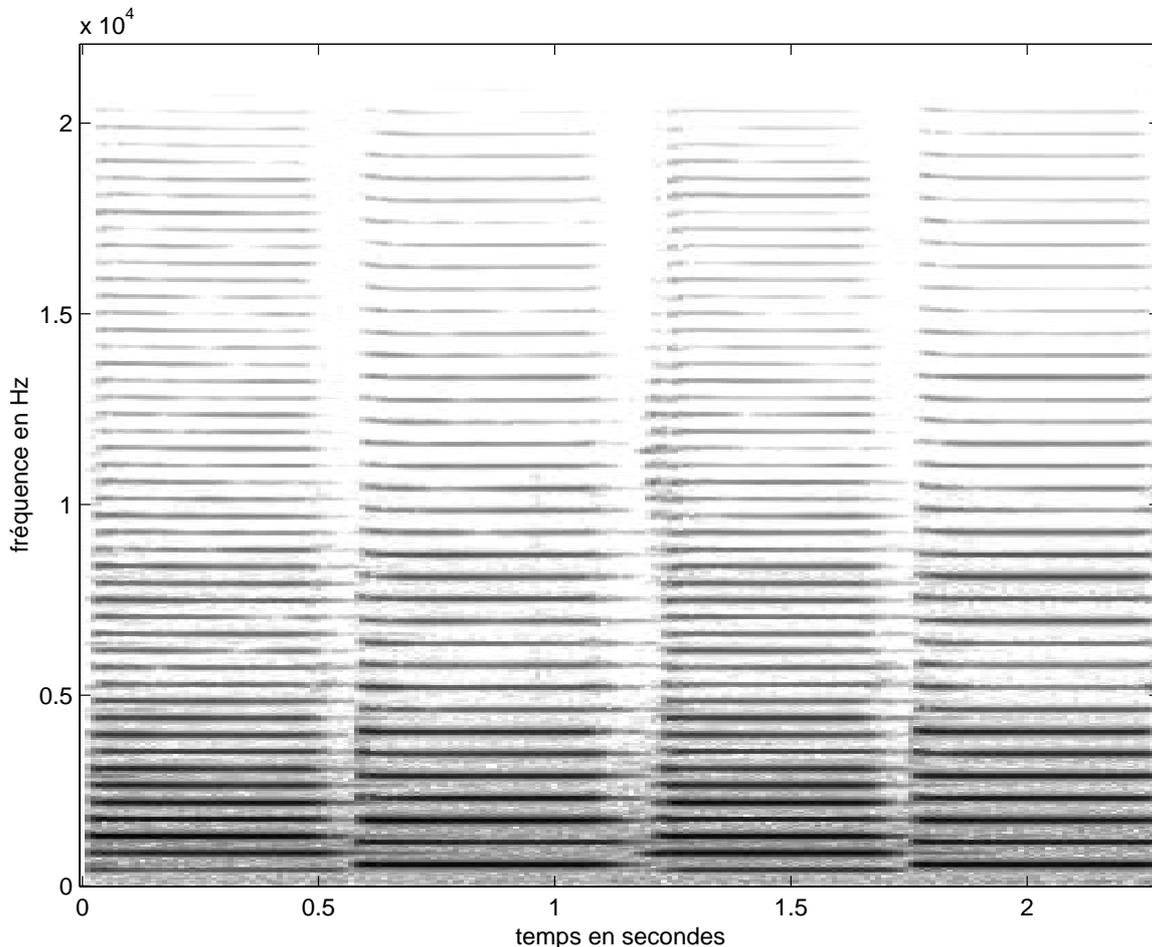


Figure 1-3 : représentation sonographique de la même sirène de police.

Les limitations de la représentation sonographique se manifestent lorsque différents événements sonores proches - d'un point de vue dynamique et spectral - se produisent de manière simultanée : les éléments visuels se superposent et se mélangent dans le sonagramme. Dans ces circonstances, l'audition humaine est bien plus à même d'analyser des situations complexes que n'importe quel instrument de mesure, notamment pour des raisons de localisation (cf. § 2.4) grâce à une écoute binaurale.

## 2.3 Grandeurs physiques issues de la psychoacoustique

L'origine d'une approche physicaliste de la perception remonte au XIX<sup>ème</sup> siècle, lorsque Weber et Fechner se sont interrogés sur l'existence de sensations obéissant à des règles universelles, partagées par tous les individus. Ainsi naquit la *psychophysique* qui traite "des relations entre les stimulations physiques et les sensations comme états psychiques correspondants" (Botte *et al.*, 1989). La *psychoacoustique*, branche de la psychophysique, étudie les rapports entre les paramètres d'une vibration mécanique aérienne et la sensation auditive engendrée. Les premiers travaux ont permis d'établir le décibel que nous venons de définir ; les suivants se sont orientés, dans la seconde moitié du XX<sup>ème</sup> siècle et jusqu'à nos jours, vers de nouvelles grandeurs afin de décrire au mieux les sensations subjectives.

### 2.3.1 Le décibel linéaire : dB(SPL)

L'unité de mesure intensimétrique du niveau sonore est le décibel, unité sans dimension puisqu'il s'agit d'un rapport de pressions, ou d'intensités. Le niveau de référence, 0 dB, est donné par le seuil d'audition d'un son pur de 1 kHz, ce qui correspond à une pression acoustique de 20  $\mu$ Pa. Ce seuil d'audition a été normalisé de différentes manières, à partir d'études menées sur des sujets humains, en écoute monaurale ou binaurale, sur casque ou haut-parleur (Botte, 1989, p.17-20). La Figure 1-4 représente le seuil absolu d'audition en fonction de la fréquence et qui correspond à une moyenne effectuée sur un échantillon de population de plus de 100 sujets âgés de moins de 25 ans (Zwicker, 1982, p.33).

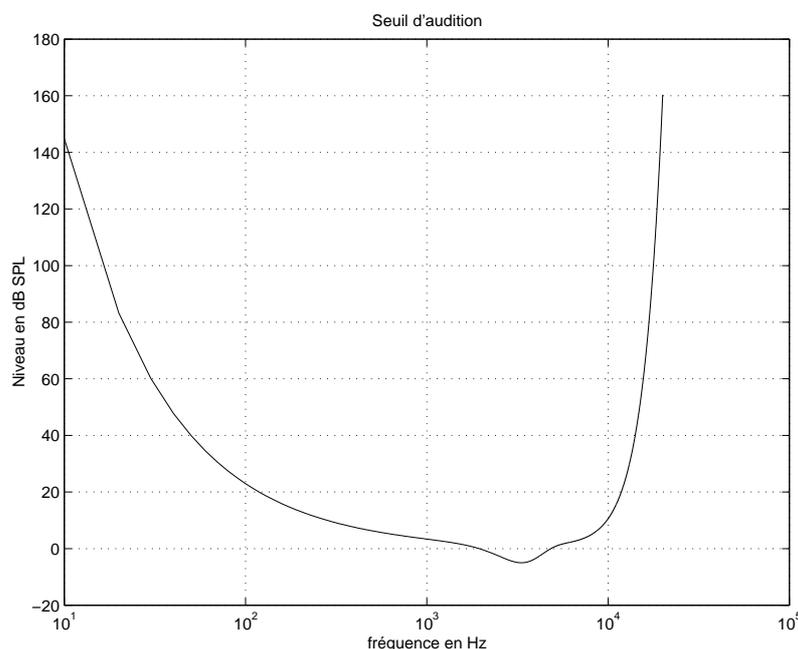


Figure 1-4 : seuil absolu d'audition (50% des sujets), d'après Zwicker (1982).

Ainsi, les grandeurs fondamentales de l'acoustique physique sont-elles déjà définies en fonction de la perception humaine ! L'échelle des niveaux sonores est logarithmique, suite à la loi de Weber-Fechner - qui exprime le fait que la sensation croît comme le logarithme de l'excitation - et ce, de manière à couvrir l'ensemble de la zone d'audition, du seuil d'audition au seuil de douleur, sur une échelle comprise entre 0 et 120 dB SPL (Sound Pressure Level) environ.

Afin d'avoir une idée plus précise des caractéristiques physiques d'un son, le niveau mesuré peut être "découpé" en une série de niveaux sonores sur tout le spectre audible. Ainsi, on analyse généralement un spectre par bandes d'octaves, ou mieux, par bandes de tiers d'octave (Figure 1-5). En effet, le système auditif humain fonctionne approximativement comme une série de filtres mis en parallèle et dont les largeurs de bandes - les *bandes critiques* de l'oreille - sont proches de celles des bandes de tiers d'octave. Néanmoins, toutes les largeurs de bandes sont possibles : demi-octave, douzième d'octave... Enfin, une analyse par bandes fines, encore plus précise, permet de visualiser des pics à certaines fréquences bien déterminées.

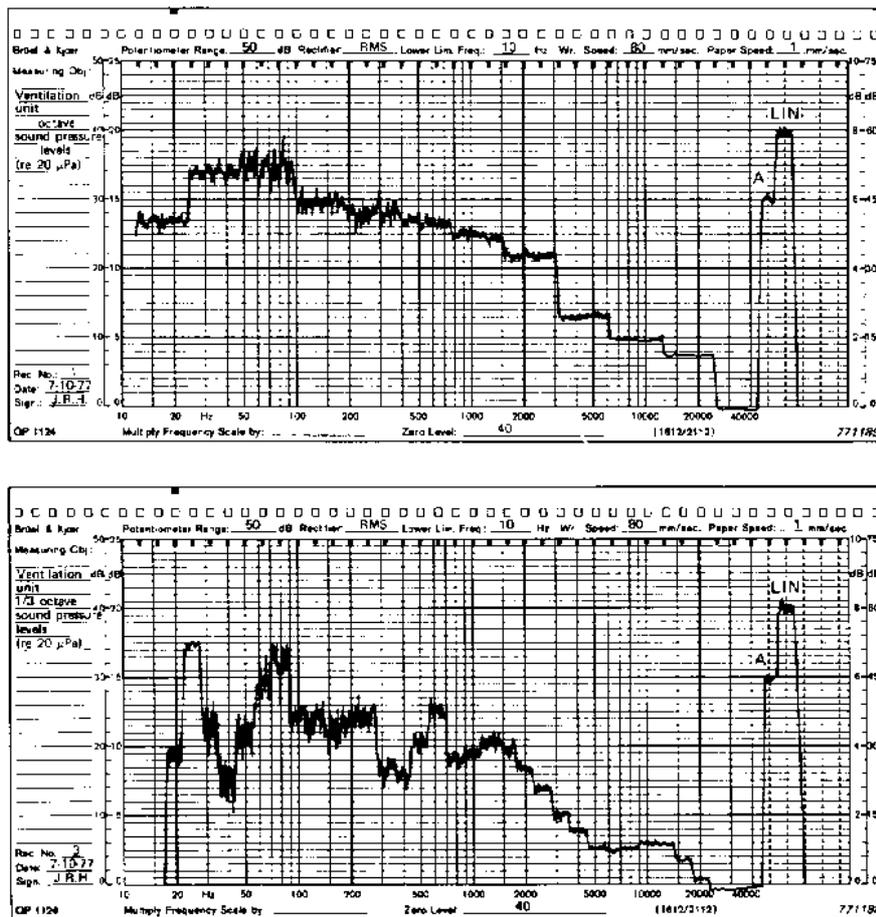


Figure 1-5 : analyse spectrale d'un bruit de ventilation par bandes d'octave et par bandes de tiers d'octave (Hassall et Zaveri, 1988). A droite des courbes sont donnés les niveaux globaux en dBA et en dBSPL (linéaire).

### 2.3.2 Les courbes de pondérations fréquentielles

Conscients de l'inadéquation entre les échelles de mesures physiques - le décibel linéaire, mesuré en dBSPL - et le sentiment subjectif d'intensité, les psychoacousticiens ont tenté de développer des unités intensimétriques proches de cette intensité subjective. Pour ce faire, ils ont réalisé des *lignes isosoniques*, courbes de niveaux sonores physiques de sons purs à différentes fréquences, qui procurent la même sensation d'intensité qu'un son pur à 1 kHz. Ils ont ensuite considéré trois de ces lignes pour pondérer le dBSPL, celles correspondant au : dBA, pour un son pur valant 40 dBSPL à 1 kHz (son de faible intensité) ; dBB, pour un son pur de 70 dBSPL à 1 kHz (son moyennement fort) : dBC, pour un son pur de 100 dBSPL à

1 kHz (son intense). Une quatrième courbe de pondération, nommée D et amplifiant fortement les fréquences comprises entre 2 et 10 kHz, est utilisée dans la mesure des bruits d'avions.

Ces pondérations (cf. Figure 1-6), et d'autres encore (Kryter, 1970), ont été normalisées et sont généralement intégrées dans les instruments de mesure, les sonomètres. Ces derniers sont homologués selon quatre classes en fonction de leur précision : classe 0 ( $\pm 0,4$  dB), classe 1 ( $\pm 0,7$  dB), classe 2 ( $\pm 1,0$  dB), classe 3 ( $\pm 1,5$  dB) (Botte et Chocholle, 1984).

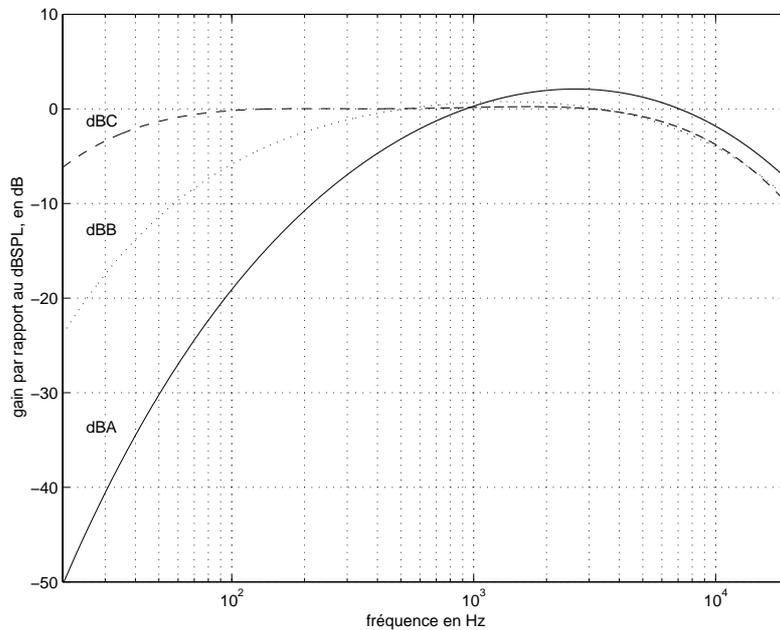


Figure 1-6 : courbes de pondération A, B et C du décibel linéaire (dB SPL).

Aujourd'hui, essentiellement le dBA est utilisé pour mesurer les bruits, notamment du fait de sa bonne prédictibilité concernant les conséquences de fatigue auditive suite à une exposition au bruit (Botte, 1989). Mentionnons enfin un autre indicateur très répandu : le LAeq, dont l'unité est le dBAeq, niveau mesuré avec une pondération A et moyenné sur la durée des échantillons, de quelques secondes à plusieurs heures selon les besoins des études (mesure de niveaux du trafic routier, de locaux industriels bruyants...).

### 2.3.3 L'intensité subjective : la sonie

L'intensité ressentie par l'individu est appelée *sonie*. Si elle est bien évidemment fonction du niveau physique d'un signal acoustique, elle ne dépend pas moins du spectre du signal, de sa durée et d'un masquage éventuel par un bruit de fond (Canévet, 1994). La sonie a fait l'objet de nombreuses études afin de pouvoir la calculer pour tous types de bruits. Le principe du calcul consiste à utiliser certains résultats de la psychoacoustique en décomposant le signal sur des bandes de fréquences de largeurs correspondant aux caractéristiques de l'audition - bandes proches du tiers d'octave, appelées bandes critiques. On réalise ainsi une analyse fréquentielle comparable à celle de l'oreille, quoique notablement simplifiée. A partir de cette décomposition, il s'agit de sommer les contributions des différentes bandes critiques à l'intensité perçue et d'appliquer un certain nombre de facteurs correctifs, notamment dûs au masquage fréquentiel, pour aboutir à la valeur de la sonie (Zwicker et Feldtkeller, 1981).

### 2.3.4 Critères psychoacoustiques de "qualité" sonore

D'autres critères psychoacoustiques usuels ont été développés, corrélés au sens de l'analyse multidimensionnelle à la sensation sonore. Ainsi, la *modulation* correspond à des variations sinusoïdales de l'amplitude d'un son (Zwicker et Feldtkeller, 1981), et la *raucité*, ou *rugosité*, correspond à une plage fréquentielle spécifique de modulation, allant de 20 à 250 Hz environ (Zwicker, 1982, p.146). La *hauteur (tonale)* correspond à la fréquence fondamentale d'un son (Botte *et al.*, 1989), et l'*acuité* est définie comme la position du centre de gravité de la distribution spectrale de sonie (Bismarck, 1974a et 1974b), définition identique à celle de la *brillance* pour une distribution spectrale d'énergie. Ces indicateurs, plus spécifiquement associés à la dimension fréquentielle du signal acoustique, sont notamment utilisés pour étudier un autre type d'objets psychologiques, basés sur un jugement de valeur, tels que la gêne ou le confort acoustique dans les véhicules. Souvent corrélés entre eux, ces indicateurs sont parfois même regroupés dans de nouveaux critères de valeur, telle que l'*euphonie*, développée par Aures (1985a et 1985b) qui estime que "la sensation du caractère agréable d'un son est sous l'influence des perceptions élémentaires de rugosité, acuité, sonorité et sonie".

Les travaux en psychoacoustique ont donc été sujets à un transfert progressif d'une analyse de traits perceptifs vers une recherche d'indicateurs de qualité, objets psychologiques très différents, qui nécessitent une prise en compte des stratégies de réponse des individus et, par conséquent, une modification de l'approche méthodologique.

## 2.4 La localisation des sons

Botte (1989) considère la sonie, la hauteur et la *localisation* comme les attributs élémentaires de la perception auditive - perception au sens de "sensation". La localisation a toutefois un statut particulier, car elle ne donne pas lieu à de nouveaux indicateurs liés à la qualité d'un son ou qui lui sont corrélés. Nous venons de voir que ce n'est pas le cas de la sonie et de la hauteur. Canévet (1989) donne la définition suivante de la localisation : "localiser une source [sonore], c'est d'abord identifier son azimuth et sa hauteur, donc sa direction, puis déterminer la distance à laquelle elle se trouve dans cette direction". Cette définition présuppose évidemment qu'une source sonore et le son qui lui est associé coïncident dans l'espace perceptif. La psychoacoustique a permis de relier avec précision ces différents paramètres au concept de localisation.

Theile (1980) étudie la perception de l'orientation et de la distance de sources fantômes en écoute stéréophonique, la source fantôme étant une source sonore virtuelle perçue comme un événement auditif unique, par association de plusieurs *événements sonores* (cf. § 4.2.1) localisés physiquement à d'autres endroits. Pour cela, il définit la localisation, comme la loi d'association entre le lieu d'un événement auditif extérieur à la tête d'un auditeur et des caractéristiques déterminées d'un ou plusieurs événement(s) sonore(s).

Ainsi, la localisation d'un son peut être comparée, en termes de fonction première de la perception auditive de l'individu, à la *reconnaissance* du son, alors que la sonie et la hauteur s'inscrivent plutôt dans une démarche de *qualification*. La localisation est pertinente pour l'individu dans la mesure où elle guide l'action, s'insère dans la narrativité d'une scène sonore (cf. la notion de validité écologique § 5.1).

### 3. Psychoacoustique des signaux d'avertissement

Les unités de mesure issues de l'acoustique de l'audible et de la psychoacoustique sont utilisées dans de nombreuses recherches appliquées et notamment, pour décrire les *comportements* des individus qui se trouvent face à des situations de danger. Ainsi, certaines études concernent la détection des signaux dans un contexte sonore donné, tandis que d'autres abordent la notion d'urgence liée à des signaux acoustiques.

#### 3.1 Une approche behavioriste

L'approche psychoacoustique (cf. § 2.3), consiste à étudier les *effets* que produit un signal sonore, émis par une source physique, sur les auditeurs. Elle s'inscrit dans le courant de pensée du *behaviorisme* qui postule qu'on ne peut pas accéder aux états mentaux des individus et qu'il est nécessaire de décomposer les conduites humaines en unités plus élémentaires (Weil-Barais, 1993). Cette *psychologie du comportement*, notamment défendue par Watson, se fonde sur des modèles associationnistes qui, à l'instar de la physique atomique du XIX<sup>ème</sup> siècle, construisent les objets ou faits perceptifs - ou l'univers physique - à partir de "briques élémentaires" : les sensations élémentaires - ou les atomes. Elle cherche à mettre en relation des caractéristiques physiques de stimulation parvenant à l'individu avec les caractéristiques observables des comportements, sans hypothèses sur le fonctionnement interne, inobservable, de l'individu.

En psychoacoustique, il s'agit de maîtriser un maximum de paramètres acoustiques, afin de pouvoir déterminer l'influence de chacun d'eux sur la perception des sujets. Le signal sonore consiste en un ensemble de "briques" acoustiques, comme peut l'illustrer la synthèse additive des signaux. Il en résulte un son artificiel, association de fréquences pures, mais qui peut proprement modéliser une sirène d'urgence ou un klaxon de véhicule. Le contexte, le "bruit" en physique, est désigné sous forme de bruit blanc ou rose : spectres "plats", couvrant toute la zone d'audition et stationnaires. Là encore, ces bruits, mathématiquement bien définis, ont l'avantage d'être très facilement décrits par des paramètres physiques. Ils demeurent néanmoins très loin des ambiances fluctuantes auxquelles nous sommes habitués : circulation urbaine, parcs, marchés... L'utilisation principale de ces bruits "artificiels" concerne l'étude des phénomènes de masquages fréquentiel et temporel des signaux par le bruit de fond (Zwicker, 1982). Les études considérant un bruit de fond réel portent, quant à elles, essentiellement sur les niveaux seuils de détection des signaux par rapport au bruit, ce dernier étant rendu le plus stable possible afin de faciliter les mesures acoustiques.

#### 3.2 Une approche par la détection de signal

##### 3.2.1 Signaux sonores pour malvoyants

Szeto, Valerio et Novak (1991) ont étudié des signaux sonores d'avertissement pour malvoyants, au niveau des passages piétons, afin de déterminer leur contribution à rendre la traversée plus sûre. L'étude porte sur un produit commercialisé, particulièrement répandu aux Etats-Unis et au Canada, et qui consiste en un boîtier fixé en hauteur, à côté du signal visuel. Il émet, uniquement lorsque le feu est vert, deux sons différents selon la direction de traversée, l'un correspondant à un "coucou" électronique de 600 ms, répété toutes les 1,5 s, à deux

fréquences fondamentales de 1170 Hz et 950 Hz (environ un intervalle de tierce majeure) et quatre harmoniques à dominantes impaires ; l'autre à un "gazouilli" électronique de 140 ms, répété toutes les 1 s, variant continûment de 2600 Hz à 1500 Hz et de même structure harmonique (Szeto *et al.*, 1991b). Afin de les tester dans leur contexte, ces signaux ont été mixés avec un bruit de circulation enregistré sur les lieux et mis en boucle, puis ont été restitués sur casque d'écoute dans une pièce assourdie. Dix-neuf sujets, ayant tous une audition normale et répartis selon trois groupes ont effectué le test : sept sujets âgés de 22 à 35 ans et sept sujets âgés de 61 à 78 ans, tous ayant une vue normale, et cinq sujets aveugles âgés de 62 à 84 ans.

Les résultats indiquent que la détection correcte à 100% du signal "bruité" survient à partir d'un rapport signal/bruit de - 15 dBSPL, mais devient significative (> 50%) pour les deux signaux dès - 25 dBSPL (Szeto *et al.*, 1991c). Par ailleurs, le groupe des sujets âgés avec une vue normale sont moins performants - d'environ 5 dB - que les jeunes sujets et les sujets âgés et aveugles, qui ont un sens auditif plus aigu.

A partir de considérations physiologiques, les auteurs (1991c) proposent ensuite un signal "optimal" pour piétons, riche en harmoniques, dont la fréquence fondamentale varierait continûment entre 2 et 3 kHz sur une durée de 500 ms. Ceci, afin d'éviter les effets de masquage basses fréquences de la circulation ainsi que les zones de pertes auditives à hautes fréquences, souvent présentes chez les personnes âgées, qui représentent la moitié des malvoyants. Il semblerait par ailleurs qu'une durée inférieure à 500 ms réduit fortement la détection de signaux bruités.

### 3.2.2 Détection "filtrée" des motocyclistes...

Une autre étude, effectuée par Binnington, McCombe et Harris (1993) et concernant les motocyclistes, avait pour objectif premier d'analyser l'effet du port de protections auditives sur la détection de signaux d'avertissement. Après avoir mesuré (en dBA) et enregistré les séquences sonores pour différentes vitesses de déplacement, à l'aide d'un microphone placé dans la conque d'une oreille, les auteurs ont réalisé un test d'écoute en mixant ces bruits de fond avec quatre types de signaux d'avertissement : un klaxon de voiture, une sirène à deux tons, un signal de passage pour piétons ("pelican crossing") et un bruit de circulation très rapide. Dix-neuf sujets, âgés de 16 à 28 ans et ayant une audition normale, écoutèrent, en chambre assourdie et sur un haut-parleur, les séquences sonores correspondant aux vitesses et conditions de circulation (sans casque, avec casque, avec casque et protections auditives). La consigne était de déterminer s'ils percevaient ou non les différents signaux d'avertissement qu'ils devaient ensuite identifier dans un rapport signal/bruit de 0 dB.

Les résultats montrent qu'avec ou sans les protections auditives, le Klaxon, le passage piéton et la sirène ont été correctement identifiés à 99%, tandis que la circulation rapide l'a été à 73% et 80% respectivement. Par ailleurs, si les niveaux minimum de détection croissent avec la vitesse, ils deviennent inférieurs, à partir de 50 km/h, avec le casque et les protections auditives qu'avec le casque seul. Sans casque, les niveaux minimum de détection deviennent rapidement très élevés, dès 20 km/h, du fait du bruit aérodynamique sur les oreilles. Une analyse par tiers d'octave permet aux auteurs de conclure que les niveaux minimum de détection les plus bas correspondent aux signaux dont le centre de gravité spectral - la brillance - est le plus éloigné du centre de gravité du bruit de fond (environ 500 Hz). Il s'agit du signal de passage piéton (environ 3000 Hz) qui est perçu à l'arrêt et sans casque à 39 dBA pour un bruit de fond de 76 dBA (rapport signal/bruit de - 37 dB), tandis que la sirène et le Klaxon ne sont perçus qu'aux environs de 53 dBA (rapport signal/bruit de - 23 dB) et la circulation rapide vers 62 dBA (rapport signal/bruit de - 14 dB).

### 3.2.3 ... et des automobilistes

Citons enfin une étude réalisée par Ebata et Baba (1997) qui concerne la détection d'une sirène d'ambulance par des automobilistes durant la conduite. L'expérience consistait à simuler un bruit intérieur de voiture, correspondant à une vitesse de 60 km/h, et plusieurs extraits musicaux, à différents niveaux d'écoute, dans une voiture à l'arrêt. Un haut-parleur placé derrière la voiture diffusait le son d'une sirène à plusieurs niveaux sonores. Quatre étudiants ont participé au test, ayant pour consigne d'indiquer l'instant où ils percevaient la sirène. Les niveaux ont tous été mesurés en LAeq, et notamment les différents extraits musicaux. Malheureusement, la description des conditions expérimentales n'est pas suffisante pour nous permettre de comparer ces résultats avec ceux des deux précédentes études.

## 3.3 Une approche par la notion d'urgence

Un autre type d'étude a été mené par Edworthy, Loxley et Dennis (1991) : la recherche avait pour but d'améliorer le design des signaux d'avertissement sonore en cherchant les relations entre les paramètres acoustiques des signaux physiques et l'estimation de la notion d'urgence qui peut être associée à ces signaux. Il s'agit donc, là encore, d'une étude en rapport avec un jugement de valeur - l'urgence - et non avec la sensation issue d'un stimulus sonore, comme le suppose à l'origine la psychoacoustique. Reprenant une procédure d'élaboration de signaux d'avertissement proposée par Patterson (1982), Edworthy *et al.* ont construit des signaux physiques à partir de paramètres acoustiques définis, dont ils ont ensuite étudié l'influence sur le sentiment d'urgence produit. Ces signaux doivent être constitués, selon Patterson, d'un *train d'onde* d'une durée de 100 à 150 ms, d'une fréquence fondamentale, comprise entre 150 et 1000 Hz, et de plusieurs harmoniques, notamment entre 1 et 4 kHz. Ce train d'onde doit ensuite être répété au moins cinq fois selon des schémas temporels distincts, c'est-à-dire en modifiant les intervalles de temps entre chaque train d'onde (moins de 150 ms pour les signaux "urgents" et plus de 300 ms pour les signaux "non-urgents") afin d'éviter les confusions entre différents signaux. Il en résulte une *salve* sonore d'environ 2 s.

Edworthy *et al.* ont construit un train d'onde "type", de 200 ms, dont 20 ms d'attaque et 20 ms d'extinction, ayant une fréquence fondamentale à 200 Hz et quatorze harmoniques, tous de même amplitude (cf. Figure 1-7).

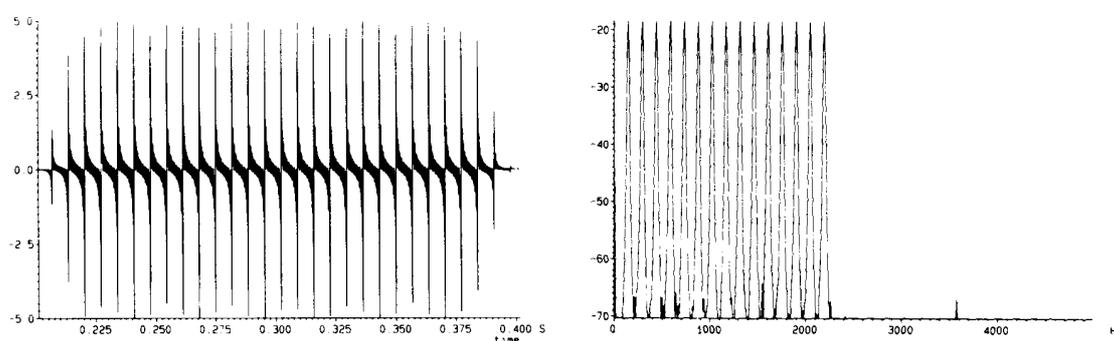


Figure 1-7 : profils temporel et spectral d'un train d'onde type (Edworthy et al., 1991).

Ce train d'onde a ensuite été modifié selon différents paramètres acoustiques : hauteur du son (150, 200, 350 ou 530 Hz), enveloppe dynamique (croissante, constante ou décroissante), régularité des harmoniques (régulière, 10% ou 50% d'irrégularité sur les harmoniques pairs, ou aléatoire entre harmoniques adjacents), ou encore retard des harmoniques de rang élevé, à partir

du rang 7 (100 ms ou aucun). Notons que les stimuli ont tous une fréquence de coupure à 4 kHz, ce qui signifie que le son de hauteur 350 Hz n'aura que dix harmoniques et celui de 530 Hz, six. Chaque stimuli consiste en une salve de six trains d'onde successifs, identiques, et sans pause entre chaque train d'onde (cf. Figure 1-8).

Deux séries d'expériences ont été réalisées, avec des sujets âgés de 18 à 40 ans, tous étudiants ou enseignants en psychologie. La première série de deux expériences consistait à tester les paramètres acoustiques des trains d'onde. Seize sujets, puis dix-huit, devaient effectuer d'une part, une tâche de comparaison, en sélectionnant le signal jugé le plus urgent parmi quatre trains d'onde, puis parmi les trois restant et enfin parmi les deux restant ; d'autre part, une tâche d'estimation de magnitude en estimant, sur une échelle allant de 0 à 100, le degré d'urgence de chacun de ces signaux, restitués deux fois dans un ordre aléatoire. La seconde série de cinq expériences, réalisée avec douze sujets, concernait les paramètres des salves sonores. Les tâches des sujets étaient les mêmes que dans la première série, mais appliquées aux salves complètes. Une dernière expérience, réalisée avec quatorze sujets, était constituée d'un corpus de treize nouveaux signaux d'avertissement, construits à partir des résultats des deux premières séries de tests et dont le degré d'urgence peut être prédit.

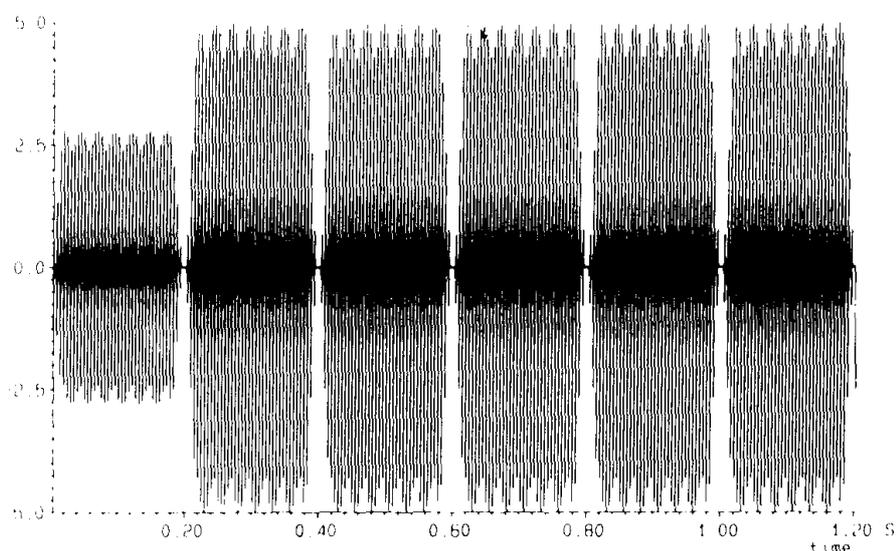


Figure 1-8 : profil temporel de salve à six trains d'onde (Edworthy *et al.*, 1991).

La classification des stimuli ayant été effectuée de manière cohérente par l'ensemble des sujets, les auteurs considèrent donc que la notion d'urgence est une construction psychologique significative et saillante pour les auditeurs, même avec des stimuli très simples. L'analyse des données par une ANOVA (analyse de variance) indique des effets significatifs pour la fréquence fondamentale, l'enveloppe dynamique, les séries harmoniques et les retards harmoniques, ainsi qu'une interaction forte entre fréquence fondamentale et enveloppe dynamique. L'effet de la fréquence fondamentale n'est pas très fort, mais la hausse de fréquence semble renforcer le caractère d'urgence du signal. En revanche, les stimuli à enveloppe dynamique constante sont considérés comme plus urgents que ceux à enveloppe croissante, ceux à enveloppe décroissante étant les moins urgents. L'explication donnée par les auteurs est le niveau global supérieur lorsqu'aucune atténuation n'est introduite, mais également la sensation d'approche pour le cas de l'enveloppe croissante. Cette dernière explication ne nous semble pas convaincante pour la raison que l'enveloppe croissante (et les autres) est reproduite cinq fois dans un intervalle d'une seconde, ce qui n'est pas du tout cohérent avec le déplacement d'un objet. Enfin, l'irrégularité des séries harmoniques semble être interprétée comme un facteur d'urgence, la disposition aléatoire étant liée à une forte urgence, suivie de 10% d'irrégularité, puis de 50% (qui crée de ce

fait une nouvelle relation entre composantes harmoniques) et enfin, une harmonicité régulière. Ainsi, la complexité et la prévisibilité peuvent être corrélés au caractère urgent d'un signal.

D'autres expériences partant des mêmes trains d'ondes, mais modifiant les salves sonores, avaient pour but de tester l'influence du rythme, de la vitesse, des répétitions et des changements de hauteur au sein d'une salve de trains d'ondes. Douze sujets devaient déterminer le stimulus le plus urgent parmi trois, puis parmi les deux stimuli restant. Tous les paramètres testés ont été trouvés comme significatifs dans leur effet sur le sentiment d'urgence. Là encore, la cohérence des réponses des sujets laisse supposer la notion d'urgence comme construction psychologique pertinente pour ces salves sonores. Parmi les paramètres temporels, la vitesse a un grand effet sur le caractère urgent des stimuli, de même que le nombre de répétitions de la salve. Par ailleurs, le caractère régulier semble plus approprié à l'urgence qu'un rythme syncopé. Quant aux paramètres mélodiques, leurs effets sont moins importants, mais de grandes variations de hauteur (8 demi-tons) augmentent le caractère urgent des stimuli, plutôt que 3 ou 5 demi-tons. Il en va de même du caractère atonal des sons.

Edworthy *et al.* de conclure qu'un signal d'avertissement considéré comme très urgent pourrait être décrit par un train d'onde avec une fréquence fondamentale élevée, une série inharmonique, pas d'harmoniques retardés et une enveloppe dynamique constante ou croissante, comme élément constitutif d'une salve rapide, régulière, avec un large changement aléatoire et atonal de hauteur. La dernière expérience, constituée de treize stimuli construits à partir des critères obtenus au préalable, a permis de montrer que le degré d'urgence prédit correspond bien aux estimations des auditeurs. Les auteurs ont ainsi donné une définition, en termes physiques, de ce qui est interprété comme urgent dans une *classe de signaux testés* et pour une *population particulière*, délimitation essentielle dans le choix du *corpus sonore* et du *type de sujets*.

### 3.4 Les temps de réponse

Une autre approche, liée à l'urgence perçue de signaux d'avertissement, est proposée par Haas et Casali (1995). Constatant l'absence de travaux qui concernent le temps de réponse à des signaux et à l'urgence perçue dans des ambiances très bruyantes, les auteurs ont mesuré ces données dans du bruit rose, un spectre proche de celui que l'on retrouve dans des environnements industriels et militaires. Les signaux testés ont été développés, comme pour Edworthy *et al.* (1991), selon la procédure définie par Patterson (1982). Trois types d'expériences ont été effectués : trente-six sujets, âgés de 18 à 22 ans et ayant tous une audition normale, devaient tout d'abord estimer librement le degré d'urgence de chaque signal d'avertissement en présence du bruit de fond, sur une échelle graduée de 0 à 10. Puis, ils devaient répondre de manière tactile à un signal auditif dès qu'ils percevaient ce dernier, en effectuant simultanément une tâche visuelle. Enfin, ils devaient comparer les avertisseurs bruités, par paires, selon le jugement d'urgence.

Les résultats montrent que pour chaque signal testé, le jugement d'urgence est maximal lorsque le rapport signal sur bruit est maximal et lorsque l'intervalle de temps entre les trains d'ondes est minimal. Par ailleurs, les trains d'ondes composés de sons purs séquentiels (mais harmoniques : 500, 1000, 2000 et 3000 Hz) sont perçus comme moins urgents que les trains d'ondes constitués des mêmes sons purs simultanés ou de sons modulés en fréquence (selon une "dent de scie" croissante de 500 à 3000 Hz). Enfin, les temps de réponse confirment de manière significative les conclusions précédentes, même si la différence des temps de réponse entre les deux niveaux sonores est de 60 ms seulement, et celle entre les types de trains d'ondes d'environ 30 ms. Seules les différences d'intervalles de temps entre trains d'ondes ne sont pas

significatives. Ces différences, très faibles à l'échelle du temps de réaction de l'être humain - de l'ordre de 500 ms -, peuvent cependant s'avérer vitales dans des situations d'urgence.

Notons toutefois que la méthode des temps de réaction pose des problèmes lorsque le corpus sonore est constitué de signaux de durées différentes et dont l'évolution temporelle est essentielle à l'identification (cf. § 6.1.1).

## 4. Reconnaissance et identification de formes sonores

Les recherches décrites au § 3 abordent les seuils de perception de signaux acoustiques dans divers bruits de fond, ainsi que les paramètres physiques liés à la perception des signaux acoustiques. Ces études ne tiennent compte, ni de la *manière* dont ces signaux sont *perçus* par les sujets, ni de leur *reconnaissance* en tant que signaux chargés de signification, comme le caractère d'avertissement ou celui d'urgence. Ces notions nécessitent d'autres théories psychologiques que le béhaviorisme. Il s'agit là d'une première étape vers des signaux qui font sens pour les individus, grâce notamment aux apports théoriques et méthodologiques de la Gestalttheorie (§ 4.1) et de la catégorisation prototypique (§ 4.3).

### 4.1 Les apports de la Gestalttheorie

La reconnaissance des formes a connu un nouvel essor au début du XX<sup>ème</sup> siècle, grâce à l'apport de la *Gestalttheorie*, dont nous allons présenter ici quelques principes. Toutefois, son support d'origine et de prédilection est la vision, et son impact sur les études concernant la perception auditive est resté relativement faible. Une approche embryonnaire, du point de vue de l'effet de contraste, a été proposée par Leipp (1965). Schaeffer (1966) s'est beaucoup plus inspiré du gestaltisme, mais également - et surtout - de la linguistique. Nous aborderons par conséquent ses recherches au § 5.3.3, après avoir introduit la linguistique (§ 5.3.1) et la psycholinguistique (§ 5.3.2).

#### 4.1.1 La Gestalttheorie

Un premier apport fondamental à la compréhension et à la prise en considération de la perception humaine des sons est issu de la *Gestalttheorie*, ou *psychologie de la forme*, née en Allemagne au début du XX<sup>ème</sup> siècle et dont les protagonistes ont été Wertheimer, Koffka et Köhler. Ce courant théorique s'appuie sur un certain nombre de critiques faites au béhaviorisme (cf. § 3.1) : d'une part, les états internes des individus, tels que les intentions, les désirs et les croyances, influencent fortement les conduites dans des situations de la vie courante (Weil-Barais, 1993) ; d'autre part, si une écoute analytique d'un objet sonore nous fait découvrir de nouveaux aspects ou détails de celui-ci - richesse harmonique, intensité subjective, évolution temporelle etc., décrits par la psychoacoustique -, elle ne nous en communique pas moins une perception différente ; psychologiquement, il s'agit d'un objet sonore totalement différent de celui habituellement perçu (Guillaume, 1937, p.20).

La Gestalttheorie peut être considérée comme un courant de pensée différent, précurseur du *cognitivism*, qui s'intéresse à la structure et au fonctionnement du système cognitif, à cette "boîte noire" dont le béhaviorisme ne se préoccupe pas. Elle cherche à comprendre la manière dont le système humain organise le monde et non plus la manière dont l'environnement agit sur le système humain (Weil-Barais, 1993). Ces recherches ont abouti à un certain nombre de lois dont :

- le tout est différent de la somme des parties qui le constituent,
- une partie dans un tout est autre chose que cette partie isolée ou dans un autre tout,
- les formes sont transposables et peuvent présenter une articulation intérieure (Guillaume, 1937).

Par ailleurs, certaines lois de structure (ressemblance et proximité de formes) supposent une organisation spontanée du processus physiologique résultant d'une stimulation et cela, indépendamment des significations de l'expérience de l'individu. Ainsi, certaines formes, régulières, simples et symétriques, sont privilégiées. Enfin, il y a un isomorphisme entre notre

expérience phénoménologique qui donne à tout objet une forme, la forme physiologique engendrée dans notre cerveau, et entre cette configuration physiologique et la configuration de percept auquel elle donne naissance.

Se basant sur le traitement des représentations visuelles, Guillaume distingue *figure* et *fond*, la première possédant une forme, un contour et une organisation, le second étant une continuité amorphe, indéfinie et inorganique, aux propriétés fonctionnelles différentes. Cette distinction, que nous réalisons à tout instant, nous permet de hiérarchiser notre champ de perception entre des choses et un milieu neutre.

#### 4.1.2 L'effet de contraste

En 1965, Leipp pose le problème de la perception des signaux d'avertissement sous l'angle de "l'effet de contraste", *émergence psychologique* d'une forme sonore sur un fond (Leipp, 1965) et établit une "typologie des signaux d'avertissement et des bruits de fond usuels" en utilisant le sonagramme. Il classe les signaux d'avertissement selon trois formes, en fonction de leur degré de complexité, les trois formes étant des combinaisons de formes élémentaires simples : rythmique (variations en fonction du temps), dynamique (variations en fonction de l'intensité) et fréquentielle ou mélodique (variations en fonction de la fréquence). Les bruits de fond sont également classés selon trois types dans la représentation sonographique visuelle : amorphes (allure de bruit blanc), organisés horizontalement (hauteur définissable) ou verticalement (bruits impulsifs) et hétérogènes (variables dans le temps). Leipp montre ainsi qu'un signal émerge physiquement d'autant plus facilement que sa forme visualisée au sonographe est :

- originale par rapport à celle des signaux constituant le bruit de fond,
- plus organisée géométriquement sur un fond inorganisé (ou inversement).

Par ailleurs, il constate que l'attente et l'expertise des auditeurs font baisser les seuils de perception des signaux dans du bruit de fond, c'est-à-dire l'émergence psychologique des signaux.

Fort de ce classement, Leipp étudie différents types de sifflets dans leurs contextes d'utilisation. Il montre que le signal sonore issu de chaque sifflet est adapté au bruit de fond dans lequel le sifflet émet : "chaque type correspond à un bruit de fond bien défini. Si le bruit de fond change, le sifflet s'adapte" - il est évidemment sous-entendu que ce sont les humains qui adaptent spontanément le sifflet au bruit de fond. Les représentations sonographiques mettent en évidence le fait que la forme visuelle du son d'un sifflet s'oppose à la forme du contexte de situation : évitement des zones spectrales à forte concentration d'énergie comme les basses fréquences de la circulation automobile, ou encore sons continus dans des contextes riches en événements brefs, comme les ambiances de jardins publics ou de marchés (cris, pas, oiseaux...). Six types de sifflets sont étudiés : le sifflet oral, la clef forée, le sifflet "S.N.C.F.", ainsi que les sifflets à deux bourdons, à deux tons successifs et à roulette. Nous analyserons au chapitre 6 trois d'entre eux afin de les comparer avec les signaux testés dans nos expériences.

Finalement, Leipp aborde les sirènes d'alerte et les cloches, avertisseurs qui doivent porter à très grandes distances. Cette condition impose l'émission de fréquences fondamentales basses afin de pallier le phénomène d'absorption des fréquences élevées lors de leur propagation aérienne. Concernant les sirènes civiles, Leipp propose des formes acoustiques plus originales, en créneau par exemple, pour augmenter leur efficacité. Quant aux cloches, Leipp insiste sur le fait que c'est "l'allure évolutive d'une forme temporelle couvrant une très large bande de fréquence" qui est significative et il critique ainsi l'approche purement spectrale, qui ne fait pas état des fluctuations de fréquences et d'intensités relatives après la frappe de la cloche.

## 4.2 Événements sonores, reconnaissance et identification

La Gestalttheorie, au contraire du béhaviorisme, décrit les objets du monde à partir d'une forme perçue dans son ensemble. Dans le domaine sonore, ces objets ont à la fois une dimension spectrale et une dimension temporelle, puisqu'ils ont une certaine durée. Dès lors, il devient nécessaire d'introduire la notion d'*événement sonore*. Par ailleurs, si le Gestaltisme permet d'isoler des formes - ou des événements - pertinents pour les individus, il est essentiel de savoir également quel type de forme - ou d'événement - les individus ont *reconnu* ou éventuellement *identifié*.

### 4.2.1 La notion d'événement sonore

Schaeffer (1966, cf. § 5.3.3) définit l'*objet sonore* comme un "élément sonore fini ayant un début et une fin, et de forme émergeant sur un fond". Cette définition peut être également associée à la notion plus générale d'*événement sonore*.

Theile (1980) tente d'être un peu plus précis en ajoutant à la structure de l'événement sonore, définie ci-dessus, un lieu et une source : c'est "la partie du son provenant d'une source sonore unique et qui détermine ou influence les propriétés de l'événement auditif associé, à savoir un lieu et une structure".

### 4.2.2 Reconnaissance et identification

Les concepts d'*identification* et de *reconnaissance* varient selon les théories psychologiques. Les définitions les plus consensuelles aujourd'hui sont dérivées de la Gestalttheorie. Pour Delorme (1982, p.79), "un objet peut être considéré comme ayant une identité perceptive lorsqu'il est perçu comme appartenant à une certaine catégorie d'objets et non pas à d'autres, ce qui le rend dès lors mémorisable lorsqu'il peut être associé à une action quelconque ou être nommé". L'identification nécessite beaucoup plus d'informations que la discrimination. La reconnaissance, se base quant à elle, sur une "analyse détaillée des caractéristiques de la forme" et sur "l'existence éventuelle d'un schéma ou d'un modèle interne auquel on compare l'objet à reconnaître" (Delorme, 1982, p.103). Ainsi, reconnaissance et identification s'effectuent par regroupement, sous forme de catégories, des représentations en mémoire d'événements signifiants (Guyot, 1996, p.50-54 ; Rosch, 1978), l'identification associant en outre une désignation lexicale - une "étiquette" - à chaque objet identifié.

La psychophysique et la psychologie expérimentale considèrent généralement le processus d'identification comme un cas particulier de la reconnaissance. Bonnet (1986, pp.18-19) nous donne les définitions suivantes des procédures expérimentales relatives à l'identification et à la reconnaissance : "une tâche d'*identification* consiste à établir une correspondance bijective entre un ensemble de stimuli et un ensemble de réponses qui sont en général les étiquettes, les noms, de ces stimuli. [...] Le terme de *reconnaissance* sera réservé à des paradigmes comportant deux phases successives et dans lesquels la tâche du sujet consiste à dire si chaque item, ou chaque stimulus, présenté au cours de la phase 2, avait été, ou non, antérieurement présenté au cours de la phase 1". Ces procédures sont donc bien en accord avec les concepts évoqués ci-dessus.

### 4.2.3 Résultats sur la reconnaissance

McAdams (1994) dresse un état des lieux concernant la recherche sur la reconnaissance de sources et d'événements sonores. Il présente essentiellement des travaux issus de la *théorie psychologique du traitement de l'information*, partie de la psychologie cognitive, initiée par

Newell et Simon en 1972 et qui considère à l'origine le système cognitif comme un système manipulant les symboles (Weil-Barais, 1993). Selon McAdams, la reconnaissance est basée sur la comparaison des "informations sensorielles traitées avec une représentation stockée en mémoire à long terme dans un lexique de formes sonores" (p.158). La branche *structurale* du cognitivisme, la *psychologie écologique* (cf. § 5.1), développée par Gibson (1966) et issue de la Gestalttheorie, n'est pas vraiment abordée dans le récapitulatif proposé par McAdams.

Les études sur les instruments de musique, présentées par McAdams, mettent en évidence un certain nombre de caractéristiques essentielles à la reconnaissance des sons de ces instruments : les transitoires d'attaque, l'enveloppe spectrale et son évolution temporelle, et de faibles variations aléatoires des fréquences des composantes. Tous ces résultats concordent avec ceux établis par Schaeffer (1966, cf. § 5.3.3). En ce qui concerne les études effectuées sur des événements acoustiques autres que la parole et les sons musicaux, McAdams les résume par le fait que le système auditif se réfère à des propriétés spectrales et spectro-temporelles et à des configurations globales de changement pour reconnaître les événements sonores complexes.

A partir des études décrites, McAdams recense des propriétés acoustiques qui entreraient en jeu lorsque l'on cherche à comparer et à reconnaître des sources et événements sonores. Il distingue d'une part, des *micropropriétés*, de durées très courtes (quelques dizaines à quelques centaines de millisecondes) : enveloppe et contenu spectraux, enveloppe dynamique, transitoires inharmoniques de faible amplitude et changements de forme du spectre en fonction du temps, et d'autre part, des *macropropriétés* (quelques centaines de millisecondes à quelques secondes) : rythmique et texture d'un événement, résonances coordonnées (rebond d'objets) ou désorganisées (rebonds des morceaux d'un objet brisé), qui semblent correspondre à des propriétés spectrales et temporelles caractéristiques (invariants transformationnels).

Outre les paramètres de hauteur, de durée, d'intensité et de localisation, essentiels lors de la perception des sons, Castellengo (1994) relate l'importance capitale des *transitoires d'attaque* dans l'identification des signaux sonores, transitoires qui constituent par ailleurs un des aspects essentiels, avec la sonorité, du *timbre* d'un son (cf. Guyot, 1996).

#### 4.2.4 Résultats sur l'identification

Si l'on considère généralement que l'être humain est capable d'identifier avec précision les sons issus de son environnement, peu d'études cherchent à estimer la qualité de cette aptitude. Dans cette optique, Lass *et al.* (1982) ont étudiés l'identification de quarante sons : vingt sons d'animaux (vaches, chats, chiens, cochons et moutons), douze sons mécaniques (machines à écrire, klaxons et téléphones), quatre sons d'instruments de musique (flûte, piano, guitare et batterie) et quatre sons humains (la syllabe /pΛ/ prononcée par deux hommes et deux femmes). Trente étudiants, vingt-sept femmes et trois hommes, ont écouté les sons et ont dû indiquer par écrit les noms des sources correspondant aux sons perçus. Les résultats indiquent que le taux d'identification est très élevé, de l'ordre de 90% ou mieux pour la majorité des quatorze différents types de sources sonores de l'étude, sauf pour le son de batterie (76%) et ceux de cochon et de mouton (de l'ordre de 50%). En comparant les classes de sons entre elles, les auteurs concluent à une meilleure identification des sons humains, puis des sons musicaux et des sons mécaniques et enfin, des sons d'animaux. Ils évoquent la possibilité d'une influence culturelle liée aux sujets : des sujets issus du milieu rural, par exemple, pourraient plus facilement identifier les sons d'animaux que des étudiants en psychologie de la parole ! Toutefois, il est à noter que le choix du corpus sonore peut également fortement influencer les résultats. Par exemple, si nous ne considérons que les deux ou trois sons les mieux identifiés dans chaque classe de sons, aucune conclusion de ce type ne peut être formulée.

Une autre étude menée par Lass *et al.* (1983) concerne l'identification de sons d'animaux imités par une voix humaine. Vingt étudiants - dix femmes et dix hommes - ont imité des sons

de vaches, chats, chiens, cochons et moutons. Trente sujets, vingt-quatre étudiantes et six étudiants, ont ensuite écouté les cent différents sons (plus vingt pour tester la reproductibilité intra-individuelle des résultats) et ont dû déterminer par écrit le nom de l'animal imité. Les résultats indiquent une très bonne identification des animaux imités, de 84,2% pour les "cochons" à 98,4% pour les "chiens". Par ailleurs, les imitations des femmes sont mieux identifiées que celles des hommes, à plus de 95% quelque soit l'animal imité. En comparant ces résultats avec ceux obtenus précédemment (Lass *et al.*, 1982), les auteurs constatent que l'identification des imitations humaines d'animaux est supérieure à celle des "mêmes" animaux enregistrés directement. Là encore, la variété du corpus sonore peut être à l'origine de ces différences *a priori* étonnantes, ainsi que la consigne, puisque les sujets sont prévenus du statut d'imitation des sons de la seconde expérience et adoptent donc une stratégie différente de celle de la première expérience.

Dans le but de comprendre ce que les gens entendent, Gaver (1993) relate une expérience d'identification libre, au cours de laquelle les sujets doivent exprimer verbalement les sons qui leur sont diffusés. Les résultats montrent que les sujets décrivent presque toujours les sons identifiés en termes de sources sonores et que leur précision d'identification est très élevée. Par exemple, ils peuvent différencier des pas montant rapidement un escalier de pas descendant rapidement un escalier. En revanche, ils confondent la fermeture d'un tiroir avec le roulement d'une boule de bowling sur sa piste. Gaver rejoint ainsi les résultats obtenus par Vanderveer (1979) à partir de trente sons quotidiens tels que des coups de marteau, un froissement de papier ou un claquement de doigts. Les sujets identifient très bien les sons et ce n'est que lorsque ils ne reconnaissent pas la source qu'ils essaient de la qualifier.

Les difficultés expérimentales soulevées par les travaux de Lass *et al.* (1982, 1983) notamment, qui concernent les corpus sonores et les consignes, seront examinées au chapitre 2 pour l'élaboration d'une méthodologie respectant la notion de validité écologique (cf. § 5.1) des expériences.

### 4.3 La catégorisation prototypique

Le concept de catégorisation prototypique a été introduit par Rosch en 1978 (Rosch et Lloyd, 1978) et suppose que les traitements perceptifs relatifs à des stimuli aboutissent nécessairement à l'élaboration de catégories. Deux principes fondamentaux régissent la formation des catégories : d'une part, l'*économie cognitive* soutient le fait que l'être humain cherche à obtenir un maximum d'informations de son environnement avec un minimum d'efforts cognitifs ; d'autre part, la *structure du monde perçu* suppose l'existence de corrélats entre les propriétés des objets qui composent le monde physique, ce qui conduit à une réduction notable du nombre de caractéristiques nécessaires à l'individu pour "comprendre" le monde : nous parlerons de propriétés *pertinentes*.

Les catégories s'organisent autour de trois niveaux d'abstraction (cf. Figure 1-9) : un niveau privilégié, le *niveau de base*, qui distingue les objets les plus facilement discriminés par l'individu dans ses besoins d'adaptation à son environnement ; en amont se trouve le *niveau superordonné* qui sépare des catégories d'objets trop différents pour leur trouver des propriétés distinctives ; en aval se situe le *niveau subordonné* qui rassemble au sein de chaque catégorie des objets au contraire très similaires, une sorte de "niveau d'expert".

A l'intérieur de chaque catégorie existe une "forme moyenne", le *prototype*, qui réunit les propriétés communes aux objets. Cette forme prototypique se définit par l'intermédiaire de ses rapports de *ressemblance* et de *dissemblance* avec les autres objets du monde perçu et permet d'introduire la notion de *typicalité* d'un objet, de par sa relation au prototype (Dubois, 1993a, pp.280-281).

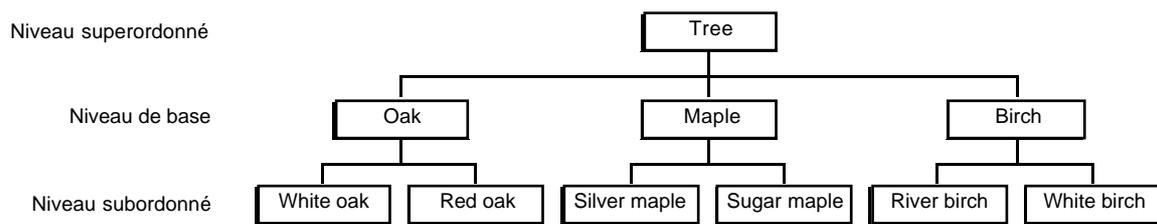


Figure 1-9 : exemple d'organisation hiérarchique des trois niveaux de catégorisation, d'après Rosch (1978).

Ainsi, les prototypes sont les membres d'une catégorie qui rendent le mieux compte de la structure redondante de l'ensemble de la catégorie, puisqu'ils possèdent le plus grand nombre d'attributs en commun avec les autres membres de la catégorie (Rosch, 1978).

## 5. Vers des signaux qui font sens...

Si les différentes théories psychologiques décrites jusqu'ici tiennent de plus en plus compte de la perception de l'individu, elles ne se réfèrent pas encore au sujet en tant qu'être humain *impliqué* dans un environnement naturel et ayant des *intentions* et des *activités*. La *psychologie écologique* tente de répondre à la question de la prise en considération de l'environnement de l'individu - le contexte. Par ailleurs, les travaux sur la reconnaissance et sur l'identification, décrits précédemment (cf. § 4.2.2), apportent des informations concernant le *sens* des signaux sonores perçus, lui-même dépendant du vécu des individus et des données disponibles en mémoire. Ces informations sont contenues dans la production verbale des sujets et l'extraction de la signification nécessite des connaissances issues de la *linguistique* et de la *psycholinguistique*.

### 5.1 Les apports de la psychologie écologique

#### 5.1.1 La psychologie écologique

Certaines expériences classiques de perception placent les sujets dans des conditions décalées par rapport à une réalité quotidienne, comme par exemple leur interdire de bouger la tête. Partant de ce constat, Gibson (1979) considère que la perception doit s'étudier dans des conditions plus naturelles - plus *écologiques* -, en accord avec nos mouvements dans notre environnement. Il postule que l'essentiel de l'information se trouve dans le stimulus et qu'il existe un certain nombre de règles d'invariance de mouvement de l'observateur ou des objets du milieu dans lequel l'observateur se trouve. Ainsi, chaque système sensoriel est muni d'une composante *propriospécifique*, c'est-à-dire spécifique à l'observateur, et d'une composante *extéropsécifique*, c'est-à-dire spécifique aux objets de l'environnement (Streri, 1993).

Par ailleurs, Gibson introduit le concept d'*affordance*, qui correspond à "une information potentiellement utile et pertinente pour l'action d'un individu d'une espèce donnée". Ainsi, l'individu utilise un certain nombre de ces capacités pour les mettre en relation avec son environnement. Les *affordances* représentent ces interactions (Streri, 1993). L'accent est donc porté sur les descriptions de la structure des événements *pertinents pour* les auditeurs, d'où émergent certaines propriétés perçues comme invariantes.

#### 5.1.2 L'approche écologique des événements sonores

S'inspirant de l'approche écologique développée par Gibson essentiellement dans le domaine visuel, Gaver (1993) tente de jeter les bases d'une acoustique écologique. Selon lui, l'écoute peut s'effectuer de deux manières distinctes : d'une part, l'*écoute musicale*, fondée sur les caractéristiques psychoacoustiques du son et qui consiste à analyser le son et à percevoir des propriétés acoustiques liées à ce son, et d'autre part, l'*écoute quotidienne*, qui se réfère à des événements producteurs de son et à leur environnement, plutôt qu'au son lui-même et à ses dimensions acoustiques élémentaires. La distinction ne s'effectue donc pas entre des sons, mais entre des *expériences* relatives à des sons. Cette approche originale a pour objectif de caractériser des propriétés de sources plutôt que des propriétés acoustiques, liées au son. Sans remettre en cause les notions issues de la psychoacoustique, il constate le manque de connaissances actuelles pour caractériser les événements entendus et propose de nouvelles dimensions perceptives. Il part du principe qu'un son fournit des informations sur des

interactions entre des matériaux, en un lieu et dans un environnement donnés, ce qui implique une nécessaire prise en compte du contexte d'écoute dans les stimuli.

Gaver propose une description hiérarchique des événements sonores, plutôt qu'une description opposant le son au contexte (cf. Figure 1-10). Ainsi, ces événements correspondent à des "matériaux qui interagissent" et peuvent être distingués selon trois catégories d'interactions qui dépendent du type de matériaux composant les objets sonores : corps solides, aérodynamiques et liquides. Un niveau encore subordonné au précédent est constitué par des événements tels que "bruits d'impacts", "grattements", "explosions", "éclaboussements"... et qui constituent le niveau de base. Ce dernier correspond d'ailleurs au niveau de base trouvé par Guyot (1996, p.120) pour des bruits domestiques à partir de l'hypothèse de la *catégorisation prototypique* développée par Rosch (1978) (cf. chapitre 4).

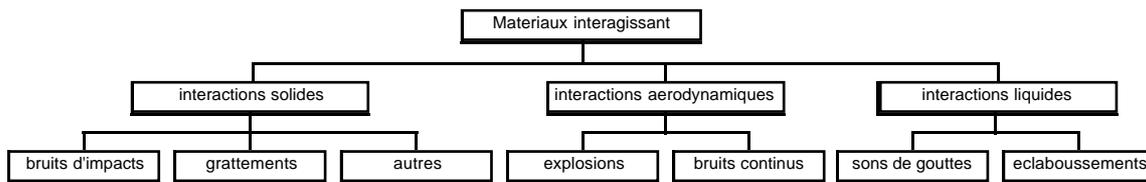


Figure 1-10 : description hiérarchique d'événements sonores, d'après Gaver (1993).

Gaver décrit ensuite ces événements classifiés en mettant en relation des caractéristiques élémentaires d'une source sonore (type et force d'excitation, structure du matériau, forme et taille de la source) et des effets sur l'onde sonore (amplitude, variations temporelles, spectre, fréquence de résonance).

## 5.2 Des études centrées sur les sujets

### 5.2.1 La notion d'urgence recentrée sur les auditeurs

Quelques années après avoir présenté des relations entre paramètres acoustiques de signaux d'avertissement et l'estimation d'urgence liée aux signaux (Edworthy *et al.*, 1991, cf. § 3.3), Edworthy et Stanton (1995) proposent une nouvelle méthodologie d'élaboration et d'évaluation des signaux sonores d'avertissement. Cette fois-ci, l'approche est centrée sur l'utilisateur et non sur des paramètres acoustiques liés à un signal physique donné. Elle s'inspire des travaux sur la validité écologique développée par Gibson dans le domaine visuel et de leur adaptation par Gaver à l'audition. La méthodologie est dérivée d'une procédure standardisée pour évaluer les systèmes publics d'information visuelle. Cette procédure, décrite dans la Figure 1-11, comporte deux phases essentielles : obtention d'une description verbale standard du contenu de l'image et réalisation d'un test d'adéquation des images proposées avec leur représentation symbolique.

Edworthy et Stanton insistent sur le fait qu'un "son peut signifier quelque chose" et que ce sens doit être pris en compte dans le design des signaux d'alerte. Ainsi, ils suggèrent de décrire le *rapport entre un son et sa signification* à l'aide d'un système de *classification* en s'appuyant sur une taxinomie proposée par Gaver. Selon ce dernier, le rapport peut être *symbolique*, *iconique* ou *métaphorique* : une relation symbolique est arbitraire (c'est l'hypothèse de la linguistique pour tout rapport objet-signification - cf. § 5.3.1 et Figure 1-13), un rapport iconique est basé sur la causalité physique, tandis qu'une relation métaphorique repose sur des similarités entre différents domaines (Ballas, 1993). Edworthy et Stanton redéfinissent

légèrement cette classification : une *correspondance (mapping) symbolique* considère un son associé à une même représentation mentale ; une *correspondance relative à un nom (iconique)* considère le son réel et ne nécessite donc pas d'apprentissage ; une *correspondance métaphorique* nécessite un son aux qualités similaires à celles de l'événement qui s'est produit.

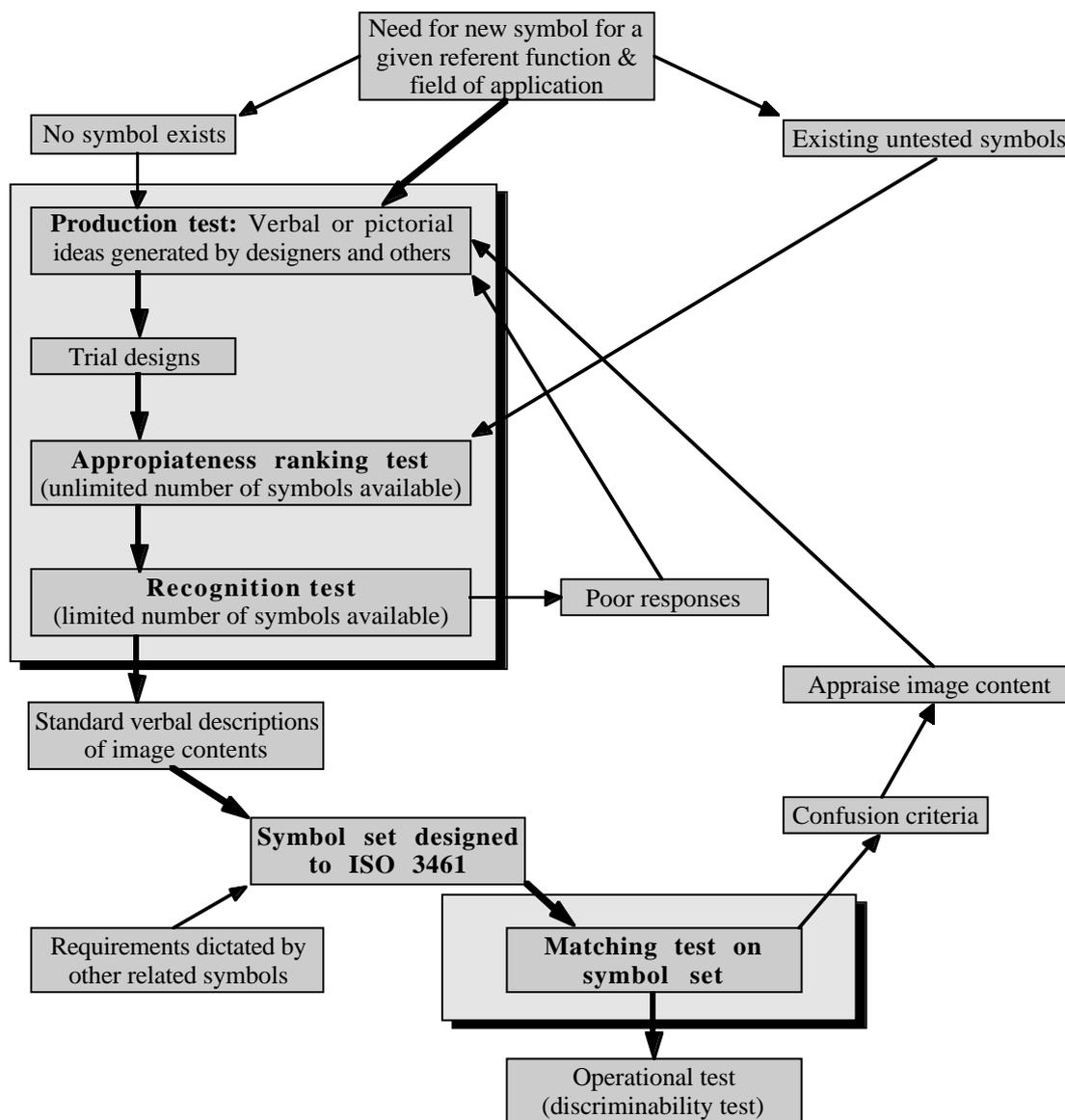


Figure 1-11 : procédure standardisée ISO pour le développement de symboles publics d'information (d'après Zwaga et Easterby, 1984).

La procédure développée par Edworthy et Stanton est constituée de onze étapes (cf. Figure 1-12) :

1. Recensement des situations nécessitant des signaux d'avertissement. Un des critères essentiels pour déterminer la nécessité ou non d'un avertisseur sonore est la vitesse à laquelle les sujets doivent réagir à la situation.
2. Etudier les éventuels signaux existants, afin de les modifier, si nécessaire, selon certains paramètres acoustiques et les rendre ainsi plus facilement détectables ou mieux adaptés.
3. Créer de nouveaux signaux à tester.
4. Recherche des signaux les mieux appropriés (test de classement, d'estimation ou de comparaison).

5. Modifier les stimuli selon certains paramètres acoustiques liés à l'acceptabilité des signaux, tout en conservant les propriétés qui sont à l'origine de leur distinction et leur reconnaissance.
6. Tester la confusion possible entre signaux ainsi que la rapidité d'apprentissage.
7. Vérifier l'adaptation entre l'urgence perçue du signal proposé et l'urgence perçue de la situation.
8. Constituer un ensemble de signaux "types".
9. Tester la reconnaissance de chaque signal et son adéquation avec une des situations étudiées.
10. Générer des description verbales standardisées, séparées en deux groupes : descriptions générales ou descriptions plus détaillées proches de propriétés acoustiques.
11. Tester le caractère opérationnel des signaux par l'intermédiaire d'une simulation.

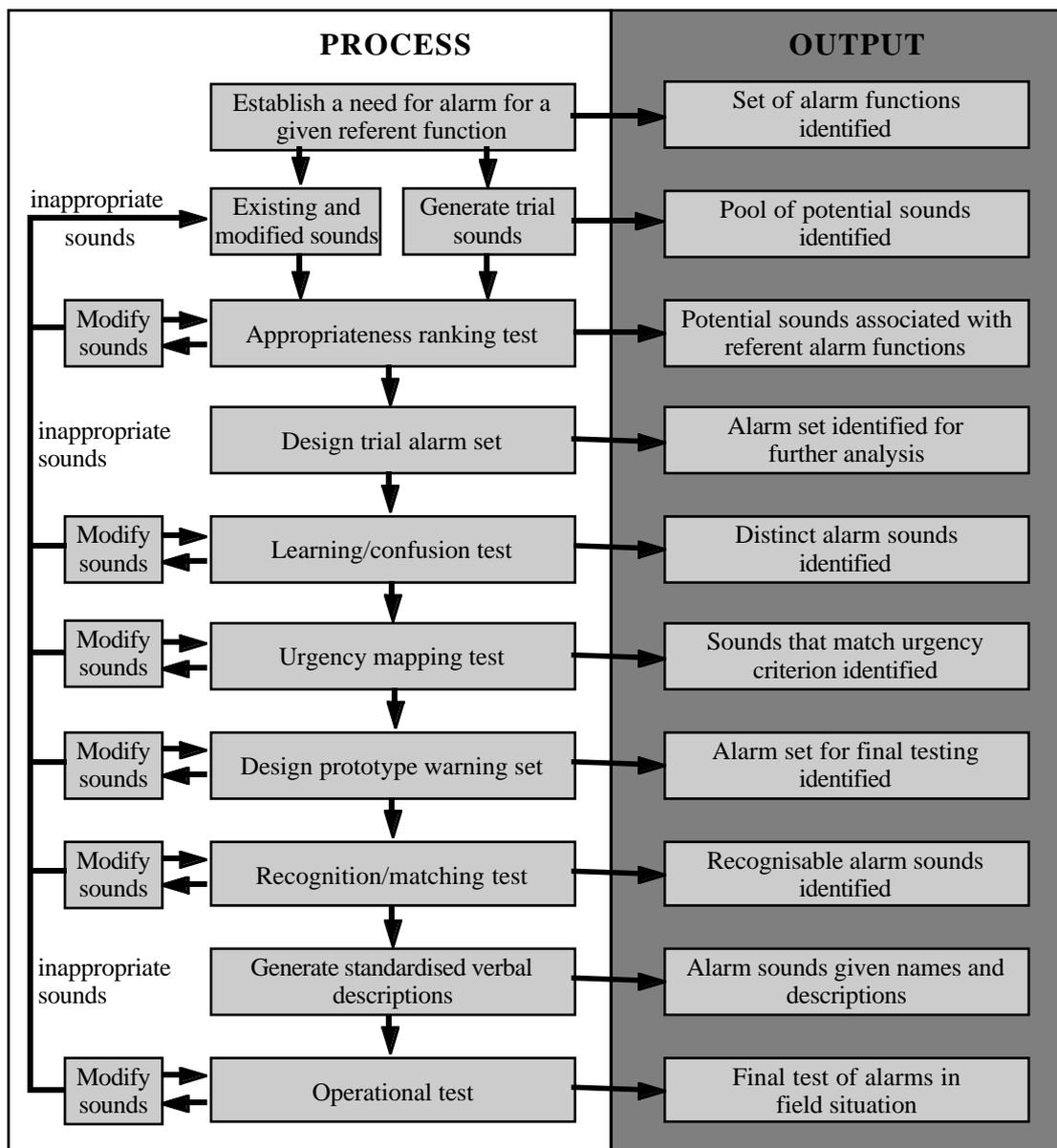


Figure 1-12 : diagramme de la procédure de design de signaux d'avertissement, d'après Edworthy et Stanton (1995).

Pour les auteurs, trois points conceptuels sont essentiels : l'implication de l'utilisateur dans le design des signaux d'avertissement, l'importance de sa connaissance et son interprétation de l'information auditive, et la prise en compte de l'ensemble des signaux pouvant

survenir afin d'éviter de possibles confusions. Cela répond d'ailleurs à une forme de validité écologique du corpus d'échantillons sonores considéré.

### 5.2.2 Le vécu des sujets : les cartes mentales sonores

Une étude réalisée par Amphoux (1991) avait pour objectif de caractériser l'identité sonore des villes. L'idée est de proposer un outil d'analyse interdisciplinaire de la qualité sonore des espaces urbains qui tient compte des dimensions acoustique, topologique et socioculturelle des villes. L'approche d'Amphoux comporte trois grandes parties : l'écoute *mémorisée*, où l'on fait appel à la mémoire et à l'expérience sonore des usagers ou de "professionnels" de la ville ; l'écoute *réactivée*, où l'on utilise la perception directe d'espaces sonores enregistrés sur les terrains sélectionnés dans la première partie ; l'écoute *qualifiée*, qui est une approche spéculative de mise en ordre théorique des critères dégagés.

Pour notre sélection des échantillons sonores, nous avons appliqué une partie de la méthode de l'écoute *mémorisée*. Celle-ci consiste à demander à des sujets de dessiner une "carte mentale sonore". Le protocole d'entretien défini par Amphoux est le suivant :

1. dessiner une carte sonore de la ville, en répondant à une question du type : "qu'est-ce que le Paris sonore pour vous ?",
2. commenter brièvement la carte,
3. préciser un certain nombre de lieux possédant des qualités acoustiques particulières, établir un répertoire des terrains ou itinéraires présentant de la richesse sonore,
4. confronter la carte sonore au plan de la ville.

Cette expérience a permis à Amphoux (p.28) de proposer trois formes de représentations : le plan général de la ville, l'itinéraire, et un ensemble d'unités fragmentaires et discontinues. Par ailleurs, il détermine trois modes spécifiques de représentations : la symbolisation codifiée du lieu, l'expression d'un certain type d'espace, et une icône représentant le type de sources sonores. Il en déduit quatre axes d'interprétations : expression-représentation, symbolique-icône (abstrait-concret), référentiel-générique et statique-dynamique.

Enfin, Amphoux (p.61-65) définit trois types de formes sonores selon une approche gestaltiste : le *fond sonore*, continu, qu'on n'entend pas, mais qu'on oit (cf. le son ouï chez Schaeffer), et qu'il faut écouter pour l'entendre ; l'*ambiance sonore*, signature du lieu, caractérisée par sa dynamique, qu'on peut entendre sans l'écouter mais qu'on ne peut écouter sans l'entendre (cf. le son entendu chez Schaeffer) ; les *signaux sonores*, discontinus, qui sont toujours entendus, pour être écoutés (cf. le son écouté chez Schaeffer).

## 5.3 Signaux, langage et communication

Les expressions verbales des individus décrivant un événement sonore impliquent un changement de représentation : le passage d'une représentation mentale auditive, issue du son perçu, vers une représentation verbale, codée dans la langue du sujet qui s'exprime. Nous avons vu que les recherches sur l'identification des signaux sonores (Lass *et al.*, 1982) et sur la classification des événements sonores (Gaver, 1993 ; Edworthy et Stanton, 1995) se basent sur de telles dénominations. Il est donc primordial de comprendre comment les individus procèdent pour lexicaliser des sons perçus et comment le sens est associé au son. Ces connaissances ne peuvent évidemment provenir que des théories psychologiques du langage et de la signification - la linguistique et la sémiologie - et doivent être mises en relation avec la psychologie de la communication.

### 5.3.1 La linguistique : de la sémiologie à la sémantique

#### 5.3.1.1 Saussure et la sémiologie

Dans l'introduction de son *Cours de linguistique générale*, Saussure (1916) situe le langage en relation avec plusieurs disciplines : la physique (ondes sonores), la physiologie (audition et phonation) et la psychologie (images verbales du son perçu et concepts), appartenant tant au domaine individuel (parole - acte volontaire et intelligent) qu'au domaine social (langue - fonction du sujet parlant). Selon Saussure, la *langue* est un système abstrait de signes exprimant des idées et où n'est essentielle que l'union du *sens* et de l'*image acoustique*. La langue est un système de valeurs pures et son rôle est de servir d'intermédiaire entre la pensée et le son, dont l'union impose des délimitations réciproques d'unités. Et c'est la *distinction* d'un signe qui en fait le caractère, comme elle en fait la valeur et l'unité. Il en découle une science qui étudie la vie des significations au sein de la vie sociale, qui décrit les signes et les lois qui les régissent : la *sémiologie*.

Saussure énonce deux principes : d'une part, le signe unit *arbitrairement* un *signifiant* (images acoustiques) et un *signifié* (concepts). Ainsi, le *son* issu du mot "sirène" (le signifiant) a un rapport arbitraire avec son signifié, le *sens* du mot "sirène". Cette relation résulte d'un choix arbitraire du signifiant et d'un apprentissage culturel du rapport du signe (signifiant-signifié) d'urgence sonore au monde qui nous entoure (cf. Figure 1-13). D'autre part, le second principe associe au signifiant de nature acoustique une unique dimension, temporelle, celle de la "chaîne parlée" - alors que des signaux visuels autorisent plusieurs dimensions spatiales.

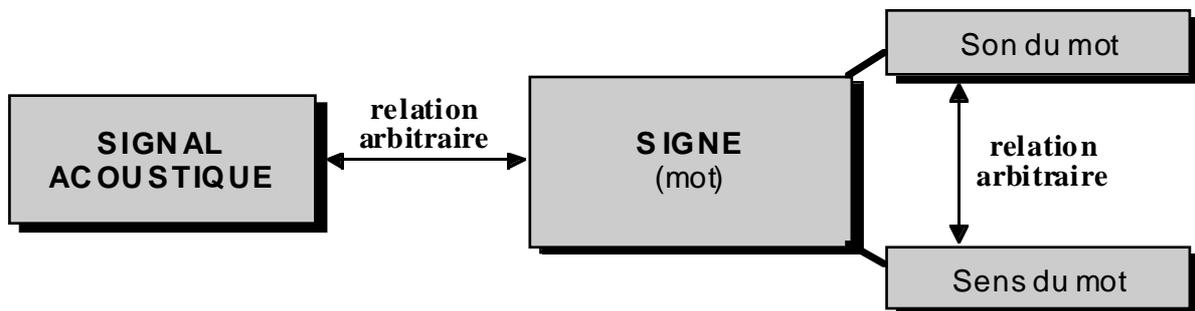


Figure 1-13 : l'arbitraire du signe (relation signifiant-signifié) et l'arbitraire entre le signe et le monde extérieur.

Par ailleurs, Saussure distingue deux sciences de la langue : la linguistique statique, ou *synchronie*, qui s'intéresse au système complet de la langue et dont toutes les parties doivent être considérées, et la linguistique évolutive, ou *diachronie*, qui étudie les altérations, au cours du temps, de certaines parties d'une langue. Nous nous restreindrons à décrire la synchronie, en rapport direct avec notre propos. En effet, celle-ci repose sur deux principes : d'une part, l'entité linguistique n'existe que par l'*association du signifiant et du signifié*, et d'autre part, l'entité linguistique n'est complètement déterminée que lorsqu'elle est *délimitée*. Une entité est donc déterminée par le sens et le rôle que nous pouvons donner à une partie de la chaîne - temporelle - de la parole. La méthode de délimitation linguistique peut s'effectuer, soit en fonction du sens, soit en fonction du son, et ne permet pas de déterminer d'emblée des entités. Saussure montre d'une part, que les enchaînements linéaires des mots de la langue créent des combinaisons qui s'associent, par deux ou plus, en éléments de valeur appelés *syntagmes* ; d'autre part, que certains mots évoquent toute une série d'autres termes chez l'individu, créant des *rappports associatifs*. Cette distinction constitue la base d'un mode de classement qui s'impose de lui-même, le seul que l'on puisse mettre, selon Saussure, à la base du système grammatical.

### 5.3.1.2 Jakobson et la phonologie

Jakobson (1976), dans un cours donné en 1942-1943, s'intéresse notamment à la structure de l'union signifiant-signifié qui constitue le signe et à la façon dont les sons remplissent la fonction de véhiculer du sens. Il considère que la théorie saussurienne décrit correctement le signe, ainsi que le rôle des propriétés des entités d'être à la fois significatives et vides de sens.

En revanche, il remet en cause les deux principes de Saussure, à savoir, d'une part, le *caractère arbitraire du signe*, car le lien entre signe et réalité extérieure est, au contraire, *nécessaire*, tant par symbolisme phonétique (évocations en rapport avec des sensations musicales, olfactives, tactiles...), que par la dualité des opposés (le caractère aigu ne saurait exister sans le caractère grave, discontinu sans continu, long sans bref...); d'autre part, le *caractère linéaire du signe*, car un signe peut accumuler *simultanément* plusieurs signifiés différents. En effet, si l'on ne peut pas émettre à la fois deux phonèmes - unités phonétiques élémentaires, dépourvues de sens - on peut parfaitement émettre à la fois plusieurs qualités distinctives, signifiantes, dont même la durée peut être une valeur. Si Jakobson estime que Saussure a justement compris que les phonèmes et la délimitation des sons de la chaîne parlée présupposent l'existence inconsciente d'une donnée extrinsèque, il se démarque de lui par le fait que c'est la *valeur linguistique* d'une donnée acoustique - son sens - qui permet de découper cette chaîne, et non la donnée acoustique du signe linguistique en soi - le phonème. Pour Jakobson, ce n'est pas le phonème qui est une entité irréductible et purement oppositive, mais chacune des *propriétés distinctives* dans lesquelles se décompose le phonème. En effet, les phonèmes se dissocient en qualités distinctives indécomposables, les *éléments différentiels*, qui permettent d'accéder au sens. Le nombre de qualités distinctives des phonèmes est considérablement plus restreint que le nombre de différences entre phonèmes.

### 5.3.1.3 Greimas et la sémantique structurale

La sémantique a pour objet d'étude les langues naturelles et s'inscrit donc dans la sémiologie. Greimas (1970) propose une description des langues en réunissant les moyens conceptuels nécessaires et suffisants. Il considère la perception comme le lieu non linguistique où se situe l'appréhension de la signification. Et la sémantique de se reconnaître comme une tentative de description de monde des qualités sensibles.

Plus prudent que Jakobson, Greimas affirme que la seule façon d'aborder le problème de la signification consiste à affirmer l'existence de *discontinuités*, sur le plan de la perception, et celle d'*écarts différentiels*, créateurs de signification, sans se préoccuper de la nature des différences perçues.

La *structure* sémantique résulte de la présence de deux termes et de la relation entre eux. Cette relation implique à la fois conjonction du premier terme - ressemblance - et disjonction du second terme - différence. Les éléments de signification, appelés éléments différentiels par Saussure et traits distincts par Jakobson, sont nommés *sèmes* par Greimas.

Greimas se distancie de Jakobson du fait qu'il suppose l'existence d'un troisième sème - neutre ou complexe - en plus des deux sèmes en opposition (par exemple : grand vs moyen vs petit). Cependant, il conserve l'articulation en deux sèmes élémentaires pour les structures élémentaires.

Partant de ces principes, Greimas développe une description des langues naturelles, description qui, toutefois, ne saurait entrer dans le cadre de ce mémoire. Toujours est-il que le parallèle entre sèmes - ou phonèmes - et propriétés acoustiques significatives du son, aussi risqué soit-il, n'en demeure pas moins tentant (cf. Schaeffer, 1966, § 5.3.3).

## 5.3.2 Communication et psycholinguistique

### 5.3.2.1 La communication

Une définition élémentaire de la *communication* est donnée par Dubois (1993a, p.202) : "la communication est la transmission entre *individus* des "informations" - contenus, messages, connaissances - par l'intermédiaire d'un *code*, d'un *système symbolique*, à travers un certain *média*, dans l'*intention* de parvenir à différents *buts*". Cette définition fait intervenir, outre l'idée de transmission d'informations entre un émetteur et un récepteur, la notion forte et essentielle d'*implication* de l'individu dans sa démarche de communiquer. Par ailleurs, chaque étude, dans un domaine de recherche spécifique, doit préciser les différents termes de la définition écrits en italique et cela, grâce à l'apport d'autres sciences - ici en l'occurrence, la linguistique, la sémiologie et l'acoustique. Ainsi, la description des formes de communication "sensorielle" (visuelle, auditive, olfactive...) permet d'établir de nouveaux concepts liés aux propriétés signifiantes de la perception.

### 5.3.2.2 Systèmes de communication

Un système de communication peut être formé sur un seul niveau d'articulation, à partir des *indices* et des *signaux* (en tant que relation symbolique, selon Gaver (1993, cf. § 5.2.1). Dans ce cas, le signal - ou l'indice - est explicite, conventionnel et produit de manière intentionnée (Dubois, 1993a). Ainsi, les signaux d'avertissement des véhicules, par convention, doivent prévenir d'un danger imminent relatif à leur passage. Un autre type de système de communication est formé par les *signes* et les *symboles* et dans lequel interagissent deux niveaux d'articulation : le premier regroupant les entités pourvues de signification et le second concernant les phonèmes, dépourvus de sens (cf. Jakobson, § 5.3.1.2).

Les langues étant elles-mêmes des systèmes de signes, deux approches distinctes tentent de les décrire : l'une, *structuraliste*, considère les relations pertinentes par contraste, opposition, substitution ou concaténation des éléments du lexique (cf. Saussure et Jakobson) ; l'autre, *générativiste*, suppose l'existence de règles qui régissent ces relations et qui permettent d'engendrer une infinité de phrases possibles, selon la compétence du locuteur (Dubois, 1993a, p.221).

### 5.3.2.3 La psycholinguistique

La psycholinguistique - fondée en 1953 par des linguistes et des psychologues - a fourni un apport essentiel à la compréhension du langage, et notamment, aux théories cognitives. Celle-ci considère deux sources d'information : les *données perceptives*, en périphérie du système - les récepteurs sensoriels - et les *représentations mentales*, centrales ; elles sont métaphoriquement placées "en bas" et "en haut" respectivement. Les informations "entrent" par les récepteurs sensoriels périphériques et sont traitées par des contraintes ascendantes, "*bottom up*", pour être interprétées. Les représentations en mémoire se projettent sur les formes d'entrée par l'exercice de contraintes descendantes, "*top down*" (Dubois, 1993a, p.228).

La description du langage, comme science du traitement de l'information, s'opère selon quatre niveaux de segmentation et d'organisation : phonologique, pour les entités élémentaires du langage oral (cf. § 5.3.1.2), syntaxique, pour l'organisation des mots en séquences grammaticalement correctes, sémantique, pour donner du sens aux phrases, et pragmatique, pour mettre les formes linguistiques en relation avec le contexte de l'auditeur.

### 5.3.3 Une approche originale : les objets sonores

Dans le *Traité des objets musicaux*, Schaeffer (1966) propose une nouvelle approche de la musique, en cherchant à synthétiser les connaissances musicales et les connaissances scientifiques dans "des structures permanentes de la pensée et de la sensibilité humaine". Il part d'une analogie entre la musique traditionnelle et la linguistique pour définir des *objets sonores*, éléments sonores constitutifs, équivalents musicaux des phonèmes de la langue. Remarquant que l'appareil phonatoire est figé depuis des millénaires, tandis que les moyens de créer du son musical n'ont cessé d'évoluer, Schaeffer n'en relève pas moins la notion de choix et de conventions, commune aux deux sources sonores et nécessaire à la signification.

Poursuivant sa comparaison avec la phonétique, Schaeffer prévient le chercheur qui essaie d'extraire les mots parlés d'une représentation sonographique : "ce qu'il vise à travers cette démarche, c'est une méthode qui permettrait de relier les valeurs sémantiques à des structures physiques.[...] Jusqu'à présent en tout cas, la reconnaissance acoustique des mots [...] doit se borner à procéder par reconstitutions syllabiques - donc à partir d'éléments qui en eux-mêmes ne sont pas sémantiques [...]. Un tel dépouillement ne saurait correspondre naturellement à un contenu sémantique" (p.126).

Schaeffer étudie les propriétés acoustiques d'extraits musicaux, pertinentes aux oreilles des auditeurs, ceci afin de trouver des corrélations entre les signaux physiques et les objets musicaux. Concernant la hauteur des sons, il montre que la perception des intervalles de hauteur est une donnée culturelle, liée à la pratique et à des conventions, et que les relations qui structurent les hauteurs résistent aux réductions de largeurs de bandes passantes fréquentielles. Par ailleurs, les transitoires d'attaque de sons brefs ou percussifs jouent un rôle essentiel pour l'identification de la source sonore et la caractérisation de son timbre. Leur influence est cependant moindre pour les sons filés et devient presque négligeable pour des sons entretenus avec vibrato, l'évolution du son prenant alors plus d'importance. Schaeffer d'en déduire deux lois des perceptions de la raideur des attaques : d'une part, pour les sons entretenus, l'oreille serait sensible à la façon dont l'énergie sonore apparaît dans le temps, et d'autre part, pour les sons percussifs ou les sons pincés suivis d'une résonance, l'oreille serait sensible à la façon dont l'énergie disparaît plutôt que n'apparaît. Enfin, il montre que la durée musicale est fonction directe de la densité d'information - en termes physicalistes - qui parvient à l'auditeur.

Après avoir défini *a priori* et de manière symétrique quatre types d'écoute (écouter, ouïr, entendre, comprendre), Schaeffer (pp.155-156) propose trois fonctionnements de l'écoute : la première dirigée vers l'origine du son (indices), la seconde tournée vers son sens (valeurs), et la troisième, négligeant les deux premières et orientée vers soi-même (tel l'instrumentiste), percevant l'*objet sonore*. Ce dernier est un objet de communication collective, contenu dans notre conscience perceptive - représentation mentale, différente selon les individus pour un même extrait musical - et qui "représente la synthèse de perceptions d'habitude dissociées" (p.156). Il est à la rencontre d'une action acoustique et d'une intention d'écoute, mais dénué de sens, d'origine ou de valeur culturelle. Les qualités du son proviennent alors de la mise en relation de l'écoute musicale réduite, d'un langage lié à ce contexte d'écoute et des mesures physiques.

Toutefois, Schaeffer ne cherche pas à aborder la "signification musicale" et se limite au "musical élémentaire" (p.294). Les comparaisons avec la linguistique se cantonnent donc aux niveaux phonétique et phonologique, et la seule tentative de confrontation aux niveaux supérieurs du langage le conduit à postuler que "les signes musicaux sont faits pour être entendus, et autrement que les signes linguistiques" (p.305). Ces derniers ont essentiellement pour fonction de donner à *comprendre* des mots, des phrases, un sens, tandis que les premiers sont destinés à être *entendus* comme valeurs. Il n'en va évidemment pas de même pour les sons non musicaux et en particulier pour les signaux sonores qui, à l'instar des signes linguistiques, doivent avant tout être reconnus, identifiés et "compris", avant d'être "entendus" et qualifiés.

## 6. Discussion

De ce résumé relatif aux études sur la perception auditive de signaux sonores, deux tendances principales se dégagent : d'une part, une *approche physicaliste*, centrée sur la description physique du son et sur les effets que ce dernier produit sur des individus (béhaviorisme, Gestalttheorie et psychologie écologique) ; d'autre part, une *approche cognitive*, centrée sur l'individu percevant (sémiologie et psycholinguistique). Reprenons ces différents travaux afin de les confronter entre eux et d'en déduire un certain nombre d'éléments conceptuels qui peuvent contribuer à une approche cohérente de la perception des signaux sonores d'avertissement en contexte urbain.

### 6.1 Les approches physicalistes

#### 6.1.1 Des considérations spectrales...

Szeto *et al.* (1991), Binnington *et al.* (1993), tout comme Ebata et Baba (1997) adoptent une approche centrée sur la détection du signal acoustique dans un bruit de fond contextualisé. Dans les trois études, la détection est étudiée en termes de masquage fréquentiel et de répartition fréquentielle de l'énergie du signal par rapport à celle du bruit de fond. Il s'agit là d'une notion psychoacoustique importante, relative aux seuils de perception de l'être humain. Plusieurs difficultés majeures apparaissent dans ce type d'étude qui concernent notamment les choix méthodologiques suivants : *unités de mesure* du signal physique, *contexte d'écoute* et *sujets*.

Tout d'abord, les niveaux mesurés doivent être en adéquation avec les intensités perçues par les auditeurs. Or, nous avons vu (cf. § 2.3) qu'aucun indice intensimétrique ne rend parfaitement compte de l'intensité subjective. Toutefois, les courbes de pondérations en sont assez proches, faciles à utiliser et surtout très répandues, ce qui facilite considérablement les comparaisons entre différentes études. Ceci justifie donc raisonnablement leur utilisation dans la plupart des travaux. Deux points surprenants subsistent cependant : pourquoi utiliser le dBA pour des niveaux sonores qui nécessiteraient du dBC ? En effet, même le dBSPL - non pondéré - est bien plus proche qu'une pondération A des lignes isosoniques pour des signaux sonores intenses - c'est notamment le cas des objets concernés ici. Et comment interpréter des mesures en dBAeq, de séquences musicales moyennées dans le temps (Ebata et Baba) ? Nos oreilles effectuent-elles des moyennes temporelles pour apprécier l'intensité sonore perçue ?

Ceci nous conduit directement au second type de problème cité, en l'occurrence le choix du contexte d'écoute. Szeto *et al.* et Binnington *et al.* ont pallié le problème intensimétrique en bouclant le fond sonore sur lui-même (boucle de 10 s pour Szeto *et al.*, durée non précisée par Binnington *et al.*) en évitant de la sorte des variations contextuelles trop importantes. Mais ce faisant, les auteurs prennent le risque de "décontextualiser" le contexte, de transformer un fond sonore urbain *signifiant* en un bruit large bande quelconque, certes proche de l'ambiance réelle d'un point de vue spectral, mais très dégradé d'un point de vue temporel. Ce point est donc très important à considérer si l'on étudie la signification de signaux sonores en contexte.

Enfin, le choix des sujets est subordonné à la réalisation d'audiogrammes qui permettent de ne retenir que des individus à audition "normale". Le critère de sélection n'a pas la même exigence selon les études (niveaux tolérés de pertes auditives, plage fréquentielle testée). Toutefois, une détection différentielle de stimuli relativement à un fond sonore est bien moins sensible à d'éventuelles pertes auditives que ne l'est une détection absolue. En effet, tous les sons présentés se trouvent largement au-dessus de la courbe de seuil absolu et seuls les phénomènes de masquages, bien plus saillants dans le son en soi que dans de légères pertes

auditives, entrent en compte. Robinson et Casali (1995) ont ainsi montré que des sujets ayant jusqu'à 40 dB de pertes auditives de 500 à 2000 Hz, détectent aussi facilement que des individus sans pertes auditives des signaux d'alarme (dont l'énergie se concentre entre 1000 et 2000 Hz) dans un bruit de fond. Ces résultats, que nous vérifierons également pour nos échantillons sonores (cf. chapitre 4, § 2.3), légitiment l'omission de l'étape des audiogrammes, préalable aux tests d'écoute de signaux sonores dans un bruit de fond relativement intense.

Les différents seuils de détection trouvés dans la littérature seront comparés avec les résultats obtenus dans les expériences décrites plus loin (chapitre 6). Nous remarquerons toutefois qu'aucun traitement sémantique ne permet de vérifier l'identification du signal physique émis en tant que "sirène" dans les représentations mentales des sujets.

### 6.1.2 ...aux considérations spectro-temporelles

Les travaux d'Edworthy *et al.* (1991), concernant la recherche de paramètres acoustiques liés à la notion d'urgence de signaux sonores, ont permis d'établir des critères d'urgence selon la hauteur du signal, son enveloppe dynamique, sa structure mélodique... Ces résultats ont été obtenus par comparaison de stimuli sonores très proches, issus de la modification d'un signal prototype. L'approche consiste donc à *affiner* l'efficacité d'un signal donné, à procéder à un réglage de finition. Cependant, si cette méthode est cohérente par rapport à une approche psychoacoustique, elle pose un certain nombre de problèmes essentiels relativement aux traitements cognitifs des sujets : le signal prototype est-il pertinent pour les auditeurs ? Le cas échéant, l'est-il dans tout contexte ?

Par ailleurs, Edworthy *et al.* proposent d'utiliser les différents signaux étudiés pour différents niveaux d'urgence donnés. Mais, sommes-nous capables de mémoriser des différences minimales pour ensuite estimer, en situation, le degré d'urgence d'un signal apparaissant de manière isolée, sans possibilité de comparer ? La notion d'urgence n'a, en effet, aucune raison d'être purement instinctive, non cognitive. Il paraît donc beaucoup plus raisonnable d'aborder le problème plus "en amont", avec des signaux très différents et, plutôt que de créer une gamme de signaux proches variant selon quelques paramètres acoustiques, de proposer des signaux clairement *identifiables* et *univoques*, car particuliers et différents entre eux, pour pallier les risques de confusion. Cela est d'autant plus valable pour un public non spécialisé, comme les usagers de la ville.

Enfin, l'étude ne tient pas compte de la qualité subjective des sons, sauf pour le caractère agaçant de certaines répétitions, car elle vise un public spécialisé dans une activité particulière, tel que les pilotes ou les médecins, et non les citoyens en général, simples usagers de la ville.

L'expérience de Haas et Casali (1995) est proche de celle d'Edworthy *et al.* (1991) puisqu'elle part du même type de stimuli sonores. Elle s'en distingue toutefois par son approche méthodologique qui consiste à mesurer les temps de réponse des sujets lors de l'écoute de signaux "contextualisés" par du bruit rose. Si cette méthode donne des résultats convaincants pour des stimuli peu différents, elle est néanmoins difficilement applicable à des signaux urbains. En effet, comment peut-on mesurer des temps de réponse pour des signaux dont le sens est donné essentiellement par leur évolution temporelle ? Si le chronomètre est lancé dès l'apparition du son, certains avertisseurs tels que celui à deux tons des pompiers, très efficace, seraient fortement désavantagés. Si on le déclenche à la fin de la séquence sonore, les sujets auront estimé bien plus tôt le degré d'urgence et les temps de réponse ne signifient plus rien. La méthode peut donc être convenablement appliquée à un corpus de stimuli très brefs (quelques centaines de millisecondes), mais difficilement à des échantillons d'une seconde ou plus, pour lesquels la structure temporelle est souvent primordiale pour provoquer la reconnaissance (cf. Cohen, 1999).

Le récapitulatif de la recherche expérimentale sur la perception de sources et d'événements sonores (McAdams, 1994) montre que les techniques de reconnaissance et/ou d'identification par appariement, de catégorisation et/ou de verbalisation des stimuli sonores, permettent de trouver des invariants de perception au sein des catégories de stimuli et d'accéder à différents éléments des représentations mentales des sujets. Ces invariants sont fortement associés à des propriétés temporelles, ce qui confirme les résultats obtenus par Leipp (1965) et par Schaeffer (1966).

A partir des mêmes études et dans une perspective écologique, Guyot (1996, pp.31-35) se positionne par rapport à la psychologie du traitement de l'information défendue par McAdams. Elle s'interroge notamment sur la pertinence des paramètres acoustiques trouvés dans ces travaux. En effet, les diverses expériences utilisent des stimuli sonores de nature et de durées très différentes d'une part, et d'autre part, pour les rares sons musicaux comparables, elles mettent en évidence des performances d'identification très inégales. Par ailleurs, de nombreux stimuli ont été modifiés au risque d'en altérer l'identification par les sujets, et par conséquent, de perdre leur signification. Or, Dubois (1993a) a montré que la signification donnée à un corpus de stimuli influence les stratégies de résolution de tâches perceptives.

### 6.1.3 L'homme "dimensionnel"

Les travaux présentés par McAdams (1994) sont essentiellement issus de la psychophysique, avec recherche de dimensions - c'est-à-dire d'un continuum - liées à la perception. Il en est ainsi des expériences de discrimination et d'évaluation des similitudes et des méthodes d'analyse multidimensionnelle (ANOVA, analyse factorielle, AMD...) qui leur sont liées et qui présupposent un caractère dimensionnel de la représentation auditive. Et McAdams de conclure en insistant sur l'importance des dimensions continues dans la représentation auditive. Toutefois, le courant écologique conteste l'hypothèse selon laquelle l'être humain perçoit selon des dimensions continues. Dubois (1997) s'interroge d'une part, sur les *a priori* dimensionnels de l'espace psychologique : la métaphore de l'espace - au sens dimensionnel - est-elle pertinente pour les représentations mentales - et sémantiques ? D'autre part, elle questionne la dépendance de cette représentation psychologique avec le modèle physique : l'espace psychologique - et sémantique - est-il vraiment homomorphe à l'espace physique ? Par exemple, si la phonétique - physicaliste - peut construire sans aucune difficulté un continuum physique entre les syllabes /ba/ et /ga/, il n'en demeure pas moins que nous ne saurons percevoir que /ba/ ou /ga/ et rien d'intermédiaire, conformément aux hypothèses de la phonologie linguistique. Ce sont ces deux catégories qualitativement distinctes, et non le continuum quantitatif créé entre ces deux syllabes, qui correspondent à nos représentations sémantiques associées au "continuum physique" entre /ba/ et /ga/ (Harnad, 1987). Les *catégories sémantiques* constituent un aspect important de cette *perception catégorielle* que nous venons d'illustrer, concept qui répond aux principes énoncés par Rosch (1978, cf. § 4.3) pour la catégorisation prototypique.

## 6.2 L'homme "écoutant"

Deux approches originales de l'écoute humaine ont été présentées car elles apportent des éléments nouveaux dans la recherche de paramètres liés à la signification de signaux sonores en situation. La première, développée par Gaver (1993), a notamment comme point fort, la prise en considération du *contexte d'écoute*, l'environnement sonore. Celle-ci s'effectue lors d'une *écoute quotidienne*, à la différence de l'*écoute musicale*, liée à des paramètres acoustiques du son instrumental perçu. En revanche, le sujet reste auditeur et n'est pas impliqué dans

l'événement sonore, son activité n'est donc pas prise en compte dans cette approche. La seconde théorie, développée par Schaeffer (1966), s'inspire largement du structuralisme issu de la linguistique (Saussure, Jakobson) et de la Gestalttheorie. Schaeffer distingue trois formes d'écoutes : la première dirigée vers l'origine du son (l'écoute *musicale* chez Gaver), la seconde tournée vers son sens (une forme de l'écoute *quotidienne* chez Gaver), ces deux écoutes constituant l'*écoute banale*, et la troisième, s'opposant aux deux premières, l'*écoute réduite*, orientée vers soi-même et percevant l'objet sonore. Ainsi, l'activité du sujet est liée à la nature même de l'événement sonore - musique, bruit, voix - et donc du type d'écoute.

L'approche des écoutes de la ville par Amphoux (1991) prend comme point de départ le monde sonore et le divise ensuite en éléments ou critères plus précis - approche inverse de celle de la psychophysique. Cela permet de conserver l'intégralité des notions (architecturales, topologiques et socioculturelles) nécessaires à la description de certains phénomènes. En ce sens, cette approche des écoutes est voisine des attitudes adoptées par Gaver et par Schaeffer, en considérant les objets sonores comme éléments de base pour décrire le monde sonore. Toutefois, les trois formes d'écoutes proposées par Amphoux - mémorisée, réactivée et qualifiée - ne distinguent pas tant les façons d'écouter les sons - comme l'entendent Gaver et Schaeffer - que les différentes méthodes d'accès aux représentations mentales des individus. Ces méthodes sont à la base de la définition des tâches et des consignes à donner aux sujets dans les expériences.

### 6.3 La cognition et la psycholinguistique

Nous avons vu que les études concernant la perception auditive et centrées sur l'individu sont très rares dans la littérature en acoustique. Si la Gestalttheorie et la psychologie écologique ont permis de réorienter les débats sur des considérations plus "naturelles" à la perception humaine, elles ne tiennent pas compte du fait que le sujet n'est pas nécessairement passif, peut être acteur d'une situation et peut avoir des intentions. Ces approches demeurent physicalistes en supposant l'existence de lois d'invariance de structures et de transformations.

Il en va de même pour les travaux de Edworthy et Stanton (1995) qui s'appuient sur l'acoustique écologique développée par Gaver (1993). En effet, l'accent est porté sur l'individu pour recenser les situations nécessitant une urgence ou pour déterminer l'urgence de signaux existants ou nouveaux, mais l'écoute reste passive et décontextualisée pour le choix et la comparaison des signaux. Le fait de demander aux sujets des descriptions verbales standardisées met ces derniers dans une situation de contrainte, de "normalisation" des individus, sans leur laisser la possibilité de s'exprimer selon leurs propres représentations mentales.

Schaeffer (1966), au contraire, implique le sujet dans une démarche volontaire et intentionnée. Cette approche a été possible grâce notamment à un parallèle audacieux avec les travaux issus de la linguistique structurale. Toutefois, la recherche de l'*objet sonore* - équivalent musical du *phonème* linguistique - nécessitant une écoute réduite, dénuée de sens et de valeur, renvoie Schaeffer à une approche gestaltiste de la *structure* sonore - équivalent acoustique de la *forme* visuelle. Si Schaeffer marque la différence entre musique et linguistique en minimisant le rôle du sens des signes musicaux devant les signes linguistiques, il en va tout autrement des signaux d'avertissement qui nous concernent, puisque ces derniers, à l'instar des mots ou des phrases, doivent être *compris* et pas seulement entendus (cf. § 5.3.3). Cette distinction entre "musique à entendre" et "signaux à comprendre" met en évidence la différence des objets visés par les différentes perspectives. De la même manière, la parole s'inscrit plutôt dans une approche de compréhension, tandis que la voix peut être entendue plutôt comme un instrument.

## 7. Conclusion

Si les représentations physicalistes sont essentielles et nécessaires à la description des signaux d'avertissement perçus en contexte urbain, elles ne suffisent pas à en donner une description exhaustive, car elles visent un objet physique et non sa représentation psychologique par l'individu. C'est là que la psychologie de la perception joue un rôle essentiel pour sortir les approches physicalistes de l'impasse méthodologique dans laquelle elles se trouvent. Et pour cause : c'est le sujet, et non la source sonore, que la psychologie met au centre de la problématique. Ainsi, de la psychologie de la forme à la psychologie cognitive et la linguistique, en passant par la psychologie écologique, il devient possible de décrire la perception des sons en liaison avec les *représentations mentales* des auditeurs. Ces représentations en mémoire d'événements sonores signifiants pour les individus sont propres à chaque personne, mais sont souvent partagées, en partie, par une collectivité, de même culture notamment.

Cette démarche devrait permettre, entre autres, de mettre en relation une partie des interprétations des sujets avec les représentations physiques du son. Ces considérations, fondées sur le sens donné par les individus aux événements sonores, conduisent donc à un nécessaire renouveau méthodologique. La méthodologie que nous avons adoptée fait l'objet du chapitre suivant.



# Chapitre 2 : Approche méthodologique

Personne ne peut peindre des êtres humains tels qu'ils sont  
et avoir le sentiment de pouvoir réaliser  
quelque chose qui exprimera le réel  
*Mark Rothko*

## 1. Introduction

Le chapitre précédent avait pour objectif de recenser des travaux relatifs à la perception des signaux sonores dans un contexte donné, en faisant appel, à la fois aux sciences physiques et aux sciences humaines. Il en résulte une bibliographie à caractère fortement pluridisciplinaire. Ces différentes approches doivent être prises en compte dans la recherche que nous proposons de mener, afin de pouvoir mettre en relation les approches psychologiques, liées au sujet et à son expérience personnelle, avec les approches physicalistes, rapportées au signal.

Le présent chapitre aborde la mise en perspective de nos travaux, d'un point de vue méthodologique cette fois-ci. Nous allons nous efforcer de justifier, un à un, les nombreux choix méthodologiques issus de cette pluridisciplinarité, nécessaires à l'élaboration d'une approche originale et rigoureuse des signaux sonores d'avertissement en milieu urbain. Ces choix concernent la qualité des échantillons sonores à tester (§ 2), les procédures de réalisation des tests d'écoute (§ 3), centrées sur les sujets, et le traitement des données recueillies, de manière à objectiver la signification des signaux dans les représentations mentales - en mémoire et signifiantes - des individus (§ 4).

## 2. Les échantillons sonores

Dans tout travail expérimental, c'est avant tout le choix du corpus de stimuli qui détermine l'orientation et la qualité de la recherche. En effet, la variété des échantillons considérés conduit les sujets à adopter des stratégies de réponses plus globales ou plus analytiques, selon le cas. Par ailleurs, des échantillons de qualité insuffisante peuvent gêner les sujets et susciter des réponses décalées par rapport à l'objet étudié. Il est par conséquent primordial de correctement définir un corpus de stimuli en fonction du type de recherche menée. En ce qui concerne les signaux sonores d'avertissement rencontrés dans une ville, les choix seront déterminés par la validité écologique des échantillons sonores (§ 2.1) et par les situations dans lesquelles ces signaux apparaissent généralement (§ 2.2).

### 2.1 Des échantillons sonores écologiquement valides

La première condition à laquelle les stimuli sonores sont soumis concerne leur qualité d'écoute lors des tests effectués en laboratoire. Seules des configurations expérimentales permettant aux sujets de répondre conformément à leurs attitudes en situation urbaine peuvent être retenues ici, puisqu'il s'agit d'une de nos hypothèses de recherche. Tous les éléments pouvant intervenir dans cette démarche doivent par conséquent faire l'objet d'une grande attention.

#### 2.1.1 Du quotidien au laboratoire

La première difficulté à laquelle on est confronté dans une étude environnementale est essentielle : comment peut-on transposer des conditions naturelles et quotidiennes à une situation "*in vitro*" de laboratoire, *a priori* artificielle ? En effet, l'expérimentation "*in vitro*" a pour conséquences une *abstraction* considérable du monde normalement perçu, ainsi qu'une implication de l'individu totalement différente quant à sa lecture du milieu. Dans le cas d'une étude acoustique, le fait d'enregistrer des séquences sonores revient à *abstraire* des objets sonores du monde réel ; il y a décontextualisation. On isole ces objets des autres modalités sensorielles qui contribuent à une perception globale de l'environnement. Par conséquent, lors d'une écoute en laboratoire, il faut supposer que les auditeurs ont effectué ce travail d'abstraction. Il s'agit de justifier le découpage que l'on effectue et de rechercher un compromis entre la qualité de la reproduction (le niveau de décontextualisation) et la préparation des sujets à l'écoute des sons. Ce dernier point sera abordé au § 3. Dans un premier temps, il faudra constituer des échantillons sonores pertinents pour les individus. En effet, si la structure même de l'échantillon n'est pas valide expérimentalement, les sujets peuvent basculer dans une logique représentationnelle différente de celle souhaitée dans l'expérience et ainsi, remettre en cause les résultats et la validité même de l'étude.

#### 2.1.2 Les apports de la psychologie écologique

Le concept de validité écologique, développé par la psychologie écologique (Gibson, 1979 ; cf. chapitre 1, § 5.1.1), nous donne les moyens méthodologiques pour effectuer certaines abstractions du monde sensible sans pour autant altérer l'accès aux représentations en mémoire des individus. En effet, l'idée principale réside dans le fait que les sujets, écoutant les échantillons sonores en laboratoire, doivent réagir "comme si" ils se trouvaient en situation

réelle. La tâche de l'expérimentateur ne consiste toutefois pas à restituer ces stimuli sonores de la façon la plus réaliste possible, mais de donner les informations nécessaires pour atteindre les mêmes représentations en mémoire que les individus avaient établies lors d'une perception directe, en situation. Sans cette exigence, il est fortement probable que nous ne travaillions pas avec les mêmes objets cognitifs que ceux que nous voulons étudier ; et l'étude de perdre tout fondement.

S'assurer de l'adéquation entre la réaction des sujets en laboratoire et celle au quotidien consiste alors à tester l'ensemble de la chaîne sonore, de l'enregistrement à la restitution. Le système le plus efficace sera celui qui permet aux individus de s'intégrer dans la séquence sonore avec l'équipement technique le plus léger possible. Il fera l'objet de tests d'écoute pour valider expérimentalement les techniques de prise de son, ainsi que les configurations de diffusion sonore dans un espace d'écoute donné. Par ailleurs, la validité écologique nécessite également des séquences sonores pertinentes pour les auditeurs ; c'est l'objet du paragraphe 3.

### 2.1.3 Ecologie et psychologie cognitive

Si la psychologie cognitive a beaucoup analysé les processus perceptifs, conçus essentiellement comme des processus "bottom up" (cf. chapitre 1, § 5.3.2.3), la perspective écologique mise en oeuvre ici impose de prendre en compte également les processus "top down". En effet, il faut s'assurer d'une part, que les sujets accèdent aux représentations *préalables* telles qu'elles ont été construites par l'expérience quotidienne (cf. § 2.1.1) et d'autre part, que tous les effets de contexte qui ont orienté les traitements perceptifs sont bien pris en compte (cf. § 2.2). La perception - au sens cognitif - des sujets et leurs interprétations sont le résultat de ce double processus "bottom up" et "top down". Cette conception de la perception, qui intègre donc un certain nombre de traitements cognitifs - "top down" -, se différencie de celle utilisée en psychophysique (cf. chapitre 1, § 2) et qui ne considère que le processus "bottom up" pour la définir.

## 2.2 Importance du contexte d'écoute

La deuxième contrainte sur les échantillons sonores n'est pas indépendante de la première, la validité écologique des stimuli : le *contexte* urbain, dans lequel émergent les signaux d'avertissement qui nous intéressent, est également chargé de sens et aide l'individu à comprendre la "scène sonore". Il s'intègre par conséquent dans le concept de validité écologique, cher à la qualité des stimuli.

### 2.2.1 Le contexte d'écoute dans les études recensées...

Parmi les travaux décrits dans le chapitre précédent, certains prennent en compte le contexte d'écoute pour étudier des signaux sonores. Ainsi, les psychoacousticiens modélisent le contexte par un bruit stationnaire large bande, blanc ou rose, afin d'en maîtriser parfaitement les paramètres acoustiques (Haas et Casali, 1995). D'autres auteurs se sont aventurés dans l'utilisation de bruits de fond réels, mais mis en boucle sur quelques secondes, afin de conserver un caractère relativement stationnaire (Binnington *et al.*, 1993 ; Szeto *et al.*, 1991c). D'autres encore n'ont pas hésité à enregistrer des bruits de fond très fluctuants, mais ont pris le parti de moyenniser ces fluctuations sur l'ensemble de chaque séquence en mesurant des niveaux en dBAeq (Ebata et Baba, 1997). Toutefois, certaines études émettent l'hypothèse de l'influence

du contexte d'écoute sur la reconnaissance et l'identification des signaux (Gaver, 1993 ; Leipp, 1965), mais n'en testent pas explicitement l'importance. Par exemple, si le son d'un klaxon de voiture peut paraître extrêmement pénible dans une ambiance de jardin public, même animée, il pourra être aisément accepté dans une rue, même tranquille.

## 2.2.2 ... et dans notre étude

Nous avons déjà mentionné l'effet lié à l'*extraction* de son hors du monde physique (cf. § 2.1.1). En fait, il est particulièrement difficile de déterminer, *a priori*, dans quelle mesure des éléments appartenant au soi-disant "contexte" sont parties prenantes de la représentation individuelle - ou collective - des phénomènes sonores. C'est cette influence contextuelle qu'il nous a paru essentiel de prendre en considération pour la perception des signaux d'avertissement en situation urbaine. Le moyen d'y parvenir est d'analyser, dans la représentation mentale des sujets, l'identification et la reconnaissance des signaux selon différentes situations contextuelles rencontrées.

Un autre effet concerne la *manipulation* de paramètres physiques définis *a priori* lors de la prise de son ou lors du choix des séquences sonores. L'expérimentateur décide de sélectionner tel événement sonore plutôt qu'un autre, sur des considérations généralement physicalistes. Cette manipulation par rapport à l'ambiance globale ne doit pas "détourner" les sujets vers d'autres représentations mentales, non acquises dans l'expérience quotidienne, mais activées (ou sollicitées) de manière élective dans les situations de laboratoire.

## 2.2.3 Des signaux contextualisés

Dubois (1993a) insiste sur le fait que, "pour qu'il y ait dialogue, il faut un *contexte commun* aux interlocuteurs". Appliqué aux signaux d'avertissement en situation urbaine, ce contexte comprend des objets et des événements sonores connus, codifiés et en relation avec un système de valeurs, des intentions... qui permettent à l'individu de *répondre* au signal communiqué. Cet échange d'informations signifiantes soulève le problème du caractère "social" ou "naturel" des stimuli dans les expériences (Dubois, 1996) : dans quelle mesure la situation expérimentale correspond-elle à une "fiction" collective, dont la validité écologique dépend en grande partie du "contrat" entre l'expérimentateur et les sujets ? Ces derniers sont-ils prêts à "jouer le jeu", à accepter les contraintes de la situation expérimentale ?

Une manière relativement simple de constituer un corpus d'échantillons sonores écologiquement valides consiste à enregistrer directement les séquences en y intégrant les signaux d'avertissement. Le problème d'une telle démarche est l'absence totale de maîtrise de la structure sonore de l'environnement au moment de l'apparition du signal et les difficultés de mesure de niveaux qui en découlent. Par ailleurs, il est particulièrement ardu de comparer des signaux entre eux dans un contexte donné, puisque ce dernier, fluctuant par nature, ne peut être enregistré qu'à des instants différents. Une solution à ce problème peut être apportée en enregistrant le contexte *indépendamment* des signaux, puis en mixant ces derniers avec la *même* séquence contextuelle pour tous. L'intérêt de mettre des signaux en contexte par l'intermédiaire d'un montage/mixage est, non seulement de conserver des références relatives aux différents signaux et de faciliter ainsi les mesures physiques, mais aussi et surtout de maîtriser les éléments contextuels qui peuvent s'intégrer dans les représentations en mémoire des sujets. Il en découle néanmoins deux autres difficultés : d'une part, le montage doit être réalisé de manière très précise et s'intégrer dans le déroulement classique d'un scénario sonore - schéma *narratif* de la situation -, afin de procurer à l'auditeur l'illusion d'une scène réelle ; d'autre part, les contextes doivent être choisis de manière à être *représentatifs* des ambiances mémorisées par les

individus. Cette seconde hypothèse de travail est traitée au chapitre 3, § 2.1, à partir d'une procédure développée par Amphoux (1991).

Quant à l'acceptabilité de la situation expérimentale par les sujets, elle dépend fortement des *consignes* qui leur sont données et qui orientent leurs activités cognitives. La procédure expérimentale, dont les consignes font partie, est l'objet du paragraphe suivant.

### 3. Les procédures des situations d'écoute

#### 3.1 La validité écologique par rapport aux sujets

Si la validité écologiques des échantillons sonores est essentielle dans la réalisation de tests d'écoute, il faut également s'assurer que les processus mis en oeuvre en réponse à des stimuli soient identiques à ceux développés dans l'environnement quotidien. Ainsi, l'artefact dû à une abstraction des objets sensibles du monde peut être contrebalancé par la consigne qui oriente différemment l'activité des sujets. Le choix de la consigne repose sur une connaissance préalable des traitements effectués et des stratégies adoptées par les sujets pour aboutir à une réponse (cf. Dubois, 1996). Nous nous plaçons du point de vue théorique de la *catégorisation prototypique* (cf. chapitre 1, § 4.3) qui considère que les concepts humains sont basés sur l'expérience des individus et que la pensée a une valeur écologique, c'est-à-dire sachant s'adapter aux contraintes extérieures. Nous nous intéresserons avant tout au *prototype*, cette "forme moyenne" qui réunit les propriétés communes aux objets d'une catégorie.

Le cadre théorique étant précisé, notamment par rapport à la catégorisation prototypique développée par Rosch (cf. § 2.1.3 et chapitre 1, § 4.3), il nous reste à déterminer les consignes à donner aux sujets lors des expériences en laboratoire, afin d'atteindre les objets étudiés. En effet, la consigne détermine l'attitude adoptée par les sujets lors de la tâche cognitive qu'ils ont à effectuer durant le test.

#### 3.2 Le choix des sujets

Les études présentées au chapitre 1 ont mis en évidence que les résultats sont, non seulement liés au choix du corpus sonore, mais également au choix des sujets, selon leur âge, leur sexe, leur culture, leur niveau d'expertise... L'âge est souvent relié à des problèmes de pertes auditives (Szeto *et al.*, 1991 ; Robinson et Casali, 1995), mais ces dernières s'avèrent de faible importance pour les types de stimuli qui nous concernent et dans la mesure où ces pertes ne dépassent pas 40 dB (cf. chapitre 1, § 6.1.1). Par ailleurs, Lass *et al.* (1983) émettent l'hypothèse que l'identification de la voix parlée masculine ou féminine peut dépendre du sexe des auditeurs, mais sans vraiment explorer la question. Toutefois, cet effet ne semble pas intervenir pour des stimuli sonores autres que des voix humaines. Enfin, concernant la culture et l'expertise des sujets, Lass *et al.* (1982) remarquent que les étudiants en psychologie qu'ils questionnent ne répondent certainement pas de la même manière que des personnes issues du milieu rural à des stimuli d'animaux de la ferme. Et il est frappant de constater que presque toutes les études recensées sont réalisées avec des sujets - généralement des étudiants, - spécialisés dans les domaines de recherche liés aux expérimentateurs (psychologie et acoustique). Seuls Szeto *et al.* (1991) ont distingué les sujets, en demandant à des personnes non-voyantes et/ou âgées d'effectuer un test.

Partant de ce constat, nous ne sélectionnerons que des sujets usagers de la ville et donc, habitués au type de stimuli de nos expériences. Nous essayerons, dans la mesure du possible, de trouver à la fois des sujets acousticiens, "experts" du sonore, et d'autres issus des milieux sociaux et culturels les plus variés possibles, "représentatifs" de la population urbaine. Cette représentativité est évidemment limitée du fait du nombre restreint de sujets, une approche sociologique, avec plusieurs centaines de sujets, étant impensable dans le cadre de cette étude.

### 3.3 Tâches demandées aux sujets

Comme nous avons déjà pu le voir (§ 3.1), ce sont les tâches qui impliquent des traitements cognitifs. Données sous forme verbale, les consignes sont généralement supposées neutres par rapport à l'interprétation donnée à la tâche par les sujets. Toutefois, elles imposent souvent des contraintes *a priori* qui limitent les possibilités de réponse des sujets (Dubois, 1996).

#### 3.3.1 Les questionnaires à choix forcés

La grande majorité des travaux expérimentaux demandent aux sujets de comparer des objets entre eux, d'en choisir un parmi plusieurs, de les caractériser selon un certain nombre d'indices ou dimensions prédéfinis... Dans toutes ces situations, les sujets sont contraints de répondre par "vrai ou faux" ou de cocher une case parmi d'autres, sur des catégories de dimensions données. Il s'agit de "choix forcés" sur des dimensions et critères définis *a priori*. Cette démarche méthodologique, qui facilite considérablement le traitement des données, suppose que l'expérimentateur a déterminé *préalablement* des concepts et des critères qui, en conséquence, seront les seuls traités dans l'expérience. Seule une - ou plusieurs - des données mises en entrée du système pourra - pourront - apparaître à la sortie du système. On valide donc l'adéquation du modèle du sujet avec celui du chercheur.

#### 3.3.2 La tâche de détection

Les expériences de détection du seuil minimal de perception d'un signal dans un bruit de fond, telles que les ont menées Szeto *et al.* (1991) et Binnington *et al.* (1993) notamment (cf. chapitre 1, § 3.2), permettent de calibrer le niveau perçu d'un signal par rapport à un contexte sonore donné. Il s'agit d'une approche psychophysique tout à fait classique qui ne questionne pas la signification des signaux à leurs niveaux seuils perçus. Nous adopterons la même démarche pour calibrer nos signaux, prérequis à l'étude des significations, afin de nous assurer d'être au-dessus des niveaux physiologiques de perception. Mais il ne s'agit ici que d'une précaution méthodologique et non d'une étude finalisée en soi, la procédure essentielle étant l'identification des signaux entendus à leurs niveaux seuils, sans que les sujets n'aient connaissance des signaux utilisés. Cette démarche permet d'appréhender la reconnaissance des signaux dès leur émergence perceptive et selon le contexte d'écoute.

#### 3.3.3 La tâche d'identification

Pour identifier des objets, les sujets ont souvent le choix entre plusieurs possibilités, mais restreintes à celles proposées par l'expérimentateur. Dans notre cas au contraire, pour appréhender la signification des signaux d'avertissement, nous demandons aux sujets d'identifier les signaux entendus dans les séquences sonores. Les sujets n'ont aucune contrainte langagière, mais doivent s'exprimer, soit oralement, soit par écrit, selon les expériences. La tâche "ouverte" à réaliser consiste donc en un travail d'accès aux représentations en mémoire sans orienter l'activité des sujets vers des traitements analytiques, comme l'exige une tâche de jugement sur un critère ou une dimension.

### 3.4 Les réponses : la verbalisation

Les productions verbales, en guise de réponses aux expériences, peuvent se présenter sous différentes formes : un discours, un énoncé, un mot... Par ailleurs, elle dépend d'un processus de décision du locuteur, qui peut choisir, sur le plan lexical : les mots, en fonction de leur disponibilité dans la langue apprise et de leur familiarité ; sur le plan syntaxique : l'ordre des mots ; sur le plan pragmatique, le type d'acte de parole (Dubois, 1993a).

Les productions verbales des individus fournissent un accès à leurs représentations mentales, à condition d'explicitement la relation entre langage et cognition à l'aide d'une théorie psycholinguistique. La psychologie cognitive a largement utilisé les "verbalisations" pour accéder aux représentations et aux activités cognitives lors de réalisations de tâches cognitives complexes. Les verbalisations des sujets diffèrent selon que ces derniers doivent les effectuer simultanément à la tâche principale ou *a posteriori*, par oral ou par écrit.

Dans le cadre des tests d'écoute, la verbalisation aura lieu, soit par oral pour être immédiatement retranscrite par l'expérimentateur, soit directement par écrit, ce qui laisse un temps de réflexion plus important - et un choix différent des productions verbales. Dans notre démarche, qui prend en compte la diversité des mots, d'organisation tant linguistique que cognitive (David *et al.*, 1997), il est essentiel de n'imposer aucune contrainte lexicale - mots imposés. Ainsi, quelle que soit la consigne, les sujets doivent disposer de leur liberté langagière, à la différence des tâches de type différenciateur sémantique - "choix forcé" entre deux ou plusieurs mots, cf. § 3.3.1 - qui imposent aux sujets des *a priori* lexicaux.

## 4. Le traitement des données

### 4.1 Respect des formes linguistiques produites par les sujets

Les productions issues des verbalisations ont le grand avantage d'être riches puisqu'elles contiennent toutes les informations données par le sujet, avec leurs nuances langagières. Ceci n'est pas le cas pour des procédures de type "vrai ou faux", pour lesquelles les termes sont imposés par l'expérimentateur, et donc nécessairement limités à des expressions "standardisées" et restreintes *a priori*. Cependant, une conséquence des verbalisations est la grande quantité de données recueillies. La première étape du traitement de ces données consiste à *trier*, sans élimination ni interprétation, les éléments utiles à l'analyse, issus des travaux de la linguistique : mots, ordre des mots, aspects grammaticaux... Les sujets ayant pour tâche d'identifier les sons perçus, nous nous intéresserons avant tout aux mots et groupes de mots formés, sans tenir compte des discours des sujets et notamment des marques de la personne (cf. David, 1997). A partir de ces différentes méthodes de tri et des théories issues de la psycholinguistique, il nous est ensuite possible d'établir un certain nombre de conclusions concernant les représentations mentales des sujets. Nous limiterons notre recherche à une analyse lexicale et syntaxique des données recueillies.

### 4.2 Les organisations cognitives

Le traitement des verbalisations s'opère en plusieurs parties afin d'en extraire un maximum d'informations. Dans un premier temps, il s'agit de dénombrer les différentes occurrences produites. Cela nous donne des informations quant à l'ensemble du répertoire lexical disponible pour le groupe des sujets interrogés. Dans un deuxième temps, il faut compter le nombre d'occurrences pour chaque type de mot différent produit. Cette donnée nous permet d'estimer les désignations les plus fréquentes et les plus disponibles dans le répertoire lexical collectif. Une troisième étape enfin consiste à chercher les désignations les plus fréquentes associées à chaque signal étudié, afin d'établir le degré de typicalité d'une désignation associée à un signal sonore donné.

### 4.3 L'acoustique en aval

Cette mise en relation de productions verbales pertinentes pour l'individu et d'objets sonores mis dans divers contextes constitue le point essentiel de la recherche de propriétés acoustiques liées à la signification d'un signal. En effet, l'élément le plus prototypique d'une occurrence linguistique produite est également l'élément le plus significatif de la représentation mentale d'un objet lexical donné. Les propriétés signifiantes du son, si elles existent et ne sont pas complètement arbitraires, d'un prototype associé à une expression verbale, doivent donc se distinguer des autres prototypes associés à d'autres productions langagières. L'objectif est donc de repérer ces propriétés signifiantes par relation de *différences* entre les contenus spectro-temporels des signaux acoustiques prototypiques d'expressions distinctes, et par relation de *ressemblance* entre les structures sonores des signaux associés à un même élément linguistique (cf. chapitre 1, § 4.3).

## **5. Conclusion**

La grande originalité de l'approche méthodologique que nous venons de décrire réside dans sa très nette orientation vers le sujet et sa perception du monde sonore. Trois points essentiels lui confèrent sa nouveauté dans le domaine de la recherche en acoustique : dans un premier temps, la validité écologique des échantillons sonores nous assure de travailler avec des stimuli conformes aux représentations mentales des individus ; dans un second temps, les tâches spécifiques données aux sujets nous permettent d'orienter leurs stratégies de réponse vers des comportements conformes aux situations réelles que l'on veut tester en laboratoire ; dans un troisième temps, le traitement des données verbales recueillies doit s'opérer dans le respect des formes linguistiques produites, ce qui demande en contrepartie un long travail de tri.

La mise en oeuvre expérimentale de cette approche constitue la suite de ce mémoire. Elle consiste tout d'abord à élaborer une chaîne complète d'enregistrement et de restitution d'échantillons sonores, dans le respect du concept de validité écologique (chapitre 3). Deux expériences de perception des signaux d'avertissement en contexte urbain seront ensuite réalisées et analysées. L'effet d'émergence perceptive des signaux sur le fond sonore et l'effet du contexte urbain seront notamment étudiés (chapitre 4 et 5). Enfin, à partir des conclusions relatives aux deux expériences mises en commun, nous tenterons de trouver des corrélations entre les significations données aux signaux par les sujets et les propriétés acoustiques des signaux (chapitre 6).

# Chapitre 3 :

## Expériences préliminaires relatives à la validité écologique

les mots n'ont plus un sens parce que les choses ont un être  
mais parce que les sujets leur accordent une signification  
*Dubois (1996)*

### 1. Introduction

L'enregistrement des ambiances sonores et leur restitution dans des conditions écologiquement valides constituent des étapes essentielles à la réalisation de tests d'écoute. Aussi s'agit-il de porter une attention particulière au matériel, mais également aux techniques de prise de son (§ 3) et de restitution sonore (§ 4) utilisés. La validité écologique des écoutes signifie alors qu'il faut donner à l'auditeur l'illusion d'un *paysage sonore* - en tant que reconstruction esthétique d'une scène sonore - tel qu'il peut le rencontrer dans la vie quotidienne (cf. chapitre 2 § 2.1 et 3.1). Par conséquent, nous avons réalisé plusieurs tests de prise de son (§ 5 et 6) et de restitution sonore (§ 4) afin de trouver les configurations microphoniques et d'écoute permettant de créer cette illusion auditive (cf. Vogel *et al.*, 1997a et 1997b).

Un autre point relatif à la validité des échantillons sonores concerne le choix des ambiances représentatives pour les individus. Ces environnements sonores sont toujours associés à des lieux mémorisés par les sujets, lieux qu'il s'agit de repérer afin de pouvoir y enregistrer ces ambiances représentatives (§ 2).

Cette étude préliminaire a été réalisée conjointement avec Valérie Maffiolo (1999) dans le cadre de sa thèse sur la caractérisation de l'environnement sonore urbain.

## **2. Choix des lieux d'écoute**

L'étude de l'environnement sonore nécessite l'établissement d'un corpus d'échantillons sonores qui s'inscrit dans le concept de la validité écologique. Il s'agit donc, dans un premier temps, de sélectionner les ambiances sonores à enregistrer. Notre approche repose sur la représentativité des ambiances sonores urbaines dans la représentation mentale des usagers de la ville étudiée. L'objectif est de développer une méthode de sélection des ambiances sonores que nous considérerons comme représentatives de cette ville. La méthode que nous allons décrire a été appliquée de manière exploratoire à Paris.

### **2.1 Méthode**

Des études s'inspirant de la sociologie ont mis en évidence la possibilité de localiser des ambiances sonores à l'aide de techniques d'enquête auprès des usagers de la ville concernée, que nous complétons ici par des contributions de la psychologie cognitive et de la linguistique. Pour choisir les échantillons sonores à enregistrer dans Paris, nous nous sommes intéressés en particulier à la méthodologie développée par Amphoux (1991), qui tient compte des dimensions acoustique, topologique et socioculturelle des villes (cf. chapitre 1, § 5.2.2). Cette méthodologie a pour but d'objectiver un mode de sélection de terrains qui puissent être considérés comme représentatifs de l'identité sonore de la ville étudiée, en faisant appel à la mémoire et à l'expérience sonores d'usagers de la ville. Nous avons volontairement restreint notre étude à la mémoire sonore et plus particulièrement à la réalisation de ce que Amphoux nomme les "cartes mentales sonores". Pour Amphoux, l'opération consiste dans une première étape à demander à un sujet de dessiner une carte sonore, ce qui oblige ce dernier à changer de logique représentationnelle, en effectuant une cartographie, c'est-à-dire une représentation visuelle, des sons. Dans une deuxième étape, le sujet s'exprime sur des phénomènes sonores vécus - et donc mémorisés.

### **2.2 Procédure et sujets**

Pour constituer un corpus représentatif des terrains qui présenteraient une richesse sonore particulière, nous avons procédé à une enquête dans Paris. Elle consiste à demander à différents usagers de la ville :

- d'établir une représentation graphique de l'espace sonore parisien par la réalisation d'un croquis, en répondant à une question du type : "qu'est-ce que le Paris sonore pour vous ?", puis de commenter le dessin obtenu,
- de citer des lieux ou itinéraires possédant, à leurs oreilles, des qualités sonores particulières.

Pour nos entretiens, nous avons utilisé l'expression "représentation de l'espace sonore" plutôt que celle de "carte mentale sonore" utilisée par Amphoux, afin d'éviter d'inciter le sujet à utiliser une stratégie spécifique d'analyse de la ville, en l'occurrence de type cartographique. Il s'agit en fait de laisser le sujet libre d'aborder le dessin comme il le désire et de représenter les sons ou lieux sonores comme il "l'entend".

Vingt-neuf sujets, usagers de la ville et relativement représentatifs de la population parisienne - âge, sexe, milieu social et culturel - ont été interrogés par deux expérimentateurs. La durée de l'entretien est conçue en principe pour être brève et donner lieu à des représentations graphiques spontanées. D'une manière générale, elle avoisine les trente minutes, mais elle peut s'étaler sur des temps plus longs en fournissant un matériau verbal beaucoup plus riche.

## 2.3 Résultats

### 2.3.1 Le "Paris sonore"

Les données verbales permettent de constater que le Paris sonore est avant tout un ensemble de **bruits de circulation** (90% des personnes interrogées). Ainsi, la circulation est décrite comme : "flot ininterrompu de voitures, bus et camions (arrêt, démarrage, compresseurs qui déchargent), klaxons, bruit de fond nul le dimanche matin (calme, agréable), mobylettes, roulements des pneus sur les pavés (très sonore), scooters, motos, démarrage aux feux, moteurs, bruits sur la chaussée, bruit de fond permanent troué de temps en temps par une sirène."

Associés aux bruits de circulation, nous trouvons également mentionnés :

- le **Boulevard Périphérique** (41%), décrit en tant que : "bruit permanent, rumeur, flot de voitures incessant, roulement ininterrompu, sourd, obstinant, monotone, on perd ses sens, assourdi."
- les **signaux d'avertissement** (41%), décrits comme : "klaxons, pinçons (bruits aigus), sifflet de l'agent de police à un carrefour, signal pour aveugles aux passages piétons, bruits de travaux (marteaux-piqueurs, grues, échafaudages), passage des poubelles."

D'autres descriptions concernent des sons liés à l'activité humaine ou des sons naturels :

- les **marchés** (45%) : "cris (seulement sur les marchés de primeurs), accostages divers, beaucoup de vie, arabisant, conversations, piétinements, agréable, bruits des cageots, pose des tentes, installation, nettoyage, bruit des poteaux ramassés..."
- les **gens** (31%) : "discussions, banc de touristes discutant haut et fort, jardins publics (conversations pour meubler, pleurs, cris, bousculades, poussettes), commerces de proximité, bars, gens sortant du restaurant tard le soir, ambiance musicale, vitrier, cafés (très riche, percolateur, bruits de verres et de couverts, conversations, caisse, téléphone, flippers, terrasses avec des gens qui font de la musique), bruits et musique qui sortent des appartements en été, manifestations (piétinement, slogans, chants), sorties le soir (gens qui bougent, qui parlent, qui s'amusent), foule, grands magasins..."
- les **Quais de Seine** (41%) : "piétons (calme), vagues, bateaux-mouches, péniches (bruit atypique, bruit régulier grave d'un moteur), calme relatif de la Seine par rapport aux axes routiers, moins de bruits continus, ambiance sonore résiduelle différente, agréable car lié à l'eau, plus aéré, bruit continu de voitures, flux continu de voitures dans un sens, voitures (assourdissant)..."
- les **Oiseaux** (34%) : "vol et roucoulements des pigeons, chant, gênants à cinq heures du matin..."

Enfin, le **Métro** (41%) tient une place à part, puisque généralement souterrain. Il est illustré par : "roulements, pneumatiques (portes, freins), signaux de fermeture, couinements dans les virages, grincements, frottements, tourniquets, métro aérien (Bir-Hakeim), voûte (caractéristique, discussions des gens d'en face), baladeurs, engueulades entre clochards, gens qui parlent tout seuls, pas, bruits de la pluie sur le métro aérien, annonces (plan vigipirate), vente ambulante, musique..."

### 2.3.2 Des ambiances représentatives de Paris

Les données verbales permettent aussi d'identifier les ambiances sonores des lieux cités le plus souvent dans les récits des sujets. En particulier, nous retrouvons les ambiances : du boulevard périphérique, des boulevards Saint-Germain et de l'Hôpital, de la place de la Bastille, du jardin des Plantes, des buttes Chaumont et du bois de Vincennes, des marchés d'Aligre et de la rue Mouffetard, des animations de Montmartre, des Halles et du parvis de Beaubourg, des promenades du Marais et de la Coulée Verte. Ces ambiances sont retenues comme les plus représentatives de Paris.

Une des difficultés de cette méthodologie est la traduction d'un percept sonore en une représentation graphique, ce qui suppose d'articuler ces recherches avec les travaux de la psychologie cognitive. En effet, quatre types d'interprétations dans le rapport à la représentation de l'espace sonore peuvent être mis en évidence (Figure 3-1) :

- des cartographies de Paris (20%),
- des lieux spécifiques avec des noms propres ou génériques (23%),
- des sources non organisées dans un site (50%),
- des schémas abstraits (13%).

Certaines personnes adoptent plusieurs stratégies ce qui explique un total de 106%.

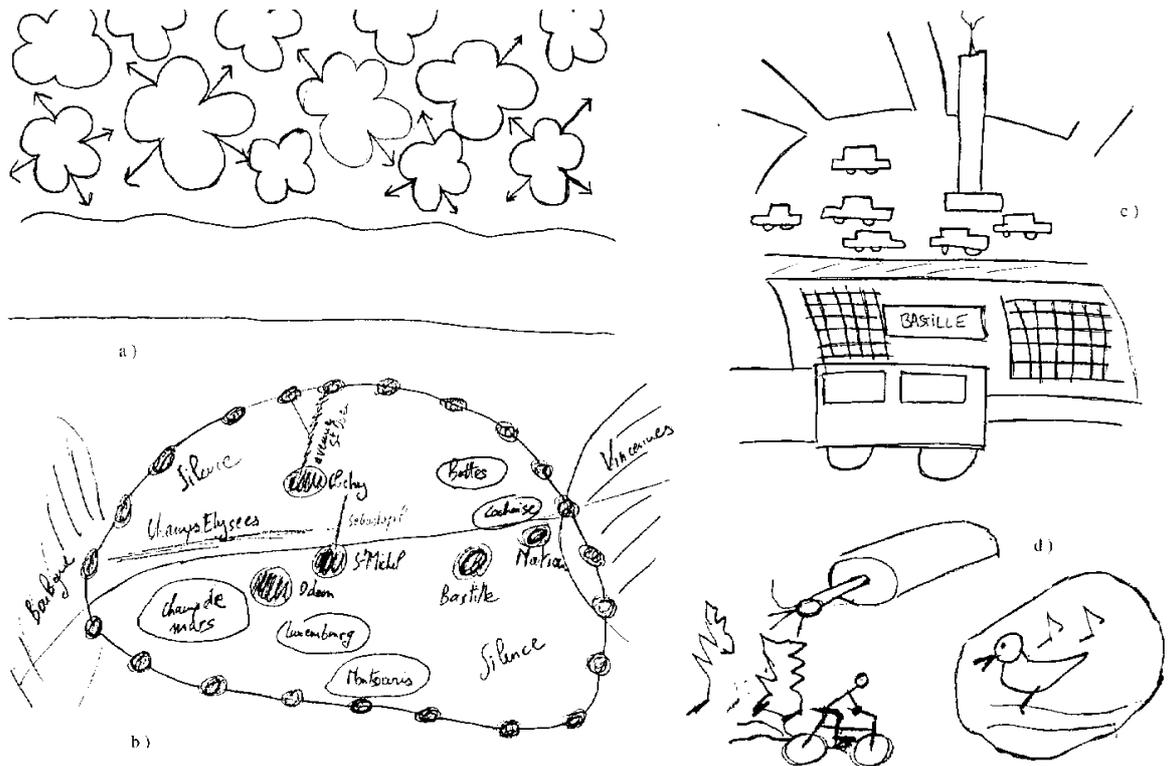


Figure 3-1 : 4 types de représentation de l'espace sonore : a) schéma abstrait, b) cartographie, c) lieux spécifiques, d) sources non organisées dans un site.

Cette différenciation entre les modes de représentations se distingue de celle réalisée par Amphoux (1991) selon laquelle il considère trois formes générales de représentations (plan général, itinéraire et ensemble d'unités fragmentaires et discontinues), ainsi que trois modes spécifiques de représentations (symbolisation codifiée du lieu, expression d'un certain type d'espace et icône représentant le type de sources sonores). L'une des raisons de cette différence provient du fait que Amphoux donne comme consigne à ses sujets de dessiner une "carte

sonores". De ce fait, il les incite à utiliser une stratégie cartographique et ne leur laisse guère la liberté de représentation par schémas abstraits, sources sonores ou lieux spécifiques.

A travers l'analyse de la verbalisation employée par les sujets lorsque nous leur demandons de commenter leurs représentations graphiques ou d'établir un répertoire des terrains présentant de la richesse sonore, nous constatons que l'interlocuteur parle du lieu de différentes manières :

- en termes de description topologique ("c'est un marché en longueur", "c'est un marché sur une place"),
- en termes de données temporelles ("oiseaux gênants à cinq heures du matin", "le matin la circulation est moins pesante, le soir elle est plus segmentée"),
- en termes d'appréciation qualitative ("c'est agréable", "c'est bruyant"),
- selon le type de lieu, la place que le sujet occupe dans ce lieu et le type d'activité qu'il peut avoir. En effet, l'analyse linguistique (David, 1997) montre que les qualificatifs utilisés nous permettent de mettre en évidence plusieurs comportements du sujet. Est-il à l'intérieur ou à l'extérieur du lieu ? S'il est à l'intérieur, est-il actif ou pas ? S'il est actif, est-il ou non producteur de bruits ? Nous distinguerons ainsi l'observateur extérieur, l'observateur immiscé, l'intervenant et l'interagissant. Il apparaît donc que le sujet structure et décrit le lieu en fonction de lui-même.

En prenant en compte toutes les façons d'appréhender et de parler d'un lieu, il est alors possible d'une part, de comprendre ce qu'est un lieu du point de vue de l'interlocuteur et comment il se le représente et d'autre part, d'analyser comment l'espace d'une manière générale est identifié (Maffiolo, 1999).

## 2.4 Des ambiances à enregistrer

Les réponses à l'enquête réalisée avec des usagers de la ville - en l'occurrence de Paris - nous ont conduit à retenir cinq types d'environnements sonores : les ambiances de circulation, de jardins publics, de marchés, de métro et de bords de la Seine. Après plusieurs campagnes d'enregistrement dans la plupart des lieux cités par les sujets, nous avons retenu, après réécoute, les séquences sonores qui sont conformes aux descriptions qualitatives des usagers. Les objectifs de cette thèse et le temps imparti nous ont contraints à limiter notre champ d'étude aux deux principaux types d'environnements sonores urbains : la circulation et les jardins publics. Les enregistrements utilisés concernent les lieux suivants :

- la place d'Austerlitz, le quai Saint-Bernard et le croisement des boulevards de l'Hôpital et Saint-Marcel pour la circulation,
- le jardin du Luxembourg, pour les jardins publics.

Les travaux de Maffiolo (1999) prennent en compte l'ensemble des environnements sonores urbains et s'appuient en partie sur le corpus de séquences sonores enregistrées lors de nos campagnes de prise de son. Les techniques de prise de son utilisées pour enregistrer ces différents environnements sonores urbains font l'objet de la prochaine section.

### 3. Enregistrement des échantillons sonores

La réalisation de tests d'écoute en laboratoire concernant l'environnement sonore nécessite évidemment un enregistrement préalable des ambiances étudiées qui sont sélectionnées selon la méthode décrite précédemment (cf. § 2). Si le matériel de stockage des données acoustique ne pose aujourd'hui plus vraiment de problèmes - à condition de ne pas compresser le son - le résultat à l'écoute dépend fortement de la *manière* dont le son est enregistré, c'est-à-dire la technique de *prise de son* adoptée.

#### 3.1 Matériel utilisé

Toute manipulation de son nécessite un traitement du signal sonore (dynamique, contenu spectral ou temporel) sans introduire de modification perceptible pour les auditeurs. Le matériel permettant de stocker nos informations sonores, première étape avant tout traitement puis restitution, a été choisi dans ce respect de fidélité. Il s'agit d'un *enregistreur numérique* de bandes D.A.T., portable (hhb Portadat) pour une utilisation en extérieur. La fréquence d'échantillonnage utilisée a été fixée soit à 48 kHz, soit à 44,1 kHz (norme C.D. audio, D.V.D...).

Par ailleurs, les enregistrements sont réalisés à l'aide de microphones électrostatiques. En effet, ces derniers ont notamment l'avantage, par rapport aux microphones électrodynamiques, d'avoir une réponse en fréquence quasiment "plate" sur toute la plage audible (de 30 Hz à 18 kHz en moyenne) ; ainsi ils ne colorent pas le son. Par ailleurs, trois directivités différentes ont été testées : omnidirectionnelle, infracardioïde et cardioïde (Figure 3-2). Le microphone *omnidirectionnel* a une réponse en fréquence parfaitement plate et capte les sons provenant de toutes les directions. Le microphone *cardioïde* atténue légèrement les basses fréquences (< 100 Hz avec une pente de 3 dB/octave) et capte principalement le son arrivant dans l'axe de la capsule microphonique. Le microphone *infracardioïde* quant à lui, constitue un compromis entre celui omnidirectionnel et celui cardioïde, en ce sens qu'il atténue légèrement les fréquences inférieures à 100 Hz (1,5 dB/octave) et qu'il privilégie légèrement le son arrivant dans l'axe de la capsule (cf. Hugonnet et Walder, 1995 ; Rossi, 1986).

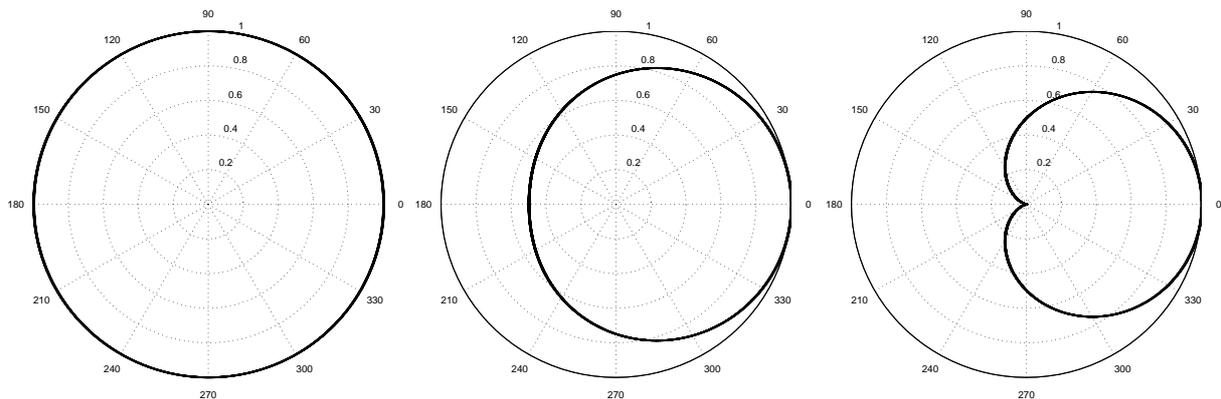


Figure 3-2 : directivités théoriques et normalisées des microphones omnidirectionnel, infracardioïde et cardioïde. L'angle 0° correspond à l'axe de la capsule microphonique.

Nous remarquerons toutefois que les dimensions des microphones impliquent une diffraction du son à hautes fréquences et donc une légère modification des courbes de réponse pour des directivités non nulles. Avec les microphones que nous avons utilisés - Schoeps MK6

(commutable en omnidirectionnel, cardioïde ou bidirectionnel), MK5 (commutable en omnidirectionnel ou cardioïde) et MK21 (infracardioïde) - ces effets deviennent sensibles (soit environ 5 dB par rapport à la théorie) à partir de 8 kHz et 90° d'angle d'arrivée du son par rapport à l'axe de la capsule.

## 3.2 La prise de son

La prise de son a, depuis ses origines - c'est-à-dire en 1877 avec Edison et Cros -, été étroitement liée à la restitution sonore. Les raisons en sont simples, puisqu'il faut généralement au moins autant de microphones que de haut-parleurs pour profiter pleinement des conditions de restitution. Ainsi, la monophonie, la stéréophonie - depuis les années 1930 - et la quadriphonie - dans les années 1970 - ont vu s'adapter la prise de son au type de restitution sonore. Aujourd'hui, ce sont les systèmes de son multicanal basés sur six haut-parleurs ou plus - Dolby surround, spatialisateur de l'IRCAM...- qui tentent de s'imposer. Dans ce cas, la technique de prise de son n'impose pas nécessairement autant de microphones que de haut-parleurs. Toutefois, les progrès liés au mixage/montage des sons conduisent souvent à des prises de sons associant vingt ou quarante microphones ! Ce n'est évidemment pas l'approche que nous allons adopter, pour des raisons matérielles et de mesure acoustique évidentes. Il s'agit en fait pour nous de trouver un système en adéquation avec le concept de validité écologique et le plus simple possible, permettant de contrôler les données acoustiques mesurables.

### 3.2.1 La stéréophonie

Parmi les possibilités offertes, nous avons opté pour la prise de son stéréophonique avec restitution stéréophonique. En effet, lorsqu'elle est bien réalisée, cette technique restitue un espace sonore proche de celui de notre perception quotidienne (2 microphones - 2 oreilles) : elle rend notamment compte de l'espace sonore dans le plan azimutal (horizontal), à la différence de la prise de son monophonique qui ne donne que des informations relatives à la distance auditive des sources, et non par rapport à un espace sonore tridimensionnel tel que nous le permet notre perception binaurale. De plus, le plan azimutal contient la quasi-totalité des sources sonores, qui peuvent donc être localisées à l'aide de la stéréophonie. Par ailleurs, la prise de son stéréophonique est relativement facile à mettre en oeuvre par rapport à une prise de son monophonique multicanaux. En effet, cette dernière consiste à placer plusieurs microphones pour capter l'environnement sonore et de leur associer un angle azimutal pour la restitution stéréophonique - compris entre -30° et +30°. Il s'agit d'une méthode lourde en besoins matériels, difficile à mettre en oeuvre quant au placement des microphones et nécessitant ensuite un mixage des séquences sonores issues de chaque microphone pour aboutir à une écoute stéréophonique.

Quelques restrictions subsistent cependant : d'une part, seule une portion du plan azimutal est restituée, de -30° à +30°, face à l'auditeur. D'autre part, selon la directivité des microphones utilisés, une partie du son arrière ou latéral est "rabattue" sur la zone d'écoute stéréophonique, déformant ainsi la scène sonore. Le plan sagittal (vertical ou élévation), qui concerne uniquement certains sons d'oiseaux ou bruits mécaniques tel le passage des avions, peut également être localisé sur l'avant. Si les nouvelles techniques de restitution sonore multicanal tentent de résoudre ces difficultés, les techniques de prise de son qui leur sont associées n'en sont encore qu'à leurs balbutiements et présentent encore beaucoup de défauts quant à la localisation des sons et surtout, quant à la validité écologique des séquences sonores.

Par ailleurs, la relation entre les phénomènes perceptifs associés à de telles configurations d'écoute et des paramètres acoustiques est très difficilement contrôlable.

La prise de son stéréophonique peut donc être un bon outil pour transmettre la validité écologique des échantillons sonores, mais elle exige un certain "savoir-faire" qui consiste à choisir et à positionner les microphones de façon à *générer l'illusion référentielle* auprès des auditeurs.

### 3.2.2 Systèmes de prise de son

La stéréophonie ayant été retenue pour la prise de son, il subsiste une multitude de possibilités pour réaliser une prise de son stéréophonique. Nous avons donc décidé de tester un certain nombre de systèmes de prise de son stéréophonique afin de déterminer le système le plus écologiquement valide pour chaque type d'ambiance sonore considérée. Trois principes de prise de son stéréophonique peuvent être envisagés (cf. Hugonnet et Walder, 1995) :

- la *prise de son stéréophonique d'intensité*, pour laquelle les capsules des deux microphones (généralement directionnels) coïncident et forment un angle physique  $\theta$  entre elles,
- la *prise de son stéréophonique de temps*, pour laquelle deux microphones omnidirectionnels sont séparés de quelques centimètres et sont parallèles entre eux,
- la *prise de son stéréophonique de temps et d'intensité*, pour laquelle les capsules des deux microphones directionnels sont séparées de quelques centimètres et forment un angle physique  $\theta$  entre elles (cf. Figure 3-6).



Figure 3-3 : une configuration de prise de son en extérieur.

Parmi les nombreuses configurations possibles, nous avons retenu la prise de son stéréophonique de temps et d'intensité, ainsi que de celle de temps. Celle d'intensité, souvent utilisée dans le cinéma, demande pour sa part un matériel spécifique. En effet, la coïncidence des microphones rend cette configuration très compacte et nécessite, en extérieur, des protections anti-vent particulières, non compatibles avec celles des autres systèmes étudiés. Par ailleurs, même si elle fournit des images bien localisées dans l'espace stéréophonique, elle ne rend pas suffisamment compte, selon Hugonnet et Walder (1995), de l'espace - azimutal - et de la profondeur.

Par ailleurs, nous n'avons pas testé un type de prise de son plus récent, constitué d'une tête artificielle, car celui-ci nécessite un dispositif de transauralisation pour une restitution convenable sur haut-parleurs, une écoute au casque n'étant pas acceptable dans le cadre de nos expériences (cf. § 4.1). En effet, cette prise de son a l'avantage de tenir compte de la diffraction de l'onde acoustique sur la tête de l'auditeur - filtrage HRTF, pour Head Related Transfer Function - et profite donc à une écoute au casque, qui restitue le son à même les oreilles. Une diffusion par haut-parleurs introduit toutefois une "pollution" du signal arrivant sur une oreille et issu du haut-parleur du côté opposé, ainsi qu'une seconde diffraction du champ acoustique par la tête de l'auditeur, ce qui nécessite un traitement du signal avant diffusion. La tête artificielle correspond à une tête de forme "moyenne", légèrement différente de chaque individu, et ne prend donc qu'approximativement en compte la diffraction pour chaque auditeur (cf. Møller, 1996). La spatialisation du son autour de ce dernier n'en sera par conséquent qu'imparfaite, voire non écologiquement valide.

Les configurations retenues feront l'objet de tests de validité écologique qui sont présentés aux § 5 et 6.

## 4. La restitution des échantillons sonores

### 4.1 Choix du mode de diffusion des échantillons sonores

Deux modes de restitution des séquences sonores sont possibles : par haut-parleurs et par casque d'écoute. La diffusion par haut-parleurs permet de délivrer une image sonore stable et cohérente lorsque l'auditeur bouge légèrement sa tête. Lors d'une écoute au casque au contraire, les sources sonores virtuelles se déplacent avec les mouvements de tête, puisque le référentiel n'est plus le local d'écoute, mais la tête de l'auditeur. Ceci va à l'encontre d'une perception quotidienne et n'est pas acceptable du point de vue de la validité écologique.

Par ailleurs, la diffusion par haut-parleurs permet toujours de conserver une certaine distance aux événements sonores, distance comparable à celle des premiers plans pour la prise de son - par exemple le passage d'une voiture devant les microphones. L'écoute au casque donne souvent l'impression de percevoir les sons *dans* la tête, notamment pour ces premiers plans sonores - la voiture passe à travers la tête de l'auditeur et non devant lui. Ceci a pour effet de réduire considérablement l'illusion d'être en situation réelle et s'oppose donc également la validité écologique des échantillons sonores. La prise de son à l'aide d'une tête artificielle pallie en partie le problème (cf. § 3.2.2), mais ne résout pas la première difficulté de référentiel, sauf par ajout au système de reproduction, en plus du système de transauralisation, d'un module de suivi de la position de la tête, le "head-tracker".

Enfin, le fait de porter un casque sur les oreilles éloigne l'auditeur de l'état naturel d'écoute. Celui-ci doit faire abstraction du poids du casque et de son contact avec la tête.

Toutes ces raisons nous ont conduit à retenir les haut-parleurs, plutôt que le casque d'écoute, comme moyen de restitution sonore. Ce choix s'accompagne nécessairement d'une sélection d'enceintes de bonne qualité (§ 4.2), ainsi que d'un local d'écoute approprié (§ 4.3).

### 4.2 Le choix des haut-parleurs : un test d'écoute

Les haut-parleurs constituent le "maillon" faible de la chaîne de transmission du son. Nous leur avons par conséquent accordé une attention particulière afin de rester le plus fidèle possible du signal sonore provenant de l'enregistrement. Nous avons procédé à un test d'écoute de plusieurs systèmes de reproduction sonore afin de sélectionner les meilleurs systèmes.

#### 4.2.1 Méthode

L'expérience consiste à comparer cinq systèmes de reproduction (amplificateur + haut-parleurs) dans un même local d'écoute peu réverbérant et le plus neutre possible. Il s'agit de choisir le système le mieux adapté à l'écoute, non seulement des ambiances sonores urbaines, mais également des instruments de musique et autres sons divers et insolites, dans le cadre des besoins expérimentaux du laboratoire. La diversité des sons utilisés, de par leurs structures spectrales et temporelles, imposent un système d'écoute d'excellente qualité.

#### 4.2.1.1 Corpus sonore et dispositif

Les extraits sonores utilisés pour l'expérience ont été sélectionnés de manière à tester un maximum de caractéristiques des systèmes de reproduction. Un extrait d'orgue notamment pour la réponse à basses fréquences, un extrait de clavecin notamment pour la réponse à hautes fréquences et les transitoires d'attaque, un extrait de voix et de claquements de doigts pour la netteté de la restitution, et des extraits d'ambiances urbaines pour valider écologiquement l'équilibre tonal.

Le test d'écoute est réalisé "en aveugle", les systèmes étant dissimulés derrière un drap très fin, ceci pour éviter un trop grand filtrage des hautes fréquences. De part et d'autre de l'axe stéréophonique, les cinq types d'enceintes sont rapprochés au maximum afin de minimiser les variations de largeurs d'images stéréophoniques entre systèmes d'écoute (Figure 3-4).



Figure 3-4 : configuration du test de systèmes de reproduction (haut-parleurs + amplificateurs).

#### 4.2.1.2 Procédure et sujets

La première partie de l'expérience consiste à diffuser un même extrait sonore sur les différents systèmes en commutant toutes les dix secondes environ, dans le respect des "points de coupure" du discours musical ou narratif des scènes sonores. L'ordre de passage des cinq systèmes d'écoute est le suivant : 1-2-3-4-5-1-3-5-2-4-1. Cet ordre a l'avantage de présenter aux auditeurs toutes les combinaisons possibles de comparaisons par paires en un minimum de commutations entre systèmes de reproduction. La seconde partie consiste à diffuser quatre extraits sonores sur un même système (orchestre symphonique, orgue, bruits de circulation, voix), puis les mêmes extraits dans le même ordre sur le prochain système, et ainsi de suite (cf. annexe 1).

Onze sujets ont participé simultanément à l'expérience en répondant à un questionnaire. Ils étaient assis, répartis sur trois rangées de quatre sièges, dans le champ stéréophonique d'écoute des cinq systèmes de reproduction.

## 4.2.2 Résultats

Dans un premier temps, trois des cinq systèmes ont été retenus à l'unanimité. Puis, l'expérience a été reconduite pour les trois systèmes sélectionnés. Finalement, deux paires d'enceintes amplifiées ont été retenues puis achetées par le laboratoire. Dès lors, les mêmes enceintes (Studer A723) seront utilisées pour l'ensemble des tests d'écoute.

## 4.3 Sélection d'un local d'écoute

Le choix d'une restitution sonore par haut-parleurs étant fait, il devient primordial de trouver un local adapté à ce type d'écoute. Pour cela, nous avons utilisé la chambre assourdie du laboratoire, qui répond aux exigences de la restitution par haut-parleurs.

Dans l'optique de disposer d'une pièce dédiée à l'écoute des environnements sonores urbains, nous avons procédé à l'étude d'une autre salle disponible au laboratoire. Une première analyse des caractéristiques de cette salle avant traitement et une seconde analyse après un premier traitement qualitatif ont été réalisées par Bermond (1997). Suite à cette étude exploratoire, un nouveau traitement acoustique plus complet a été effectué pour transformer la salle en lieu d'écoute. Toutefois, le traitement n'étant toujours pas suffisant pour la réalisation de nos tests d'écoute, c'est la chambre assourdie du laboratoire qui a finalement permis d'effectuer l'ensemble des tests d'ambiances sonores urbaines.



Figure 3-5 : local d'écoute et dispositif expérimental.

## 5. Premier test : écartements entre microphones

De nombreuses études ont été menées en intérieur pour tester différentes techniques de prise de son, en particulier pour l'enregistrement d'oeuvres musicales (Streicher et Dooley, 1985). En extérieur, les conditions de prise de son s'avèrent très différentes : réverbération souvent très faible, vent, sources en mouvement... Afin de sélectionner des prises de son en extérieur qui s'inscrivent dans le concept de validité écologique, une expérience préliminaire a été réalisée.

### 5.1 Méthode

Cette expérience exploratoire a pour objectif de déterminer un ordre de grandeur des écartements optimaux entre microphones pour des angles physiques d'ouverture donnés (Figure 3-6). Trois directivités (cf. § 3.1) et quatre écartements ont été testés pour chaque type de microphone :

- **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**.

Toutefois, suite à une erreur de manipulation, les microphones infracardioïdes ont été mal orientés, à la verticale. Ces résultats erronés peuvent toutefois être réinterprétés, en première approximation, en termes de configuration omnidirectionnelle puisque la prise de son infracardioïde, placée verticalement, est omnidirectionnelle dans le plan horizontal.

#### 5.1.1 Corpus sonore et dispositif

Les enregistrements effectués en extérieur sont constitués de bruits urbains parisiens. Les séquences sonores ont été enregistrées près d'un carrefour, en milieu piétonnier (sur le trottoir, à 1,60 m du sol). Douze séquences de bruits de circulation d'environ une minute chacune ont été sélectionnées, issues d'enregistrements effectués à Paris, place d'Austerlitz et quai Saint-Bernard (cf. § 2.4).

Le dispositif de prise de son consiste à fixer les microphones sur une barre d'écartement et à les orienter de manière symétrique par rapport à la normale de la barre et selon l'angle physique d'ouverture désiré (Figure 3-6).

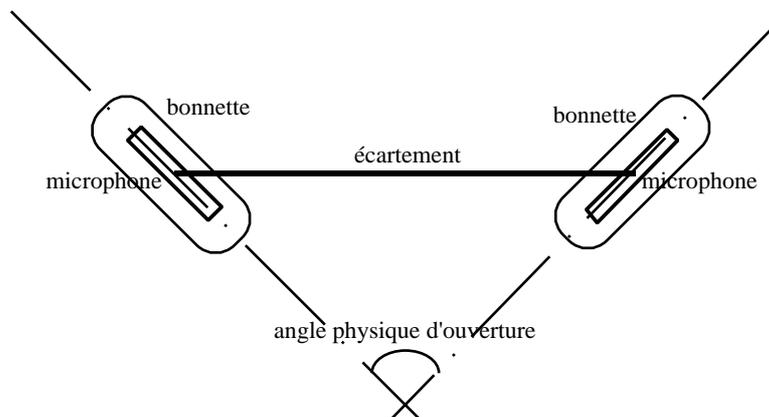


Figure 3-6 : configuration schématique de prise de son en extérieur.

Les microphones sont munis de bonnettes - cage en plastique de type Rycote recouverte éventuellement, par vent moyen, d'une peau de type Jersey et par vent fort, d'une fourrure - afin d'éviter les effets néfastes du vent sur l'enregistrement. L'ensemble repose sur un pied de microphone posé sur le sol.

La restitution sonore est réalisée par haut-parleurs, en chambre assourdie, selon une configuration d'écoute conforme à une recommandation internationale, qui propose de disposer les haut-parleurs et la tête de l'auditeur aux sommets d'un triangle équilatéral (Figure 3-7). Les séquences sonores sont diffusées à partir d'une bande D.A.T., échantillonnée à 44,1 kHz, et d'un lecteur Tascam DA-30, directement relié aux enceintes Studer A723.

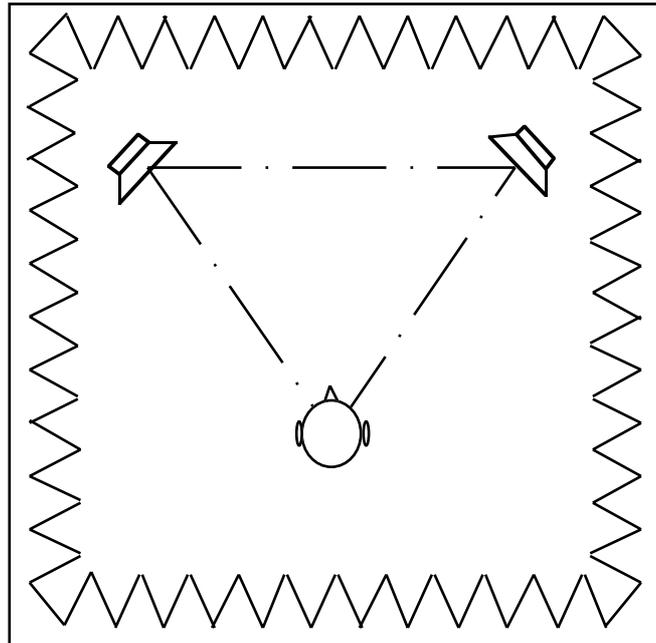


Figure 3-7 : schéma de la configuration d'écoute (triangle équilatéral).

### 5.1.2 Procédure et sujets

Le test consiste à diffuser douze séquences de bruits de circulation à douze sujets âgés de 23 à 58 ans, tous acousticiens. La consigne qui leur est donnée est la suivante : "Vous allez entendre 12 séquences sonores d'environ une minute chacune. Dans un premier temps, nous vous demandons de sélectionner parmi les 4 premières séquences (index 1 à 4), celle qui vous semble la plus proche de la réalité. De même pour les séquences d'index 5 à 8 et 9 à 12. Vous êtes libres de réécouter les séquences et de varier l'ordre d'écoute. Réécoutez ensuite les 3 séquences retenues et choisissez parmi elles la plus réaliste." Sur la bande D.A.T., les séquences sont présentées dans le désordre, en termes d'écartements entre microphones, par rapport à l'ordre d'indexation. Les sujets passent d'une séquence à une autre à l'aide de la télécommande du lecteur de D.A.T.. Ils sont libres de faire des commentaires pour chaque étape et pour l'ensemble de l'expérience (annexe 1).

Les 12 séquences de bruits de circulation correspondent aux plages 01 à 03 (4 index pour chaque plage) du C.D. accompagnant ce mémoire, conformément aux index indiqués dans la procédure ci-dessus.

## 5.2 Résultats

Les réponses des sujets mettent en évidence une nette préférence pour les configurations dont les écartements sont compris entre 60 et 70 cm (Figure 3-8). Ainsi, 23 réponses, soit 64% de l'ensemble des réponses, se rapportent à ces deux configurations.

N° de séquence	Prise de son	Nombre de séquences sélectionnées par configuration	Nombre de séquences sélectionnées parmi toutes
4	Infra. 20 cm	2	1
2	Infra. 40 cm	1	0
1	Infra. 60 cm	4	1
3	<b>Infra. 70 cm</b>	<b>5</b>	1
6	Omni. 60 cm	3	0
8	<b>Omni. 70 cm</b>	<b>5</b>	3
7	Omni. 80 cm	3	1
5	Omni. 90 cm	1	1
9	Cardio. 20 cm	0	0
11	Cardio. 30 cm	3	0
10	Cardio. 40 cm	3	2
12	<b>Cardio. 60 cm</b>	<b>6</b>	2

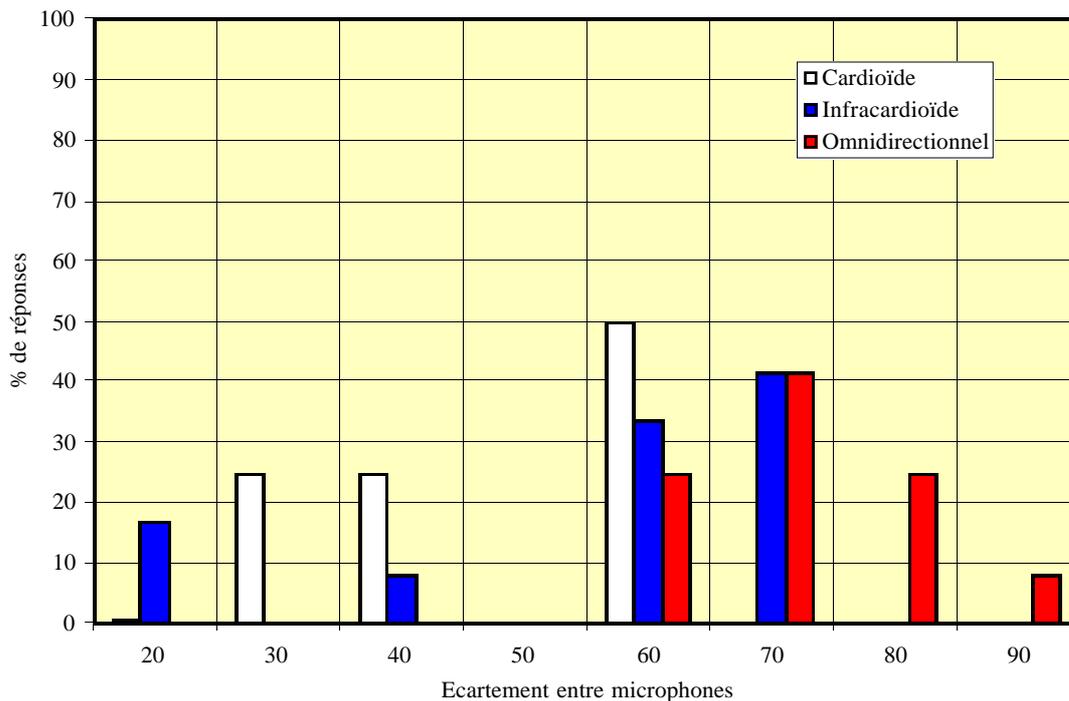


Figure 3-8 : résultats du premier test de prise de son.

En-deçà de 60 cm, l'image stéréophonique paraît étriquée aux sujets. Au-dessus de 70 cm, ils font état d'un "trou" au centre de l'image stéréophonique. En effet, la différence d'intensité des signaux reçus par les microphones, pour une source passant près du dispositif de prise de son, devient alors trop importante. Et le décalage temporel des signaux arrivant sur les deux microphones amplifie ce phénomène.

Dans le groupe de séquences cardioïdes, la prise de son à 60 cm d'écartement a été préférée par la moitié des sujets, alors que celle à 20 cm n'a été retenue par personne. Ces résultats mettent en évidence des différences importantes avec des études menées en *intérieur* et qui suggèrent une vingtaine de centimètres d'écartement pour une centaine de degrés d'angle physique d'ouverture, tels les procédés utilisés pour la radiodiffusion (Hugonnet et Walder, 1995, p.161-162).

### **5.3 Défauts de l'expérience**

Cette expérience exploratoire a soulevé un certain nombre de problèmes qu'il s'agit de résoudre pour les prochains tests. Dans un premier temps, les séquences de bruits de circulation sont toutes légèrement différentes car elles n'ont pas pu être enregistrées simultanément. Cela peut biaiser le jugement des sujets car certaines séquences sonores sont moins fournies que d'autres en détails acoustiques, sur lesquels l'auditeur fonde son jugement. Ces séquences peuvent donc être jugées mauvaises, plus par leur contenu sonore que par le système de prise de son. Toutefois, les résultats présentent une certaine homogénéité quant aux préférences pour une directivité de microphone donnée. Ainsi, la prise de son omnidirectionnelle à moins de 50 cm ne semble pas convainquante, tout comme la prise de son cardioïde 120° à moins de 30 cm. Il est cependant préférable de réaliser des enregistrements simultanés des séquences sonores afin de faciliter leur comparaison et d'en maîtriser l'interprétation des résultats.

Dans un second temps, il est apparu que les échantillons sonores ont quelques défauts : d'une part, leur durée est considérée par les sujets comme trop longue, donc à l'origine d'une perte d'attention, et d'autre part, leur niveau sonore est perceptiblement inégal.

Enfin, l'utilisation de la notion de réalisme n'est pas appropriée car elle introduit, selon certains sujets, un concept sémantique inadapté à un contexte de simulation.

## 6. Second test : configurations de prise de son

Le premier test de prise de son en extérieur a permis de déterminer les écartements entre microphones pour des angles physiques d'ouverture donnés et pour chaque type de directivité. Une seconde expérience de prise de son a été réalisée afin de préciser les configurations les mieux adaptées aux ambiances sonores de la ville. L'objectif en est de déterminer les configurations de prise de son à utiliser selon le contexte urbain, en vue d'enregistrer les séquences sonores qui constitueront des fonds sonores représentatifs des contextes d'émission de signaux d'avertissement.

### 6.1 Méthode

Le principe du test consiste à enregistrer *simultanément* des séquences sonores, de manière à faciliter les comparaisons entre séquences. Pour des raisons pratiques, il devient très difficile d'enregistrer simultanément plus de deux configurations de prise de son. En effet, ces dernières doivent être juxtaposées, pour enregistrer les mêmes événements sonores d'un même lieu, sans pour autant affecter - par masquage ou diffraction - le son enregistré par le système voisin. La procédure adoptée consiste alors en une comparaison par paires des configurations confrontées.

#### 6.1.1 Corpus sonore et dispositif

Deux types d'ambiances sonores, de circulation et de jardin public, ont été enregistrés dans Paris (cf. § 2.4), respectivement au croisement des boulevards de l'Hôpital et Saint-Marcel et dans le jardin du Luxembourg. Le dispositif d'enregistrement pour chaque système de prise de son est le même que dans l'expérience préliminaire, mais cette fois-ci, deux configurations sont systématiquement placées ensemble, séparées verticalement de quelques centimètres seulement, afin d'effectuer des enregistrements simultanés d'ambiances dans des conditions très proches. Cinq configurations différentes sont testées :

- les deux configurations retenues à l'issue du premier test, **O70** (omnidirectionnel, 70 cm d'écartement) et **C60** (cardioïde, 60 cm, 100° d'ouverture),
- une configuration intermédiaire, **I70** (infracardioïde, 70 cm, 100°),
- deux configurations "classiques", éprouvées en intérieur, **ORTF** (cardioïde, 17 cm, 110°, radiodiffusion française) et **C35** (35 cm, 90°, proche du système NOS de la radiodiffusion hollandaise).

Par ailleurs, pour pallier les problèmes mis en évidence dans le premier test, la durée des échantillons sonores est réduite de moitié, soit 30 secondes. Quant aux niveaux sonores, les paires sont égalisées dans chaque séquence par rapport au niveau maximum mesuré (en dBA) qui correspond à un événement sonore commun aux deux séquences.

#### 6.1.2 Procédure et sujets

L'expérience consiste en une comparaison par paires, méthode que nous avons déjà utilisée dans des travaux antérieurs (Vogel, 1995). Quatorze paires de séquences, chacune enregistrée simultanément, ont été présentées aux auditeurs. La consigne est la suivante : "Vous allez entendre des séquences sonores issues de notre environnement "naturel" (bruit de circulation, ambiance de jardin public). Votre tâche consiste à effectuer 8 comparaisons par

paires pour l'ambiance de circulation et 6 pour l'ambiance de jardin public. Les paires sont séparées d'une dizaine de secondes. Tout d'abord, il s'agit de choisir pour chaque paire la séquence qui vous semble la plus proche de votre expérience quotidienne. Pour cela, essayez de vous imaginer dans le lieu, de vous mettre en situation, éventuellement en fermant les yeux. Ensuite, vous devez estimer sur une échelle continue, graduée de 1 à 7, le degré de dissemblance entre deux extraits sonores." Pour éviter tout jugement esthétique de la part des sujets, il leur est demandé de choisir la séquence dans laquelle ils s'intègrent le plus facilement. Ainsi, la notion de réalisme a pu être éliminée (cf. annexe 1).

La première paire de circulation (paire 1) est replacée plus loin dans le test mais de manière inversée (paire 7). De même pour la première paire de jardin public (paires 9 et 13). Cette opération a pour but de vérifier la cohérence des sujets avec eux-mêmes et de les habituer au test (en ne tenant éventuellement pas compte de la première paire).

Quarante-huit sujets, âgés de 20 à 58 ans, ont participé au test. Vingt-quatre d'entre eux sont acousticiens, les autres, psychologues ou étudiants de diverses formations.

Les 14 paires de séquences sonores correspondent aux pages 04 à 17 (2 index pour chaque page) du C.D. accompagnant ce mémoire, dans l'ordre du test d'écoute, en commençant par l'ambiance de circulation, puis celle de jardin public.

## **6.2 Résultats**

Une première analyse des résultats, avec une vingtaine de sujets, a permis de mettre en évidence que les auditeurs ont tendance à choisir la deuxième séquence de chaque paire (Figure 3-9). Pour savoir si cette attitude est liée à l'ordre d'apparition des séquences dans chaque paire, nous avons réalisé une deuxième série du même test dans laquelle les séquences de chaque paire sont inversées. Les deux séries de l'expérience comptent le même nombre de sujets, soit deux fois vingt-quatre sujets. Les résultats de la seconde série (Figure 3-10), comparés à ceux de la première, indiquent effectivement la présence d'un effet d'ordre, ce dernier étant prépondérant dans les paires 1, 2, 3, 6, 7, 11 et 14.

Seules les paires 10 et 13 donnent des résultats vraiment fiables : 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.

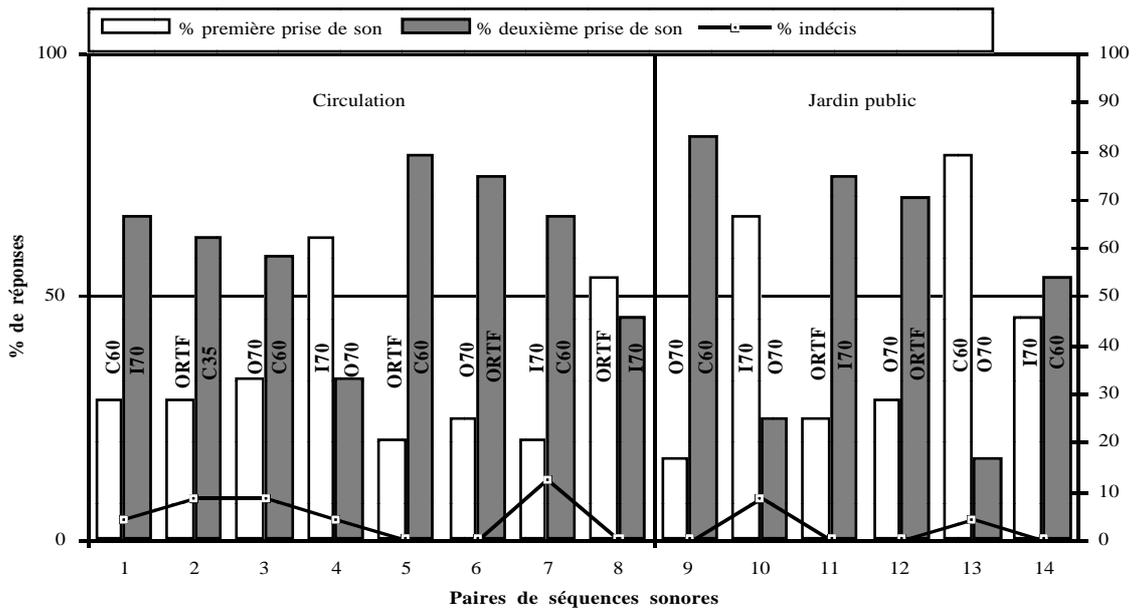


Figure 3-9 : résultats du second test, première série (24 sujets).

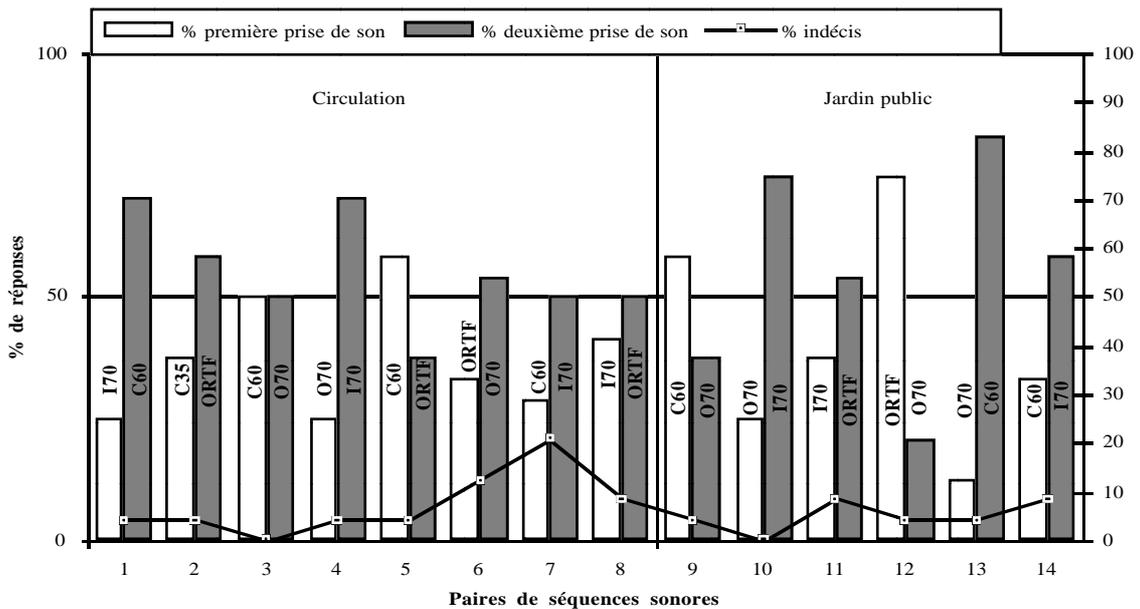


Figure 3-10 : résultats du second test, deuxième série (24 sujets).

En dépouillant les échelles de dissemblance, il apparaît que configurations estimées peu dissemblables correspondent à la paire infracardiïde/cardiïde et que celles estimées très dissemblables correspondent à la paire omnidirectionnel/cardiïde. Cela conforte les résultats précédents car la configuration qui se distingue des autres est bien celle avec les microphones omnidirectionnels. Toute comparaison avec celle-ci donne des résultats tranchés. De même, la comparaison infracardiïde/cardiïde est la plus influencée par l'effet d'ordre. Les sujets ont des difficultés pour choisir, ils notent donc la paire comme peu dissemblable.

### **6.3 Interprétation**

L'importance de l'effet d'ordre masque la moitié des résultats. En fait, il apparaît que le sujet, à l'écoute de la première séquence, est dans une phase de reconnaissance et d'identification de l'ambiance ; à l'écoute de la seconde séquence au contraire, connaissant le contenu événementiel, il s'intègre plus facilement dans l'ambiance.

La configuration avec des microphones omnidirectionnels est à éviter pour des prises de sons d'environnements urbains et dans nos conditions d'écoute. En effet, les sujets considèrent que les basses fréquences des ambiances enregistrées avec des microphones omnidirectionnels sont excessives. En revanche le système C60 convient pour les deux contextes sonores étudiés.

La durée des séquences, réduite à 30 secondes, est encore considérée comme trop longue, les sujets se plaignant de la trop grande richesse sonore des extraits et donc, des oublis qui en découlent.

## 7. Conclusion

Ce chapitre a présenté une méthode complète pour constituer un corpus d'échantillons sonores d'ambiances urbaines, dans le souci constant de préserver la validité écologique des échantillons lors de leur restitution en laboratoire. Elle intègre ainsi les choix des lieux représentatifs à faire écouter, de la technique de prise de son adaptée au milieu, ainsi que du matériel et du lieu de restitution.

Le travail de sélection d'ambiances sonores urbaines a permis de récolter de nombreuses informations à partir des descriptions graphiques et verbales des représentations de l'environnement sonore. En effet, l'analyse des données verbales issues des entretiens a permis d'une part, d'identifier les lieux les plus souvent cités et d'autre part, de montrer qu'il existe une certaine structure dans l'appréhension d'un lieu, non seulement en termes de topologie et d'appréciation qualitative, mais aussi par rapport au sujet qui en parle, la place qu'il tient dans ce lieu et l'activité qu'il peut y exercer. Le fait d'obtenir quatre types de croquis, allant des cartes aux schémas abstraits, montre l'intérêt d'une démarche pluridisciplinaire qui permet d'analyser un percept sonore à partir de son expression graphique et verbale.

Les résultats des tests de prise de son indiquent que les systèmes stéréophoniques habituellement utilisés en intérieur, pour des sources fixes, semblent moins bien adaptés que certaines nouvelles configurations à l'enregistrement en extérieur. Ainsi, les configurations O70 - omnidirectionnelle avec 70 cm d'écartement - et ORTF - cardioïde avec 17 cm d'écartement et 110° d'angle d'ouverture - ne conviennent pas pour des ambiances de circulation dans lesquelles les sources sont en mouvement. Il en va de même pour O70 dans des ambiances de jardins publics, ambiances privilégiant une écoute de scène, moins "mouvementée". 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.



# Chapitre 4 :

## Seuils d'identification en contexte urbain et identification hors contexte

### 1. Introduction

L'étude de la perception des signaux d'avertissement en contexte urbain pose la problématique de l'efficacité de ces signaux auprès des citoyens, tant du point de vue de leur *reconnaissance* (cf. chapitre 1, § 4) que de leur *localisation* (cf. chapitre 1, § 2.4). Notre propos concerne essentiellement *l'identification des signaux*, leur localisation nécessitant de nombreuses recherches complémentaires, notamment en psychoacoustique.

Ainsi, un signal efficace est-il avant tout porteur de *signification* clairement définie dans l'esprit des gens. La signification donnée au signal peut se retrouver à un niveau générique, qui indique simplement le nom d'une classe de sources connues, ou à des niveaux spécifiques, qui précisent à la fois des propriétés sonores et l'interprétation sémantique de la source émettrice. Par l'intermédiaire de la verbalisation de séquences sonores, nous aborderons l'identification des signaux dans le but d'établir des relations de proximité en fonction du sens qui leur est attribué. Ces relations doivent permettre d'obtenir une forme de propriétés acoustiques liées à la signification, elle-même dépendante de la culture des gens (cf. chapitre 6).

Une première expérience consiste à rechercher les seuils d'identification des signaux d'avertissement dans deux contextes urbains distincts. En outre, une identification hors contexte est demandée et les résultats sont confrontés avec ceux en contexte.

## 2. Méthode

Une expérience a été réalisée à partir de sons enregistrés en milieu urbain. Les sites, ainsi que la méthode de prise de son et de restitution sonore, ont été préalablement sélectionnés, selon la démarche décrite dans le chapitre précédent. L'objectif est d'accéder aux représentations mentales des gens par l'intermédiaire de leurs verbalisations et d'étudier l'indépendance ou non de l'identification d'un signal par rapport au contexte d'écoute. Pour ce faire, signaux et contextes sont tout d'abord sélectionnés de manière indépendante. Puis, à l'aide d'une station de travail audionumérique (Sonic Solutions), un montage et un mixage précis de chaque signal avec chaque ambiance sonore permet de réaliser des scénarios sonores qui, à l'écoute, doivent être écologiquement valides.

### 2.1 Corpus sonore

#### 2.1.1 Fonds sonores...

Nous avons retenu deux contextes urbains représentatifs de Paris, extraits d'enregistrements effectués dans les conditions de prise de son définies au chapitre 3. Ces conditions, rappelons-le, correspondent à une stéréophonie à microphones cardioïdes, avec 60 cm d'écartement entre les capsules et 100° d'angle physique d'ouverture. Les deux contextes correspondent à une ambiance de *circulation intense*, enregistrée au croisement des boulevards de l'Hôpital et Saint-Marcel, et une ambiance de *jardin public*, enregistrée au jardin des Plantes. Les deux fonds sonores ont été sélectionnés de manière à être les plus *neutres* possibles, des points de vue événementiel et des variations du niveau sonore. En effet, il est important d'avoir une écoute globale des fonds sonores afin de diriger l'attention des sujets sur les signaux mixés, les fonds sonores devant avoir un rôle de "fond" et non de "figure" dans les termes de la Gestalttheorie (cf. chapitre 1, § 4.1.1). Des extraits qui ne sont pas suffisamment neutres conduisent à une écoute analytique du fond sonore (Maffiolo, 1998). Ainsi, les extraits contenant des avertisseurs de véhicules ont été éliminés, tout comme ceux s'avérant trop variables et qui pouvaient provoquer d'importants effets de masquage au sein d'une même séquence.

#### 2.1.2 ...et signaux d'avertissement

Les signaux d'avertissement ont été sélectionnés à la fois pour leurs caractéristiques acoustiques et pour leur représentativité dans une ambiance urbaine. Quinze signaux ont été choisis : deux avertisseurs de la police (récent, ancien), deux avertisseurs d'ambulance (récent, ancien), un avertisseur des pompiers, deux sirènes de lance-eau (rapide, lente), deux "klaxons" de voiture (un coup long, trois coups brefs), trois avertisseurs de bicyclette ("dring dring", "ding dong", poire) et trois sifflets (à un ton, à deux tons, à roulette). Ces signaux ont des propriétés spectro-temporelles très variées en fonction des différents types d'utilisations (sirènes, klaxons, sonnettes, sifflets), mais sont également très proches pour une même utilisation donnée (ancien et nouvel avertisseur d'ambulance par exemple). Une analyse acoustique de ces signaux est proposée au chapitre 6.

Tous les signaux sont enregistrés de manière statique, sans mouvement apparent lors d'une écoute stéréophonique.

Les signaux ont été enregistrés :

- en chambre semi-anéchoïque, pour les trois sonnettes, à l'aide de deux microphones cardioïdes en configuration de prise de son d'intensité X-Y (capsules des microphones coïncidentes, 90° d'angle d'ouverture). Cette configuration permet une bonne localisation de la source dans l'espace stéréophonique ;
- en extérieur, pour les deux avertisseurs de voiture, avec notre configuration spécifique (C60, 100°) et dans un espace dégagé ;
- d'un disque compact de bruitage (Bataille, 1992), pour les autres signaux que nous désignerons par *signaux d'urgence*.

Les 15 signaux d'avertissement correspondent aux pages 18 à 32 du C.D. accompagnant ce mémoire, dans l'ordre donné ci-dessus. Les deux ambiances sonores correspondent aux pages 33 (circulation) et 49 (jardin public).

### 2.1.3 Montage et Mixage des séquences sonores

Lors du montage des séquences, nous avons systématiquement appliqué chaque signal à cinq reprises à un même fond sonore d'une durée de 20 s, 30 s ayant été estimées trop longues (cf. chapitre 3, § 6.3). Une autre technique consiste à mettre chaque fond sonore en boucle "infinie", démarche adoptée par Szeto *et al.* (1991) et Binnington *et al.* (1993), et de tester en continu les différents signaux d'avertissement. Mais ce faisant, le fond perd son "rôle" de contexte (cf. chapitre 1, § 6.1.1) et ne constitue plus qu'un bruit masquant caractérisé par ses propriétés spectrales. Nous avons donc préféré disjoindre chaque signal contextualisé, quitte à allonger fortement la durée totale de l'expérience. Ainsi, chaque signal est perçu dans "son" contexte propre, en tant qu'extrait d'une "scène sonore" écologiquement valide.

Le mixage des signaux aux contextes sonores consiste ensuite à varier l'intensité sonore des cinq itérations de chaque signal, en augmentant le niveau de ce dernier de 3 dB à chaque itération (Figure 4-1). Ce pas de 3 dB correspond à un doublement du niveau sonore entre deux occurrences successives du signal.

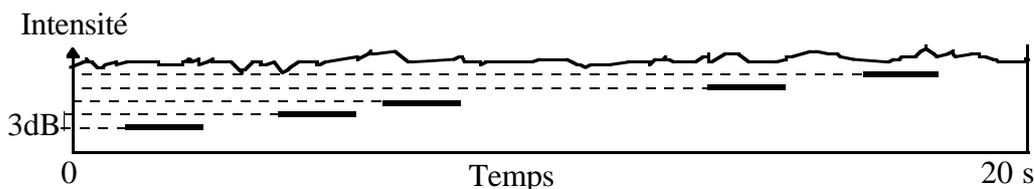


Figure 4-1 : représentation dynamique et schématique d'une séquence mixée.

Les signaux ont d'abord été montés dans la séquence de circulation, de part et d'autre d'un passage de camion qui provoque un masquage important, puis ils ont été appliqués de la même manière à la séquence de jardin public. Une seule exception : la sirène lente du lance-eau ayant une période de près de 5 s, il nous a fallu doubler la durée des fond sonore (38 s) en les bouclant une seule fois sur eux-mêmes et en vérifiant la validité écologique, afin de rester cohérent avec les vingt-huit autres échantillons sonores. Ce montage particulier n'autorise cependant que quatre itérations du signal (toujours à niveau croissant de 3 dB).

Lors du mixage, les niveaux des signaux ont été calibrés subjectivement, de manière à ce que la seconde itération constitue le seuil d'identification. Ce choix de la seconde itération résulte du fait que d'une part, le signal est connu et attendu par l'expérimentateur, ce qui lui facilite une reconnaissance à plus faible niveau, et d'autre part, une première itération non détectée nous assure un rapport signal/bruit insuffisant au début de chaque séquence sonore.

Afin de vérifier la calibration des niveaux sonores, deux sujets, non comptabilisés dans les résultats, ont effectué le test sans connaître les signaux. A l'issue de leurs réponses, quatre séquences de jardin public, pour lesquelles les signaux n'ont pas été détectés, ont dû être modifiées. Il est important de remarquer que ces calibrations sont différentes selon les contextes sonores.

#### **2.1.4 dispositif d'écoute**

Le dispositif de restitution sonore est le même que celui décrit au chapitre 3, § 5.1.1. L'ambiance sonore de circulation est restituée à 81 dBA max., soit 10 dB au-dessous du niveau mesuré lors de l'enregistrement en extérieur ; celle du jardin public à 65 dBA max., soit au niveau mesuré à l'enregistrement - les niveaux donnés sont mesurés à l'aide d'un sonomètre en mode rapide (constante de temps = 125 ms). Le fait de devoir réduire le niveau des séquences de circulation est lié à une plus grande tolérance des niveaux sonores intenses en extérieur, en partie grâce à l'activité de l'individu et de la situation naturelle (cohérence entre les différents sens), facteurs que l'on perd en écoute en intérieur, décontextualisée. L'ambiance de jardin public est considérablement plus faible et est donc acceptée par les sujets en intérieur.

## **2.2 Procédures et sujets**

En préalable au test d'écoute, il est demandé aux sujets de participer à une épreuve verbale. La consigne est la suivante : "citez 10 signaux d'avertissement et donnez-en les caractéristiques". Cette épreuve est suivie du test d'écoute dont la consigne est : "vous allez tout d'abord entendre 16 séquences sonores de circulation urbaine. La première vous présente le fond sonore urbain seul, les 15 suivantes le même fond sonore hors duquel émergent 15 signaux d'avertissement différents. Dès que vous entendrez le signal d'avertissement, vous devrez nous le signaler et la séquence sera interrompue. Nous vous demanderons alors simplement d'identifier chaque signal que vous aurez perçu. Le même protocole opératoire a été établi pour 16 séquences sonores de jardin public" (cf. annexe 1). L'ordre d'apparition des signaux est différent dans les contextes de circulation et de jardin public. Enfin, les signaux d'avertissement sont présentés de manière isolée et les sujets doivent à nouveau les identifier.

Les 32 séquences du test, munies des signaux, correspondent aux pages 33 à 64 du C.D. accompagnant ce mémoire, dans l'ordre donné ci-dessus.

L'ensemble du test dure en moyenne 35 minutes, dont dix minutes pour répondre à la question préalable à l'écoute. L'expérimentateur participe à l'écoute à côté de chaque sujet et manipule la télécommande du lecteur de D.A.T., ceci afin d'arrêter les séquences au moment de l'identification. Par ailleurs, il rédige sur une feuille de test (cf. annexe 1) les résultats de l'identification : numéros d'itérations et réponses verbales.

Vingt-six sujets ont participé à l'expérience, âgés de 20 à 59 ans, onze femmes et quinze hommes. Quatorze sujets sont acousticiens, les douze autres sont issus de divers milieux socioculturels (cf. chapitre 2, § 3.2). Tous sont usagers de Paris et donc habitués aux contenus sonores des stimuli présentés. La moitié des sujets a effectué l'expérience dans l'ordre inverse des autres sujets : en débutant par le jardin public et dans l'ordre inverse des séquences munies d'un signal. Cette démarche a pour rôle de contrôler l'influence éventuelle de l'effet d'ordre sur les réponses des sujets, à l'intérieur d'un même contexte et entre contextes.

## 2.3 Audiogrammes

Chaque sujet a été soumis à un audiogramme avant le test d'écoute. L'étude des audiogrammes en fonction des réponses des sujets n'indique aucune différence significative liée à des pertes auditives. Cette constatation est en accord avec l'étude menée par Robinson et Casali (1995 ; cf. chapitre 1, § 6.1.1). La perception des seuils se fait donc bien par émergence relativement aux fonds sonores et aucun phénomène de masquage supplémentaire, lié à des pertes auditives, ne vient perturber les résultats. Nous retrouvons également la notion de perception auditive par relation de contraste (Leipp, 1966 ; cf. chapitre 1, § 4.1.2), issue de la Gestalttheorie (cf. chapitre 1, § 4.1.1) et régie par la structure figure-fond.

## 2.4 Une petite précaution...

Avant de procéder à l'analyse des résultats, nous avons converti en numéros les noms donnés aux quinze signaux acoustiques (Figure 4-2). Cette précaution nous permet d'interpréter les résultats sans les *a priori* de l'expérimentateur sur les signaux. Cette démarche est conservée dans ce mémoire pour ne pas influencer le lecteur, mais il nous a semblé important de préciser régulièrement le type de signaux se distinguant des autres, afin d'éviter une éventuelle "frustration" lors de la lecture.

Type de signal	Nom du signal	Durée	Numéro
Signaux d'urgence (couramment appelés sirènes)	Police (nouveau, électronique)	2,2 s	<b>1</b>
	Police (ancien, mécanique)	2,1 s	<b>2</b>
	Ambulance (nouveau, électronique)	1,0 s	<b>3</b>
	Ambulance (ancien, mécanique)	1,0 s	<b>4</b>
	Pompiers (mécanique)	2,1 s	<b>5</b>
	Lance-eau (modulation lente)	4,7 s	<b>6</b>
	Lance-eau (modulation rapide)	1,5 s	<b>7</b>
Klaxons	Klaxon de voiture (un coup long)	0,5 s	<b>8</b>
	Klaxon de voiture (trois coups brefs)	0,8 s	<b>9</b>
Sonnettes	Sonnette de vélo classique ("dring dring")	1,3 s	<b>10</b>
	Sonnette de vélo à deux tons ("ding dong")	1,8 s	<b>11</b>
	Poire en caoutchouc ("pouêt")	0,7 s	<b>12</b>
Sifflets	Sifflet à un ton	0,7 s	<b>13</b>
	Sifflet à deux tons (type scout)	1,8 s	<b>14</b>
	Sifflet à roulette	1,7 s	<b>15</b>

Figure 4-2 : conversion des noms de signaux en numéros.

Ce classement des signaux d'avertissement (type de signal) n'a rien d'objectif et est réalisé en fonction des caractéristiques de la source émettrice (l'objet physique). Son unique but est de pouvoir nommer les signaux testés par rapport aux données verbales apportées par les sujets.

### 3. Résultats de l'épreuve verbale

Vingt-trois sujets ont répondu à la question : "citez 10 signaux d'avertissement et donnez-en les caractéristiques". Ils disposaient de dix minutes pour y répondre. 211 désignations de signaux ont été produites, les sujets n'ayant pas toujours cité dix signaux. Les résultats sont analysés en termes de *fréquence* des désignations produites (§ 3.1), ainsi que de leur *disponibilité* dans la langue (§ 3.2).

#### 3.1 Désignations les plus fréquentes

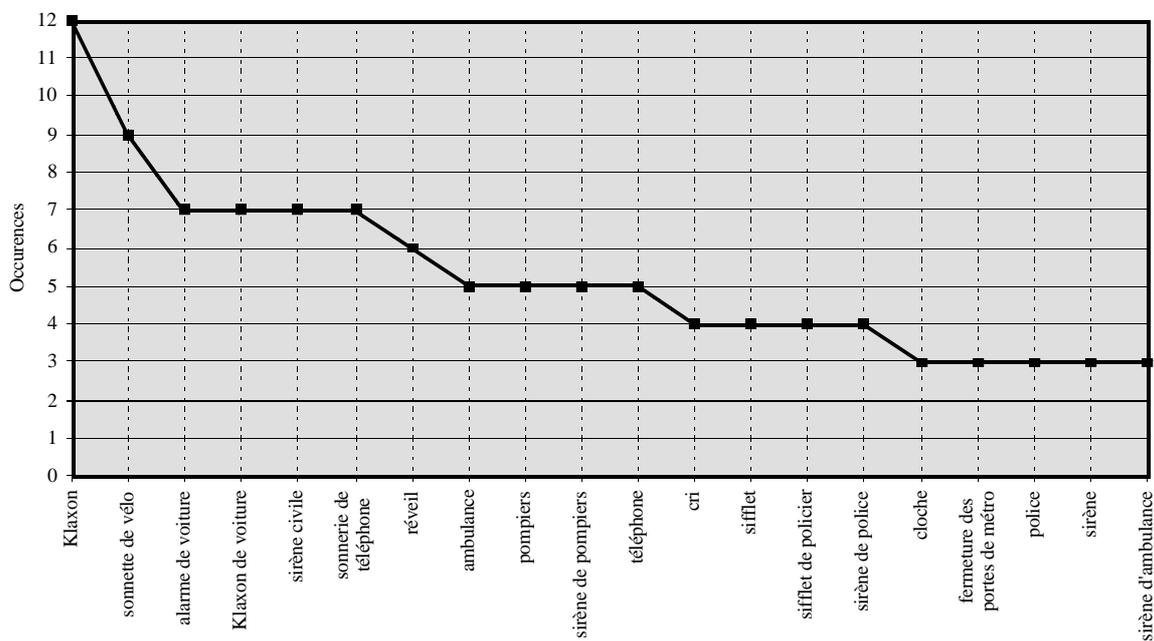


Figure 4-3 : désignations les plus fréquentes (avant écoute).

Les termes les plus fréquents sur l'ensemble des données (Figure 4-3) sont "klaxon" (12 occurrences), "sonnette de vélo" (9), "alarme de voiture", "klaxon de voiture", "sirène civile" et "sonnerie de téléphone" (7). Les caractéristiques qui leur sont associées sont les suivantes :

- "klaxon" : "fort, agressif, brutal, désagréable, gênant, stressant..." (15) ; "bref, efficace, sécurité, caractéristique, aigu..." (10).
- "klaxon de voiture" : "fort, agressif, désagréable, stressant..."(7) ; "danger, aigu, de toutes sortes..."(4).
- "Sonnette de vélo" : "agréable, chantant, mignon..."(10) ; "pas très efficace, peu perceptible..."(4) ; "typique" (1) ; "émerge bien" (1).
- "Alarme de voiture" : "agressif, trop intense, insupportable, nuisance sonore..."(8) ; "strident, souvent très aigu, répétitif..." (6).
- "Sirène civile" : "enveloppe la ville, beau, surprend tous les mois, pompiers, campagne, long, bruyant, prolongé".
- "Sonnerie de téléphone", "téléphone" : "dépend des sonneries, discret ou strident, neutre, stressant, pas assez mélodique".

- "Réveil" : "affreux, toujours pénible, aigu, strident, répétitif".
- "Ambulance" : "stressant, fort, agressif..." (4) ; "non localisable" (2) ; "efficace".
- "Pompiers", "sirène de pompiers" : "intense, strident" (5) ; "stressant, non localisable" (2) ; "efficace" (3) ; "caractéristique, répétitif, agréable et informateur, moins stressante, moins aiguë".

Les caractéristiques associées aux signaux d'avertissement cités le plus fréquemment et *a priori* sont essentiellement de type hédonique : "gênant, agressif, agréable, amical...". Celles relatant des caractéristiques acoustiques concernent l'intensité sonore des signaux, leur caractère aigu ou grave, ainsi que leur localisation (cf. annexe 2).

### 3.2 Désignations les plus disponibles

En ne considérant que les trois réponses citées en premier par les sujets, c'est-à-dire les signaux d'avertissement que l'on peut considérer comme étant les plus disponibles dans la mémoire des gens, les termes les plus fréquents deviennent (Figure 4-4) : "klaxon" (9), "ambulance", "klaxon de voiture" et "sirène de pompiers" (4), "pompiers", "sirène civile" et "sirène de police" (3), "cris", "réveil", "sifflet", "sifflet de policier", "sirène", "sirène d'ambulance", "sonnerie", "sonnerie de téléphone" (2).

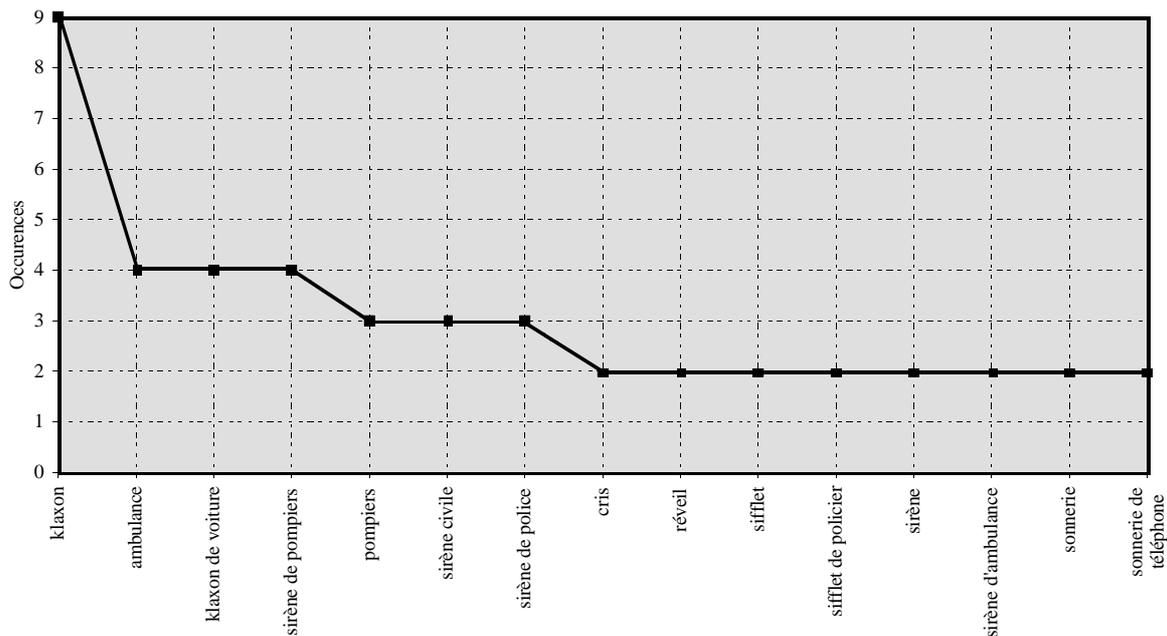


Figure 4-4 : désignations les plus disponibles (3 premières citées - avant écoute).

Il s'agit essentiellement de signaux d'avertissement de véhicules prioritaires ou de voitures. La sonnette de vélo, à la différence des désignations les plus fréquentes, n'apparaît pas dans les désignations les plus disponibles. De même, la sonnerie de téléphone ne se trouve que rarement dans les trois premiers signaux cités.

### 3.3 Interprétation

Ces données nous apportent des indications sur le répertoire lexical et les représentations cognitives disponibles (David *et al.*, 1997) *a priori*, sans écoute ni entretien préalables. Une première analyse sémantique des désignations les plus fréquentes permet de regrouper ces dernières en deux classes :

- des **propriétés physiques du son**, plus ou moins lexicalisées : "crissements, freinage, bruitages, basses, sifflement..." et onomatopées ("bip, ding ding, coin coin..."),
- des **noms de sources interprétées** : au niveau sémantique, le son est décrit à travers l'identification de la source. On peut distinguer deux types de sources interprétées : le nom de la source ou du "statut" de la source ("ambulance, pompiers, sifflet, téléphone, police, sonnette, cloche...") et le nom de l'objet sonore associé à une source plus globale ("sonnette de vélo, klaxon de voiture, sirène de pompiers, sifflet de policier..."). Dans le second cas, que nous appellerons les *noms de noms*, les descriptions sont les plus complètes et réfèrent à la fois à l'identification et aux propriétés du son, ainsi qu'à celles de la source.

Dans l'épreuve verbale, la quasi-totalité des productions verbales des sujets correspond à des sources interprétées (106 des 120 productions). Parmi elles, les *noms de noms* représentent près de la moitié des désignations (54 productions). Ces désignations, souvent précises, correspondent donc à des représentations mentales bien définies chez l'individu. Très peu de sujets produisent des expressions en termes de propriétés physiques du son (14 productions) et même dans ce cas peu fréquent, les propriétés sont associées à des sources interprétées ("bip d'ordinateur, crissement de pneus au freinage..."). L'identification *a priori* des signaux d'avertissement nous fournit ainsi des prototypes de productions langagières significantes. Nous remarquons que tous les types de signaux d'avertissement testés (klaxons, sirènes, sifflets, sonnettes) sont, non seulement cités par les sujets en tant que signifiés, mais représentent aussi la majorité des signifiés produits. Les résultats des écoutes permettront de valider ou d'invalider les correspondances entre les signifiants de nos stimuli et les signifiés qui leur sont associés, en fonction des contextes d'écoute.

Par ailleurs, les signaux d'avertissement sont avant tout associés à des sons jugés de façon négative, tels les avertisseurs de voitures et de véhicules prioritaires, avant d'être considérés parfois comme plus agréables, telles la sonnette de vélo et certaines sonneries de téléphones.

Enfin, les termes employés par les sujets correspondent généralement à des effets sur les sujets. Une analyse linguistique relative à ce phénomène a été menée par David (1997).

## 4. Résultats des seuils de perception

En écoutant les séquences, les sujets doivent indiquer dès qu'ils perçoivent un signal d'avertissement ; la séquence sonore est alors interrompue. Le numéro d'itération du signal - cinq itérations, exceptionnellement quatre, à niveaux sonores croissants (cf. 2.1.3) - est relevé par l'expérimentateur afin de déterminer le niveau seuil de perception pour chaque signal et pour les deux contextes urbains (circulation et jardin public) diffusés à des niveaux sonores donnés (81 dBA max. et 65 dBA max. respectivement). Les verbalisations effectuées par sujets permettent, dans un premier temps, de vérifier si le signal perçu par les sujets correspond au signal testé. Si ce n'est pas le cas (les sujets citent des "bruits de pas", des "voix", des "oiseaux"...), les sujets réécoutent la séquence concernée. Dans un second temps, ces productions langagières font l'objet d'une analyse verbale détaillée (§ 5 et 6).

### 4.1 Itérations et seuils

Les résultats des seuils d'identification sont présentés en Figure 4-5 et Figure 4-6. Rappelons que la correspondance des numéros avec les noms des signaux est donnée au § 2.4. Malgré un nombre de sujets un peu faible pour procéder à des analyses probabilistes, nous pouvons observer, pour deux tiers des signaux, un comportement proche d'une répartition statistique des réponses selon une forme gaussienne (loi normale centrée). Cela permet une détermination aisée des itérations correspondant aux seuils d'identification. Toutefois, seules huit des trente séquences présentent des écarts-type suffisamment faibles pour ne pas atteindre les itérations des seuils inférieur et supérieur. Les écarts-type sont mesurés séparément pour les deux groupes de treize personnes, ainsi que pour l'ensemble des vingt-six personnes (cf. annexe 2).

Ainsi, seuls les signaux 4 et 9 permettent d'obtenir clairement des niveaux seuils pour les deux ambiances, de même que les signaux 1, 2, 7 et 14 pour l'ambiance de jardin public. Cependant, l'analyse précise des pourcentages de réponses pour chaque niveau seuil, ainsi que de leur distribution pour chaque signal d'avertissement, permet de déterminer aisément les autres niveaux seuils en considérant le premier pourcentage significatif. Si l'itération du seuil suivant (plus élevé) présente un taux de réponse anormalement faible, c'est qu'un objet sonore de l'ambiance masque le signal à cet instant précis.

Les seuils retenus pour chaque signal sont donnés dans la Figure 4-7, par le numéro d'itération du signal (première colonne de chaque contexte).

L'effet d'ordre est faible et se manifeste essentiellement par une augmentation du seuil ou de l'écart-type  $\Delta$  de la première séquence de chaque ambiance sonore (cf. annexe 2). Le seuil retenu est alors celui du groupe ayant écouté cette séquence en dernier pour chaque ambiance, car déjà habitué au test.

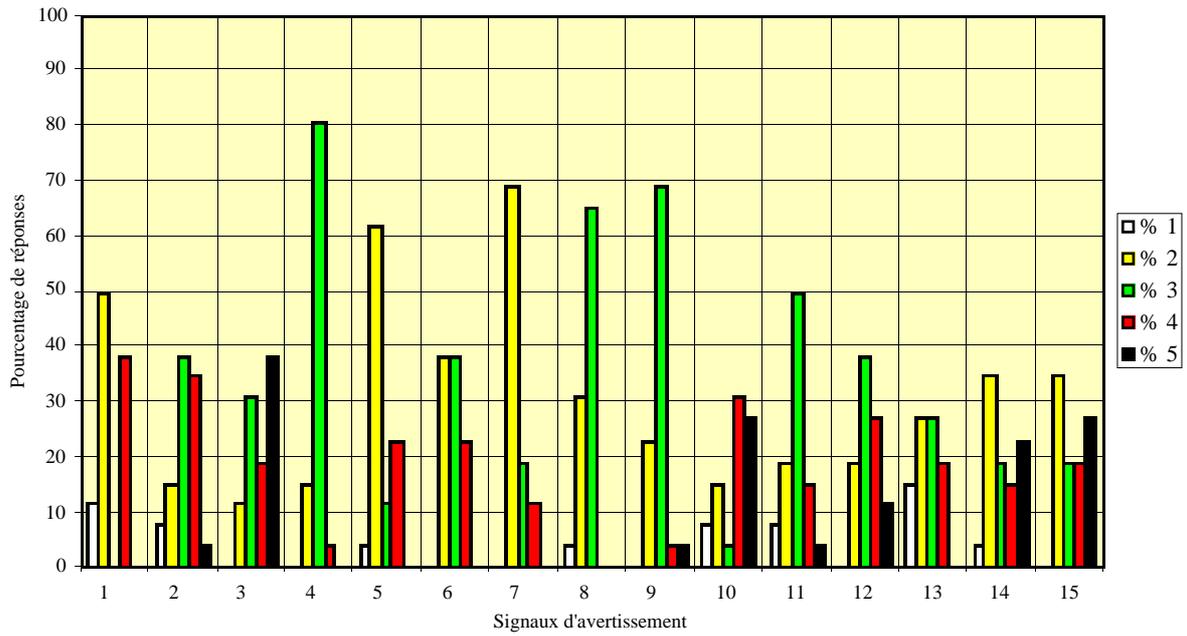


Figure 4-5 : niveaux seuils des 15 signaux en circulation.

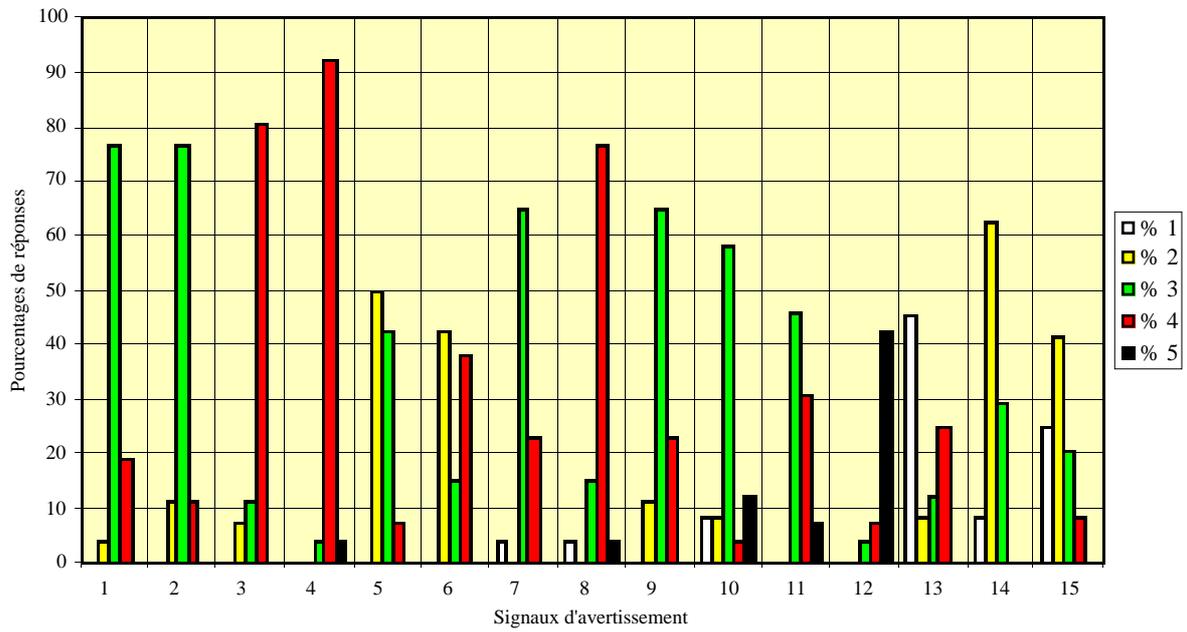


Figure 4-6 : niveaux seuils des 15 signaux en jardin public.

## 4.2 Analyse quantitative des résultats

Une première approche de ces résultats consiste à comparer les niveaux seuils obtenus en fonction des contextes d'écoute. Notre objectif n'est pas de mesurer des seuils de détection de signaux "bruités", mais uniquement de déterminer des seuils d'émergence perceptive, afin de pouvoir modifier de manière équivalente les niveaux d'émergence de tous les signaux. Or, les lignes isosoniques étant quasiment parallèles entre elles pour des variations d'intensité inférieures à 20 dB, toute unité de mesure peut convenir, exclusion faite de celles aux temps d'intégration ou de moyennage trop longs (supérieurs à la durée du signal, soit 0,5 s pour le plus court). Les pas d'intensité entre chaque itération d'un signal étant de 3 dB, cette comparaison ne pourra de toute façon qu'être empirique. Nous avons choisi le décibel pondéré A, puisqu'il s'agit de l'unité la plus courante et, de ce fait, autorise le plus de comparaisons, notamment avec les études de Szeto *et al.* (1991) et Binnington *et al.* (1993).

### 4.2.1 Calcul des niveaux sonores

La détermination des niveaux sonores nécessite plusieurs étapes. Le niveau de restitution est calibré par rapport à la situation réelle : niveau effectif moins 10 dB pour la circulation, soit 81 dBA<sub>max</sub>, et niveau effectif pour le jardin public, soit 65 dBA<sub>max</sub> (cf. § 2.1.4). Afin de ne pas avoir à modifier le volume sonore de restitution au cours du test, il faut tenir compte de ces différences de niveaux sur la bande D.A.T.. La dynamique de cette dernière, ainsi que celle de la station de montage audionumérique, et le faible niveau de bruit de fond dans la chambre semi-anéchoïque autorisent cette démarche. Les réglages intensimétriques sur la station de montage nous permettent d'atteindre le niveau maximum d'enregistrement sur la bande D.A.T. et donc, de bénéficier de toute sa dynamique. Ces réglages doivent être pris en compte lors du calcul des niveaux absolus. Ainsi, nous avons effectué sur la station Sonic Solutions :

- une atténuation de 10 dB pour la restitution des signaux isolés. Leurs niveaux mesurés correspondent à la colonne "dBA<sub>iso</sub>" de la Figure 4-7.
- un gain de +6 dB pour la restitution des stimuli en contexte de circulation. La colonne "dBA<sub>amb,c</sub>" de la Figure 4-7 donne les niveaux mesurés du fond sonore à l'instant correspondant au seuil d'identification - mesure en absence du signal.
- une atténuation de -10 dB pour la restitution des stimuli en contexte de jardin public. La colonne "dBA<sub>amb,j</sub>" de la Figure 4-7 indique les niveaux mesurés du fond sonore à l'instant correspondant au seuil d'identification - mesure en absence du signal.

Connaissant le niveau seuil de chaque signal, en niveau relatif donné par la station de montage, nous pouvons en déduire les *niveaux absolus des signaux à leur seuil d'identification* dans chaque contexte. Nous avons :

- pour la circulation :  $dBA_{sig,c} = dBA_{iso} + 6 + 10 + dB_{seuil,c}$
- pour le jardin public :  $dBA_{sig,j} = dBA_{iso} - 10 + 10 + dB_{seuil,j} = dBA_{iso} + dB_{seuil,j}$

### 4.2.2 Seuils en fonction des contextes

Pour chaque signal d'avertissement, la différence des niveaux seuils entre la circulation et le jardin public est donnée par :

$$\Delta(dBA_{sig}) = dBA_{sig,c} - dBA_{sig,j}$$

et la différence des niveaux ambiants de circulation et de jardin public au moment de l'apparition du signal seuil vaut :

$$\Delta(\text{dBAamb}) = \text{dBAamb,c} - \text{dBAamb,j}.$$

Ces deux relations permettent d'obtenir un critère de prédétermination d'un signal d'avertissement dans une ambiance plutôt que dans l'autre :

$$\mathcal{E}(\text{c-j}) = \Delta(\text{dBAamb}) - \Delta(\text{dBAsig}).$$

Nous remarquerons que  $\Delta(\text{dBAamb})$  ne présente aucune pertinence physique dans l'absolu. En effet, les niveaux mesurés sont des niveaux globaux moyennés sur tout le spectre auditif qui ne font pas état des différences spectrales entre l'ambiance de circulation, riche en fréquences basses, et celle de jardin public, plutôt riche en fréquences aiguës. Toutefois, les composantes spectrales à l'apparition des signaux seuils sont quasiment les mêmes dans un contexte donné, les fonds sonores évoluant peu sur toute la durée des séquences. Il est alors possible de comparer les quinze signaux entre eux.  $\Delta(\text{dBAamb})$  ne constitue donc qu'une référence de comparaison pour les signaux considérés. Les incertitudes sur les niveaux seuils étant de l'ordre de  $\pm 3$  dB, seuls les signaux d'avertissement se trouvant aux deux extrémités de l'échelle des  $\mathcal{E}(\text{c-j})$  peuvent être comparés (Figure 4-7).

Or, il s'avère que la moyenne des quinze valeurs  $\mathcal{E}(\text{c-j})$  est inférieure à 0,3. Nous pouvons donc raisonnablement considérer que :

- si  $\mathcal{E}(\text{c-j}) \gg 0$ , le signal est perçu plus tôt dans le jardin public que dans la circulation,
- si  $\mathcal{E}(\text{c-j}) \ll 0$ , le signal est perçu plus tôt dans la circulation que dans le jardin public.

Signal	Isolé	Circulation				Jardin public				Variations		
		n°	dBA,iso	itération	dBAamb,c	dBseuil,c	dBAsig,c	itération	dBAamb,j	dBseuil,j	dBAsig,j	$\Delta(\text{dBAamb})$
1	77	2	72	-33	60	3	59	-30	47	13	13	0
2	75	3	74	-27	64	3	59	-27	48	15	16	-1
3	80	3	72	-30	66	4	58	-24	56	14	10	4
4	72	3	72	-27	61	4	58	-27	45	14	16	-2
5	76	2	72	-30	62	2	60	-30	46	12	16	-4
6	83	2	74	-33	66	2	58	-33	50	16	16	0
7	80	2	72	-30	66	3	59	-27	53	13	13	0
8	70	3	72	-21	65	4	58	-18	52	14	13	1
9	68	3	72	-24	60	3	59	-21	47	13	13	0
10	69	4	73	-42	43	3	59	-42	27	14	16	-2
11	65	3	74	-36	45	3	59	-39	26	15	19	-4
12	80	3	74	-30	66	5	58	-24	56	16	10	6
13	80	2	72	-45	51	1	59	-39	41	13	10	3
14	74	2	72	-30	60	2	60	-27	47	12	13	-1
15	78	2	72	-36	58	2	58	-30	48	14	10	4

Figure 4-7 : seuils de perception des signaux d'avertissement selon le contexte d'écoute.

### 4.2.3 Les seuils de détection

L'utilisation de la pondération A pour mesurer les seuils de détection de signaux dans un bruit de fond relativement intense, même si elle est contestable (cf. chapitre 1, § 6.1.1), nous permet de comparer nos résultats avec ceux de Szeto *et al.* (1991) et Binnington *et al.* (1993).

La Figure 4-8 récapitule et résume les données intensimétriques de la Figure 4-7. Dans un premier temps, il est remarquable le niveau du contexte varie très peu, les écart-types étant

inférieurs à 1 dB. En revanche, les seuils de détection des signaux sont très différents. Dans un second temps, nous pouvons analyser les résultats signal par signal et constater que les sonnettes de vélo 10 et 11 nécessitent beaucoup moins d'énergie que les autres avertisseurs pour être détectés. Il en va de même pour le sifflet 13 par rapport aux sifflets 14 et 15. Les signaux 3, 6, 7, 8 et 12, au contraire, demandent un niveau plus élevé pour être perçus en contexte urbain et notamment, le nouvel avertisseur de l'ambulance, le signal 3, par rapport à l'ancien, l'avertisseur 4. Ainsi, l'ancienne sirène d'ambulance est plus efficace que la nouvelle du point de vue des niveaux sonores.

Signal n°	Circulation		Jardin public		Seuils de détection (dB)	
	dBAamb,c	dBAsig,c	dBAamb,j	dBAsig,j	$\Delta(\text{sig-amb}),c$	$\Delta(\text{sig-amb}),j$
1	72	60	59	47	-12	-12
2	74	64	59	48	-10	-11
3	72	66	58	56	-6	-2
4	72	61	58	45	-11	-13
5	72	62	60	46	-10	-14
6	74	66	58	50	-8	-8
7	72	66	59	53	-6	-6
8	72	65	58	52	-7	-6
9	72	60	59	47	-12	-12
10	73	43	59	27	-30	-32
11	74	45	59	26	-29	-33
12	74	66	58	56	-8	-2
13	72	51	59	41	-21	-18
14	72	60	60	47	-12	-13
15	72	58	58	48	-14	-10
Moyenne	72,6	59,5	58,7	45,9	-13,1	-12,8
Ecart-type	0,9	7,5	0,7	8,9	7,6	9,1

Figure 4-8 : Seuils de détection des signaux.

En comparant ces données avec celles de Szeto *et al.* (1991) et Binnington *et al.* (1993), il apparaît que les seuils que nous avons mesurés sont plus élevés d'environ 10 dB que ceux de leurs études. En effet, Binnington *et al.* ont trouvé un rapport "signal sur bruit" de -23 dB pour leur sirène et leur klaxon, tandis que Szeto *et al.* ont mis en évidence une détection, par plus de 50% des sujets, à partir de -25 dB pour les deux signaux étudiés.

### 4.3 Variation des seuils selon le contexte

Les valeurs du critère  $\mathcal{E}(c-j)$  en fonction des signaux indiquent les tendances de ces derniers à être perçus plus tôt dans une ambiance de jardin public ou de circulation. Elles sont présentées dans la Figure 4-9.

Il apparaît que les sonnettes de bicyclettes 10 et 11 sont perçues à un seuil plus bas dans le contexte de jardin public que dans la circulation. Il en va de même pour les signaux d'urgence des pompiers, 5, et l'ancien avertisseur mécanique des ambulances, 4.

En revanche, le signal 3 (nouvel avertisseur électronique des ambulances) est perçu à un seuil plus bas dans la circulation que dans le jardin public. Il en est également ainsi pour l'avertisseur 12 (poire) et les sifflets 15 et 13, mais pas pour le sifflet 14.

Pour les autres signaux, les écarts sont inférieurs à 5 dBA et les incertitudes de mesures ne permettent pas de dégager d'autres tendances.

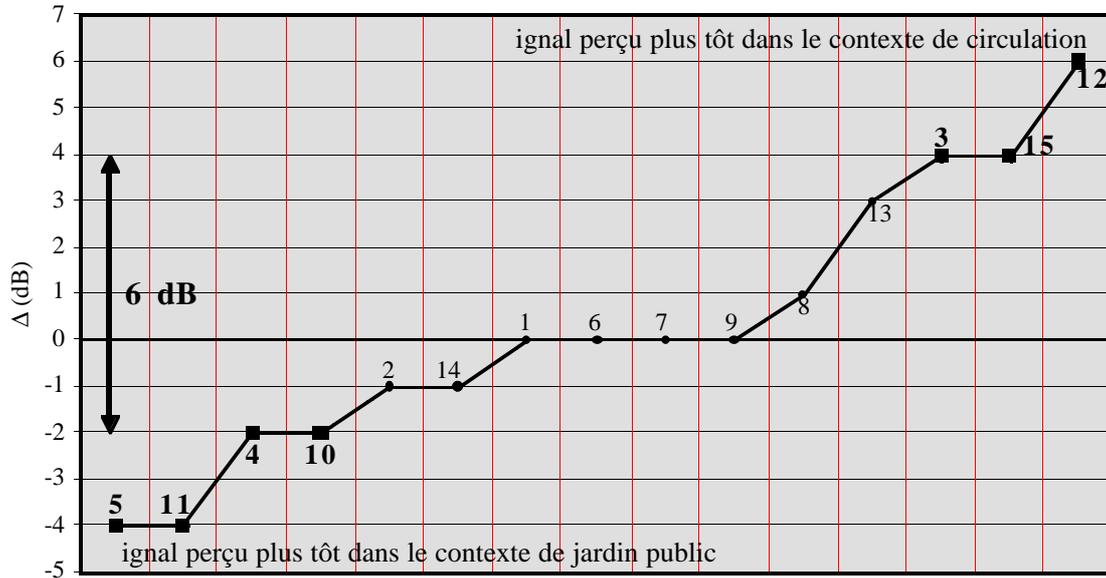


Figure 4-9 : effets des contextes sur la perception des niveaux seuils.

Retenons donc que les caractéristiques acoustiques du signal 5 et des sonnettes de vélo semblent plus efficaces dans une ambiance sonore de jardin public, tandis que celles des signaux 12, 3 et 15 le sont plutôt dans une ambiance de circulation.

Indépendamment des contextes d'écoute, les signaux 10, 11 et 13 ont des seuils de détection plus bas que les autres signaux. Mais la question reste entière, à cette étape de la recherche, de savoir si leur identification est aisée ou non. C'est d'ailleurs l'objet de la suite de cette étude.

Ces résultats seront mis en relation avec les résultats des verbalisations des sujets, au § 7 de ce chapitre. L'analyse acoustique présentant les spectres des signaux et des contextes comparés entre eux sera, quant à elle, présentée au chapitre 6.

## 5. Résultats de la verbalisation des sujets

### 5.1 Répertoire lexical disponible sur l'ensemble des données

Le mode de traitement des données est le même que celui présenté dans les résultats de l'épreuve verbale (§ 3). L'ensemble de toutes les données réunies constitue un corpus de 1170 données (26 sujets, 3 contextes d'écoute et 15 signaux d'avertissement). Ces 1170 réponses des sujets sont réparties sur 559 types différents d'expressions linguistiques dont seulement 85 produites au moins deux fois.

- Parmi ces 559 désignations recueillies, nous retrouvons 474 (soit 85%) d'*hapax*, expressions qui n'ont été produites qu'une seule fois. Par exemple, une seule personne a mentionné "trompette de match de football" pour un son de klaxon de voiture hors contexte urbain.

- Parmi les 85 désignations produites au moins deux fois, les plus fréquentes sont "ambulance" (74 occurrences), "klaxon" (73), "sifflet" (47), "pompiers" (46), "?" (37 occurrences sans réponse), "sonnette de vélo" (36 occurrences) et "police" (34). La Figure 4-10 représente les 17 désignations les plus fréquentes ( $f \geq 10$ ) et permet d'identifier les distributions de ces productions verbales sur les trois contextes d'écoute - isolé, circulation et jardin public.

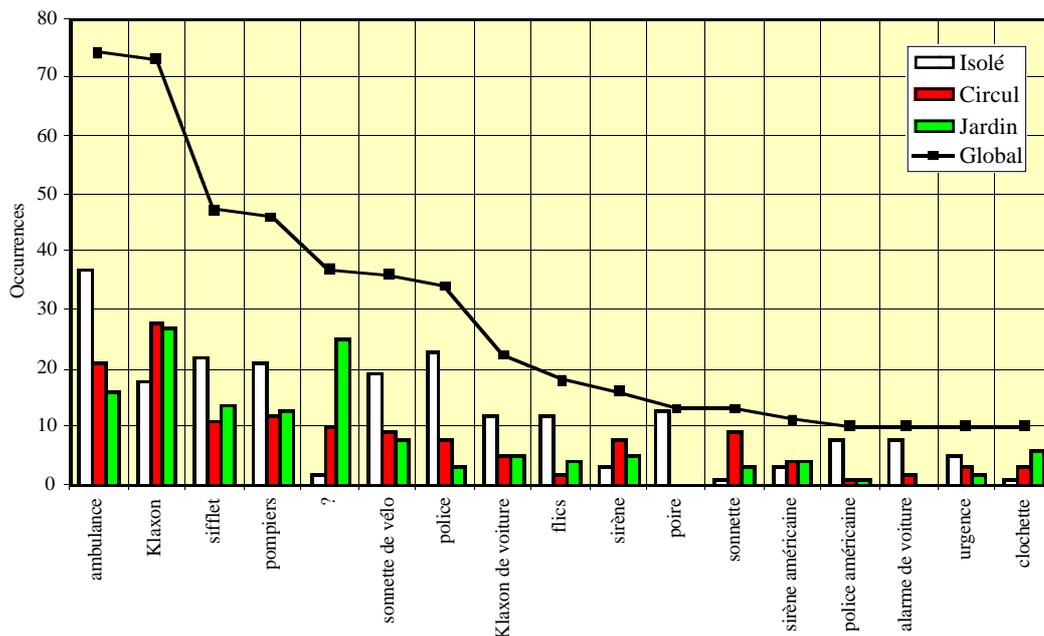


Figure 4-10 : désignations les plus fréquentes (après écoute).

En considérant les données mélangées et compilées sur l'ensemble des contextes, nous obtenons des indications sur le répertoire lexical disponible, cela de manière indépendante par rapport aux signaux et aux contextes. Les désignations les plus fréquentes ( $f \geq 2$ ) permettent d'identifier celles qui correspondent à :

- des **propriétés physiques du son** : 15% d'onomatopées ("pin pon, pouêt, bip, ding...") et de termes génériques ("sonnerie, sifflement..."),
- des **noms de sources interprétées** : 85% ("ambulance, klaxon, sifflet, pompiers, police, flics, sirène, poire, sonnette, clochette..."), dont les "**noms de noms**" ("sonnette de vélo,

klaxon de voiture, alarme de voiture, sifflet de flics...") représentent **33%** des désignations les plus fréquentes.

Le répertoire lexical disponible sur l'ensemble des données est très riche mais très concentré sur un type de la classification linguistique. Par conséquent, une analyse plus détaillée de ce vocabulaire s'impose afin de corrélérer l'identification des sources (et leurs significations) aux désignations des différents signaux (cf. annexe 2).

## 5.2 Description des 15 signaux dans les différentes situations d'écoute

L'ensemble des 1170 données ne se répartit pas de manière aléatoire selon les signaux, et des formes verbales différentes sont différemment associées aux différents signaux. Dès lors, il devient possible, à partir d'hypothèses formulées et évaluées par ailleurs (Dubois, 1993b), d'identifier les caractéristiques pertinentes des différents signaux, celles perçues et intégrées comme significatives pour le sujet, en fonction de ses connaissances et de sa mémoire préalable. Pour chaque signal, nous avons établi un classement de ces formes verbales selon les trois situations d'écoute.

Les noms de sources interprétés nous donnent directement accès à la signification du signal décrit et nous permettent ainsi de classer les signaux en fonction du sens qui leur a été attribué. Les propriétés physiques du son, décrites au cours de l'expérience, nous permettent de faire le lien entre le sens affecté aux signaux et des paramètres acoustiques, et de formuler les hypothèses relatives aux représentations mentales. Ces paramètres constitueront les variables pertinentes pour la réalisation des autres expérimentations, à partir de la représentation mentale du son et sans *a priori* de spécialistes du son.

### 5.2.1 Les absences de réponse ("?")

Lorsqu'un sujet est certain d'avoir détecté une itération de signal mais n'arrive, ni à identifier ce dernier, ni même à lui reconnaître une structure sonore particulière (réponse du type : "je ne sais pas ce que c'est"), nous tenons compte de l'itération sans production verbale. Le sujet a dépassé le seuil de détection, mais pas encore celui de reconnaissance et encore moins celui d'identification. Or, pour la plupart des sujets, ces trois seuils sont extrêmement proches puisque l'itération précédant celle d'identification n'est pas même détectée. L'écart entre les trois seuils doit donc être généralement inférieur à 3 dB. Néanmoins, cette indécision quant à la reconnaissance du signal perçu peut également être interprétée en termes d'*ambiguïté* du signal aux "abords" de son seuil de détection. A même niveau sonore et dans un même contexte, un signal sera d'autant plus efficace que son seuil d'identification est proche de celui de détection, puisqu'il devient *univoque* dès la détection. Dès lors, il devient intéressant d'analyser les occurrences sans production verbale. La Figure 4-11 représente ces occurrences "?" en fonction des signaux et des contextes.

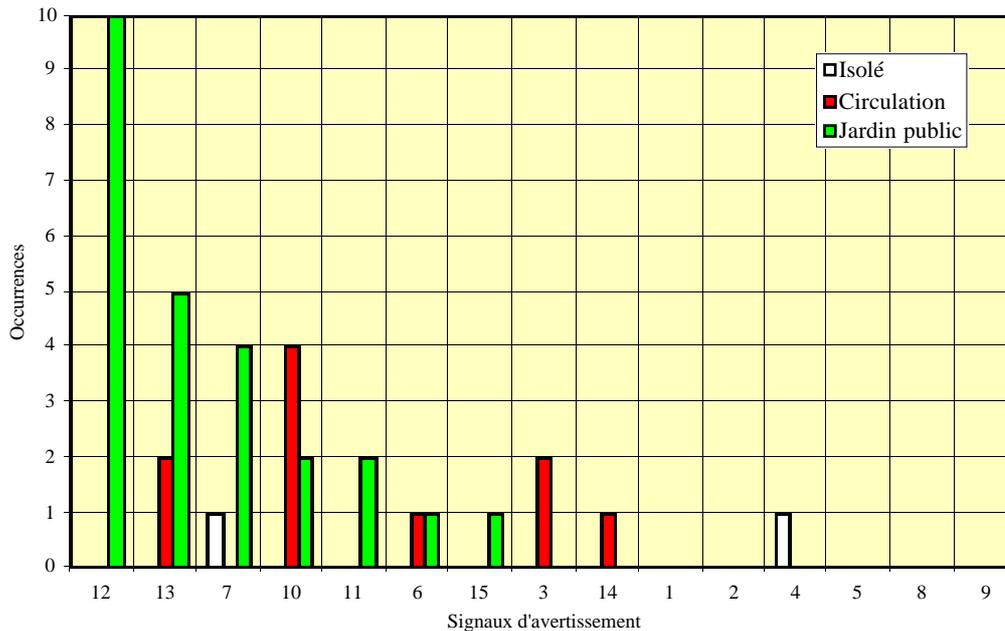


Figure 4-11 : absences de réponse.

#### 5.2.1.1 En fonction des signaux...

Les signaux de police, 1 et 2, de pompiers, 5, et les klaxons, 8 et 9, ne sont pas concernés par ce type de réponse et peuvent donc être considérés comme toujours nettement identifiés. En revanche, quatre signaux sont particulièrement affectés et représentent à eux seuls les trois-quarts des 37 occurrences "?". Il s'agit de la "poire" de bicyclette 12, du sifflet à un ton 13, de la sirène rapide du lance-eau 7, qui ne pose cependant problème que dans le jardin public, et de la sonnette "classique" de bicyclette 10, dont les occurrences "?" apparaissent plutôt dans la circulation. Outre les différences de verbalisation, sur lesquelles nous reviendrons plus loin, il apparaît clairement que les signaux d'urgence les plus pertinents, auxquels les sujets sont quasi-quotidiennement soumis, donnent toujours lieu à une production verbale.

#### 5.2.1.2 ...et en fonction du contexte

Sur les 37 occurrences d'absence de réponse, 95% concernent l'écoute en situation contextuelle. L'écoute en contexte de jardin public en compte 68%, les 27% restant concernant l'écoute en contexte de circulation.

#### 5.2.1.3 Interprétation

La différence entre l'écoute isolée et celle en situation s'explique du fait des conditions d'écoute : les signaux isolés sont reproduits à un niveau confortable pour être perçus sans ambiguïté, tandis qu'en contexte urbain, ils sont restitués à leurs niveaux seuils de perception dans une ambiance bruitée. Il est évidemment beaucoup plus difficile, dans nos conditions d'écoute, de distinguer les signaux "bruités" des signaux isolés.

En revanche, la différence entre circulation et jardin public peut s'expliquer par deux raisons principales : d'une part, il est plus difficile d'interpréter clairement un signal dans une ambiance de jardin public, riche en événements sonores très différents (cris, pas, sifflements,

conversations...), que de circulation, dont le contenu sonore est plus homogène (passage de véhicules, bruits de roulement et de moteurs) ; d'autre part, les signaux d'avertissement de véhicules peuvent paraître incongrus dans un contexte de jardin public, ce qui expliquerait la différence contextuelle en raison d'un phénomène de congruence sémantique. Mais cela n'empêche pas certains signaux d'être perçus à un seuil plus bas dans une ambiance de jardin public que de circulation (cf. § 4).

## 5.2.2 Productivité et variabilité des désignations

### 5.2.2.1 Productivité des 15 signaux

Le nombre de désignations pour chaque signal d'avertissement permet d'apprécier la productivité verbale des sujets, liée au signal (Figure 4-12).

Sur l'ensemble des contextes, quatre signaux sont riches en désignations différentes : les sifflets 14 et 13, la sirène lente du lance-eau 6, et la sonnette de bicyclette 11. La multiplicité des désignations peut avoir différentes origines : 14 et 13 sont les deux signaux qui contribuent aux plus grands nombres d'onomatopées (autres que "pin pon" et "pouêt"). Par ailleurs, ils sont très ambigus et sont confondus avec des klaxons et des sirènes. Le signal 11 est, lui aussi, très ambigu et est perçu comme "un carillon, une sonnette de vélo, une porte d'entrée, un ascenseur...". Quant au signal 6, il est identifié à une "sirène" ou à une "alarme", mais ne dispose pas de terme attesté dans la langue car il n'est quasiment jamais rencontré au quotidien. Il n'est donc pas clairement identifié dans la représentation mentale des gens.

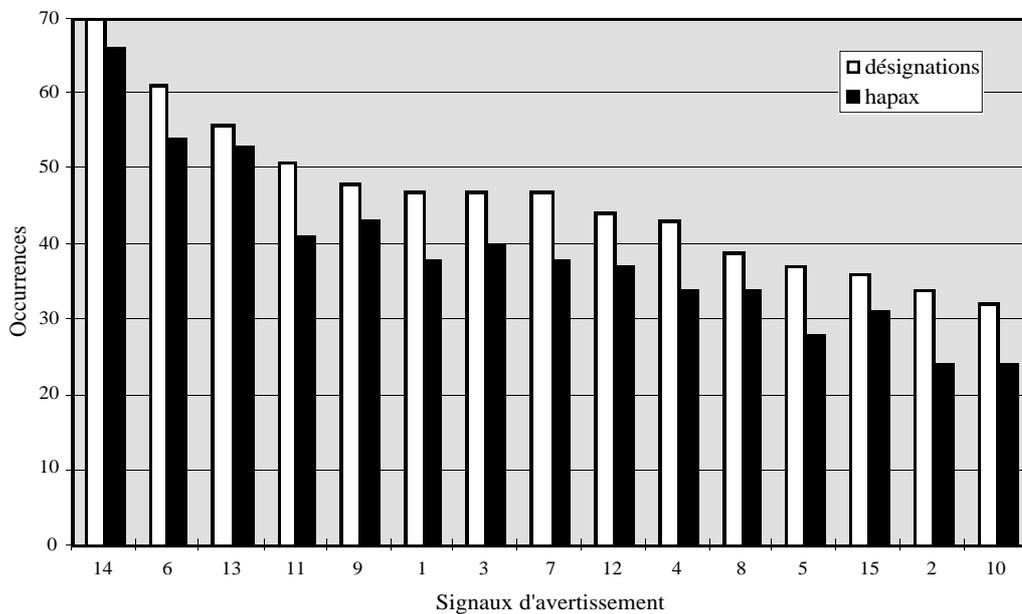


Figure 4-12 : désignations et hapax - global.

A l'opposé, la sonnette de bicyclette 10, le sifflet 15 et les sirènes 2 et 5 sont les plus pauvres en nombre de désignations. Ce sont des termes univoques, ayant une signification très forte. Le signal 10 est associé très clairement à une "sonnette de vélo", 15 à un "sifflet d'agent de la police", 2 et 5 à des avertisseurs de véhicule d'urgence ("sirène"), mais avec une certaine confusion à un niveau moins générique, entre sirène de police, de pompiers et d'ambulance.

### 5.2.2.2 Effets de contexte

En observant chaque contexte séparément (Figure 4-13, Figure 4-14 et Figure 4-15), il apparaît nettement plus de désignations en contexte urbain que de manière isolée. Par ailleurs, les signaux aux désignations les moins fréquentes présentent moins d'occurrences hors contexte (8 désignations) qu'en situation (16 et 13 désignations).

Pour les signaux les plus riches en désignations différentes, 14 et 6, le nombre de désignations est indépendant du contexte. Toutefois, 14 est identifié à un sifflet de façon univoque en écoute isolée et devient ambigu en situation urbaine ; 6 reste toujours aussi peu identifiable, même si, en contexte urbain, il perd son statut "d'alarme" au bénéfice de "sirène". En situation isolée, les termes les plus pauvres en désignations correspondent à des "prototypes" de signaux, dont la signification est complètement intégrée. Les termes employés correspondent essentiellement à des noms de sources. Les termes les plus riches correspondent, quant à eux, à des signaux bien plus ambigus, mais non nécessairement dépourvus de sens.

En revanche, 11 est le seul signal d'avertissement qui "s'appauvrit" en désignations lorsqu'il est en contexte urbain. Ainsi, le contexte, et notamment le contexte de jardin public, abaisse son niveau d'ambiguïté : l'ascenseur et la porte d'entrée n'ont plus lieu d'exister ! La sonnette 11 fait donc bien sens et devient beaucoup moins ambiguë en contexte urbain. Quant au signal 13, son ambiguïté est légèrement levée en contexte de jardin public.

Deux signaux ont des comportements remarquables en fonction des contextes. Le premier, la poire 12, a nettement moins de désignations en situation de jardin public qu'en contexte de circulation. Tout comme le sifflet 13, c'est un signal plus efficace, d'un point de vue perceptif, dans une ambiance de jardin public que dans la circulation. Inversement, la sirène 3, qui a peu de désignations en écoute isolée, en a plus dans le jardin. Mais son statut univoque d'ambulance n'est remis en cause dans aucun contexte.

Les sifflets 14 et 13 contribuent aux plus grands nombres d'onomatopées (autres que "pin pon" et "pouêt"). Par ailleurs, s'ils sont bien identifiés comme sifflets en écoute isolée, ils deviennent très ambigus en contexte urbain et sont confondus avec des klaxons et des sirènes.

Cette première analyse permet déjà de classer les signaux, non pas en termes de caractéristiques acoustiques, mais en termes de *signification* et de *polysémie*. Ainsi, les sifflets, les sonnettes et les signaux d'urgence se retrouvent dans deux classes opposées :

- sifflet 15, sonnette 10, sirènes 2 et 5 et avertisseur 8 produisent une faible polysémie et peuvent s'apparenter, du fait de leur caractère univoque, à des prototypes dans les représentations mentales des gens,
- sifflets 13 et 14, sonnette 11 et sirènes 6 et 7 produisent une forte polysémie et présentent par conséquent une certaine ambiguïté aux oreilles des sujets.

### 5.2.2.3 Variabilité : nombre d'hapax

L'analyse des hapax - expression verbale produite une seule fois - est très fortement corrélée à celle des désignations. Sur les données globales, aucune information supplémentaire ne peut être extraite des réponses des sujets, mais les hapax confirment ces analyses.

Si le signal 11 n'est pas, sur l'ensemble des données, particulièrement riche en hapax, il l'est cependant en écoute isolée. Cela conforte le constat d'ambiguïté trouvée pour ce signal. Pour les signaux 15 hors contexte et 5 en situation de jardin public, le faible nombre d'hapax conforte l'hypothèse de leur caractère univoque et indique une certaine aisance à trouver une désignation commune.

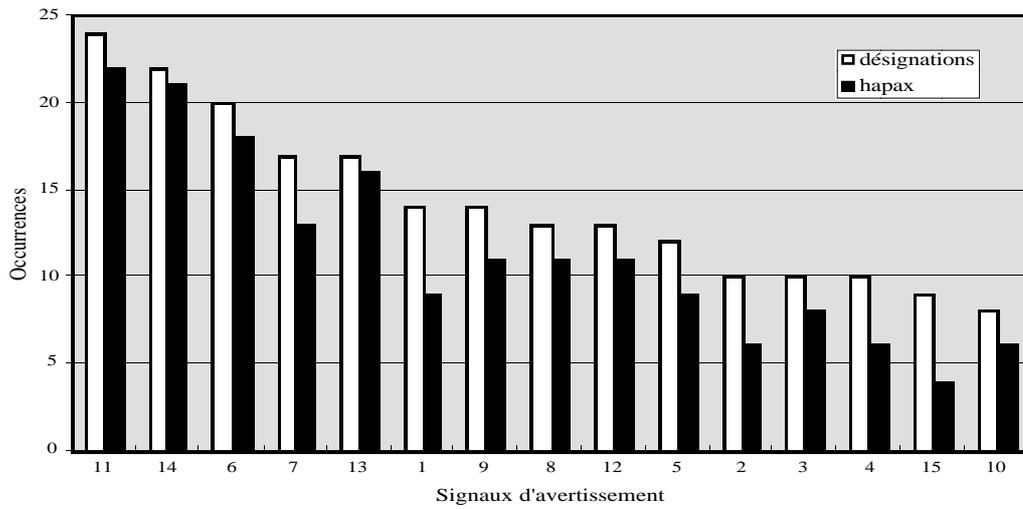


Figure 4-13 : désignations et hapax - isolé.

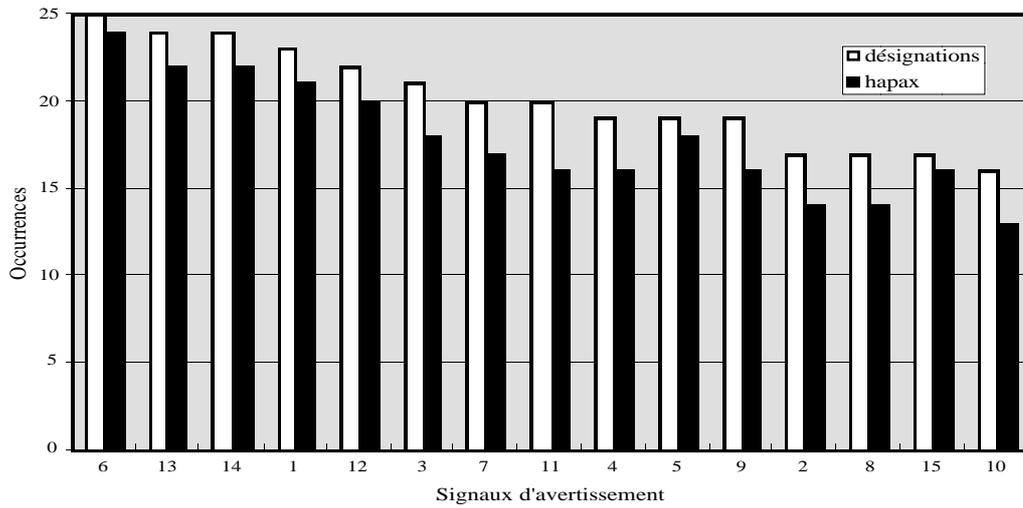


Figure 4-14 : désignations et hapax - circulation.

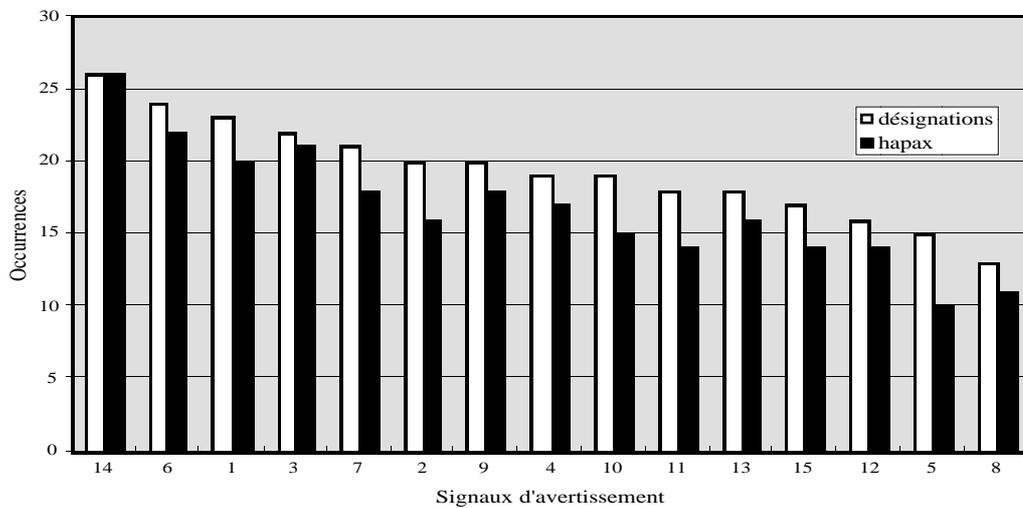


Figure 4-15 : désignations et hapax - jardin public.

### 5.2.3 Fréquence la plus élevée

La fréquence la plus élevée correspond à la désignation qui a été produite le plus grand nombre de fois par les sujets. La Figure 4-16 présente ces désignations sur l'ensemble des données, en écoute isolée et dans les deux contextes urbains. Les fréquences maximales peuvent correspondre à différentes désignations selon les contextes d'écoute, ce qui explique que les valeurs sur l'ensemble des données ne correspondent pas nécessairement aux sommes des valeurs dans les trois contextes.

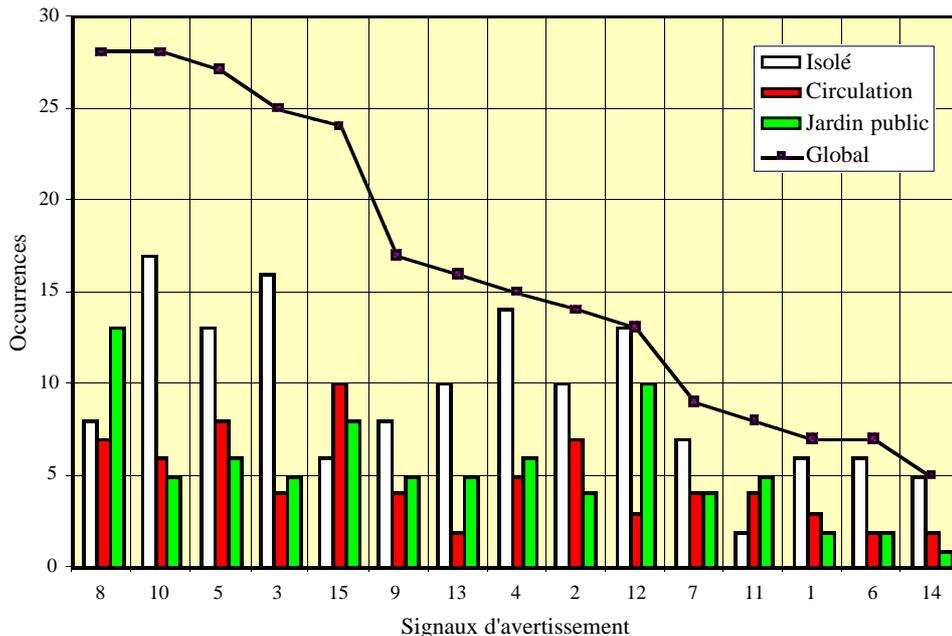


Figure 4-16 : Fréquence la plus élevée.

#### 5.2.3.1 Sur l'ensemble des données...

Les signaux 8, 10, 5, 3 et 15 ont les fréquences les plus élevées globalement et correspondent donc aux signaux qui mettent le plus de sujets en accord. Là encore, ces signaux peuvent représenter un reflet de prototypicité pour les sujets.

#### 5.2.3.2 ...et en fonction du contexte

En fonction des contextes, il apparaît que le signal 8 contribue au meilleur accord entre sujets en contexte de jardin public. Pour la circulation, il s'agit du sifflet 15, alors que c'est un des termes les moins fréquents en situation isolée.

Les fréquences maximales correspondent à une même désignation quelque soit le contexte d'écoute pour les six premiers signaux. Ainsi, la désignation "klaxon" est systématiquement la plus fréquente pour les signaux 8 et 9, "sonnette de vélo" pour le signal 10, "pompiers" pour le signal 5, "ambulance" pour le signal 3 et "sifflet" pour le signal 15.

## 5.2.4 Interprétation des désignations

L'analyse des types de termes utilisés pour chaque signal consiste à comparer les proportions de citations de propriétés physiques du son et de celles de noms de sources interprétées. Elle devrait permettre d'établir des classes sémantiques de signaux.

### 5.2.4.1 Effets de contexte

Quelle que soit la situation d'écoute, les sujets produisent principalement des désignations correspondant à des noms de sources (Figure 4-17). Les sources interprétées représentent globalement plus de quatre fois le nombre de désignations de phénomènes acoustiques. Toutefois, la mise en contexte des signaux et leur écoute à un niveau seuil de perception augmentent fortement la proportion de propriétés acoustiques liées au son, qui passent de 4% en écoute isolé à 33% en écoute contextualisée. Ces valeurs restent indicatives et ne présentent qu'une tendance, les incertitudes (écarts-type  $\Delta$ ) étant très élevées.

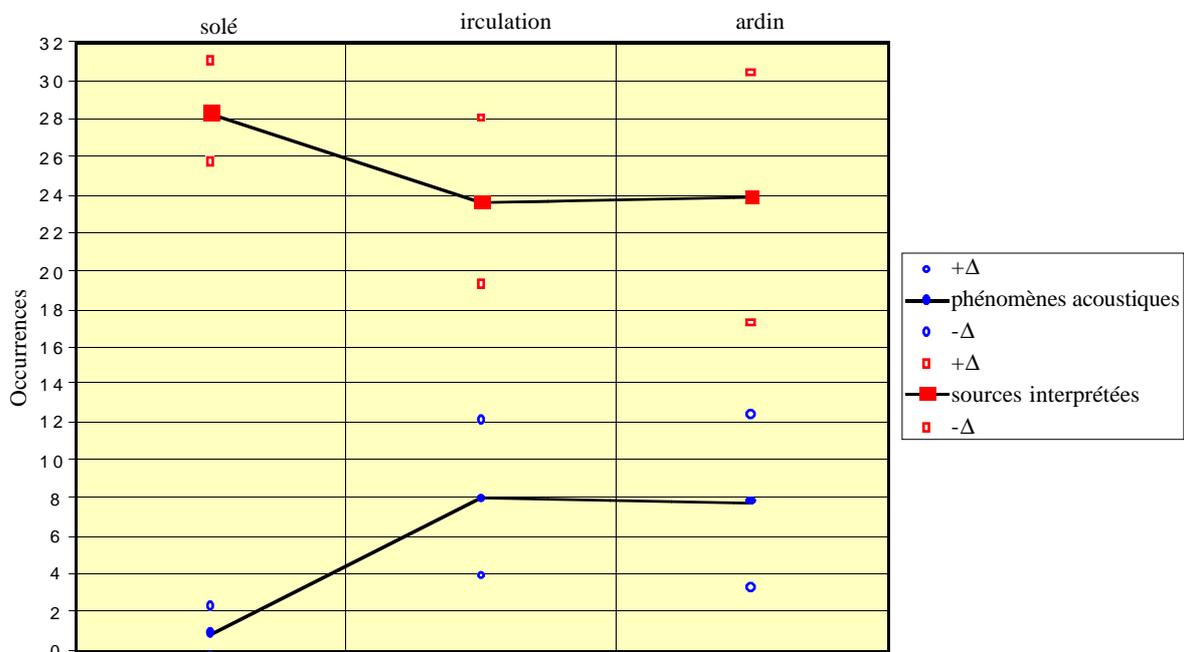


Figure 4-17 : effets du contexte.

### 5.2.4.2 Analyse par classes de signaux

Les courbes détaillées par signal - Figure 4-18, Figure 4-19 et Figure 4-20 - sont obtenues à partir d'un tri décroissant principal des occurrences de phénomènes acoustiques et d'un tri croissant secondaire des sources interprétées, suivi d'éventuelles corrections de "placement" de certains signaux afin de regrouper les avertisseurs en classes homogènes. Il apparaît que les faibles différences entre signaux ne permettent pas de mettre en évidence des classes de signaux d'avertissement. Tout au plus est-il remarquable qu'en contexte, les signaux 12, 13 et 14 présentent le plus de propriétés physiques liées au son. Une part importante de ces propriétés sont décrites par des onomatopées. En situation isolée, cette classe disparaît sous l'écrasante prédominance des sources interprétées.

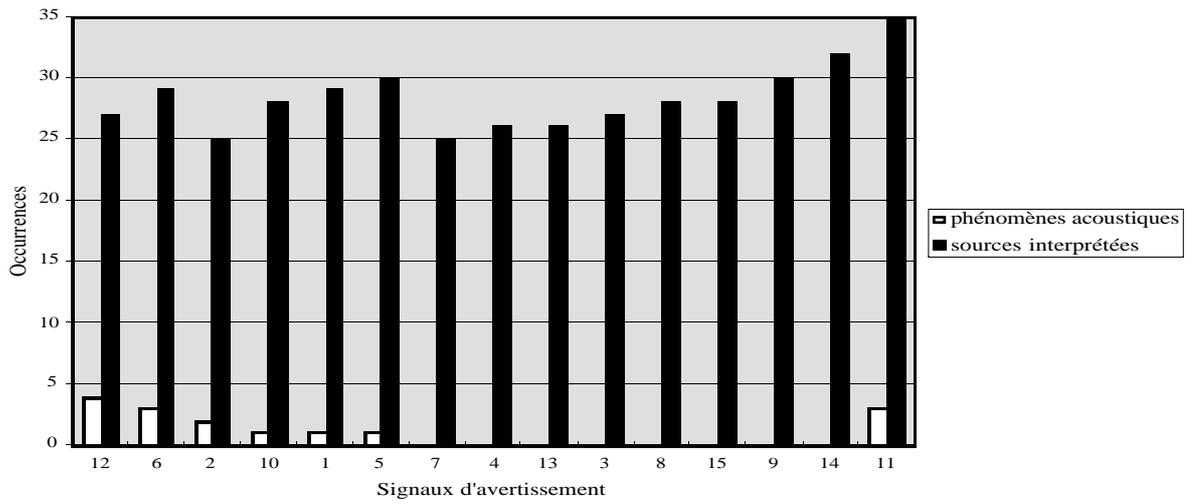


Figure 4-18 : acoustique et sémantique - isolé.

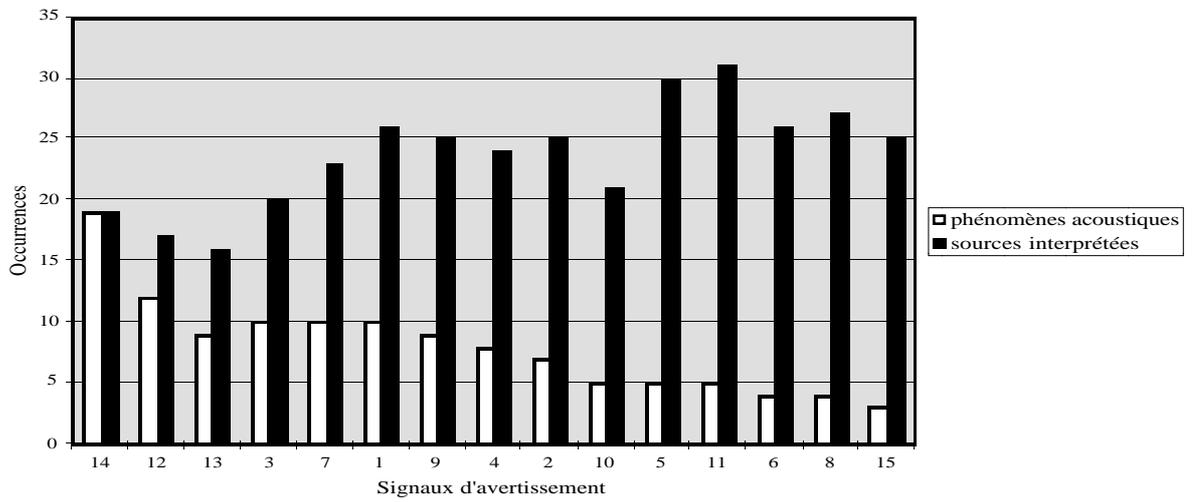


Figure 4-19 : acoustique et sémantique - circulation.

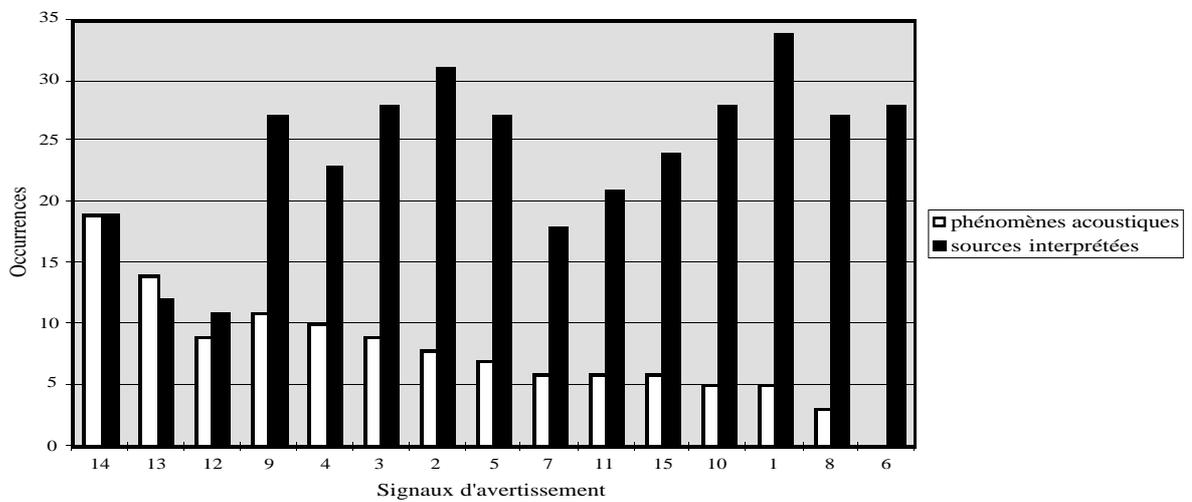


Figure 4-20 : acoustique et sémantique - jardin public.

Une première analyse avait été réalisée en considérant des termes tels que "klaxon", "sonnette" et "sirène" comme étant des propriétés physiques du son. En effet, leur description peut se faire aussi bien en termes d'objet matériel émettant un son caractéristique (source interprétée), qu'en termes de son caractéristique en soi, non associé à un objet. Par ailleurs, les *noms de noms* n'étaient pas comptabilisés dans les sources interprétées. Les résultats en contexte avaient alors fait apparaître quatre classes de signaux :

- 15, 11, 10 et 5 : très majoritairement décrits par des sources interprétées dans les deux contextes (rapport sources interprétées / propriétés du son supérieur à 2). Il s'agirait de signaux perçus comme tels, proches d'une représentation mentale prototypique des sujets ;
- 2, 3, 1 et 13 : majoritairement décrits par des sources interprétées, et de manière plus nette dans le jardin public que dans la circulation. Ainsi, la représentation mentale de ces signaux resterait précise, notamment dans le jardin public ;
- 8, 12, 7 et 6 : majoritairement décrits par des propriétés du son dans les deux contextes, mais de manière plus claire dans la circulation. Ces signaux présenteraient donc plutôt des indices décrivant le signal sonore ;
- 4, 14 et 9 : décrits essentiellement par des propriétés du son pour les deux contextes urbains (rapport propriétés du son / sources interprétées supérieur à 4). Il s'agirait, là aussi, de signaux non prototypiques, perçus en tant qu'indices acoustiques.

Toutefois, en situation isolée, ces classes sont partiellement détruites et seuls résistent les avertisseurs 8 et 9 d'une part, et le signal 15 d'autre part. Par ailleurs, nous avons pu nous rendre compte de l'extrême fragilité de ces classes en variant les critères de comparaison : rapports propriétés / sources, % entre propriétés et sources, % par rapport au total (prise en compte de "noms de noms"...). Enfin, ces quatre classes n'apparaissent pas du tout dans la seconde analyse et le seul regroupement obtenu dans cette dernière se trouve dispersé dans trois classes distinctes de la première analyse. Les classes formées ne sont donc pas résistantes à des modifications de données qui ne devraient pas, *a priori*, influencer sensiblement les résultats.

Les difficultés rencontrées pour analyser les résultats selon cette approche d'opposition entre propriétés du son et sources interprétées ne nous permettent pas de décrire, même *grosso modo*, la prototypicité des signaux. Par conséquent, nous allons directement considérer l'approche prototypique.

## 6. Signaux univoques : les prototypes

Si l'établissement de classes de signaux d'avertissement à partir des sources interprétées s'est révélé infructueux à cette étape d'analyse, la mise en évidence de signaux univoques dans les représentations mentales des gens peut s'avérer très utile pour les classer. En effet, ces signaux prototypiques devraient permettre de déterminer des caractéristiques acoustiques susceptibles de transmettre des informations comportant un contenu sémantique (cf. chapitre 6), à l'instar des unités de sens linguistique, les éléments différentiels (cf. chapitre 1, § 5.3.1.2). Cette approche est basée sur la théorie de la catégorisation prototypique (Rosch, 1978, cf. chapitre 1, § 4.3). Ce sont les données verbales, par l'intermédiaire des désignations les plus fréquentes, qui permettent d'accéder aux signaux d'avertissement prototypiques.

### 6.1 Signaux prototypes des désignations les plus fréquentes

Dans cette section, les termes les plus fréquemment produits par les sujets (cf. § 3) sont analysés en fonction de leur répartition sur les quinze signaux d'avertissement étudiés et selon les contextes d'écoute. Ainsi, pour chaque signal et chaque contexte d'écoute, le signal sonore prototypique correspond à la production verbale fournie par le plus grand nombre de sujets. Il s'agit du signal le plus univoque et le plus typique des représentations en mémoire.

#### 6.1.1 Prototypes associés à "ambulance"

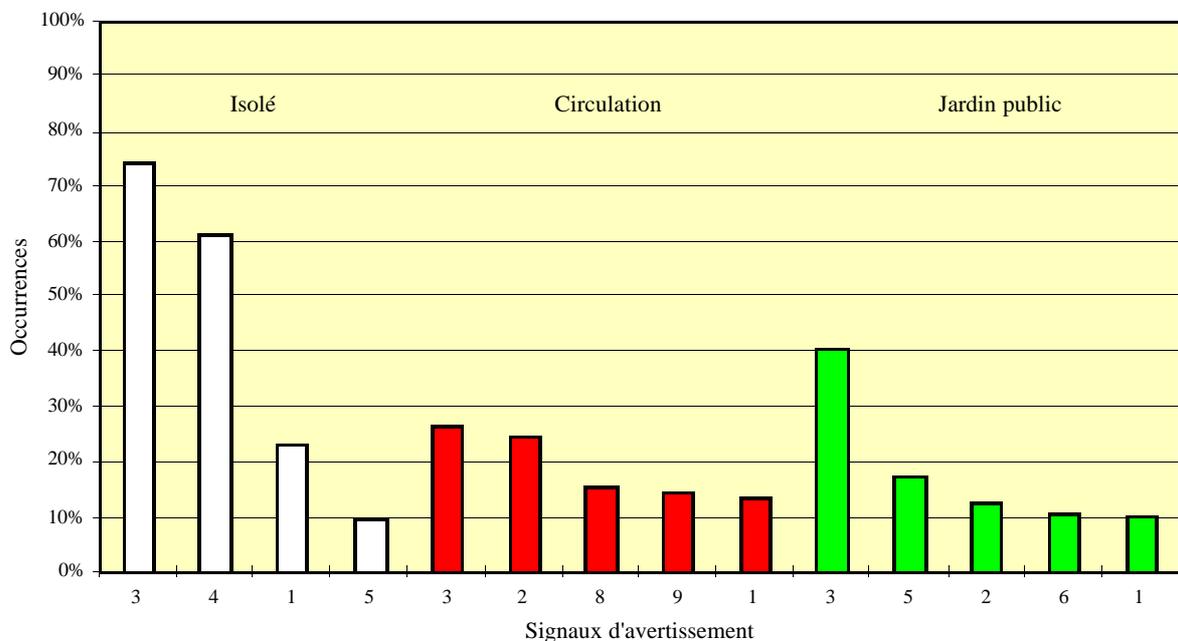


Figure 4-21 : prototypes associés à "ambulance" et "S.A.M.U.".

Les termes "ambulance" et "S.A.M.U." peuvent être considérés comme similaires dans la représentation mentale des sujets car ils décrivent la même fonction de secours. En écoute isolée, ils apparaissent le plus souvent pour les signaux 3 et 4. Ces deux signaux correspondent

effectivement à deux avertisseurs d'ambulances. En contexte, le signal 3 reste le plus prototypique "d'ambulance", mais à moindre mesure qu'en situation isolée. En revanche, le signal 4 n'est plus du tout perçu comme tel. Nous retrouverons ce dernier dans les prototypes associés à "klaxon" (cf. § 6.1.5). En contexte de circulation, seul le signal 2 (avec 3) est aussi décrit comme "ambulance". En contexte de jardin public, il s'agit plutôt du signal 5, mais dans une moindre mesure.

### 6.1.2 Prototypes associés à "police" et "flics"

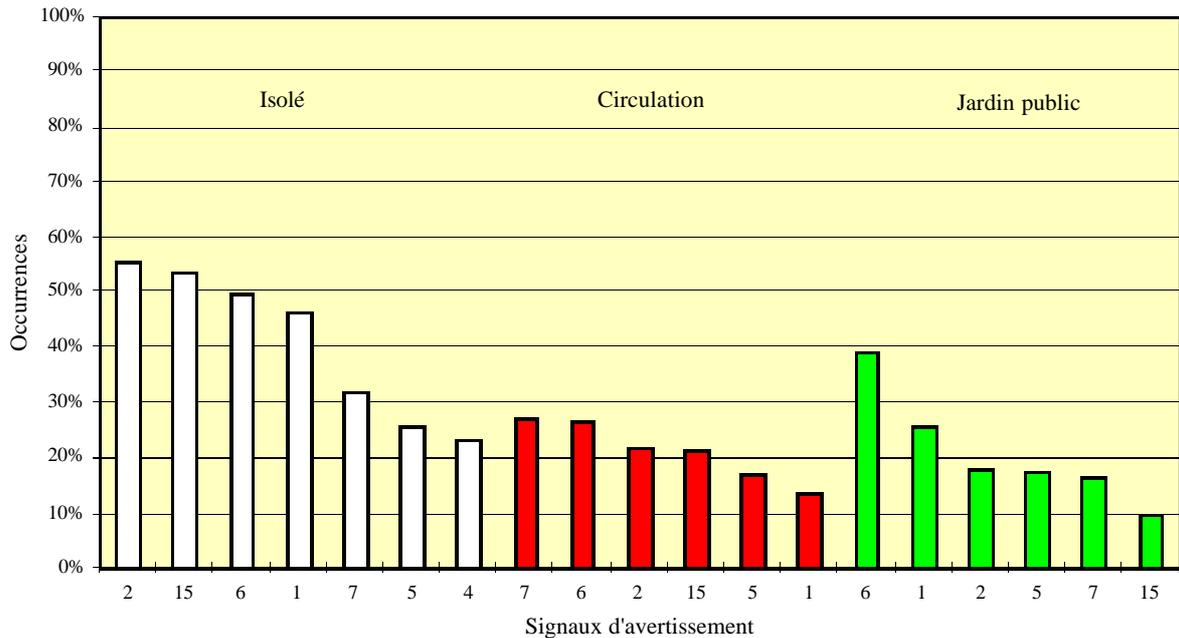


Figure 4-22 : prototypes associés à "police" et "flics".

Les termes "police" et "flics" peuvent, eux aussi, être considérés comme similaires dans la représentation mentale des sujets, même si le second terme peut, en outre, tenir compte d'un jugement de valeur. Si les sirènes 1, 2, 5, 6, 7 et le sifflet 15 sont associés à la désignation "police" et ce, quelque soit la situation d'écoute, il apparaît que la sirène 6 est plus résistante à la mise en contexte. Les signaux 15 et 5, notamment, le sont beaucoup moins. Dans ce cas précis de désignations, plusieurs signaux, aux caractéristiques acoustiques très différentes, peuvent être considérés comme prototypiques de "police".

### 6.1.3 Prototypes associés à "pompiers"

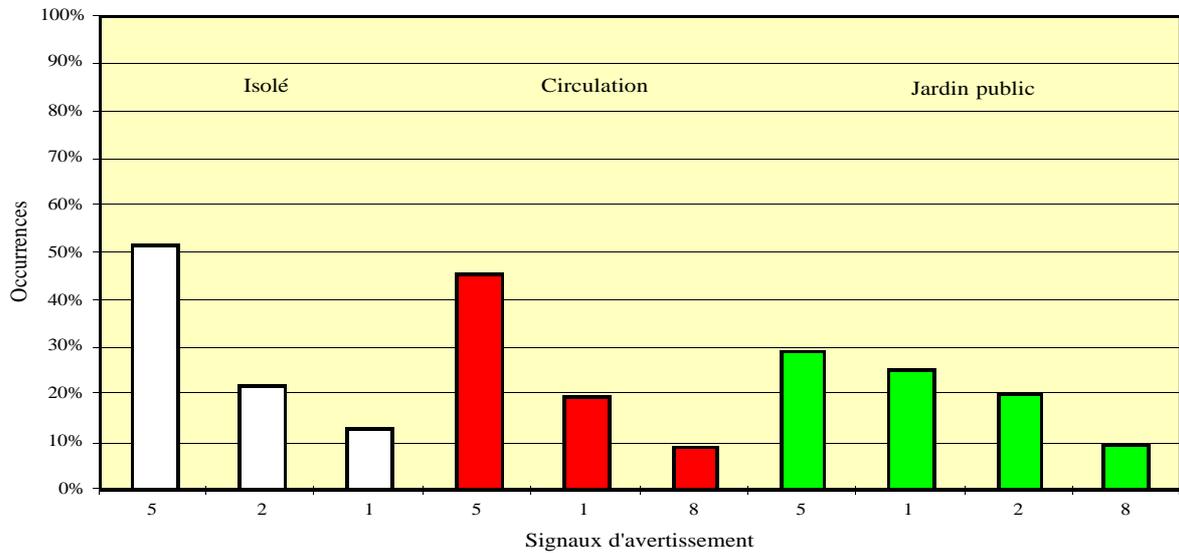


Figure 4-23 : prototypes associés à "pompiers".

La désignation "pompiers" a pour prototype l'avertisseur 5 des pompiers quelque soit le contexte. Les signaux 1 et 2 de la police y sont également associés, mais à moindre mesure.

### 6.1.4 Prototypes associés à "sirène"

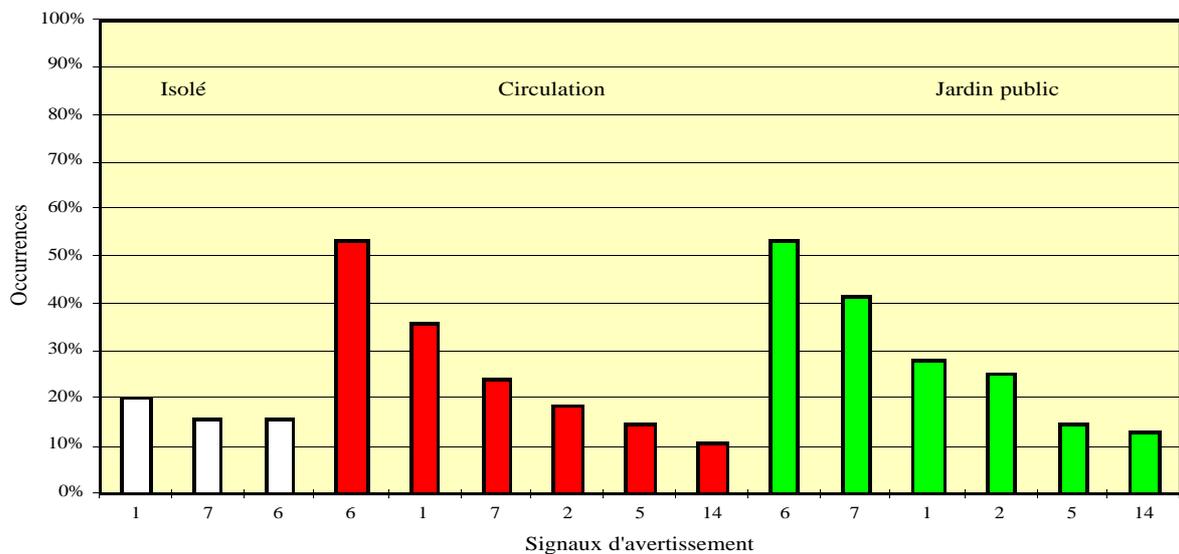


Figure 4-24 : prototypes associés à "sirène".

En situation isolée, la désignation "sirène" est très peu produite par les sujets et concerne essentiellement les signaux 1, 6 et 7. Il s'agit des trois signaux qui apparaissent dans les prototypes de "police". En contexte, les mêmes signaux restent les plus fréquemment associés aux réponses des sujets, avec des proportions plus fortes, notamment pour le signal 6.

### 6.1.5 Prototypes associés à "klaxon"

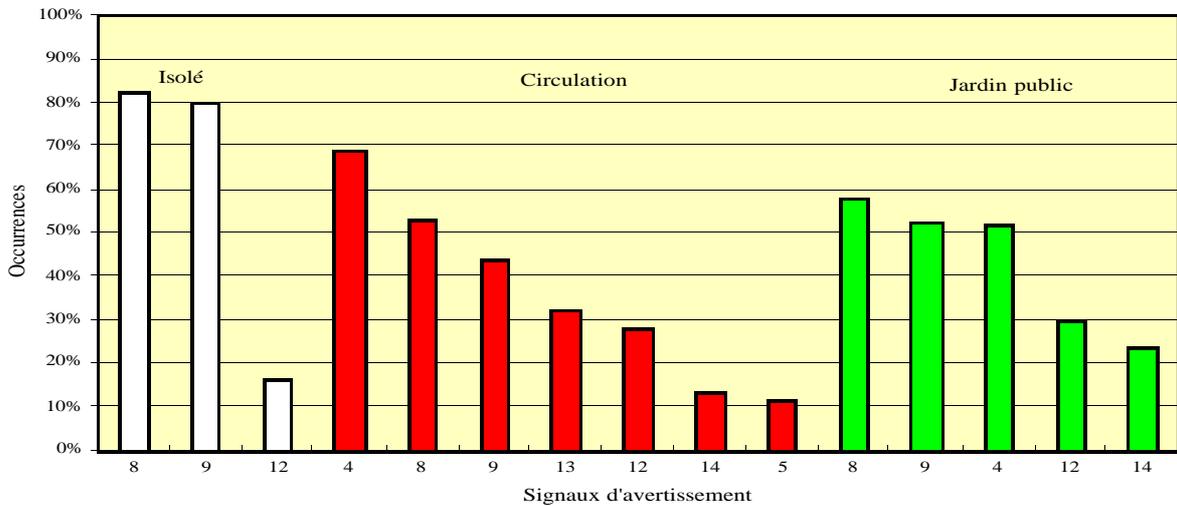


Figure 4-25 : prototypes associés à "klaxon".

Les deux avertisseurs de voiture, 8 et 9, correspondent aux prototypes de la désignation "klaxon". Toutefois, en situation urbaine, le signal 4 est fortement décrit comme "klaxon" et, dans le contexte de circulation, il obtient même un nombre d'occurrences supérieur à celui des avertisseurs 8 et 9.

### 6.1.6 Prototypes associés à "voiture"

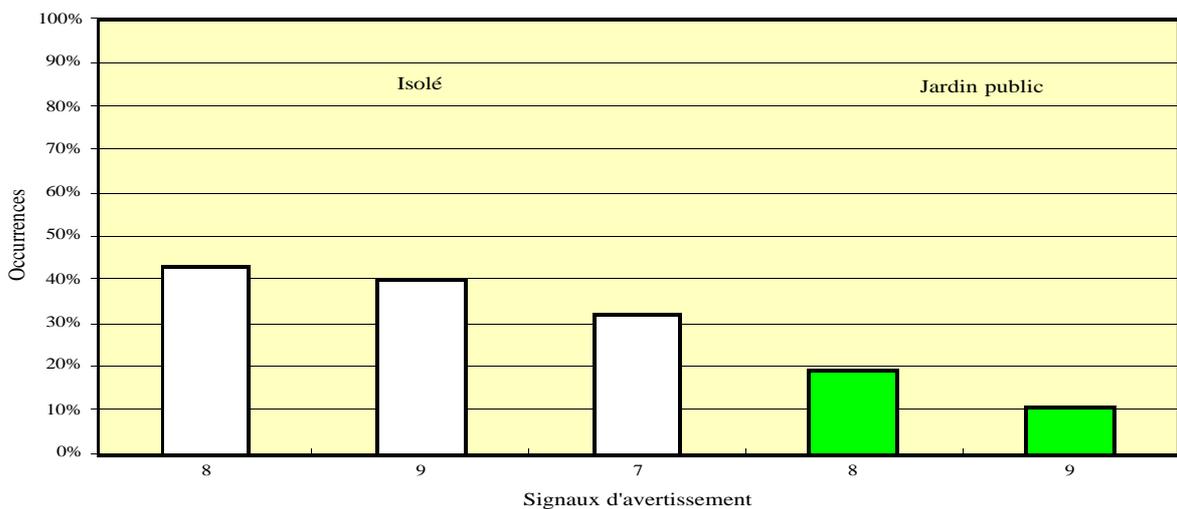


Figure 4-26 : prototypes associés à "voiture".

Les signaux 8 et 9, les plus fortement associés à "voiture", sont également ceux associés à "klaxon" (cf. § 6.1.5). En situation isolée d'ailleurs, ce fait est dû à la désignation, fréquente, "klaxon de voiture". Le signal 7, associé également à "klaxon" en écoute isolée, provient notamment de la désignation "alarme de voiture". En contexte de circulation, les occurrences "voiture" sont trop peu nombreuses pour être quelque peu significatives.

### 6.1.7 Prototypes associés à "vélo" et "bicyclette"

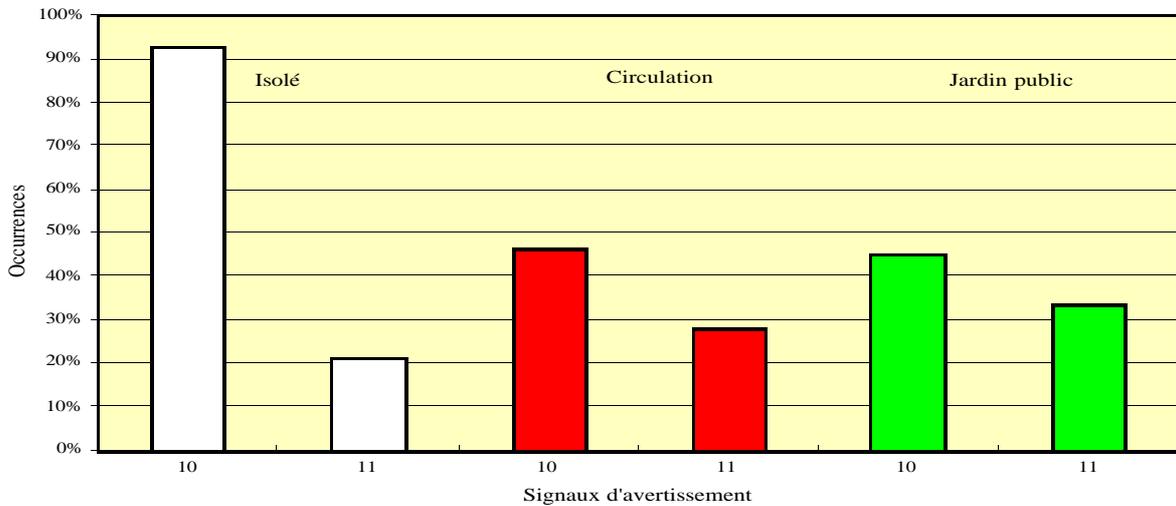


Figure 4-27 : prototypes associés à "vélo" et "bicyclette".

Deux signaux, 10 et 11, sont fortement associés à "vélo". Toutefois, si l'avertisseur 10 perd logiquement de son caractère prototypique en contexte urbain, car perçu à son niveau seuil, le signal 11, au contraire, gagne en prototypicité. En effet, la mise en contexte urbain, et notamment en contexte de jardin public, a pour effet de ramener un signal décrit plutôt comme "clochette" ou "carillon" en écoute isolée vers une description de type "sonnette de vélo".

### 6.1.8 Prototypes associés à "sifflet"

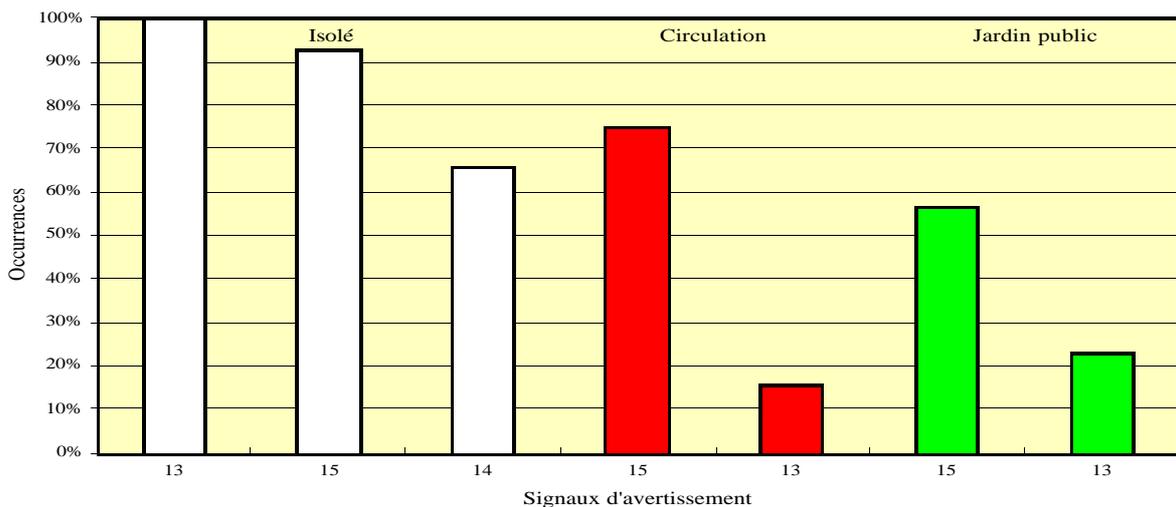


Figure 4-28 : prototypes associés à "sifflet".

Si le signal 13 est unanimement décrit par "sifflet" en situation isolée, il l'est beaucoup moins en contexte urbain. De même, le signal 14, qui est fortement associé à "sifflet" hors contexte, ne l'est pas du tout en situation urbaine. Seul le signal 15 reste bien représenté quelle que soit la situation d'écoute et résiste donc à une mise en contexte.

## **6.2 Interprétation**

Pour les signaux d'urgence, il est intéressant de remarquer que, si les sujets s'estimaient incapables de distinguer les avertisseurs de la police, des ambulances et des pompiers, les résultats de l'expérience montrent bien qu'ils savent faire la différence entre ces signaux, tout du moins par comparaison. Ainsi, nous pouvons considérer que le signal 5 est prototypique de "pompiers", le signal 3 "d'ambulance" et les signaux 1 et 6 de "police". Par ailleurs, la désignation plus générique "sirène" met en évidence les signaux 1, 6 et 7, notamment en contexte. En effet, leur niveau seuil en situation les rend plus difficiles à être distingués entre eux. Notons cependant que les avertisseurs des ambulances se diversifient de plus en plus et sont parfois identiques à ceux de la police. Cela conduit à une certaine confusion quant à l'identification des signaux. Mais est-il vraiment utile de devoir distinguer les signaux d'urgence des véhicules prioritaires ? En effet, ce signal sonore a pour unique fonction de libérer le passage à ces véhicules afin d'atteindre rapidement les lieux d'un sinistre. Par ailleurs, si la différenciation est nécessaire pour les trois corps de métiers, elle est en contradiction avec la diversification actuelle des signaux.

En ce qui concerne les prototypes associés à "klaxons", nous retiendrons comme prototype le signal 8. En effet, le signal 9 est toujours légèrement moins prototypique que 8, alors que 8 et 9 sont issus du même avertisseur de voiture, le premier étant long et continu, le second bref et répété trois fois. Par ailleurs, si le signal 4 est très représenté en contexte urbain, c'est parce que son audition au niveau seuil de perception masque la composante aiguë du signal à deux tons de l'avertisseur de l'ambulance. Les sujets ne perçoivent en fait que "deux coups de klaxon" de même hauteur. Le signal 8 est d'ailleurs également prototypique du terme "voiture".

Le signal 10 est clairement prototypique de "vélo" et "bicyclette" car il n'est pas porteur de sens ambigu comme le signal 11 qui est également associé à une "clochette" ou à une "porte d'entrée".

Enfin, le prototype associé à "sifflet" est le signal 15 car il est le seul signal qui résiste à la mise en contexte urbain au niveau seuil de perception des signaux. En effet, les sifflets 13 et 14 deviennent ambigus en contexte et sont associés à un "klaxon", un signal "long" ou décrits par des onomatopées car non identifiés.

## 7. En résumé

Cette expérience a permis, grâce à diverses méthodes d'analyse, de mettre en évidence un certains nombres de signaux d'avertissement quant à leur efficacité et à leur signification. En confrontant ces résultats, il apparaît que :

- le signal 15, prototype de "sifflet", et le signal 3, prototype "d'ambulance", sont perçus à des seuils inférieurs dans le contexte de circulation que dans celui de jardin public, et disposent de fréquences de désignations très élevées (15 en circulation et jardin public, 3 hors contexte) et d'une faible polysémie (notamment pour 15). Il s'agit donc des deux signaux les plus efficaces en contexte de circulation, tant du point de vue de leur signification que de celui de leur émergence perceptive.

- A l'opposé, le signal 10, prototype de "sonnette de vélo", et le signal 5, prototype de "pompiers", sont perçus à des seuils inférieurs dans le contexte de jardin public que dans celui de circulation et disposent également d'une faible polysémie et de fréquences de désignations très élevées (10 et 5 en isolé). Il s'agit donc des deux signaux les plus efficaces en contexte de jardin public, tant du point de vue de leur signification que de celui de leur émergence perceptive.

- Entre ces deux "extrémités" contextuelles, le signal 8, prototype de "klaxon", dispose d'une faible polysémie et de fréquences de désignations élevées pour les deux contextes urbains. Il est efficace dans tous les cas.

- Enfin, les signaux 12, 13 et 14 ne sont pas prototypiques, présentent essentiellement des propriétés physiques liées à leurs sonorités et ont une forte polysémie ou un grand nombre d'absences de réponse.

## **8. Conclusion**

Les données recueillies à l'issue de cette première expérience constituent une mine d'informations relatives à la signification des signaux d'avertissement. Dans un premier temps, nous avons accès au répertoire lexical disponible sur les signaux d'avertissement *a priori* et après écoute. Dans un second temps, les désignations les plus fréquentes servent à rechercher des signaux prototypiques qui leur sont associés. Ainsi, nous avons pu déterminer des signaux efficaces, d'un point de vue perceptif, et signifiants dans deux contextes urbains différents et de manière isolée.

Toutefois, en contexte, seuls des écoutes à des niveaux seuils d'émergence ont été prises en compte. Il s'avère donc intéressant de poursuivre cette expérience avec des niveaux d'émergence perceptive des signaux croissants par rapport aux fonds sonores, afin de disposer de données allant de la limite perceptible à l'écoute décontextualisée. Ces résultats sont exposés au chapitre 5 suivant, avant la recherche de paramètres acoustiques reliés à la signification et la pertinence des signaux (chapitre 6).

# Chapitre 5 : Identification en contexte

## 1. Introduction

La première expérience, décrite au chapitre précédent, nous a permis de caractériser au niveau sémantique des signaux d'avertissement en écoute isolée à niveau "confortable", et en écoute contextuelle à leurs niveaux seuils d'identification. D'importantes différences de reconnaissance et d'identification des signaux entre les écoutes isolée et en situation ont été mises en évidence. Il s'avère donc nécessaire de réaliser une seconde épreuve verbale qui permette de faire le lien entre ces deux situations "extrêmes" d'écoute. En effet, la recherche des niveaux d'émergence des signaux, à partir desquels les auditeurs changent de stratégie d'analyse, est essentielle à la caractérisation des signaux d'avertissement.

Ce chapitre décrit cette seconde expérience pour laquelle l'analyse des données verbales est identique à celle du chapitre précédent. Le lien entre les deux expériences, qui permet de comparer les résultats pour quatre niveaux d'émergence des signaux sur deux fonds sonores urbains, est traité à la fin du chapitre.

## 2. Méthode

Afin de pouvoir comparer les résultats de la seconde expérience avec ceux de la première, les mêmes objets sonores sont utilisés. Les séquences sonores sont montées et mixées selon deux approches : trois signaux très "différents" ou deux à quatre signaux très "semblables" par séquence. Cette démarche permet de regrouper les signaux par leurs relations de *dissemblance intercatégorielle* et de *ressemblance intracatégorielle*, concepts issus de la catégorisation prototypique (Rosch, 1978 ; cf. chapitre 1, § 4.3). Le caractère de similarité des signaux est déterminé à partir des résultats de la première épreuve verbale. En effet, les signaux fortement associés à une désignation commune sont considérés comme semblables. Ainsi, les prototypes obtenus au chapitre précédent nous donnent les six associations de "ressemblances" suivantes (cf. chapitre 4, § 6.1) : 1-2-6-7 (pour "sirène"), 1-2-5 (pour "pompiers"), 2-3-4 (pour "ambulance"), 4-8-9 (pour "klaxon"), 10-11 (pour "vélo") et 13-14-15 (pour "sifflet"). Les groupements de trois signaux "différents" sont obtenus en excluant la possibilité d'avoir deux signaux "semblables" dans une même séquence. Il en découle, par exemple, les cinq groupements de signaux suivants : 1-8-13, 2-9-10, 3-6-15, 4-7-14 et 5-11-12.

### 2.1 Corpus sonore

L'expérience reprend les quinze signaux d'avertissement et les deux mêmes fonds sonores urbains de circulation et de jardin public du premier test. Les durées et les niveaux de restitution des fonds sonores demeurent également inchangés (cf. chapitre 4, § 2.1) ; de même pour les conditions d'écoute (chapitre 3, § 5.1.1). En revanche, les niveaux des signaux d'avertissement ont été modifiés : à partir de leurs niveaux seuils de perception, déterminés dans la première expérience (cf. chapitre 4, § 4.2), ils ont été augmentés, soit de 9 dB, soit de 18 dB, mais toujours de la même quantité dans une même séquence sonore. Ces valeurs ont été choisies de manière subjective, afin d'obtenir des niveaux d'émergence perceptive qui constituent des "échelons intermédiaires" entre l'écoute aux niveaux seuils et l'écoute isolée. Ainsi, nous avons six types de séquences sonores, trois pour chaque contexte d'écoute :

- signaux différents à +9 dB d'émergence perceptive sur le fond (Figure 5-1),
- signaux différents à +18 dB d'émergence perceptive sur le fond,
- signaux semblables à +18 dB d'émergence perceptive sur le fond.

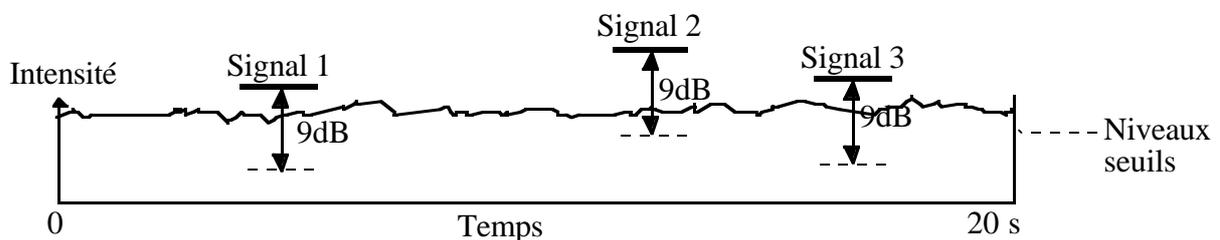


Figure 5-1 : représentation dynamique et schématique d'une séquence mixée de trois signaux différents à +9 dB d'émergence perceptive.

## 2.2 Procédure et sujets

L'expérience se déroule en deux parties. Une première étape consiste à présenter aux sujets les séquences contenant les triades de signaux "différents". Dix séquences de circulation, avec une émergence perceptive des signaux d'avertissement de 9 dB pour les cinq premières séquences et de 18 dB pour les cinq suivantes, puis les dix séquences de jardin public, disposées de la même manière. Ainsi, chaque triade apparaît à quatre reprises dans cette première partie. Nous avons systématiquement modifié l'ordre d'apparition des signaux dans chaque séquence afin de pouvoir négliger l'effet d'ordre. Par ailleurs, l'ordre même des séquences est modifié entre les deux niveaux d'émergence, mais pas entre les contextes. La consigne donnée aux sujets est la suivante : "Dans un premier temps (index 1 à 20), nous vous demandons simplement d'identifier les signaux d'avertissement que vous percevrez dans chaque séquence sonore. Puis, dans une seconde expérience (index 21 à 32), nous vous demandons de comparer les signaux de chaque séquence afin de les caractériser plus précisément". Dans cette seconde étape sont présentées les six séquences contenant les signaux "semblables", qui émergent perceptivement de 18 dB, pour la circulation puis pour le jardin public. L'ordre d'apparition des signaux dans chaque séquence est également modifié d'un contexte à l'autre.

A la différence de la première expérience, les sujets sont autonomes : ils choisissent les index à l'aide de la télécommande du lecteur de D.A.T. et rédigent leurs réponses sur feuille volante ou sur ordinateur dans un tableau comportant quatre cases par séquence pour la première partie (trois signaux étant testés par séquence) et une grande case pour le texte de la seconde partie. Une feuille de test et un exemple de fichier de réponse informatique, fichier que tous les sujets ont en fait adopté, sont donnés en annexe 1.

Vingt-quatre sujets âgés de 18 à 59 ans ont participé à l'expérience, 12 personnes de chaque sexe. Dix d'entre eux sont acousticiens, les quatorze autres sujets sont issus de divers milieux socioculturels, tous étant des usagers de Paris (cf. chapitre 2, § 3.2).

Les 20 séquences de signaux "différents" correspondent aux pages 65 à 84 du C.D. accompagnant ce mémoire, dans l'ordre donné ci-dessus ; les 12 séquences de signaux "semblables" correspondent aux pages 85 à 96.
--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### 3. Résultats de l'identification verbale

L'analyse des productions verbales issues de la seconde expérience est identique à celle réalisée pour le premier test (cf. chapitre 4, § 5). Nous allons ainsi étudier le répertoire lexical disponible sur l'ensemble des données (§ 3.1), puis décrire les quinze signaux d'avertissement en termes d'identification (cf. annexe 3) dans les différentes situations d'écoute (§ 3.2).

#### 3.1 Répertoire lexical disponible sur l'ensemble des données

L'ensemble de toutes les données réunies constitue un corpus de 1440 données (24 sujets, 4 situations d'écoute et 15 signaux d'avertissement) qui sont réparties sur 202 types différents d'expressions linguistiques, dont 88 désignations produites au moins deux fois (pour la technique d'analyse verbale, voir chapitre 4, § 3).

- Parmi ces 202 désignations, 114 (soit 56%) constituent des hapax, expressions qui n'ont été produites qu'une seule fois.

- Parmi les 88 désignations produites au moins deux fois, les plus fréquentes sont "klaxon" (225 occurrences), "ambulance" (169), "sifflet" (130), "police" (94), "pompiers" (80), "sirène" (79) et "?" (75). La Figure 5-2 présente les dix-huit désignations les plus fréquentes ( $f \geq 10$ ), ainsi que leur distribution sur les quatre situations d'écoute.

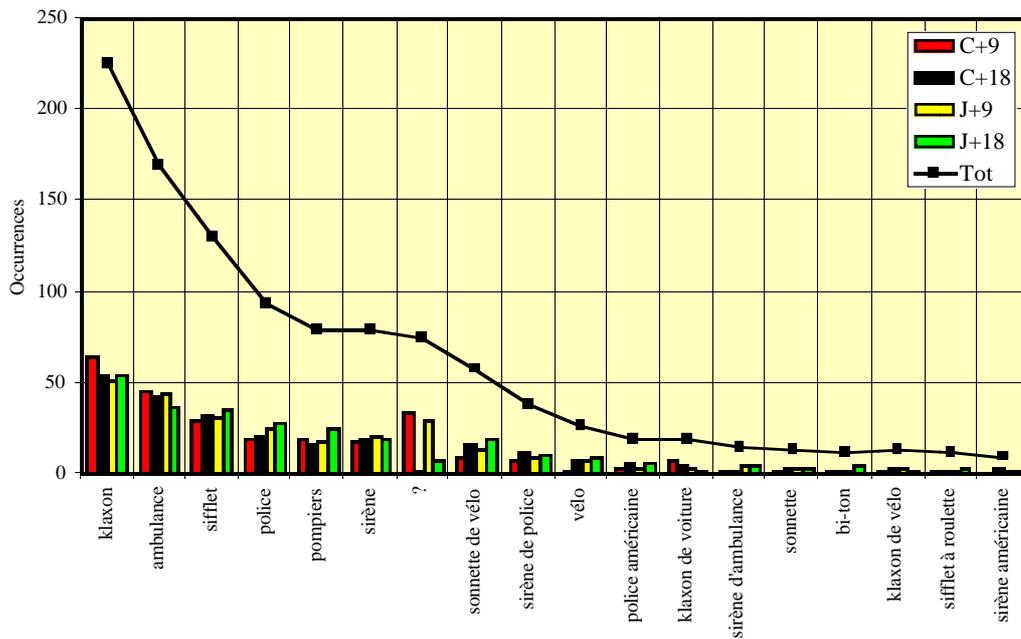


Figure 5-2 : désignations les plus fréquentes.

Les désignations les plus fréquentes ( $f \geq 2$ ) permettent d'identifier celles qui correspondent à :

- des **propriétés physiques du son** : 11% d'onomatopées ("pin pon, ding, pouêt, tuuut...") et de termes génériques ("tintement"),
- des **noms de sources interprétées** : 89% ("klaxon, ambulance, sifflet, police..."), dont les **noms de noms** ("sonnette de vélo, sirène de police, klaxon de voiture...") représentent 31% des désignations les plus fréquentes.

Ainsi, le répertoire disponible sur l'ensemble des données est concentré sur un type particulier de la classification linguistique, à savoir les noms de sources interprétées (cf. chapitre 4, § 5.1).

Une étude des désignations les plus fréquentes par situation d'écoute montre qu'il y a très peu de différences entre les quatre types de stimuli. Aucun effet d'émergence ni de contexte n'est donc observable à ce niveau d'analyse.

### 3.2 Description des 15 signaux dans les différentes situations d'écoute

#### 3.2.1 Les absences de réponse ("?")

A la différence de la première expérience, les sujets rédigent eux-mêmes leurs réponses dans un fichier informatique. Par ailleurs, tous les stimuli se situent au-dessus du seuil d'identification déterminé dans l'expérience préalable. Les absences de réponses ne peuvent donc plus être liés au niveau d'émergence physique des signaux, mais correspondent plutôt à un manque de pertinence de ces signaux dans un contexte donné : les sujets n'y portent pas attention et les intègrent dans le contexte sonore.

La Figure 5-3 présente les occurrences "?" en fonction des signaux, de leur émergence (+9 dB et +18 dB) et des contextes d'écoute. Nous utiliserons la notation "C" pour Circulation et "J" pour Jardin public.

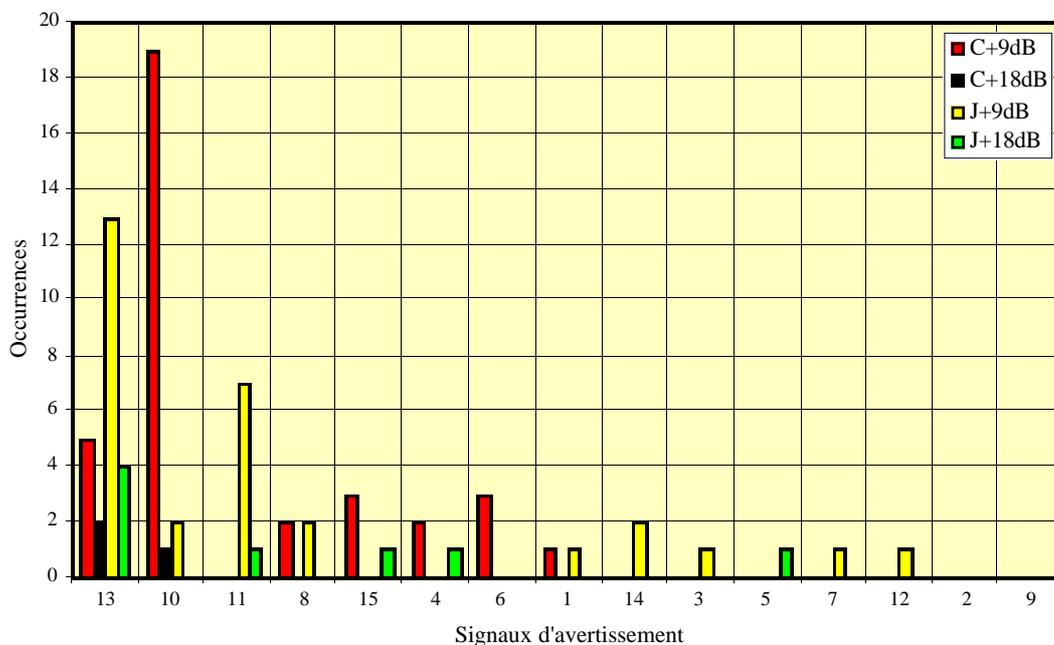


Figure 5-3 : absences de réponse.

Rappelons que les noms "officiels" correspondant aux numéros des signaux d'avertissement se trouvent au chapitre 4, § 2.4.

### 3.2.1.1 En fonction des signaux...

Deux signaux, le sifflet 13 et la sonnette 10, sont particulièrement affectés par des absences de réponse : ils représentent à eux seuls 60% des 75 occurrences "?". Par ailleurs, les signaux 2 et 9 ne sont pas du tout concernés par les absences de réponse et les signaux 1, 3, 5, 7, 12 et 14 le sont très faiblement (occurrences  $\leq 2$ ).

### 3.2.1.2 ...et en fonction du contexte et de l'émergence

Globalement, les 75 occurrences d'absences de réponse sont réparties uniformément entre le contexte de circulation et celui de jardin public. En revanche, si la sonnette 10 concerne essentiellement la circulation, le sifflet 13 et la sonnette 11 concernent plutôt le jardin public.

Quant à l'effet de l'émergence perceptive des signaux sur les fonds sonores, nous obtenons 85% des occurrences "?" à +9 dB et les 15% restants à +18 dB. Notons également que, à +9 dB d'émergence, trois des quinze signaux ne sont pas concernés par les absences de réponse, alors qu'à +18 dB, il s'agit de neuf signaux.

En résumé, le sifflet 13 présente la majorité des occurrences dans le contexte de jardin public (J+9 dB), tandis que la sonnette 10 les compte essentiellement dans le contexte de circulation (C+9 dB).

### 3.2.1.3 Interprétation

Pour l'ensemble des signaux, il apparaît qu'une émergence perceptive du signal même faible (9 dB) suffit à rendre des signaux tout aussi reconnaissables en ambiance de circulation que de jardin public. En effet, les sujets ont produit à plus de 95% des réponses verbales d'identification et ce, pour douze des quinze signaux. Toutefois, les avertisseurs 10 et 13 sont moins bien identifiés et mettent en évidence que certaines structures sonores sont plus difficilement perceptibles dans un contexte donné plutôt que dans un autre. Ainsi, le sifflet 13 et la sonnette 11 s'intègrent-ils aisément dans un contexte de jardin public, tandis que la sonnette 10 passe inaperçue dans une ambiance de trafic automobile.

Par ailleurs, l'augmentation de l'émergence des signaux a pour effet de réduire le nombre d'absences de réponse (de 9% à 1,5% sur l'ensemble des signaux). Une émergence élevée facilite donc l'identification d'un signal sonore. Mais elle n'implique pas nécessairement une signification univoque du signal perçu.

## 3.2.2 Productivité et variabilité des désignations des signaux d'avertissement

### 3.2.2.1 Productivité par rapport aux 15 signaux

La productivité verbale des sujets, liée au signal, est donnée par le nombre de désignations pour chaque signal d'avertissement. Les résultats indiquent (Figure 5-4) que, indépendamment des contextes d'écoute, deux signaux sont particulièrement riches en désignations : la sirène 7 et la sonnette 11. La première est majoritairement décrite par une "sirène" ou une "alarme", mais ne dispose pas de terme attesté dans la langue car elle n'est pas rencontrée au quotidien et n'est donc pas clairement représentée dans la mémoire des sujets. La seconde est perçue comme "clochette, carillon, sonnerie d'ascenseur, sonnette de vélo..." et a donc plutôt un caractère ambigu.

Sur la partie droite de la Figure 5-4, les signaux 15, 9 et 3 sont pauvres en désignations. Il s'agit donc de termes univoques, associés à un "sifflet de policier, de gendarme ou de gardien" pour le premier, à un "klaxon de voiture" pour le second et à une "ambulance" pour le troisième.

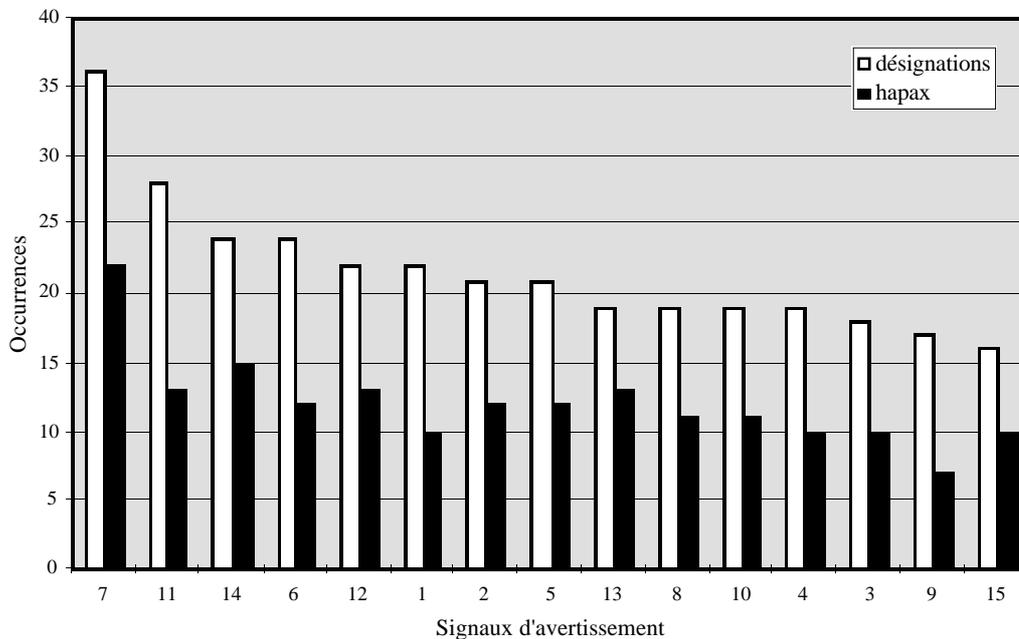


Figure 5-4 : désignations et hapax - global.

### 3.2.2.2 Effets d'émergence et de contexte

Si nous distinguons maintenant les quatre situations d'écoute, nous pouvons analyser les signaux d'avertissement en fonction des données contextuelles et de celles d'émergence des signaux. La Figure 5-5 et la Figure 5-6 présentent le nombre de désignations et d'hapax pour la circulation, la Figure 5-7 et la Figure 5-8 pour le jardin public.

Sur l'ensemble des signaux, l'émergence n'a pas d'effet significatif sur le nombre de désignations. En effet, la circulation passe de 150 à 158 désignations lorsque les niveaux des signaux augmentent, le jardin public, de 144 à 143. Il n'apparaît pas non plus de différences significatives du nombre de désignations entre les contextes d'écoute. Tout au plus pouvons nous remarquer une légère baisse du nombre de désignations en passant de la circulation au jardin public pour les signaux 6, 11 et 13.

En contexte de circulation, les signaux 6, 7 et 11 disposent d'un grand nombre de désignations. Par ailleurs, beaucoup de désignations sont produites pour le signal 5 à +9 dB et pour le signal 1 à +18 dB. A l'autre extrême des figures - peu de désignations - nous trouvons les signaux 10, 12, 8 et 9 à +9 dB et les signaux 9, 15, 13, 4 et 3 à +18 dB. Toutefois, dans le contexte de circulation, nous pouvons noter une baisse du nombre de désignations pour les signaux 5 et 7, ce qui va dans le sens d'une perte d'ambiguïté. Pour les signaux 1, 8, 10 et 12 au contraire, le nombre de désignations augmente avec l'émergence perceptive. Ce phénomène est dû à des productions langagières plus explicites - et donc plus variées aussi - qui correspondent le plus souvent à des "noms de noms" lorsque l'émergence augmente.

Pour le jardin public, nous retrouvons encore le signal 7 à l'origine d'une forte productivité verbale et dans une moindre mesure, le signal 11. A l'opposé, nous trouvons les

signaux 13 et 15 à +9 dB et les avertisseurs 10 et 13 à +18 dB. Ainsi, le signal 10 s'appauvrit en désignations lors du passage de +9 dB à +18 dB.

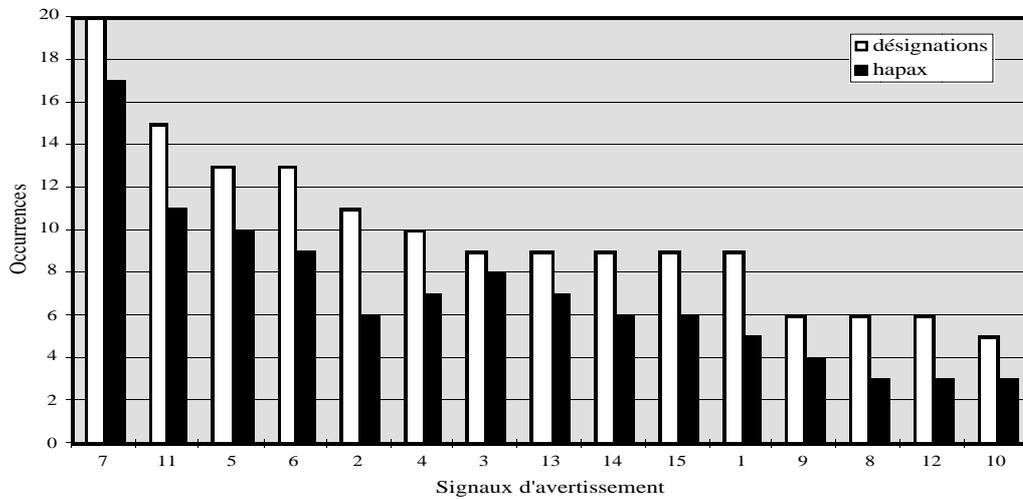


Figure 5-5 : désignations et hapax - circulation +9 dB.

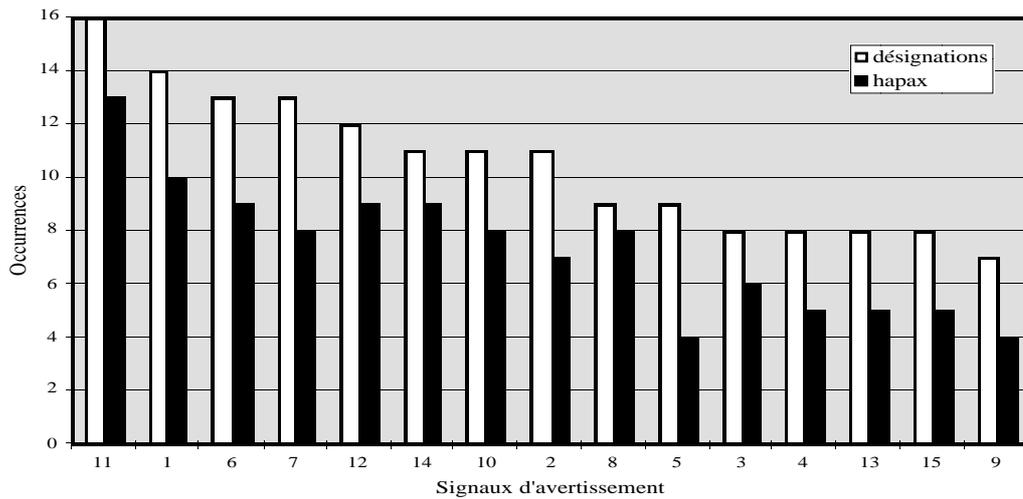


Figure 5-6 : désignations et hapax - circulation +18 dB.

### 3.2.2.3 Variabilité : nombre d'hapax

L'analyse des hapax est fortement corrélée à celle des désignations. Néanmoins, sur l'ensemble des données (Figure 5-4), il apparaît que les sujets produisent peu de variabilité pour les signaux 11, 1 et 9 par rapport au nombre de désignations produites. Ces valeurs concernent la variabilité pour chaque signal, toutes écoutes confondues et ne sauraient correspondre à la somme des quatre contributions d'hapax. D'ailleurs ces observations faites sur ces trois signaux ne se vérifient pas sur les différentes situations d'écoute considérées séparément.

La forte variabilité du signal d'urgence 7 dans toutes les situations confirme l'interprétation faite quant à sa productivité, à savoir, le fait que cette sirène rapide ne dispose pas de terme attesté dans la langue car peu rencontrée dans la vie quotidienne.

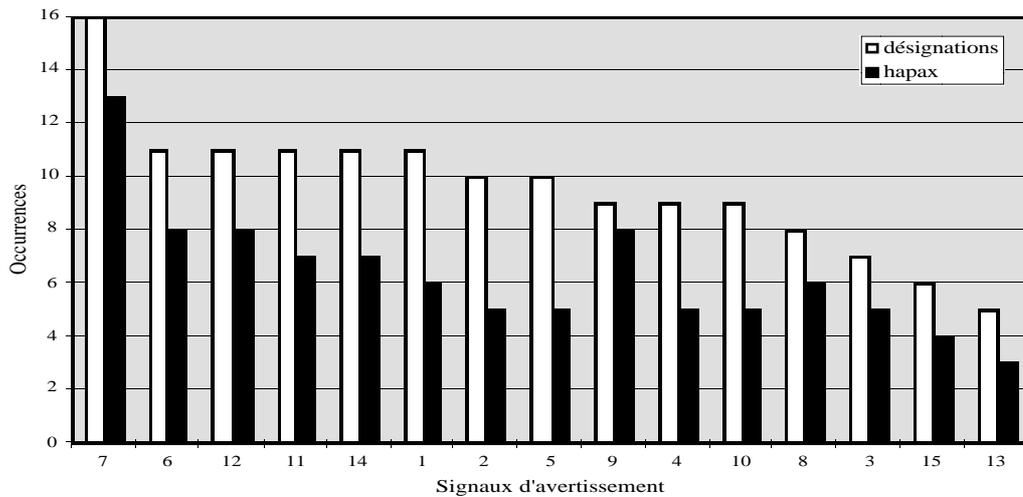


Figure 5-7 : désignations et hapax - jardin public +9 dB.

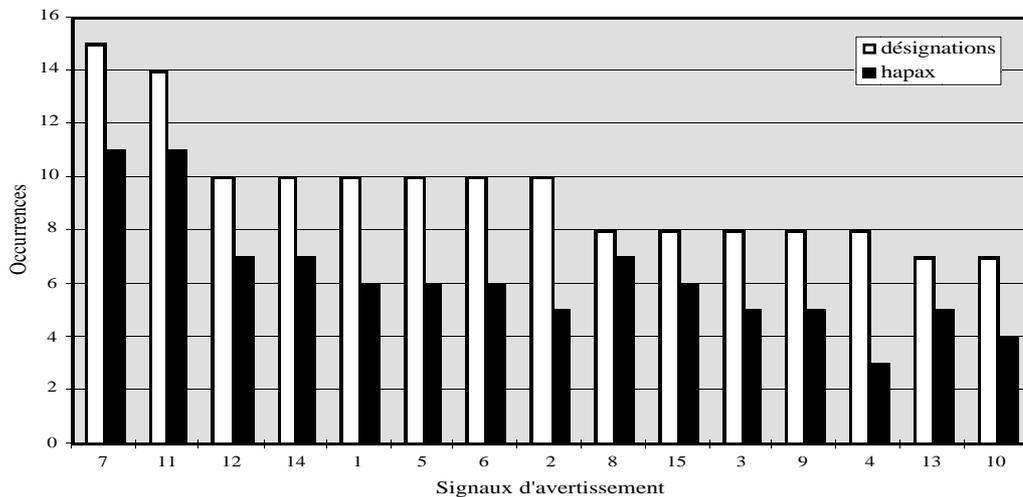


Figure 5-8 : désignations et hapax - jardin public +18 dB.

### 3.2.3 Fréquence la plus élevée

La fréquence la plus élevée correspond à la désignation qui a été produite le plus grand nombre de fois par les sujets. Elle correspond donc au degré d'accord entre les sujets et donc, à un degré de prototypicité des signaux. La Figure 5-9 présente ce critère dans les quatre situations d'écoute ainsi que sur l'ensemble des données. Notons que la fréquence sur l'ensemble des données n'est pas la somme des fréquences dans chaque situation, puisque la fréquence la plus élevée ne concerne pas systématiquement la même désignation.

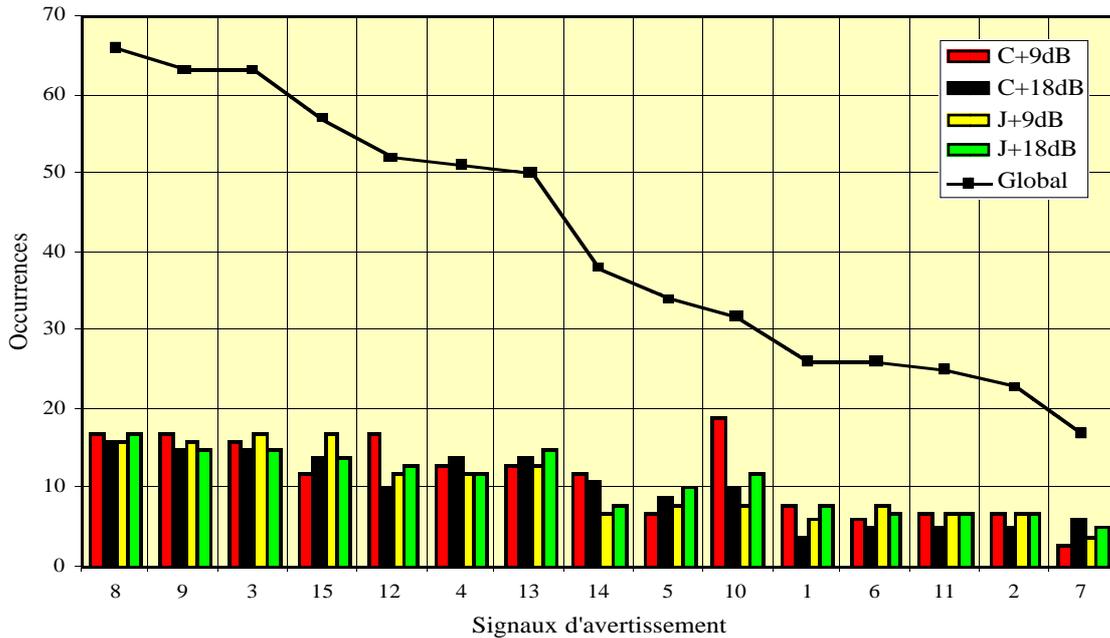


Figure 5-9 : fréquence la plus élevée.

### 3.2.3.1 Sur l'ensemble des données...

Sur l'ensemble des données, les signaux 8, 9 et 3 contribuent au meilleur accord entre les sujets, puisque un seul terme est fortement associé à chacun de ces signaux. A l'inverse, les signaux 7, 2, 11, 6 et 1 donnent les taux d'occurrences les plus faibles. Il serait toutefois un peu hâtif de les considérer comme n'étant pas perçus comme univoques.

### 3.2.3.2 ...et en fonction du contexte et de l'émergence

Le nombre d'occurrences correspondant aux fréquences les plus élevées varie très peu pour un signal donné. Seule la sonnette 10 et l'avertisseur 12 - dans une moindre mesure - présentent un effet d'émergence. Ils correspondent aux signaux contribuant au meilleur accord entre sujets dans le contexte de circulation à +9 dB, alors que cet accord est nettement moins marqué à +18 dB et dans le contexte de jardin public.

## 3.2.4 Interprétation des désignations

L'analyse des résultats en termes d'interprétation sémantique ou de description acoustique n'ayant pas donné de résultats probants lors de l'étude de la première expérience, nous ne chercherons que des différences particulièrement significatives de ce type d'analyse.

### 3.2.4.1 Effets d'émergence et de contexte

Quelle que soit la situation d'écoute, le nombre de sources interprétées est en moyenne vingt fois supérieur à celui de phénomènes acoustiques. Si le contexte ne semble pas avoir d'effet sur ces données, une augmentation du niveau d'émergence semble augmenter faiblement

la proportion de sources interprétées. Toutefois, les écarts-type calculés sont trop importants pour en déduire une variation significative.

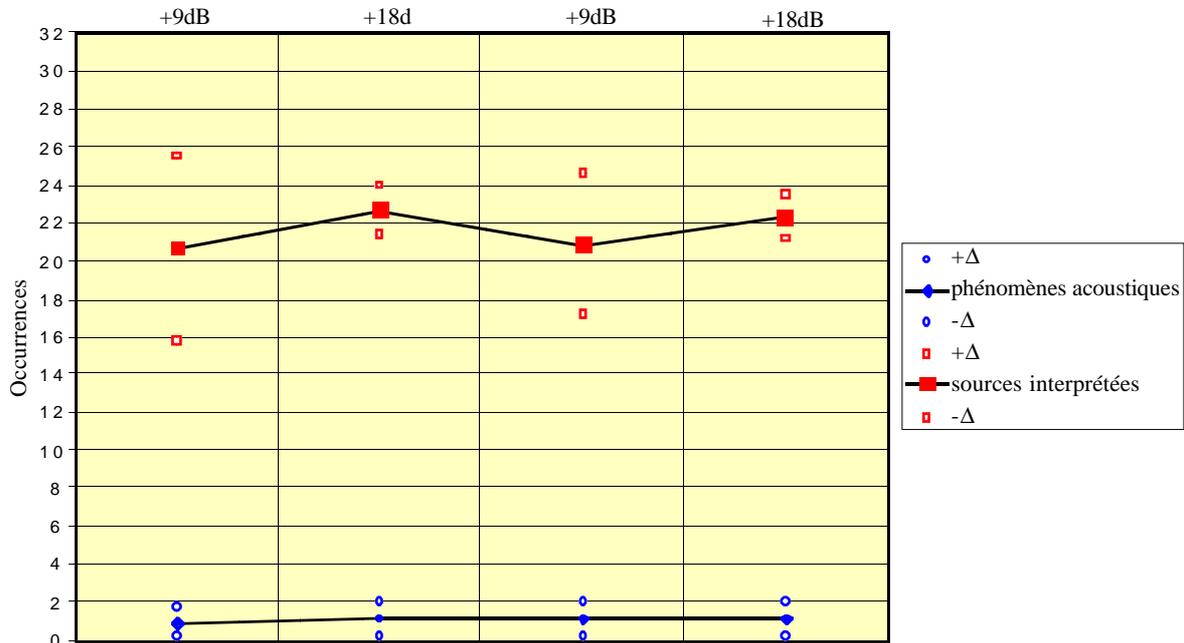


Figure 5-10 : effets de l'émergence et du contexte.

### 3.2.4.2 Analyse par classes de signaux

Une analyse détaillée des contributions acoustiques et sémantiques n'apporte pas de résultats supplémentaires : quels que soient le contexte et le niveau d'émergence, le nombre de propriétés acoustiques varie de 0 à 3 et celui des sources interprétées de 20 à 24 (Figure 5-11). Deux exceptions cependant : le signal 10 en circulation et à +9 dB d'émergence, auquel ne correspondent que quatre sources interprétées, et le signal 13 en jardin public et à +9 dB d'émergence, pour lequel il n'y a que dix sources interprétées. Dans les deux cas, les données manquantes se retrouvent en fait dans les absences de réponse "?" (Figure 5-3).

Signal n°	phénomènes acoustiques					sources interprétées					"?"				
	C+9dB	C+18dB	J+9dB	J+18dB	Total	C+9dB	C+18dB	J+9dB	J+18dB	Total	C+9dB	C+18dB	J+9dB	J+18dB	Total
1	1	2	1	2	<b>6</b>	22	22	22	22	<b>88</b>	1	0	1	0	<b>2</b>
2	2	2	2	2	<b>8</b>	22	22	22	22	<b>88</b>	0	0	0	0	<b>0</b>
3	0	0	1	2	<b>3</b>	24	24	22	22	<b>92</b>	0	0	1	0	<b>1</b>
4	0	0	1	1	<b>2</b>	22	24	23	22	<b>91</b>	2	0	0	1	<b>3</b>
5	1	1	2	1	<b>5</b>	23	23	22	22	<b>90</b>	0	0	0	1	<b>1</b>
6	1	0	1	0	<b>2</b>	20	24	23	24	<b>91</b>	3	0	0	0	<b>3</b>
7	2	0	0	0	<b>2</b>	22	24	23	24	<b>93</b>	0	0	1	0	<b>1</b>
8	0	1	0	1	<b>2</b>	22	23	22	23	<b>90</b>	2	0	2	0	<b>4</b>
9	0	1	1	1	<b>3</b>	24	23	23	23	<b>93</b>	0	0	0	0	<b>0</b>
10	1	1	0	2	<b>4</b>	4	22	22	22	<b>70</b>	19	1	2	0	<b>22</b>
11	2	2	2	3	<b>9</b>	22	22	15	20	<b>79</b>	0	0	7	1	<b>8</b>
12	1	2	3	1	<b>7</b>	23	22	20	23	<b>88</b>	0	0	1	0	<b>1</b>
13	1	3	1	0	<b>5</b>	18	19	10	20	<b>67</b>	5	2	13	4	<b>24</b>
14	2	2	2	1	<b>7</b>	22	22	20	23	<b>87</b>	0	0	2	0	<b>2</b>
15	1	0	0	0	<b>1</b>	20	24	24	23	<b>91</b>	3	0	0	1	<b>4</b>

Figure 5-11 : acoustique et sémantique - données détaillées.

## 4. Prototypes issus de l'identification verbale

Dans ce paragraphe, nous reprendrons, une à une, les désignations les plus fréquemment citées par les sujets afin de les associer, en termes de ressemblance prototypique (cf. chapitre 1, § 4.3), aux signaux sonores testés.

### 4.1 Signaux prototypes des désignations les plus fréquentes

#### 4.1.1 Prototypes associés à "ambulance"

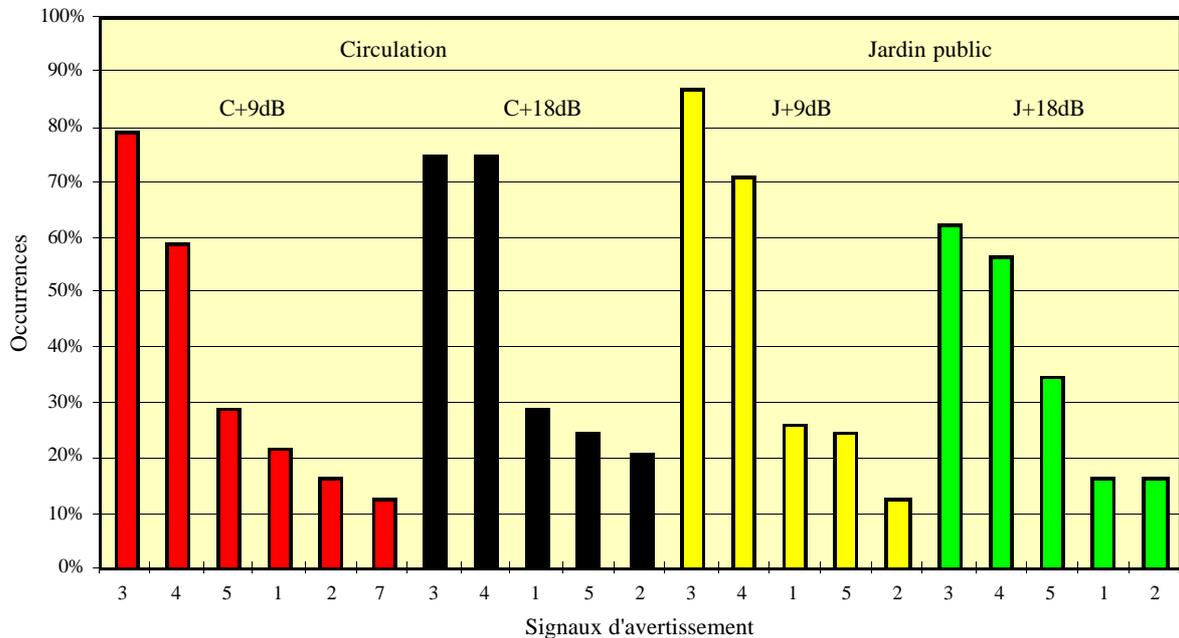


Figure 5-12 : prototypes associés à "ambulance" et "S.A.M.U.".

Dans toutes les situations d'écoute, le signal 3 est le plus prototypique, parmi les quinze signaux d'avertissement, du terme "ambulance". Remarquons que, plus les signaux émergent du bruit de fond, plus le signal 4 devient proche du signal 3 en termes de prototypicité.

### 4.1.2 Prototypes associés à "police"

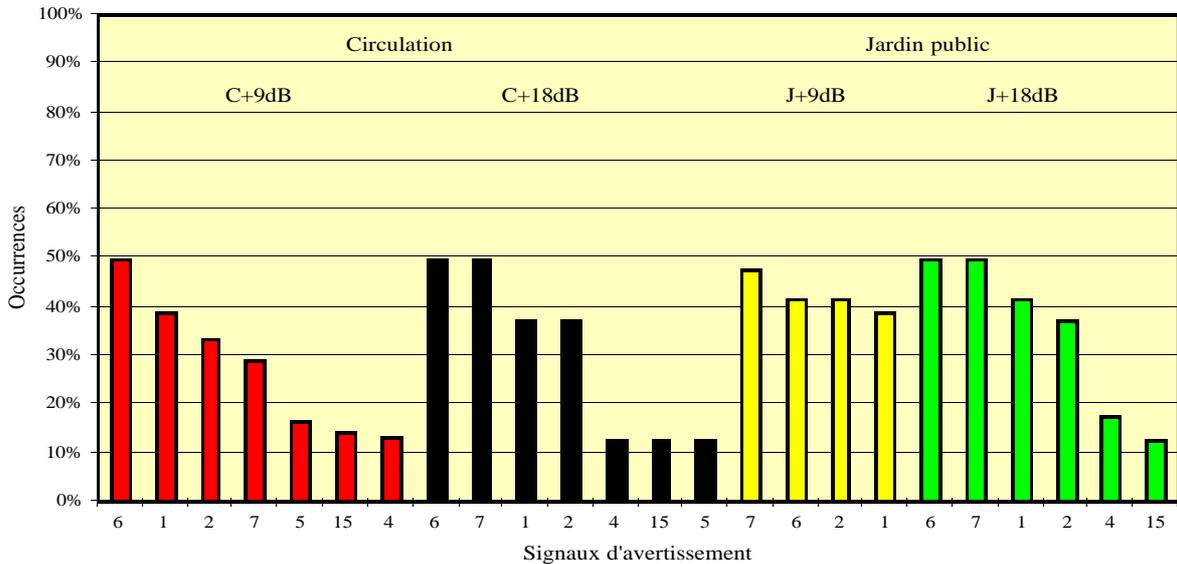


Figure 5-13 : prototypes associés à "police" et "flics".

Les signaux 6, 7, 1 et 2 sont fortement associés à "police" ou "flics" dans les quatre situations d'écoute. Le signal 6 est néanmoins le seul à dépasser 40% d'occurrences dans tous les cas. En effet, le signal 7 est moins prototypique en circulation à +9 dB, mais l'est beaucoup plus en situation de jardin public, tandis que les signaux 1 et 2 sont notamment moins prototypiques à +18 dB d'émergence dans les deux contextes d'écoute.

### 4.1.3 Prototypes associés à "pompiers"

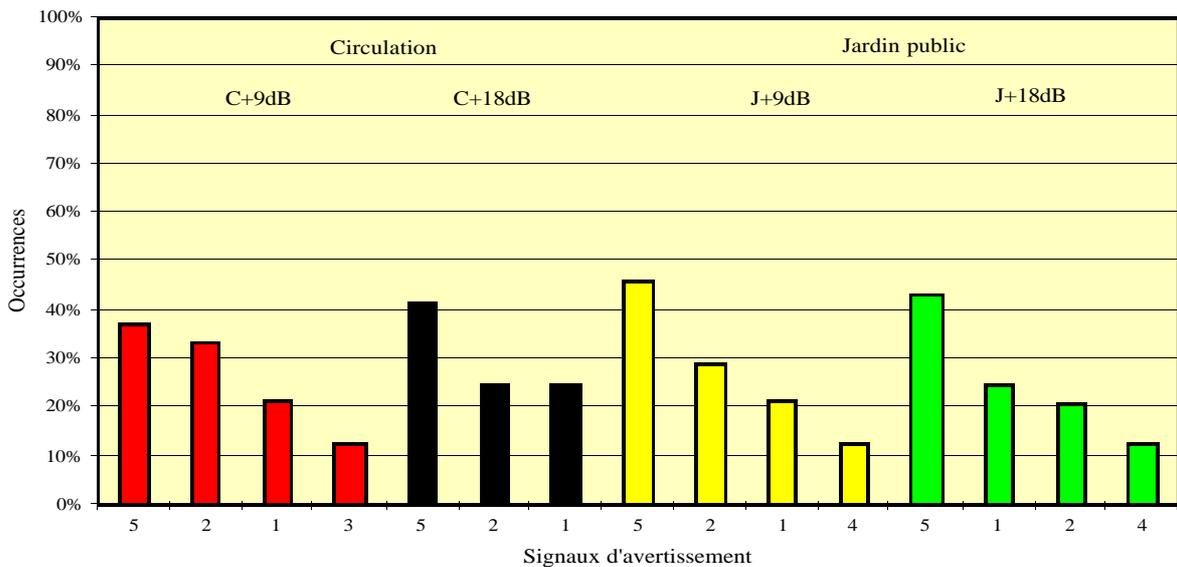


Figure 5-14 : prototypes associés à "pompiers".

Quant à la désignation "pompiers", le signal 5 lui est clairement le plus proche et ce, quels que soient le niveau d'émergence et le contexte. L'augmentation du niveau d'émergence des signaux a pour effet de réduire la prototypicité du signal 2, seul véritable "concurrent" de l'avertisseur 5 à +9 dB d'émergence.

#### 4.1.4 Prototypes associés à "sirène"

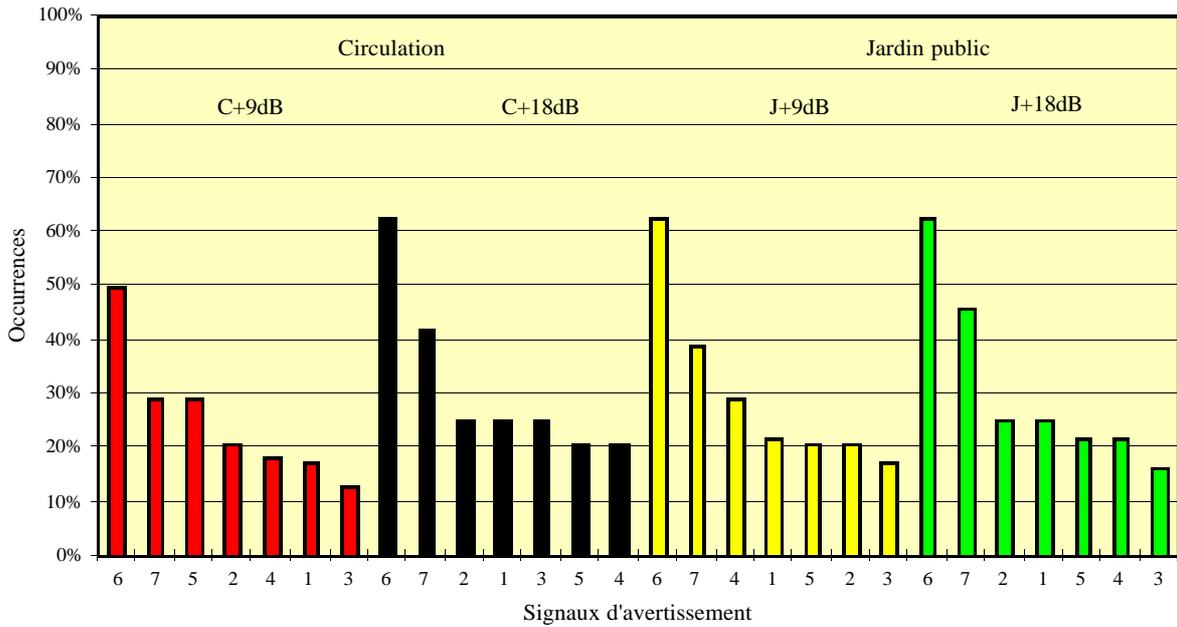


Figure 5-15 : prototypes associés à "sirène".

Les prototypes associés à "sirène" sont nombreux, mais le signal 6 se distingue très nettement des autres avertisseurs dans tous les cas d'écoute.

### 4.1.5 Prototypes associés à "klaxon"

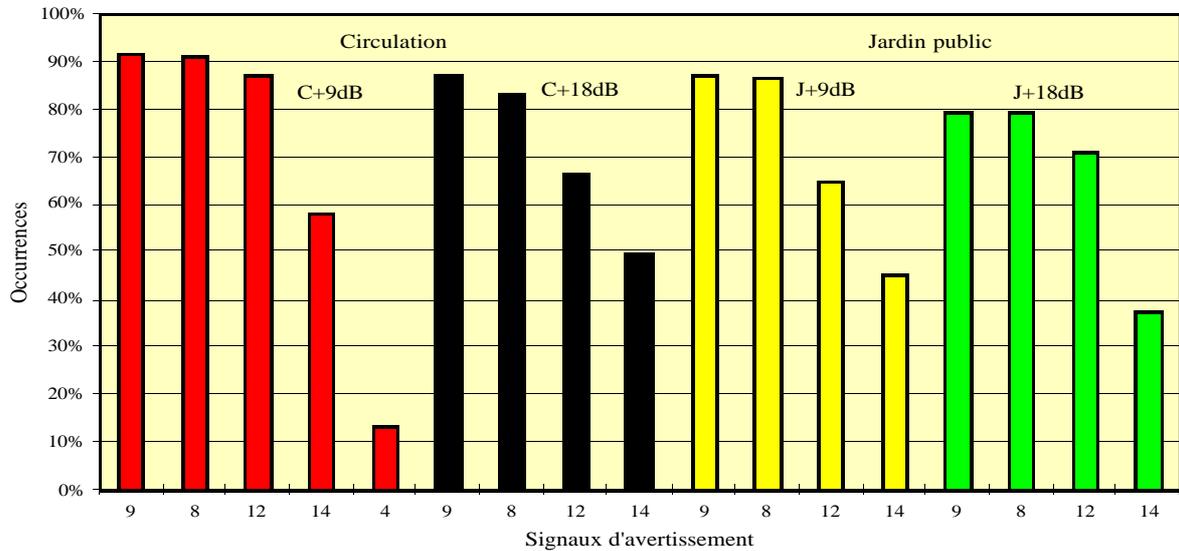


Figure 5-16 : prototypes associés à "klaxon".

Les signaux 8, 9, 12 et 14 sont associés à "klaxon" dans tous les contextes et à tous niveaux d'émergence. Mais la majorité des occurrences va aux signaux 9 et 8.

### 4.1.6 Prototypes associés à "voiture"

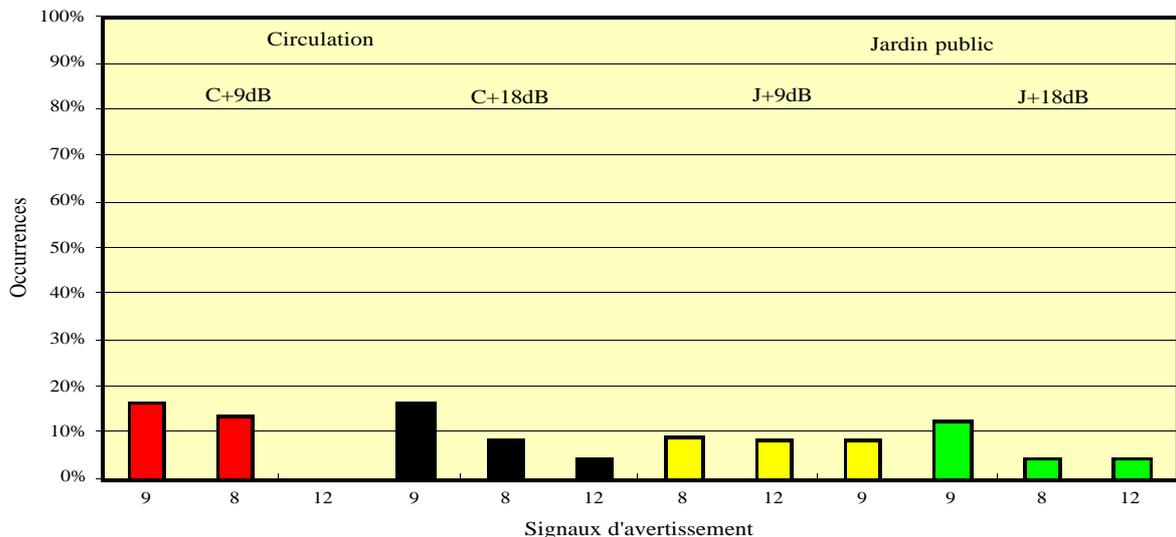


Figure 5-17 : prototypes associés à "voiture"

Les pourcentages d'occurrences associées à "voiture" sont très faibles quels que soit le contexte d'écoute et le niveau d'émergence perceptive. Nous retrouvons les signaux 9, 8 et, à moindre mesure, 12 qui sont avant tout associés à "klaxon" et qui sont essentiellement issus de "klaxon de voiture".

#### 4.1.7 Prototypes associés à "vélo"

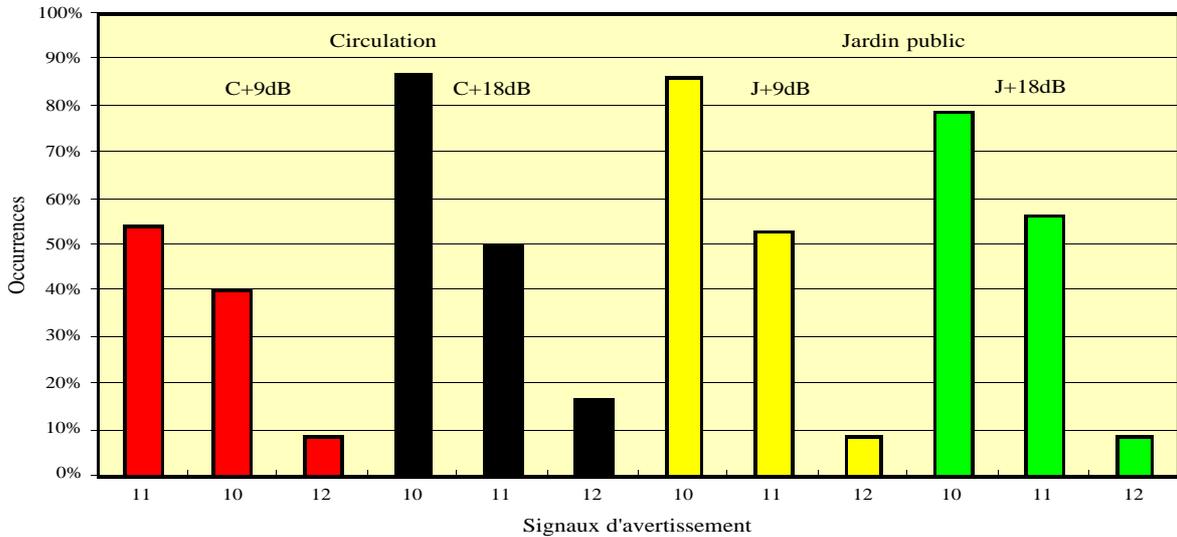


Figure 5-18 : prototypes associés à "vélo" et "bicyclette".

Le signal 10 est fortement prototypique des termes "vélo" et "bicyclette". Toutefois, en circulation et à faible émergence, ce signal est moins prototypique que l'avertisseur 11, qui a une prototypicité à peu près constante, de plus de 50%, sur les quatre situations d'écoute.

#### 4.1.8 Prototypes associés à "sifflet"

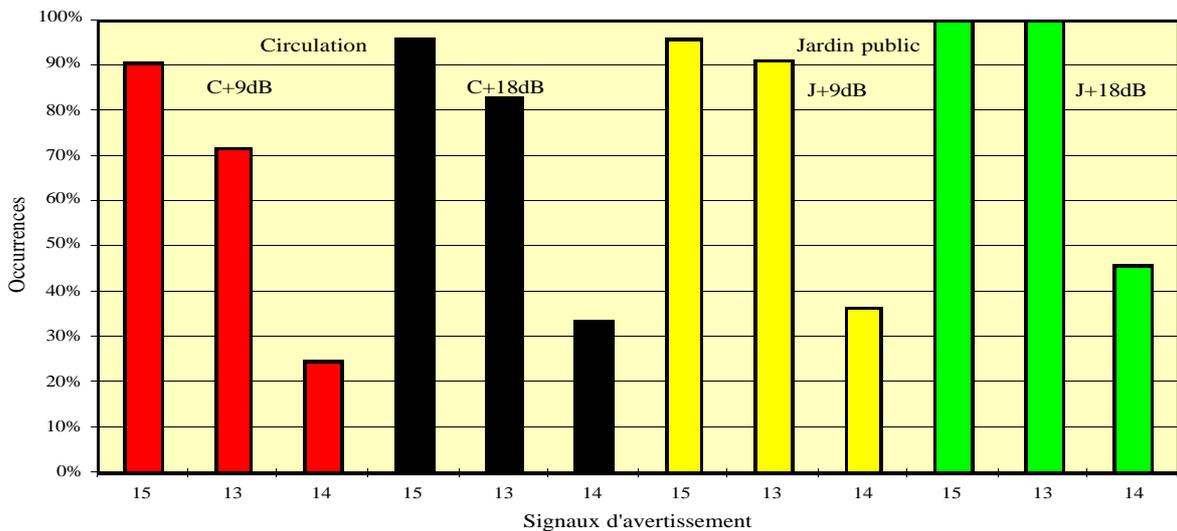


Figure 5-19 : prototypes associés à "sifflet".

Le signal prototype associé à "sifflet" est de toute évidence l'avertisseur 15. Il est cependant remarquable que lorsque le niveau d'émergence des signaux augmente, l'avertisseur 13 a tendance à se rapprocher de la typicalité de 15. De même pour le signal 14, mais dans une moindre mesure.

## 4.2 Interprétation

Parmi les signaux d'urgence, il y en a toujours un particulier qui est prototypique d'une désignation très fréquente. Ainsi, le signal 5 est prototypique de "pompiers", le signal 3 "d'ambulance" et le signal 6 de "police". Ce dernier est d'ailleurs également prototypique de la désignation plus générique "sirène".

Pour les signaux associés à "klaxon", les avertisseurs 8 et 9 sont tout aussi prototypiques l'un que l'autre de cette désignation. Par ailleurs, le signal 10 est très nettement prototypique de "vélo", sauf dans le contexte de circulation, à +9 dB d'émergence, où l'avertisseur 11 est plus représentatif. Cet effet est dû aux nombreuses absences de réponse qui affectent le signal 10 dans cette dernière situation d'écoute.

Enfin, le prototype associé à "sifflet" est le signal 15. Mais l'avertisseur 13 le devient également lorsque son émergence atteint +18 dB.

## 5. Les signaux "semblables"

La procédure expérimentale de la seconde expérience d'identification consiste en deux parties. Dans la première, que nous venons d'analyser (§ 3 et 4), les sujets doivent identifier des signaux *dissemblables*, c'est-à-dire intercatégoriels - appartenant à différentes catégories prototypiques -, en fonction du contexte et de l'émergence perceptive des signaux. La deuxième étape concerne des signaux *semblables*, intracatégoriels, avec une émergence perceptive de +18 dB, qui garantit une identification de tous les signaux d'avertissement (cf. § 3.2.1). Elle permet de déterminer, par une consigne de comparaison, les relations de similarité et de différence entre des signaux estimés proches les uns des autres. L'objectif est d'atteindre un certain nombre de caractéristiques acoustiques qui conduisent les individus à considérer un signal d'avertissement plus prototypique qu'un autre d'une signification donnée.

### 5.1 Résultats

Les réponses produites par les sujets, généralement sous forme de phrases complètes, constituent 264 données également réparties sur les deux contextes urbains. Ces productions verbales mettent en évidence les différentes manières de comparer les signaux entre eux, au sein d'une même séquence. Ainsi, nous avons pu constater que les sujets s'expriment en termes de :

- **propriétés physiques du son (64%) :**

- caractéristiques fréquentielles : hauteur et tonalité du son (19% des réponses),
- caractéristiques temporelles : rythme et répétitions (13%), durée (4%),
- émergence perceptive : niveau sonore (9%), distance (6%),
- propriétés acoustiques diverses (13%) : "sec, mat, modulation...",

- **jugements de valeur (25%) :** "agréable, désagréable, stressant, agaçant...",

- **niveau plus spécifique d'identification (4%) :** "sifflet à deux tons, simple, à roulette..."

Par ailleurs, 3% des sujets n'ont pas constaté de différences. La somme des réponses ne donne pas 100% car certains sujets n'ont pas fait de commentaires de comparaison, d'autres ont adopté plusieurs stratégies comparatives.

Mais la partie essentielle de l'analyse concerne le détail des signaux les uns par rapport aux autres. C'est ce que nous allons aborder dans les prochains paragraphes.

#### 5.1.1 Séquence des signaux 1-2-5

Dans cette séquence, les trois signaux 1, 2 et 5 ont été associés à "pompiers". Il apparaît très nettement que l'avertisseur des pompiers 5 est "plus agréable" que ceux de la police, 1 et 2. Les raisons évoquées indiquent qu'il est plus lent, plus grave et que l'intervalle des deux notes est plus faible.

A l'opposé, les sirènes de police 1 et 2 sont considérées comme "agressives" et "stressantes" car plus rapide, plus aiguës et à intervalle plus grand (quarte). Le signal 2 se distingue de 1 du fait qu'il paraît "essoufflé" et "poussif" et que la transition entre notes est détachée alors qu'elle est continue pour le signal 1.

### 5.1.2 Séquence des signaux 1-2-6-7

La comparaison entre les quatre signaux associés à "sirène" met en évidence deux groupes d'avertisseurs : 1 et 2 d'une part, pour lesquels nous retrouvons les mêmes critères de différences qu'au paragraphe 5.1.1 ; 6 et 7 d'autre part, qui sont associés à des sirènes américaines et qui se différencient par la rapidité de la modulation. Le premier groupe est généralement perçu plus dérangeant et stressant que le second, dont le son est qualifié de "plus rond" et "plus agréable" en raison de la modulation et de "moins d'attaque". Par ailleurs, le signal 2 paraît plus proche que 1, tous deux étant plus proches - ou forts - que les sirènes 6 et 7.

### 5.1.3 Séquence des signaux 2-3-4

La série des signaux 2, 3 et 4 nous donne un dernier moyen de comparaison de l'avertisseur de la police 2 avec les autres signaux "semblables". Ce dernier est qualifié de plus aigu et plus agressif que le signal 3, mais moins agressif que le signal 4. Or, 3 et 4 constituent tous deux le même signal mélodique, en hauteur comme en durée. Analysons donc les caractéristiques qui distinguent les deux avertisseurs d'ambulance : le signal 3, s'il paraît souvent plus fort et proche que les deux autres signaux, n'en est pas moins considéré comme plus agréable, car étant "plus rond" et plus riche en fréquences graves. Le signal 4 semble plus aigu et plus strident que 3.

### 5.1.4 Séquence des signaux 4-8-9

Comparé aux klaxons 8 et 9, l'avertisseur de l'ambulance 4 procure le même type de réponses que précédemment : il est décrit comme plus agressif, car plus aigu, et plus fort. En revanche, si les sujets distinguent facilement les deux avertisseurs de voiture par leurs propriétés rythmiques (un coup long pour 8, trois coups brefs pour 9), ils demeurent partagés quant à leurs interprétations de la signification de ces signaux. Pour les uns, la répétitivité du son est gage d'énervement du conducteur, pour les autres, c'est l'insistance du son continu qui prouve l'agacement de l'automobiliste...

### 5.1.5 Séquence des signaux 10-11

Concernant les deux sonnettes de bicyclette, les réponses sont beaucoup moins dispersées que pour les klaxons de voiture. En effet, la sonnette classique 10, munie de ses deux sons répétés, est plus aisément perçue que la sonnette 11, à deux hauteurs. Cette dernière est cependant jugée plus agréable et "plus douce", mais également plus grave.

### 5.1.6 Séquence des signaux 13-14-15

Les sifflets 13, 14 et 15 correspondent aux signaux pour lesquels les sujets ont produit le plus d'identifications en termes de "noms de noms" pour les comparer. En effet, les termes "sifflet à roulette", "sifflet à deux tons" et "sifflet à bouche" donnent déjà des informations quant à la sonorité du signal. Nous remarquerons toutefois que ces deux derniers sont parfois encore perçus comme klaxons. Ainsi, le sifflet à roulette 15 est le plus facilement perçu, mais également le plus strident et le plus agressif, tandis que le sifflet 13 est considéré comme le "plus lointain".

## **5.2 Interprétation**

Si l'analyse prototypique des signaux associés à "pompiers" a mis l'avertisseur 5 en exergue comme prototype, devant les signaux 1 et 2, la comparaison de ces trois signaux nous informe également que le signal 5 est le mieux accepté par les sujets, notamment du fait de ses caractéristiques rythmiques. Par ailleurs, l'avertisseur 2 est toujours considéré comme "désagréable", non seulement pour sa cadence rapide, mais également pour son caractère plus aigu et l'intervalle plus grand entre les deux fréquences fondamentales. Seul le signal d'ambulance 4 apparaît comme encore plus "stressant", car encore plus aigu. Quant à la sonnette 10, elle se voit confirmé son statut de prototype grâce à la répétition de la même note. Enfin, le sifflet 15 est également confirmé comme prototype, même s'il est souvent estimé désagréable.

L'analyse détaillée, séquence par séquence, des comparaisons entre signaux semblables peut ainsi expliquer les différences intracatégorielles qui sont constatées dans l'approche d'identification libre. Nous remarquerons qu'aucun effet de contexte n'a pu être mis en évidence dans cette seconde partie de l'expérience d'identification en contexte.

## 6. Conclusion

Cette expérience d'identification, en contexte, de signaux d'avertissement a permis de déterminer les signaux les plus prototypiques pour deux niveaux d'émergence perceptive donnés. La méthode développée dans la première expérience s'est également révélée fructueuse dans le second test, ce qui confirme son efficacité pour appréhender l'étude des signaux d'avertissement en contexte urbain par l'intermédiaire des représentations mentales des sujets. Dans une seconde partie d'expérience, la comparaison intracatégorielle des signaux a permis d'obtenir, non seulement des éléments de réponse quant aux propriétés signifiantes du signal sonore, mais également par rapport à des considérations hédoniques et de préférences de la part des sujets.

La comparaison des signaux semblables a également permis de vérifier l'importance fondamentale de la consigne. En effet, les faibles différences intracatégorielles entre signaux et la consigne de comparaison donnée aux sujets, qui avait pour but d'obtenir des descriptions de propriétés acoustiques différenciant les signaux, ont permis d'orienter les sujets vers une autre stratégie de réponse : au lieu de simplement *identifier* les signaux, ils les *qualifient*. Ce changement de stratégie n'est possible que parce que nous nous sommes assurés au préalable - en déterminant les seuils d'identification - que tous les signaux sont bien identifiables. Nous considérons en effet que l'individu ne peut qualifier un son qu'après l'avoir reconnu (Castellengo, 1994). Schaeffer (1966) avait d'ailleurs déjà constaté que les individus ne peuvent qualifier un son que lorsqu'il est mis en boucle, sans quoi ils ne cherchent qu'à l'identifier.

Ces nombreux éléments nous permettent désormais de rechercher les éventuels liens entre le sens perçu et verbalisé par les sujets et les propriétés acoustiques des signaux sonores d'avertissement. Ce rapprochement fait l'objet du chapitre suivant, après une mise en relation comparative des deux expériences.



# Chapitre 6 : Le sens et le son ; analyses acoustiques de propriétés pertinentes

Hunderte Töne waren zu einem drahtigen Geräusch ineinander verwunden,  
aus dem einzelne Spitzen vorstanden,  
längs dessen schneidige Kanten liefen und sich wieder einebneten,  
von dem klare Töne absplitterten und verflogen.  
An diesem Geräusch, ohne daß sich seine Besonderheit beschreiben ließe,  
würde ein Mensch nach jahrelanger Abwesenheit mit geschlossenen Augen erkannt haben,  
daß er sich in der Reichshauptstadt und Residenzstadt Wien befinde.

(L'enchevêtrement d'innombrables sons créait un grand vacarme barbelé  
aux arêtes tantôt tranchantes, tantôt émoussées, confuse masse d'où saillait une pointe ici et là  
et d'où se détachaient comme des éclats, puis se perdaient, des notes plus claires.

A ce seul bruit, sans qu'on pût définir pourtant la singularité,  
un voyageur eût reconnu les yeux fermés qu'il se trouvait à Vienne, capitale et résidence de l'Empire.)

*Robert Musil (1930)*

## 1. Introduction

Jusqu'à présent, tout le travail a porté sur le sens qui est associé aux signaux d'avertissement dans différentes situations urbaines. Cette signification a été abordée en termes de prototypicité, d'univocité et d'ambiguïté des signaux pour les sujets. Forts de ces enseignements, nous allons tenter de dresser un bilan de nos expériences et de chercher des corrélats acoustiques de ces significations. L'enjeu est de pouvoir optimiser l'efficacité des signaux d'avertissement en contexte urbain en réduisant la gêne occasionnée sur les individus usagers de l'espace urbain.

## **2. Le sens issu des expériences...**

Les procédures et consignes des deux expériences étant très similaires, il devient dès lors possible de mettre en relation les résultats issus de ces deux tests. L'unique différence de protocole réside dans le fait de demander une réponse par écrit dans la seconde expérience, alors qu'une réponse orale était rédigée par écrit par l'expérimentateur dans la première expérience. Les éléments d'analyse communs aux deux tests sont repris dans le même ordre des paragraphes des deux chapitres précédents.

### **2.1 Répertoire lexical disponible sur l'ensemble des données**

Sur l'ensemble des données il apparaît que les désignations les plus fréquentes sont sensiblement les mêmes dans les deux expériences (cf. chapitre 5 § 3.1 et chapitre 4, § 5.1). Ainsi, les trois désignations "klaxon", "ambulance" et "sifflet" sont les plus fréquentes dans les deux tests. Plus généralement, les huit désignations les plus fréquentes du second test se retrouvent dans les dix premières du premier test.

Les proportions de propriétés physiques du son (15% dans le premier test et 11% dans le second) et de noms de sources interprétées (85% et 89%) ne varient pas non plus de manière significative, tout comme celle des "noms de noms", interprétations sémantiques précises de sources interprétées (33% et 31%). Ainsi, les stratégies de réponse adoptées par les sujets sont identiques et confirment la similarité des deux expériences. Les résultats de ces dernières peuvent donc être comparées et ce, malgré la différence de procédure - productions orales ou écrites.

Les effets de contextes sont faibles à tous les niveaux d'émergence perceptive du signal. En revanche, des effets d'émergence peuvent être mis en évidence : l'écoute isolée se différencie des écoutes en situation par une production plus élevée des désignations "ambulance", "pompiers" et "police", mais également "klaxon de voiture", "poire" et "sonnette de vélo" (

Figure 6-1). Ces termes s'inscrivent donc dans un lexique particulièrement approprié à la description de signaux hors contexte.

Inversement, "klaxon" et "sirène" sont plus fréquents en écoute contextuelle que isolée, ainsi que "sonnette" au niveau seuil en circulation et "clochette" au niveau seuil en jardin public (

Figure 6-2). Ces quatre termes seraient donc plus appropriés au lexique utilisé par les gens en situations d'écoute contextuelle des mêmes signaux. Remarquons toutefois que les écarts-types associés aux données des deux figures sont relativement élevés -  $\Delta = 4$  (en valeurs relatives) au niveau seuil,  $\Delta = 7$  hors contexte et  $\Delta = 10$  en moyenne à +9 et +18 dB d'émergence perceptive - mais ne remettent pas en cause les tendances décrites ici.

Le changement de vocabulaire, plutôt générique en contexte et plutôt orienté vers des sources significatives hors contexte, est lié à un changement de stratégie de réponse des sujets. Cela confirme l'importance cruciale des stimuli et du corpus sonore dans les expériences.

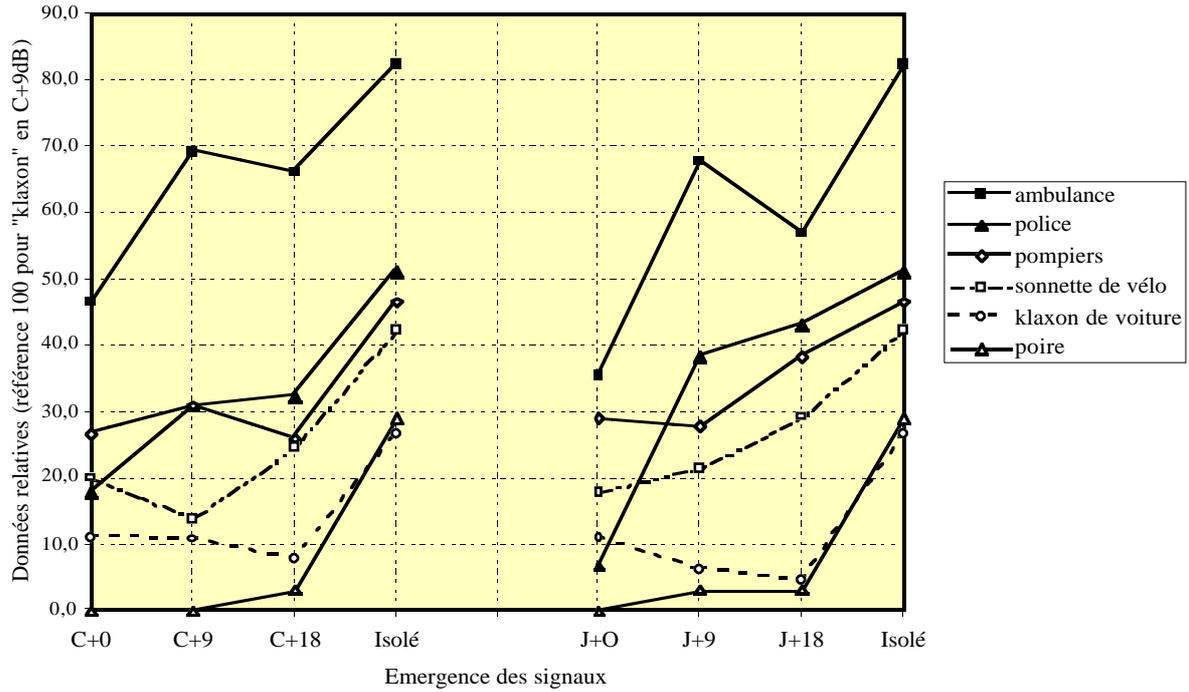


Figure 6-1 : les effets d'émergence 1 - expériences comparées.

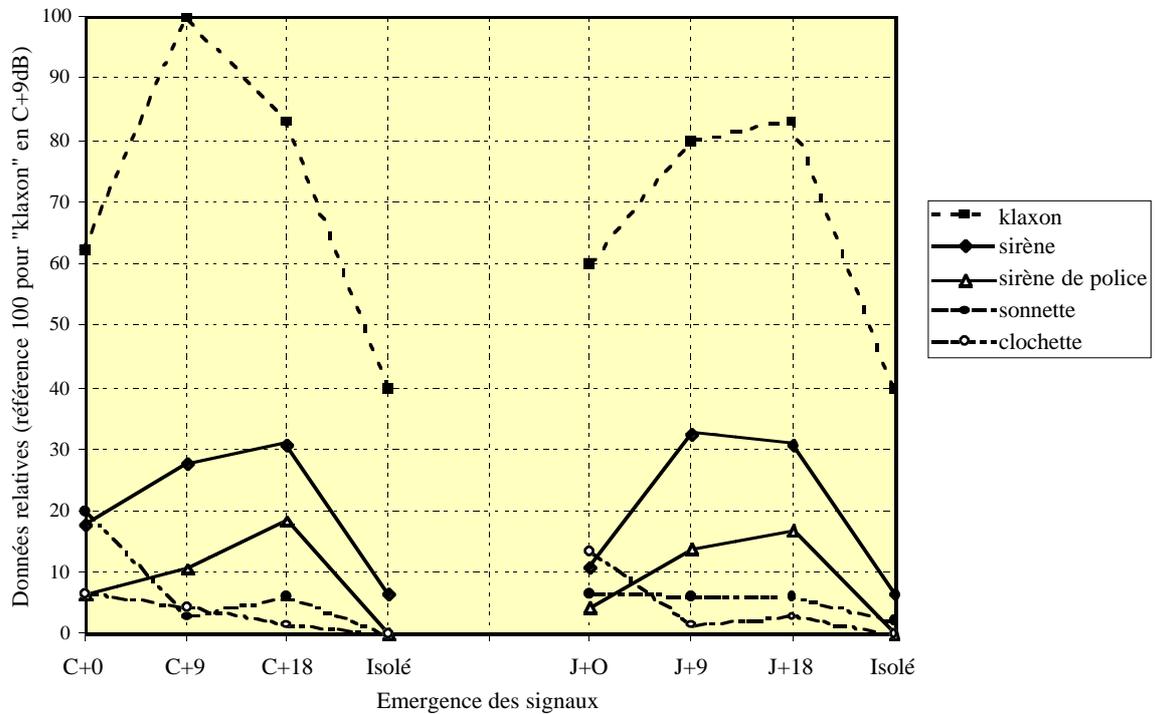


Figure 6-2 : effets d'émergence 2 - expériences comparées.

## **2.2 Description des 15 signaux dans les différentes situations d'écoute**

### **2.2.1 Les absences de réponse ("?")**

Les absences de réponse, qui caractérisent une absence d'identification des signaux, mettent en évidence, dans les deux tests, la sonnette 10 et le sifflet 13. La première a tendance à être moins bien reconnue dans une ambiance de circulation et le second dans un environnement de jardin public. En particulier, une émergence perceptive de +9 dB, quelque soit le contexte urbain, ne suffit pas à les rendre plus reconnaissables. Ce n'est qu'à +18 dB d'émergence que les absences de réponse deviennent insignifiantes.

En revanche, les signaux 12 et 7, difficilement identifiés à leur niveau seuil dans le jardin public, deviennent clairement identifiables à partir de +9 dB d'émergence.

Les absences de réponse représentent 4,5% des désignations au niveau seuil du signal, 8,9% à +9 dB d'émergence, 1,5% à +18 dB d'émergence et 0,5% en écoute isolée. La hausse entre le seuil et +9 dB est probablement liée aux différences de protocole expérimental entre les deux tests. En effet, la première procédure laisse le sujet s'exprimer oralement et l'expérimentateur rédige de la manière la plus exhaustive possible, tandis que la seconde procédure demande au sujet une réponse écrite, moins spontanée et certainement moins complète. Cette différence de procédure rend délicate la comparaison des données entre les deux tests. Toujours est-il que, si la majorité des signaux sont toujours nettement identifiés par les sujets dès le seuil, à partir de +18 dB d'émergence perceptive, même les signaux 10, 11 et 13 ne s'intègrent plus au contexte en passant inaperçus aux oreilles des auditeurs.

### **2.2.2 Productivité et variabilité des désignations des signaux d'avertissement**

En comparant le nombre de désignations pour chaque signal dans les deux expériences, il apparaît que les signaux 11 et 15 sont toujours respectivement riche et pauvre en désignations. Dans une moindre mesure, les signaux 14 et 6 restent également riches en désignations dans les deux expériences.

L'influence du fond sonore semble très faible sur la productivité des sujets. Seul le sifflet 13 comporte moins de désignations dans le jardin public que dans la circulation et ce, dans les deux expériences.

### **2.2.3 Fréquence la plus élevée**

Parmi les désignations produites le plus grand nombre de fois par les sujets dans les deux expériences, la désignation "klaxon" est systématiquement associée aux signaux 8 et 9 (mais de manière moins évidente pour 9), "ambulance" au signal 3 et "sifflet" au signal 15. Ces trois signaux manifestent donc un bon accord entre sujets et deviennent des prototypes potentiels.

## 2.3 Signaux prototypes des désignations les plus fréquentes

Dans ce paragraphe, nous allons reprendre les résultats des prototypes issus des deux expériences et les combiner dans des mêmes représentations graphiques. Ceci, afin de nous permettre d'évaluer aisément les effets de contexte et d'émergence perceptive significatifs pour chaque signal et par rapport à une désignation donnée. Nous commenterons et interpréterons ces résultats "compilés" à la suite des figures, dans le § 3.1.3.

### 2.3.1 Prototypes associés à "ambulance"

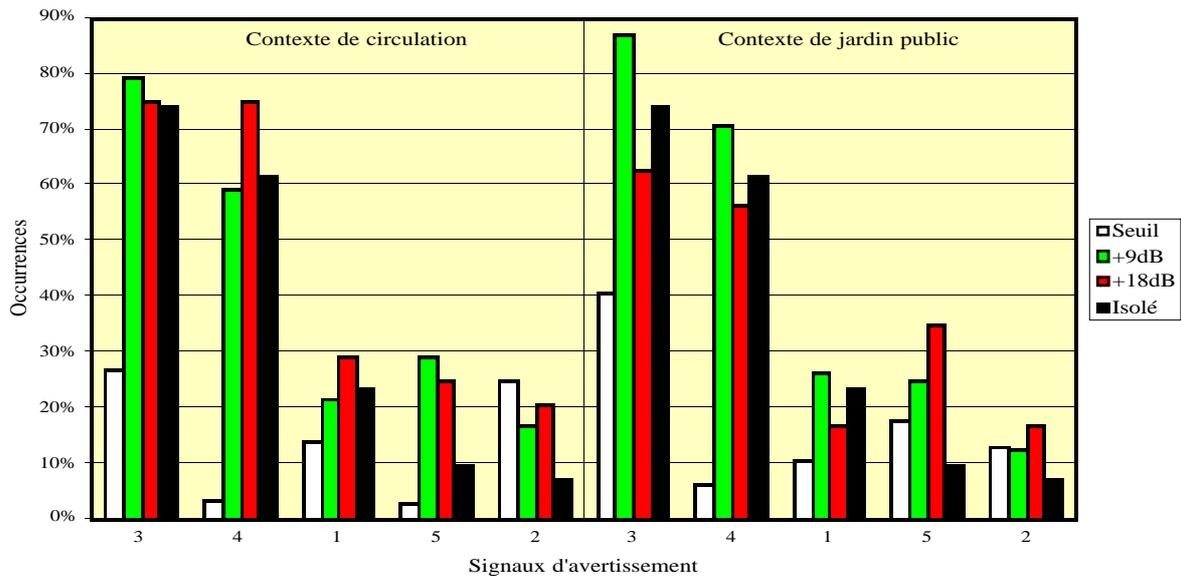


Figure 6-3 : prototypes associés à "ambulance" et "S.A.M.U." en fonction de l'émergence perceptive.

Le signal 3 est le plus prototypique de la désignation "ambulance" et ce, quels que soient le niveau d'émergence perceptive et le contexte d'écoute. Dans toutes les situations, l'avertisseur 4 est moins prototypique que 3.

### 2.3.2 Prototypes associés à "police"

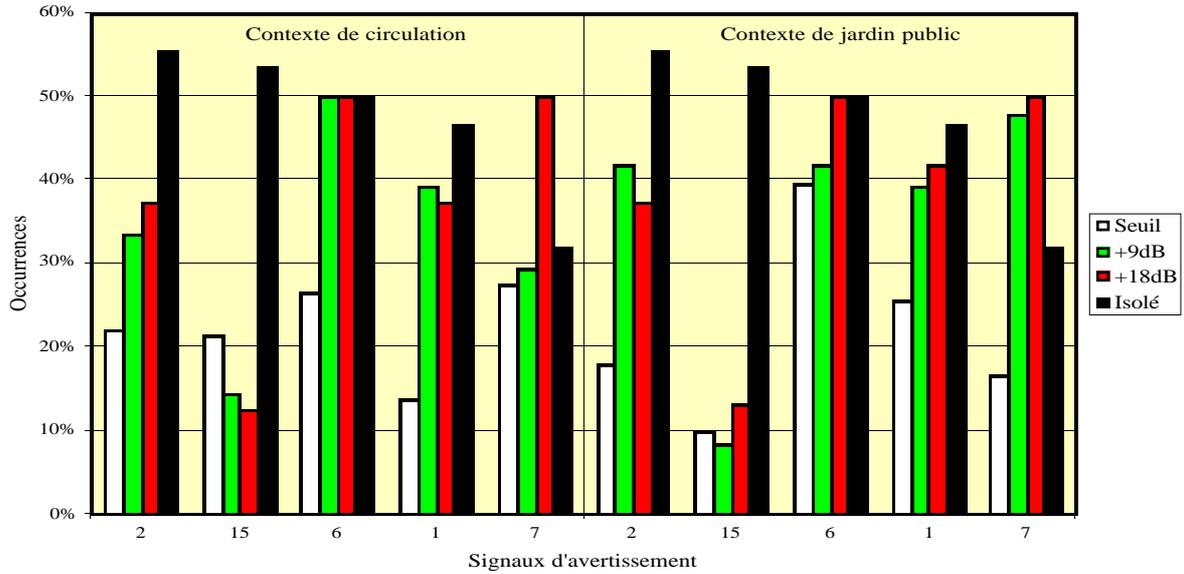


Figure 6-4 : prototypes associés à "police" et "flics" en fonction de l'émergence perceptive.

Si le signal 2 est le prototype de la désignation "police" en situation isolée, il s'avère que le signal 6 l'est bien plus en situation contextuelle et est résistant aux faibles niveaux d'émergence perceptive.

### 2.3.3 Prototypes associés à "pompiers"

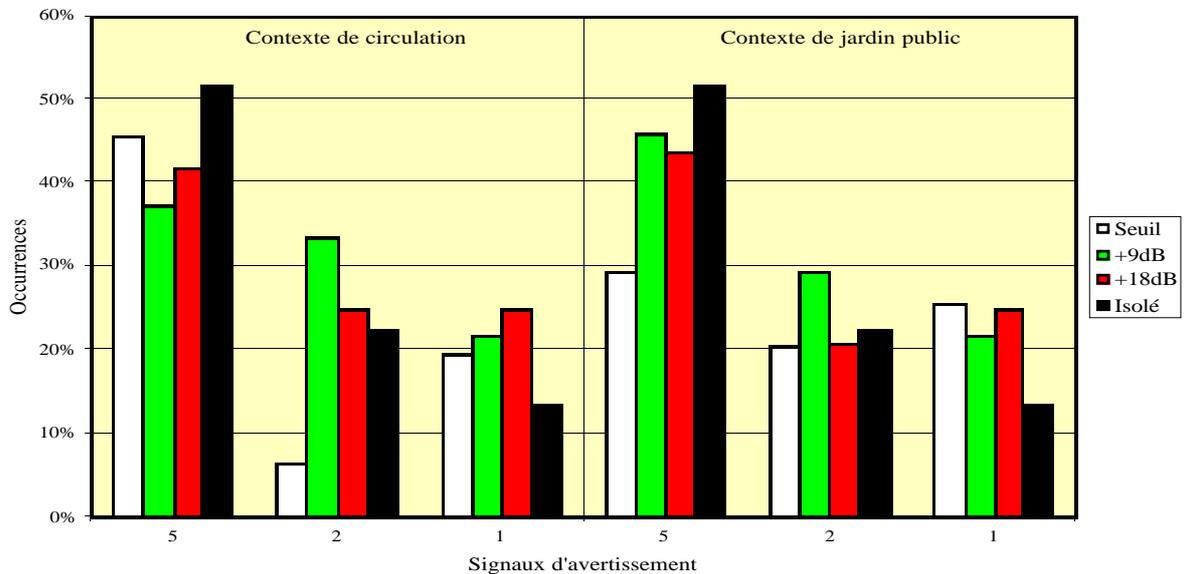


Figure 6-5 : prototypes associés à "pompiers" en fonction de l'émergence perceptive.

Pour la désignation "pompiers", il ne fait aucun doute que le signal 5 lui est associé comme prototype. Cette avertisseur est remarquablement résistant aux effets d'émergence perceptive. Toutefois, l'identification demeure inférieure à 50% en situation contextuelle.

### 2.3.4 Prototypes associés à "sirène"

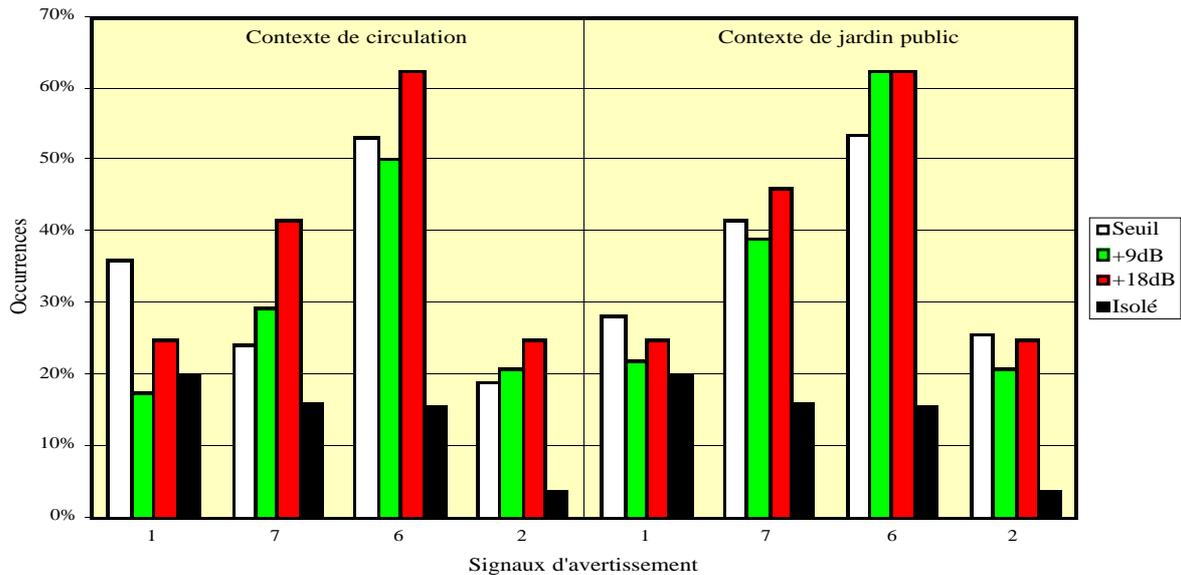


Figure 6-6 : prototypes associés à "sirène" en fonction de l'émergence perceptive.

En contexte, le signal 6 est clairement associé à la désignation "sirène". En revanche, hors contexte, cette désignation est moins souvent produite par les sujets, quelque soit l'avertisseur testé.

### 2.3.5 Prototypes associés à "klaxon"

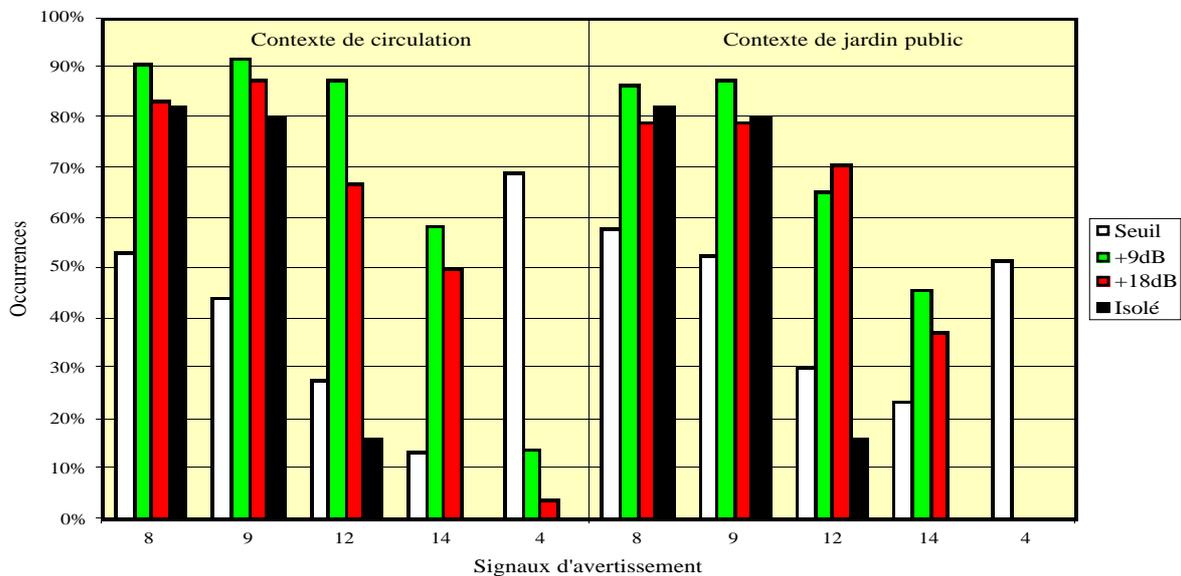


Figure 6-7 : prototypes associés à "klaxon" en fonction de l'émergence perceptive.

Les signaux 8 et 9 correspondent tous deux à des prototypes de la désignation "klaxon". Ils sont particulièrement résistants aux faibles niveaux d'émergence, notamment l'avertisseur 8 pour lequel plus de la moitié des occurrences contiennent le terme "klaxon", même aux niveaux seuils. Les signaux 13 et 14 quant à eux, dépendent fortement du niveau d'émergence perceptive, tout comme l'avertisseur 4.

### 2.3.6 Prototypes associés à "voiture"

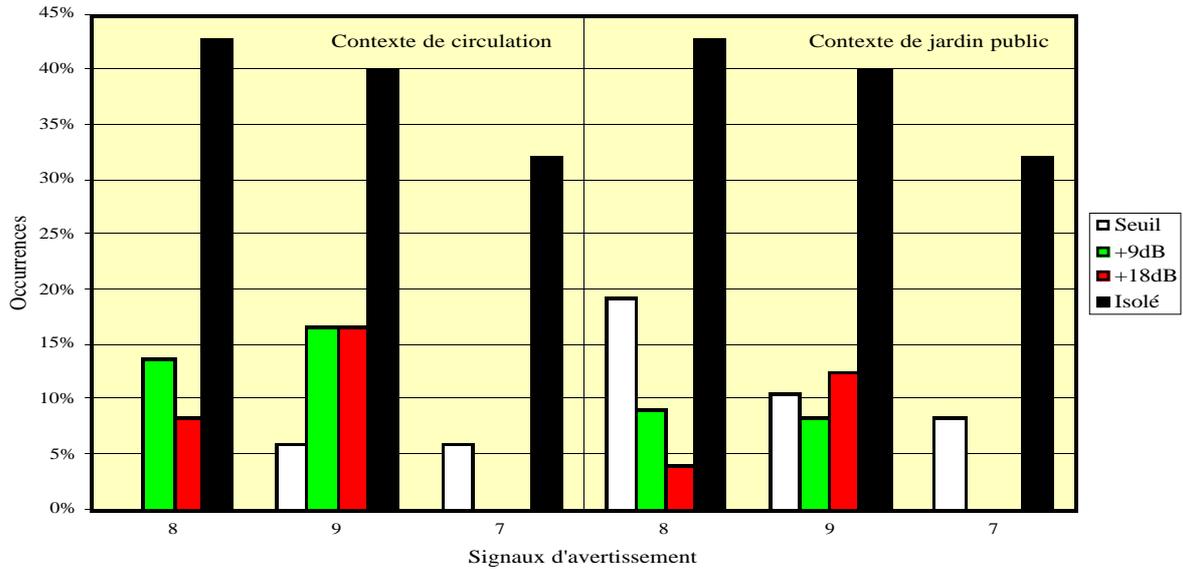


Figure 6-8 : prototypes associés à "voiture" en fonction de l'émergence perceptive.

La désignation "voiture" apparaît essentiellement en situation d'écoute isolée par l'intermédiaire de désignations du type *noms de noms*, tels que "klaxon de voiture" pour les avertisseurs 8 et 9 ou encore "alarme de voiture" pour le signal 7. A noter, l'influence du contexte pour le signal 8 au niveau seuil d'émergence perceptive.

### 2.3.7 Prototypes associés à "sonnette"

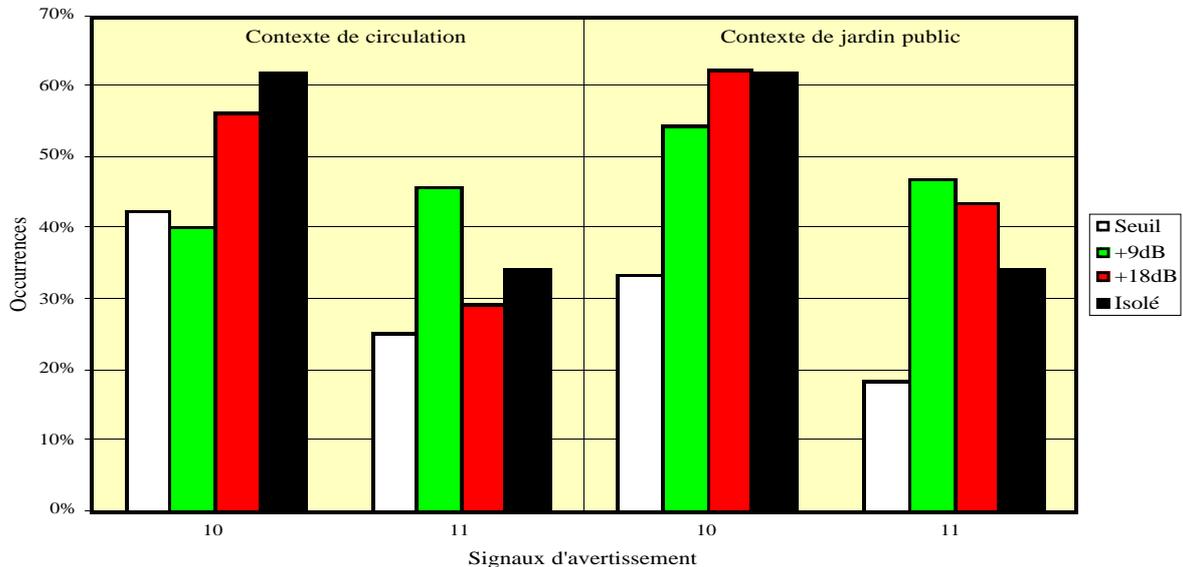


Figure 6-9 : prototypes associés à "sonnette" en fonction de l'émergence perceptive.

Les signaux 10 et 11 sont prototypiques de "sonnette" et ce, indépendamment du contexte. Néanmoins, l'avertisseur 10 l'est globalement un peu plus que 11 qui est moins influencé par l'émergence perceptive, mais à un niveau d'occurrence plus faible, du fait de sa sonorité inhabituelle.

### 2.3.8 Prototypes associés à "vélo"

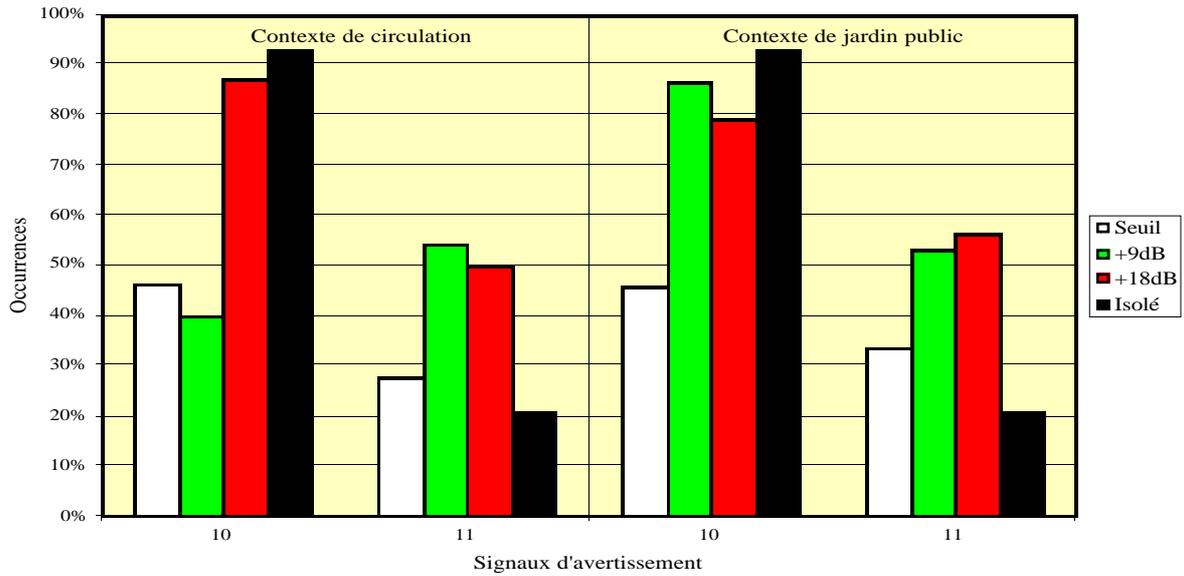


Figure 6-10 : prototypes associés à "vélo" et "bicyclette" en fonction de l'émergence perceptive.

Les prototypes associés à la désignation "vélo" sont les mêmes que ceux associés à "sonnette". Il apparaît nettement que l'avertisseur 10 est le plus souvent associé à une "sonnette de vélo", notamment lorsque l'émergence perceptive est grande (cf. § 2.3.7), tandis que le signal 11 est très peu représenté par le terme "vélo" en situation isolé, ce qui confirme son caractère atypique.

### 2.3.9 Prototypes associés à "sifflet"

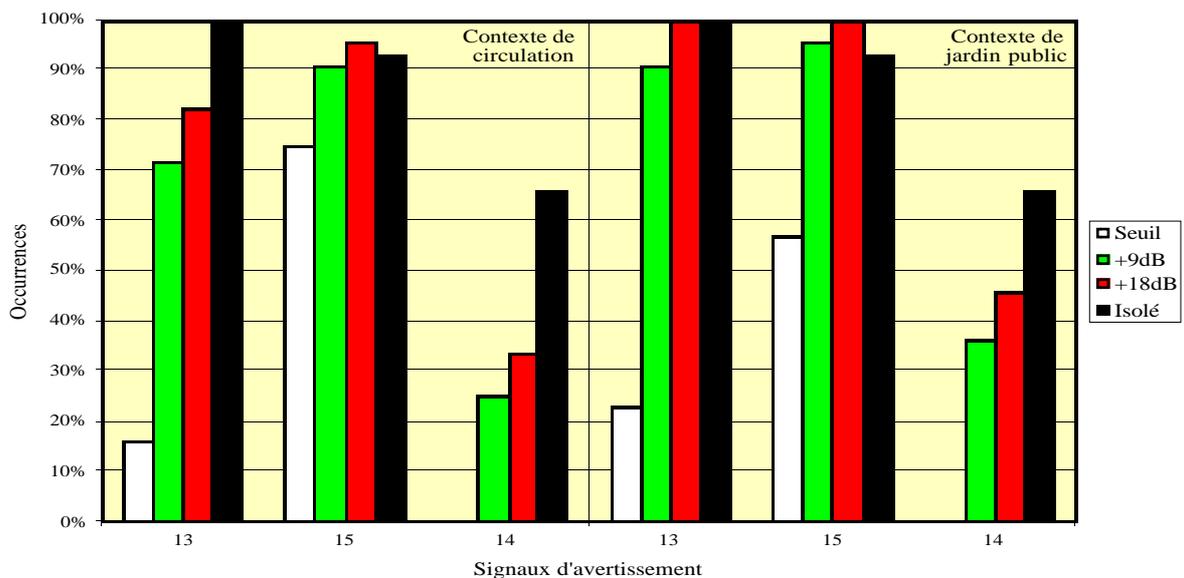


Figure 6-11 : prototypes associés à "sifflet" en fonction de l'émergence perceptive.

Trois signaux sont associés à la désignation "sifflet", mais deux d'entre eux seulement, 13 et 14, sont fortement influencés par l'émergence perceptive. L'avertisseur 15 est résistant à

l'effet d'émergence et ce, pour plus de la moitié de occurrences contenant le terme "sifflet", quels que soient le niveau et le contexte.

### 2.3.10 Interprétations

En combinant les résultats issus des analyses prototypiques des deux tests comme nous venons de le voir, il en découle sans ambiguïté les prototypes suivants :

- signal 3 associé à "ambulance",
- signal 5 associé à "pompiers",
- signal 6 associé à "sirène",
- signaux 6 et 1 associés à "police".
- signal 8 associé à "klaxon",
- signaux 8 et 9 associés à "voiture",
- signal 10 associé à "sonnette" et "vélo",
- signal 15 associé à "sifflet".

Partant de ces signaux prototypiques, il s'agit désormais d'essayer de trouver des propriétés acoustiques qui puissent être associées à la forte signification de ces signaux. L'absence de prototypicité des autres signaux doit aider à dégager clairement ces propriétés.

### 3. ...et le son des expériences

Les analyses acoustiques des stimuli testés constituent l'approche complémentaire de l'aspect sémiotique de cette étude. En effet, les résultats issus de la psycholinguistique, présentés jusqu'ici, doivent pouvoir être mis en relation avec des propriétés sonores signifiantes des signaux (chapitre 2, § 4). Dans un premier temps, nous donnerons les caractéristiques acoustiques des deux fonds sonores testés (§ 3.1). Dans un deuxième temps, nous décrirons pour chaque signal les propriétés acoustiques générales des avertisseurs, puis nous tenterons de mettre leurs significations en lumière au vu de leurs différences sonores (§ 3.2).

#### 3.1 Acoustique des contextes urbains

La représentation sonographique permet de visualiser les informations spectrales, temporelles et intensimétriques (cf. chapitre 1, § 2.2). Nous allons présenter les sonagrammes des séquences complètes du fond sonore urbain. Toutefois, la quantité des informations - 20 s réduites sur un graphique - rend l'analyse délicate. L'idée de cette démarche est uniquement de visualiser l'évolution temporelle des propriétés spectrales des fonds sonores urbains utilisés. Pour ce faire, nous allons compléter la représentation sonographique par des spectres cumulés sur deux secondes et mesurés en début de séquence ( $2 < t < 4$  s), en milieu de séquence ( $9 < t < 11$  s) et en fin de séquence ( $16 < t < 18$  s). Pour plus de précision, nous donnerons les spectres en échelle fréquentielle linéaire (détail des hautes fréquences) et logarithmique (détail des basses fréquences). La dynamique des représentations, c'est-à-dire du blanc au noir, est de 80 dB.

##### 3.1.1 Contexte de circulation

Le contexte de circulation présente une évolution temporelle relativement faible, marquée par quelques bandes verticales qui correspondent au passage de certains véhicules (Figure 6-12). En milieu de séquence (Figure 6-13), les fréquences autour de 100 Hz sont fortement amplifiées (10 dB) lors du passage d'un camion, de même que les fréquences comprises entre 4 et 8 kHz (5 dB). Pour cette raison, nous n'avons jamais disposé de signal à cet instant précis de la séquence sonore car nous aurions été confrontés à d'importants phénomènes de masquage fréquentiel. Néanmoins, la décroissance spectrale reste globalement constante et les fréquences au-dessus de 15 kHz sont fortement atténuées.

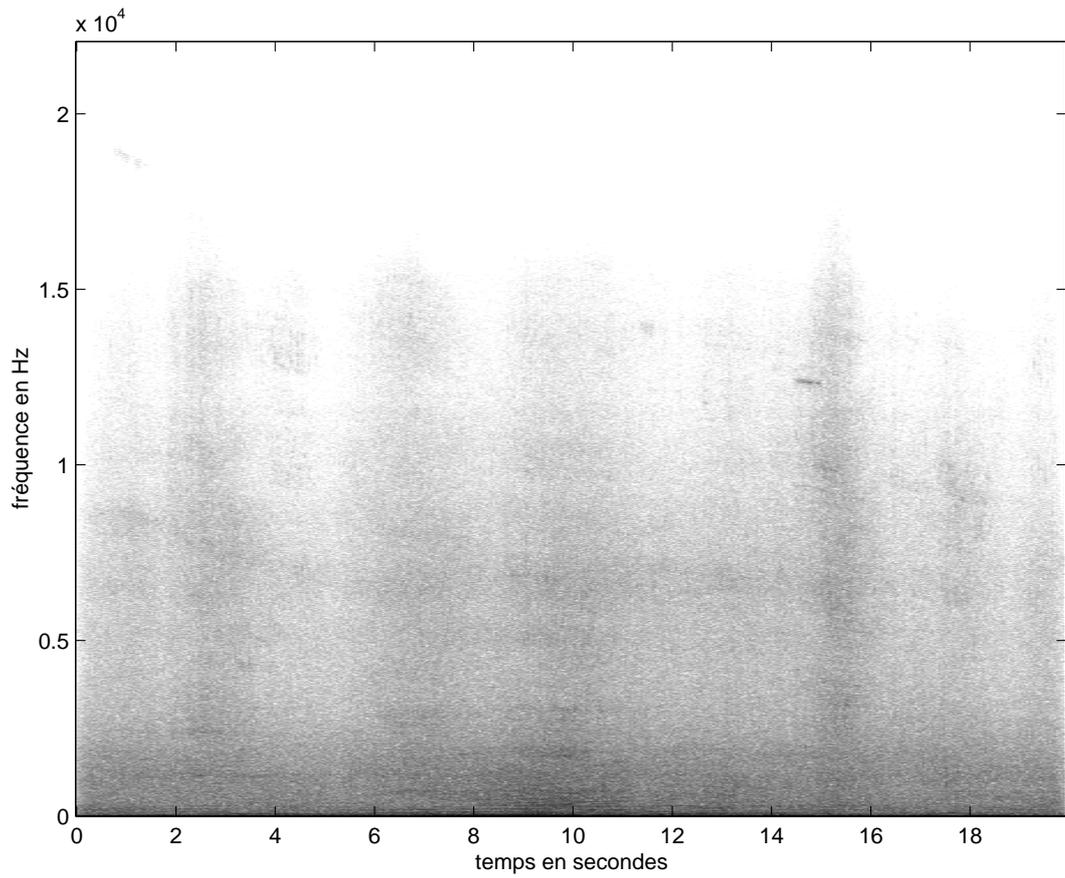


Figure 6-12 : sonagramme du contexte de circulation.

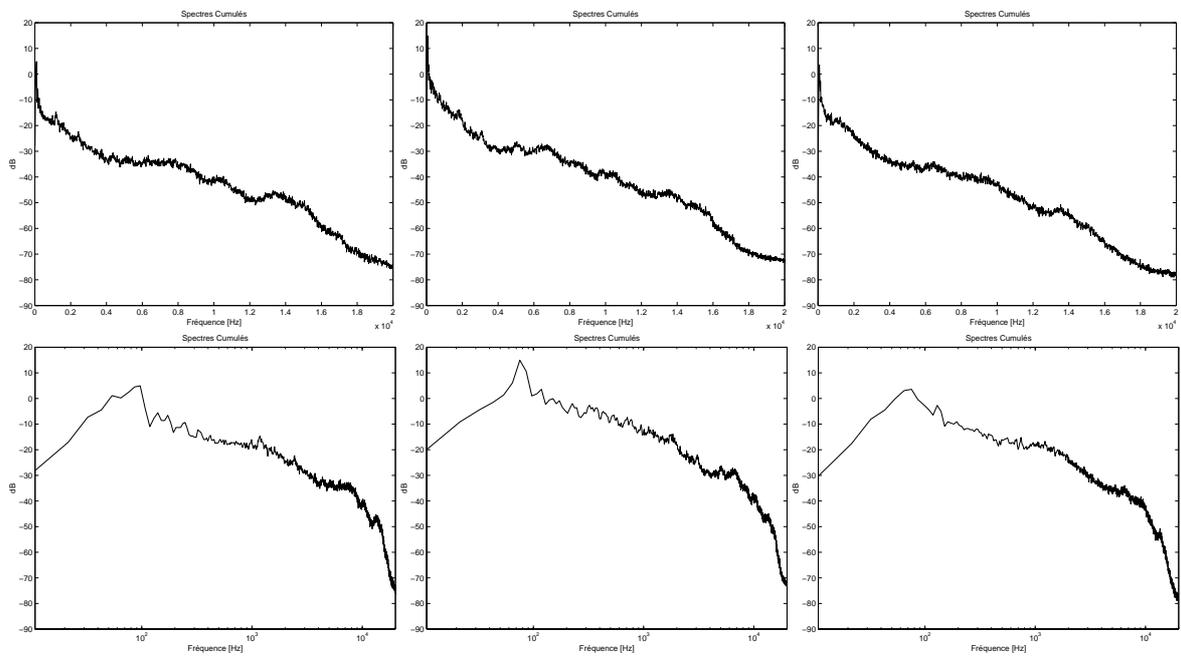


Figure 6-13 : spectres cumulés du fond de circulation :  
 échelle linéaire ( $t = [2,4]$  s ;  $t = [9,11]$  s ;  $t = [16,18]$  s),  
 échelle logarithmique ( $t = [2,4]$  s ;  $t = [9,11]$  s ;  $t = [16,18]$  s).

### 3.1.2 Contexte de jardin public

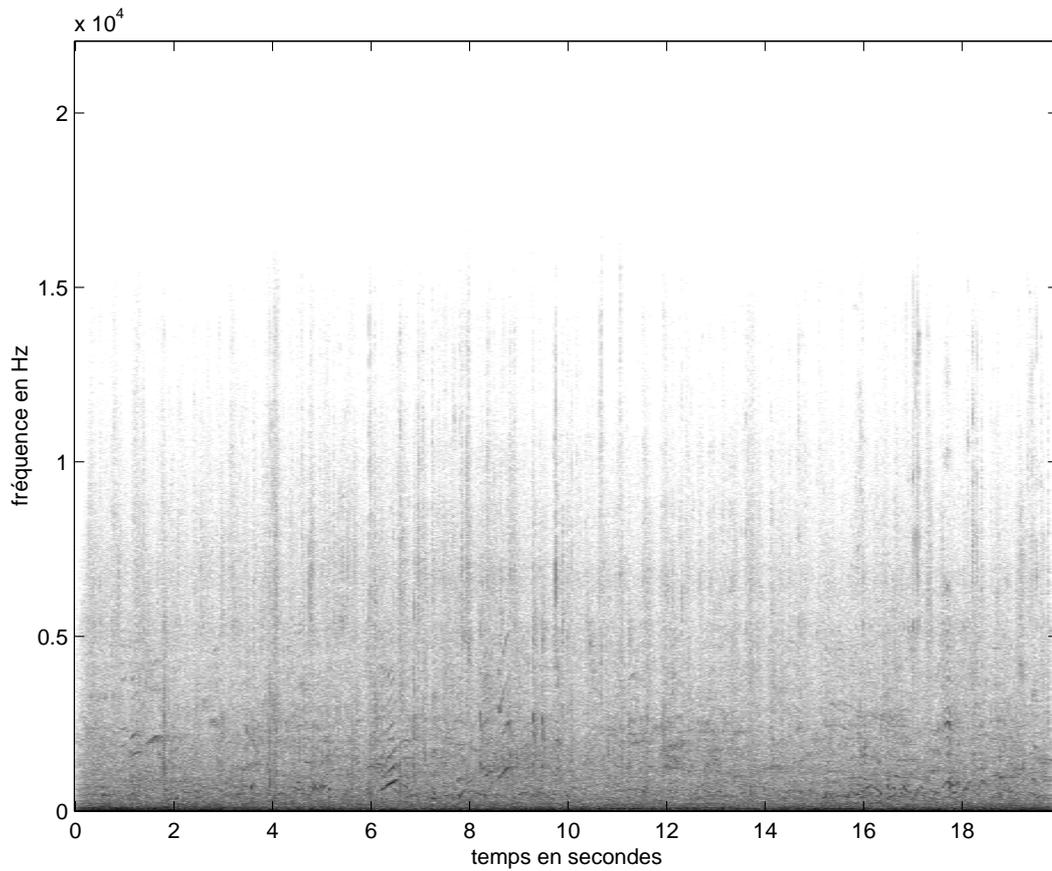


Figure 6-14 : sonagramme du contexte de jardin public.

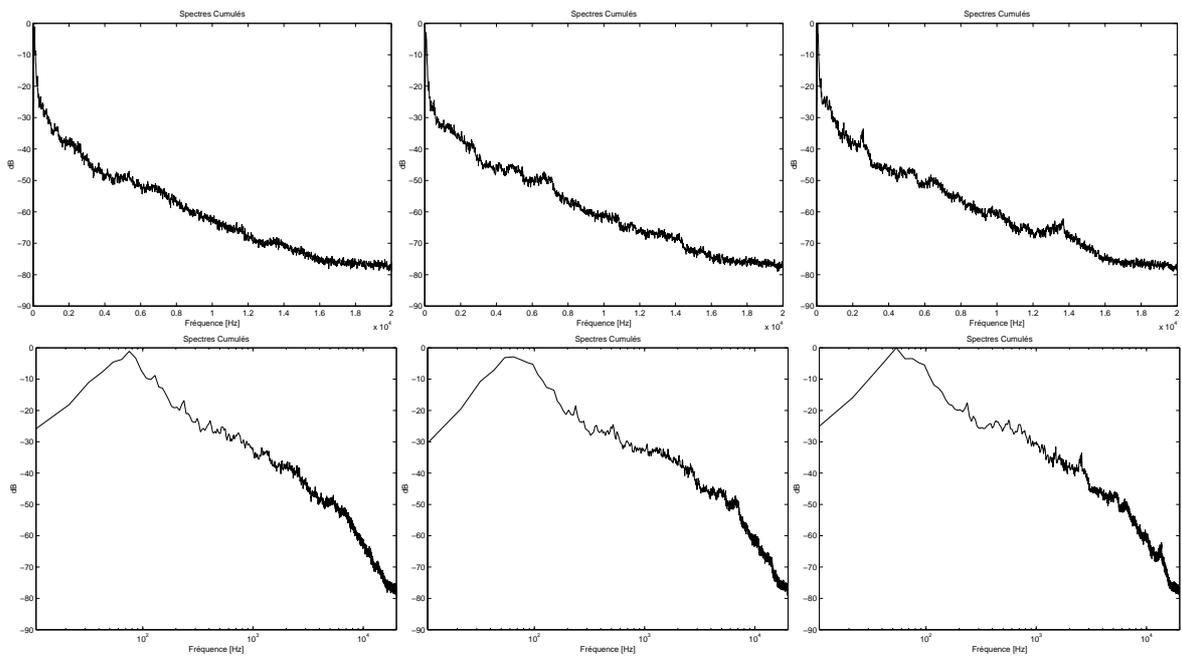


Figure 6-15 : spectres cumulés du fond de circulation :  
 échelle linéaire ( $t = [2,4] \text{ s}$  ;  $t = [9,11] \text{ s}$  ;  $t = [16,18] \text{ s}$ ),  
 échelle logarithmique ( $t = [2,4] \text{ s}$  ;  $t = [9,11] \text{ s}$  ;  $t = [16,18] \text{ s}$ ).

Le contexte de jardin public présente (Figure 6-14), à la différence de celui de circulation, de nombreuses raies verticales qui indiquent la présence de bruits impulsifs tels que des bruits de pas, des cris d'enfants, des piaillements d'oiseaux... Ces sons secs couvrent toutes les fréquences, tandis que le fond continu devient très faible au-dessus de 10 kHz. Les spectres cumulés (Figure 6-15) indiquent très peu de différences spectrales entre le début et la fin de la séquence et les écarts d'intensité maximale, autour de 70 Hz pour toute la séquence, sont relativement faibles (3 dB).

### 3.1.3 Interprétation

Les mesures intensimétriques effectuées sur les séquences contextuelles valident nos mesures de seuils d'émergence perceptive. En effet, la procédure de la première expérience (pas de +3 dB à chaque itération ; cf. chapitre 4, § 2.1) limite la précision des niveaux seuils à  $\pm 3$  dB. Nous en avons évidemment tenu compte dans nos analyses (chapitre 4, § 4) afin d'éviter toute interprétation hâtive. De ce fait, les trois niveaux d'émergence perceptive testés (0 dB, +9 dB et +18 dB) sont bien valides dans la limite des incertitudes précitées, les lignes isosoniques écartées de moins de 20 dB pouvant être considérées comme parallèles entre elles.

## 3.2 Acoustique des 15 signaux d'avertissement

Tout comme pour les contextes d'écoute, nous allons présenter les quinze signaux sur tout le spectre audible et sur toute la durée de chaque signal. Cette représentation est suffisante du fait de la courte durée des avertisseurs. Les valeurs numériques fréquentielles et temporelles sont données à partir d'analyses de Fourier par F.F.T. plus détaillées, non représentées dans ce mémoire. La dynamique de la représentation est toujours de 80 dB.

### 3.2.1 Signal 1 : Police (nouveau, électronique)

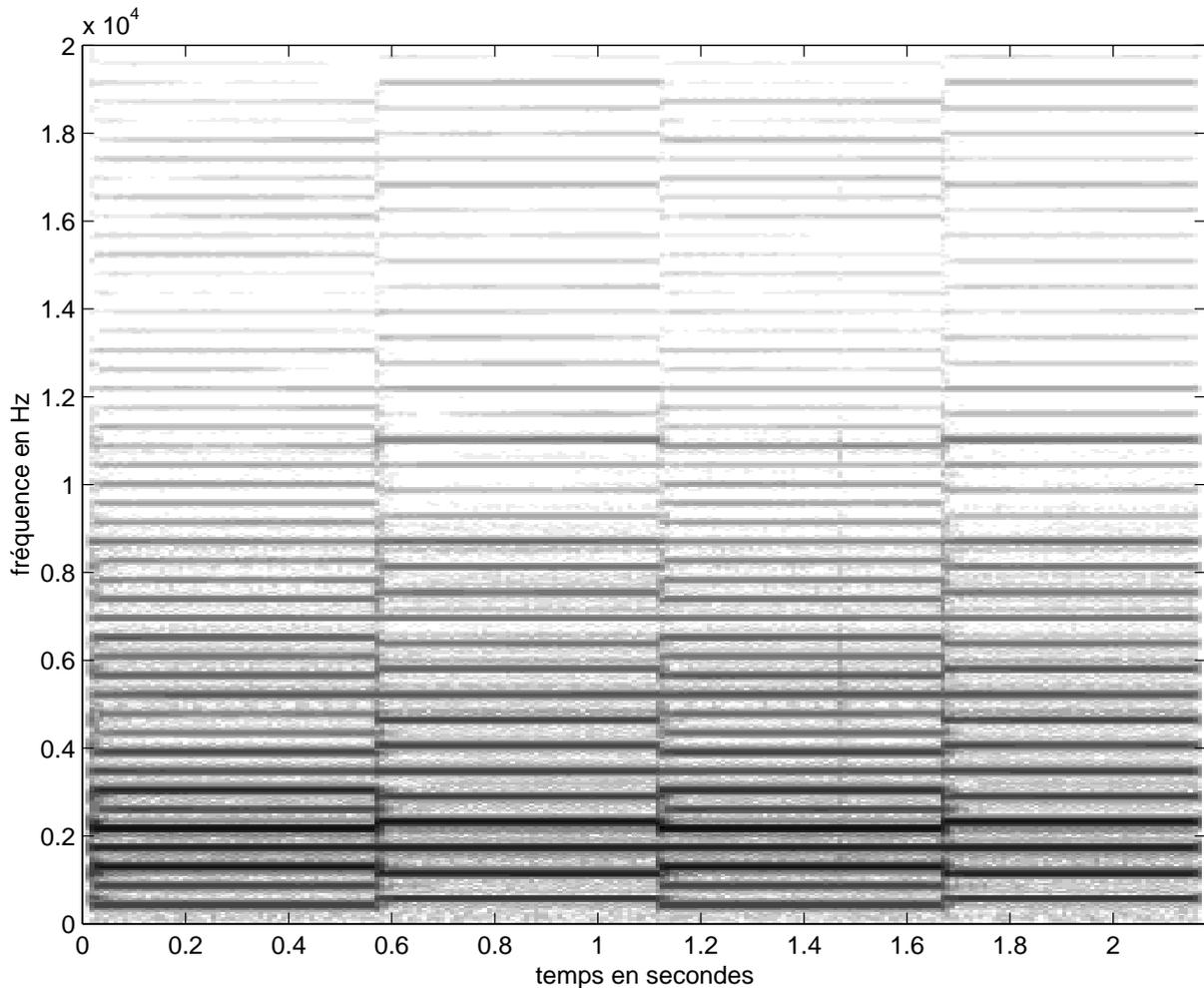


Figure 6-16 : Sonagramme du signal 1  
(Féché = 44100 Hz, 1024 pts, hanning(1024), 800 pts de recouvrement).

L'avertisseur électronique de la police présente deux notes très riches en harmoniques et dont les fréquences fondamentales sont exactement dans un rapport de quarte ( $4/3$ ). En effet, le premier fondamental se trouve à 435 Hz, soit très proche du La 440 du diapason (la richesse harmonique en plus), tandis que le second vaut 580 Hz. La précision est telle que le quarantième harmonique de la première note et le trentième de la seconde se retrouvent précisément à la même fréquence de 17400 Hz. Les neuf premiers harmoniques de chaque note sont particulièrement intenses.

Le passage d'une note à l'autre se produit avec un très léger recouvrement d'environ 20 ms. La durée de chaque note est de 0,55 s, soit 2,2 s pour deux périodes du signal.

### 3.2.2 Signal 2 : Police (ancien, mécanique)

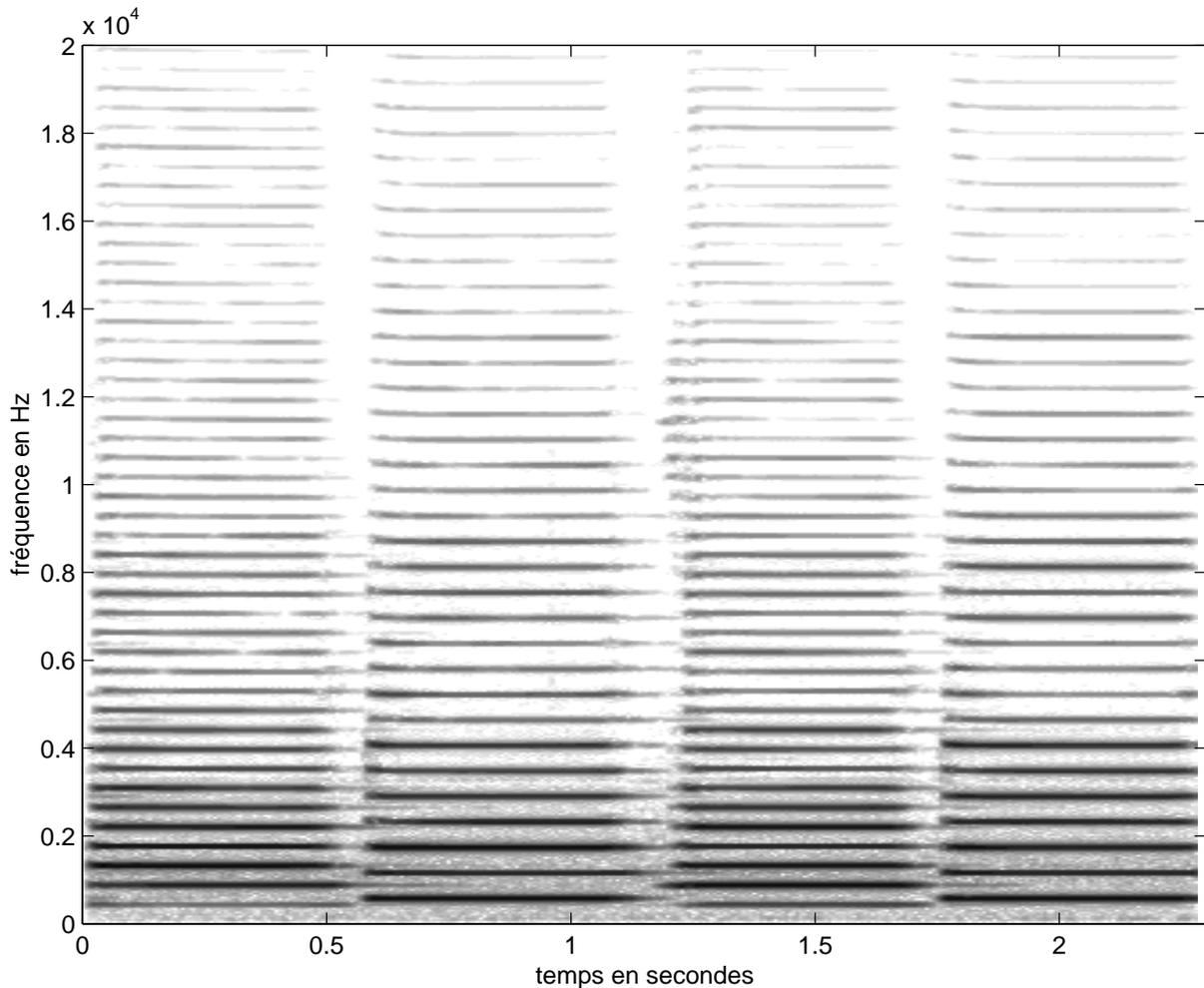


Figure 6-17 : Sonagramme du signal 2  
(Féch = 44100 Hz, 1024 pts, hanning(1024), 700 pts de recouvrement).

La nouvelle version de l'avertisseur de la police, que nous venons d'analyser, est très proche du modèle plus ancien. En effet, ce dernier comporte deux notes de hauteur 440 Hz et 580 Hz. Leur rapport est donc également d'une quarte, mais la très légère imprécision sur ce rapport conduit à ce que, non pas les harmoniques 20 de la première note et 15 de la seconde, mais 20 et 16 se retrouvent à la même fréquence. Le spectre est aussi riche que pour l'avertisseur récent.

D'un point de vue temporel, les différences sont plus marquées, notamment au niveau des transitoires d'attaque des notes ; d'une part, l'arrivée des harmoniques d'ordre élevé est retardée par rapport à celle des premiers harmoniques et d'autre part, la seconde note grave présente un transitoire particulier, dont les harmoniques 17 à 34 sont plus marqués et non retardés, qui constitue une attaque très aiguë, presque "sifflée". Un autre fait marquant est le faible recouvrement des notes, l'extinction progressive d'une note débutant avant l'apparition de la prochaine.

### 3.2.3 Signal 3 : Ambulance (nouveau, électronique)

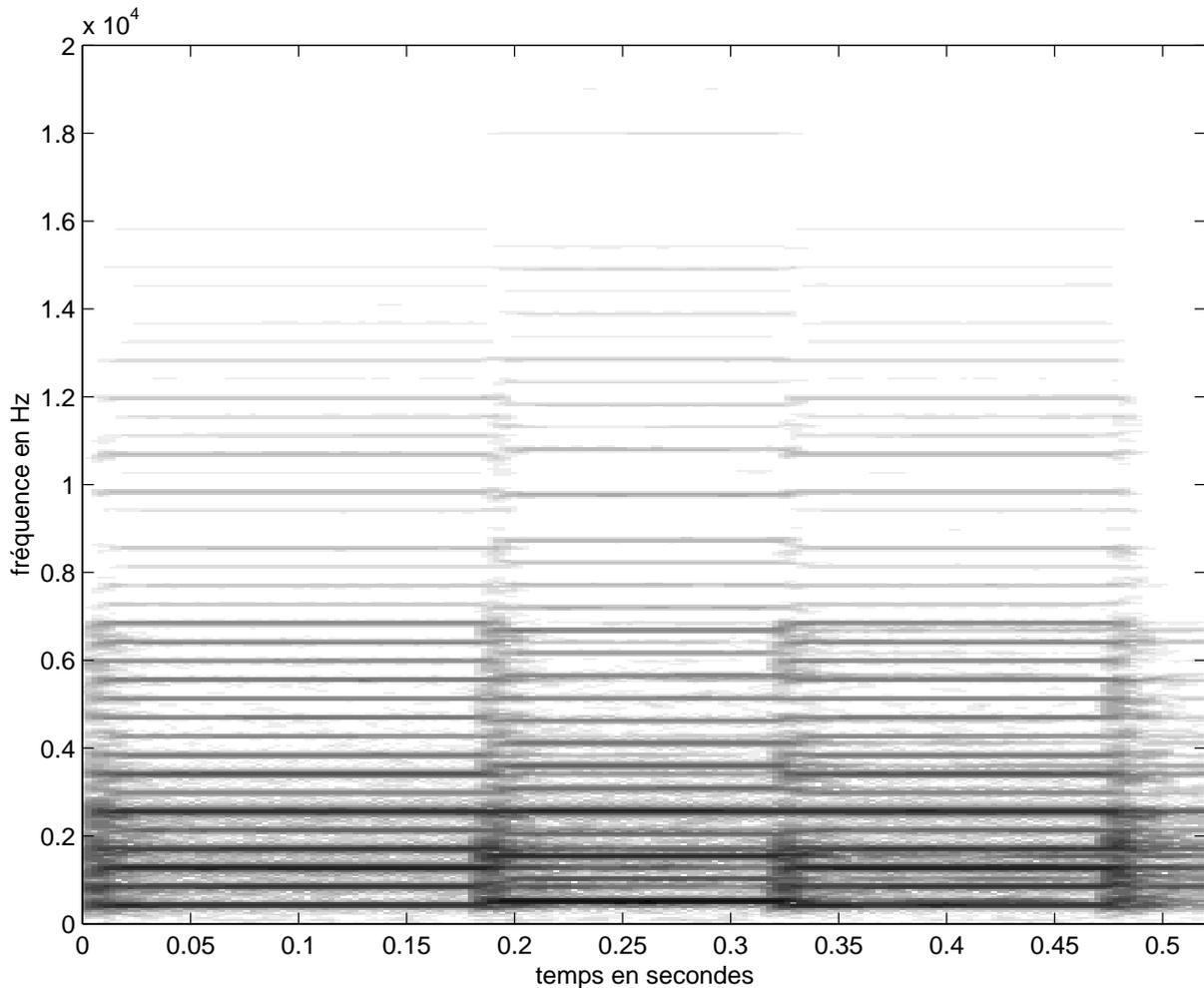


Figure 6-18 : Sonagramme du signal 3  
(Féq = 44100 Hz, 1024 pts, hanning(1024), 900 pts de recouvrement).

Le signal électronique de l'ambulance comporte deux hauteurs de notes proches de celles de la police : le premier fondamental vaut 425 Hz et le second 515 Hz. Si ces hauteurs se trouvent dans la même zone fréquentielle que pour l'avertisseur de la police, leur rapport n'en est pas moins différent puisqu'il constitue une tierce mineure (6/5). Par ailleurs, il est nettement plus pauvre en harmoniques, une forte atténuation de ces derniers apparaissant dès le seizième harmonique pour la première note et le treizième pour la seconde, soit à partir de 7 kHz. Les harmoniques 1 (fondamental), 2, 3, 4, 6 et 8 sont très marqués dans la première note, alors que la seconde note présente avant tout les harmoniques 1, 3 et 5, ce qui lui confère une sonorité particulière qui se rapproche d'un son de clarinette.

La structure temporelle du signal 3 présente une succession de trois notes rapprochées, sous forme de schéma a-b-a, réparties sur une durée inférieure à 0,5 s. Les trois notes se recouvrent sur environ 20 ms, la première dure 190 ms, la seconde et la troisième 150 ms chacune. Ce schéma est ensuite répété selon une période de 2 s.

### 3.2.4 Signal 4 : Ambulance (ancien, mécanique)

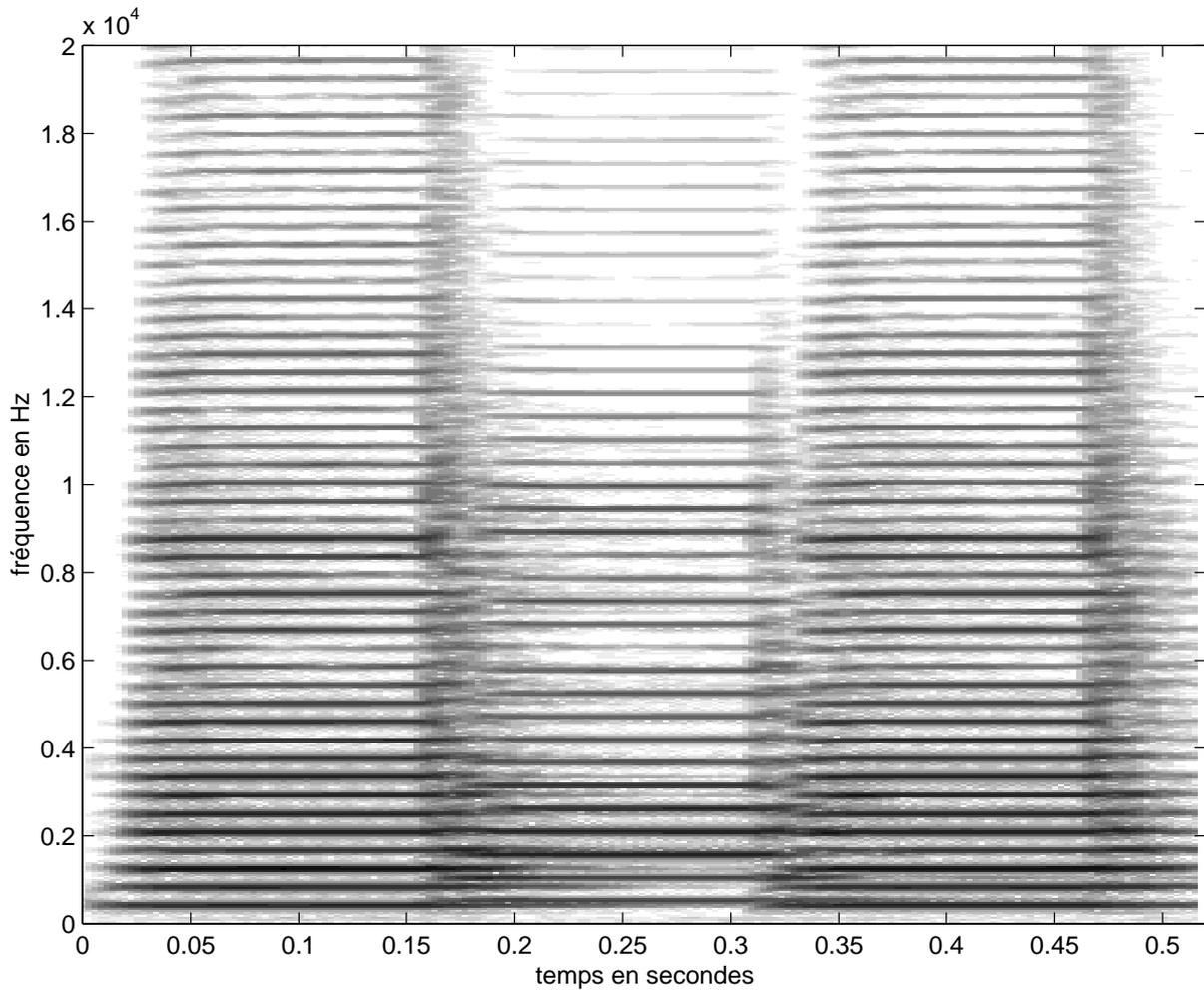


Figure 6-19 : Sonagramme du signal 4  
(Féch = 44100 Hz, 1024 pts, hanning(1024), 900 pts de recouvrement).

Si la structure temporelle de l'avertisseur mécanique de l'ambulance est proche de celle du signal 3 (schéma a-b-a de 0,5 s, répété toutes les 1,85 s), il en va différemment de la structure spectrale. En effet, les fréquences fondamentales mesurées valent 420 Hz et 525 Hz, ce qui conduit à un rapport de tierce majeure ( $5/4$ ), supérieur à celui du signal 3. Par ailleurs, les notes sont beaucoup plus riches et ce, jusqu'à la limite d'audibilité. Les deux notes présentent en outre un renforcement des harmoniques autour de 9 kHz. Le chevauchement des notes est, comme pour le signal 2, moins marqué qu'avec les signaux 3 et 1, notamment du fait des retards d'apparition des harmoniques de rang élevé.

## 3.2.5 Signal 5 : Pompiers (mécanique)

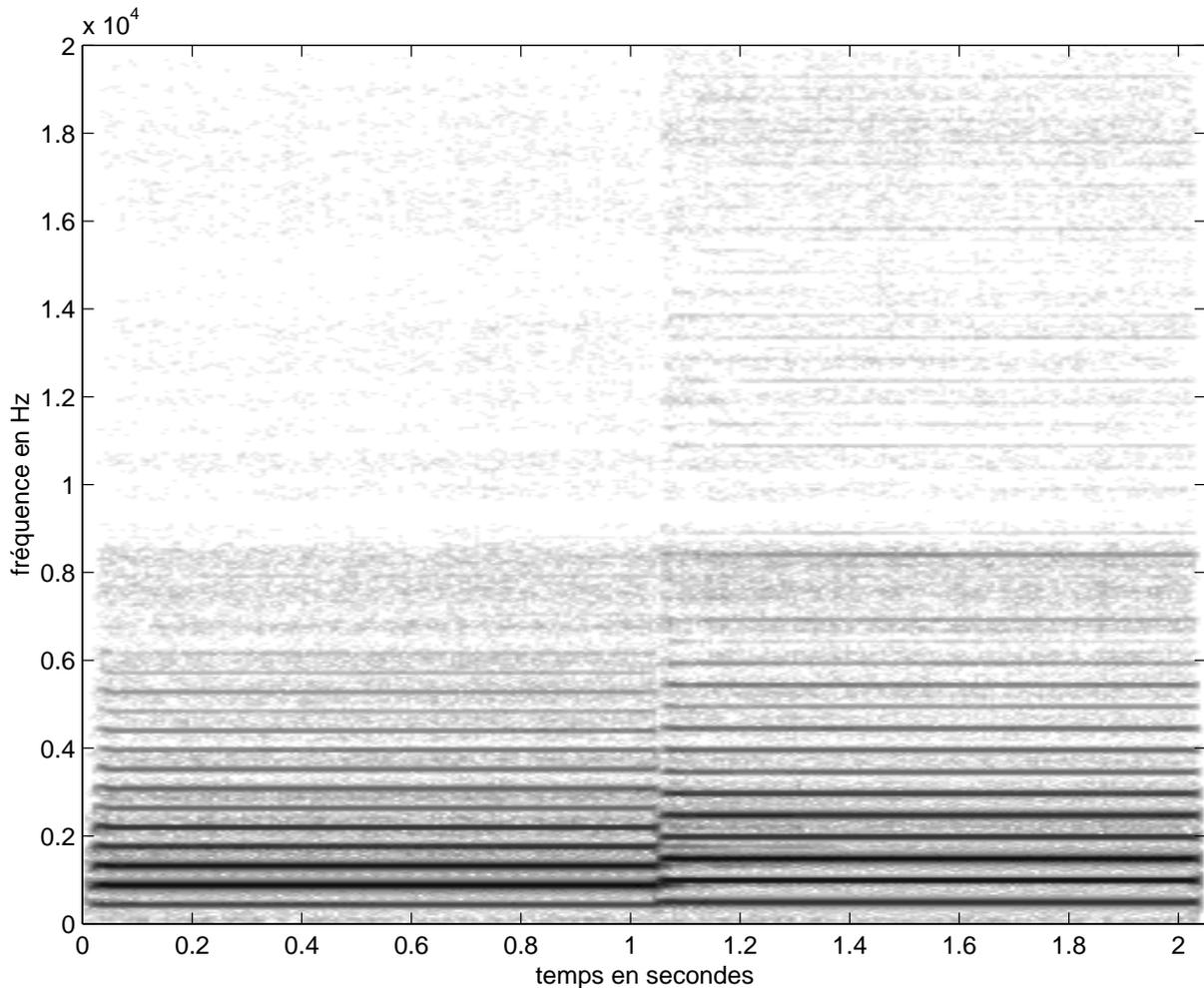


Figure 6-20 : Sonagramme du signal 5  
(Féq = 44100 Hz, 1024 pts, hanning(1024), 700 pts de recouvrement).

L'avertisseur des pompiers est encore similaire aux quatre premiers signaux du point de vue de la hauteur des fréquences fondamentales. Mais si la première note présente un fondamental de 439 Hz et la seconde de 494 Hz, leur rapport (9/8) correspond cette fois-ci à un ton, soit une seconde majeure. Les six premiers harmoniques contiennent l'essentiel de l'énergie acoustique, avec l'harmonique 2 pour composante la plus forte. Au-dessus de 7 kHz, le signal est fortement atténué.

D'un point de vue temporel, l'avertisseur des pompiers a une période de 2 s, chaque note durant une seconde. Les répétitions se succèdent donc sans interruption, mais nos résultats montrent qu'une seule période suffit pour identifier le signal.

### 3.2.6 Signal 6 : Lance-eau (modulation lente)

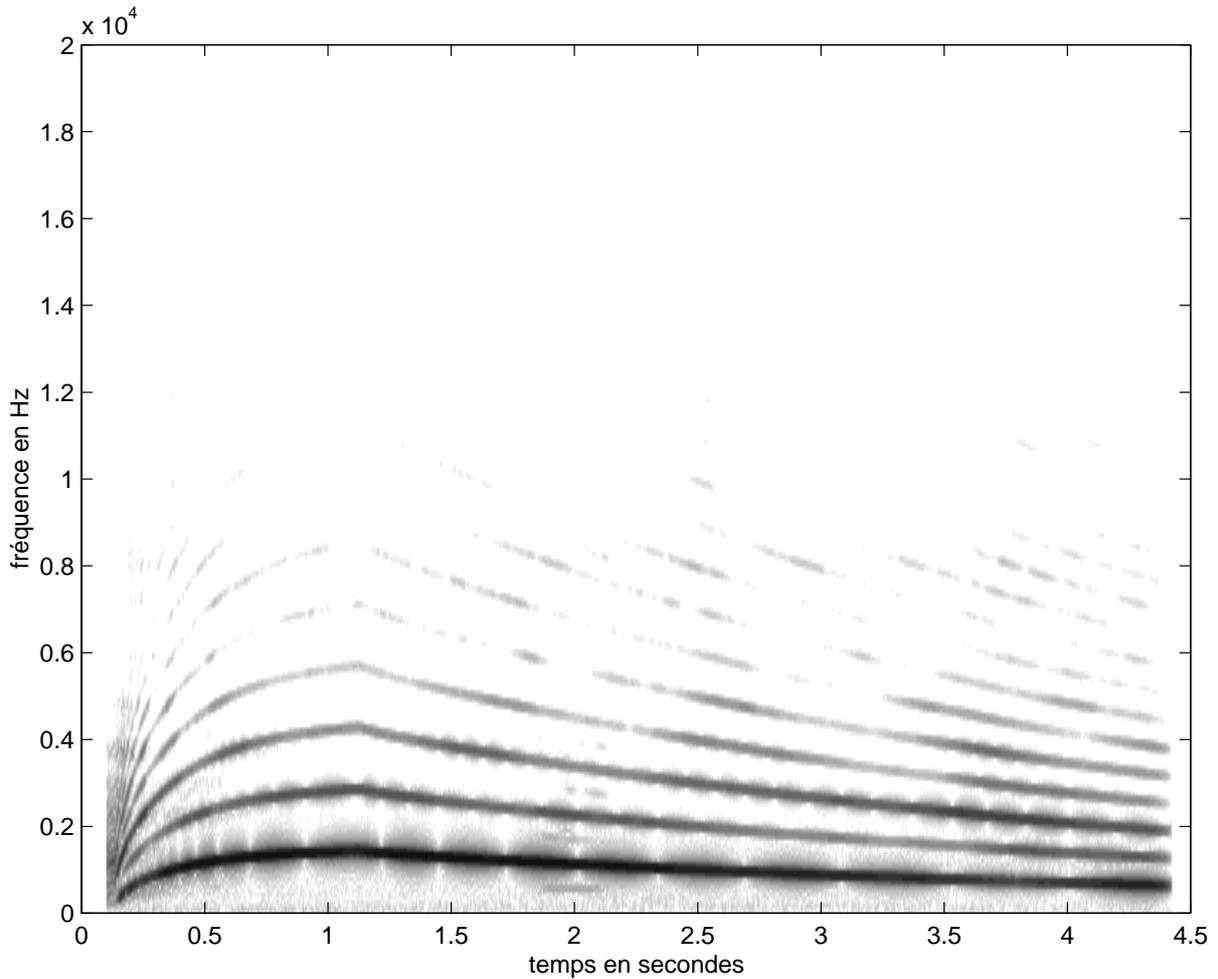


Figure 6-21 : Sonagramme du signal 6  
(Féché = 44100 Hz, 512 pts, hanning(512), 256 pts de recouvrement).

La sirène du lance-eau est très peu connue en France ; elle est d'ailleurs généralement associée à une "sirène américaine". Il s'agit en fait du seul signal d'avertissement - avec le signal 7 qui lui est associé - dont la hauteur est modulée. Ainsi, la fréquence fondamentale croît durant 1 s de 300 à 1425 Hz, pour décroître plus lentement ensuite, - durant 3,2 s - de 1425 à 630 Hz. Le fondamental est très puissant et les fréquences supérieures à 6 kHz sont fortement atténuées.

### 3.2.7 Signal 7 : Lance-eau (modulation rapide)

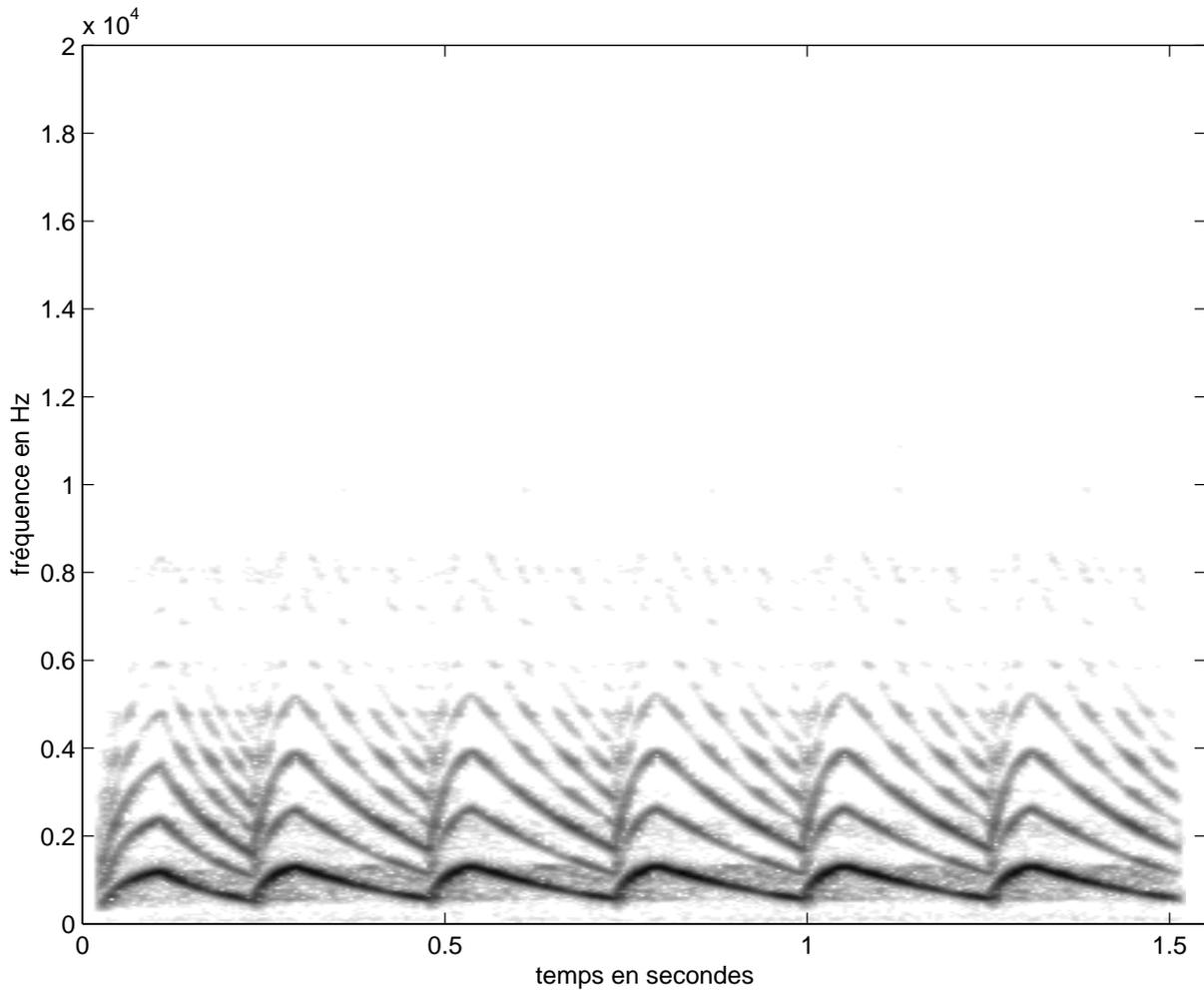


Figure 6-22 : Sonogramme du signal 7  
(Féq = 44100 Hz, 1024 pts, hanning(1024), 800 pts de recouvrement).

La version rapide du lance-eau comporte les mêmes caractéristiques spectrales. En revanche, l'évolution temporelle diffère : la période de la modulation est de 0,25 s et la gamme de variation fréquentielle est de 350 à 1325 Hz, avec un temps de croissance de 60 ms pour 190 ms de décroissance.

### 3.2.8 Signal 8 : Klaxon de voiture 1 (un coup long)

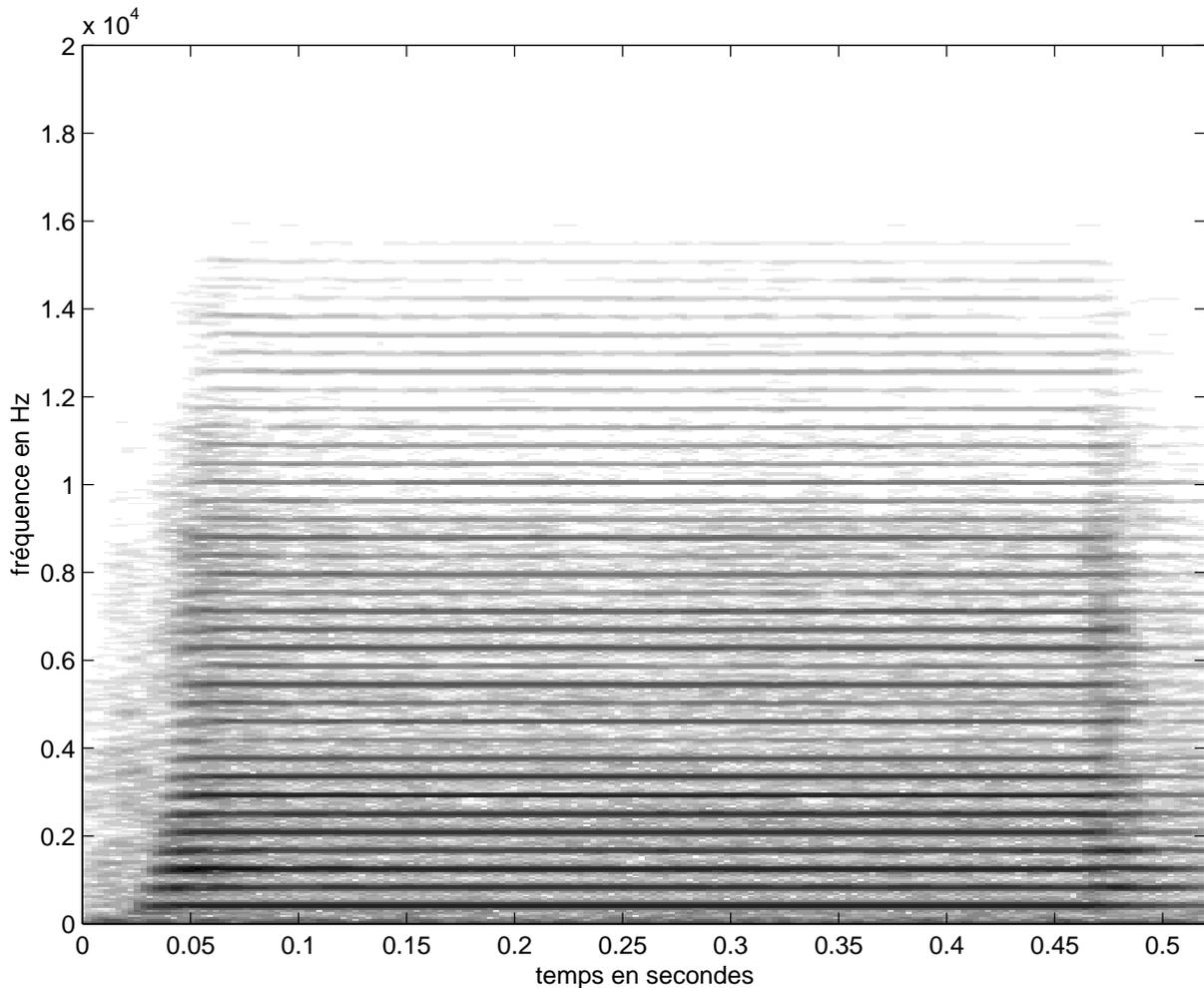


Figure 6-23 : Sonagramme du signal 8  
(Féch = 44100 Hz, 1024 pts, hanning(1024), 900 pts de recouvrement).

Le klaxon de voiture présente un spectre très riche, allant jusqu'à 16 kHz, avec une forte prédominance des neuf premiers harmoniques. La fréquence fondamentale correspond à 418 Hz.

La durée du signal est de 0,45 s. Comme pour les signaux 2 et 4, les transitoires d'attaque des harmoniques de rang élevé sont retardés d'environ 20 ms par rapport à l'attaque du fondamental. Par ailleurs, le temps d'extinction du klaxon est plus long pour les premiers harmoniques. Enfin, l'attaque et l'extinction du klaxon sont accompagnés d'un bruit large bande sur une dizaine de millisecondes, assez proche de ce que nous avons pu observer pour l'ancien signal mécanique de la police.

Le bruit de fond observable sur le sonagramme correspond au bruit du moteur, au ralenti, de la voiture.

### 3.2.9 Signal 9 : Klaxon de voiture 2 (trois coups brefs)

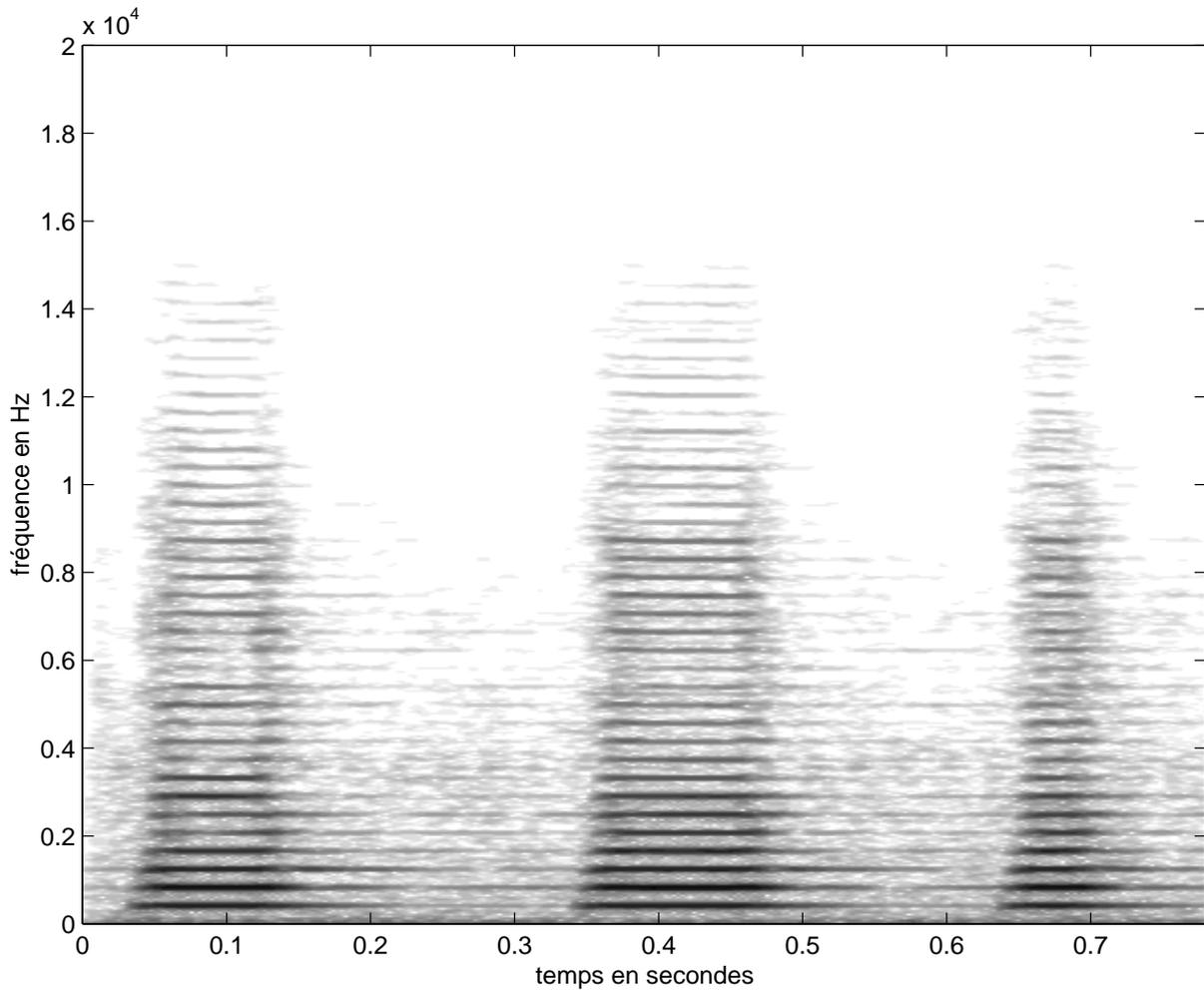


Figure 6-24 : Sonagramme du signal 9  
(Féq = 44100 Hz, 1024 pts, hanning(1024), 900 pts de recouvrement).

Ce signal correspond à la même source émettrice que le signal 8. Et nous retrouvons bien les mêmes caractéristiques spectrales, d'attaque et d'extinction du son. La différence réside simplement dans le fait qu'il ne s'agit plus d'une note continue, mais de la répétition rapide de la note. Les trois coups de klaxon "s'organisent" de la manière suivante : un son de 100 ms, suivi de 200 ms de pause, durant laquelle subsiste l'extinction de la note, puis à nouveau un coup de klaxon de 130 ms, suivi de 160 ms de "silence", et enfin, un dernier coup plus court, de 80 ms.

### 3.2.10 Signal 10 : Sonnette de vélo classique ("dring dring")

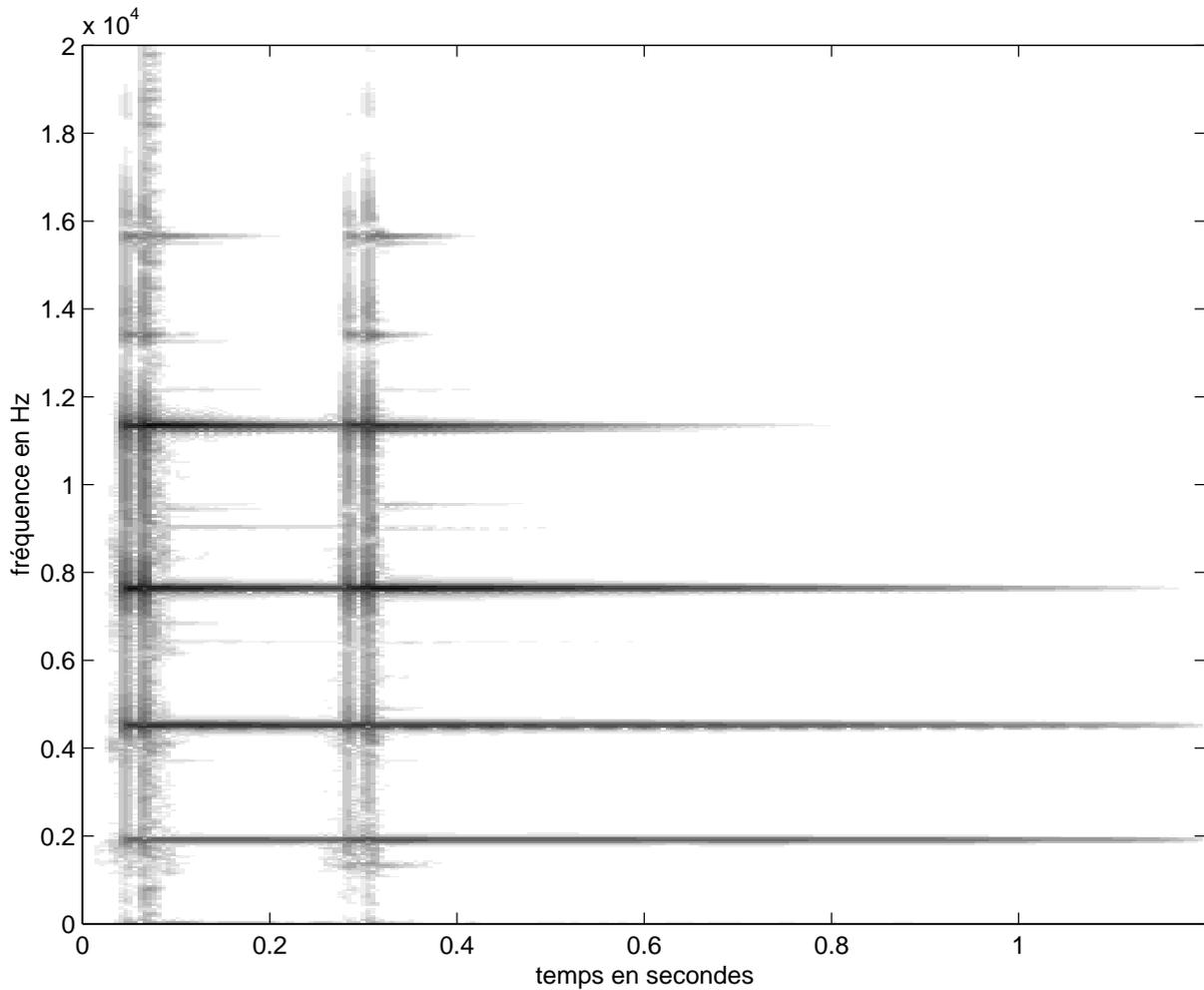


Figure 6-25 : Sonagramme du signal 10  
(Féch = 44100 Hz, 1024 pts, hanning(1024), 800 pts de recouvrement).

Les sonnettes 10 et 11 constituent les deux seuls signaux non entretenus, car frappés, et qui ne présentent par conséquent pas de relation harmonique entre les partiels du son. La sonnette classique consiste en une double percussion, espacée de 21 ms, lors de la poussée du levier, suivie de la même double percussion lors du retour mécanique du levier. Cette répétition est séparée de 287 ms. La double percussion correspond à une fréquence de 48 Hz et peut être à l'origine d'une sensation proche de celle du phénomène de *raucité* (cf. chapitre 1, § 2.3.4)

La sonnette étant percutée au même niveau à chaque percussion, nous retrouvons les mêmes partiels au mêmes intensités. Les partiels les plus importants ont pour fréquences : 1910 Hz, 4510 Hz, 7630 Hz et une double raie à 11130 et 11330 Hz.

### 3.2.11 Signal 11 : Sonnette de vélo à deux tons ("ding dong")

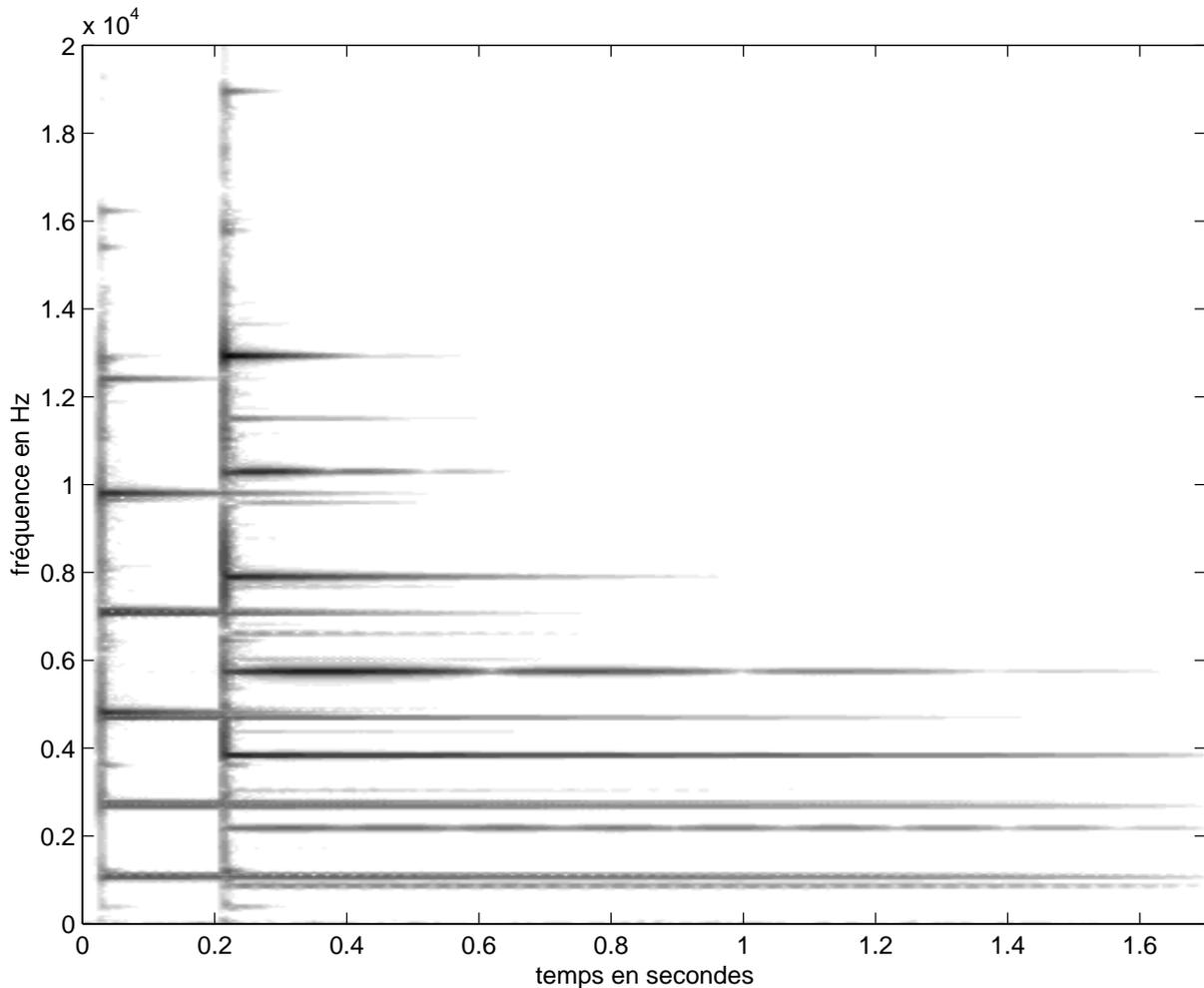


Figure 6-26 : Sonagramme du signal 11  
(Féq = 44100 Hz, 1024 pts, hanning(1024), 800 pts de recouvrement).

La sonnette 11 fonctionne différemment de la précédente. En effet, l'action du levier produit une percussion sur la partie inférieure de la sonnette, tandis que le retour du levier percute la partie supérieure, plus grande et donc plus grave. Dans notre cas, la durée entre les deux percussions vaut 188 ms.

Les partiels associés à la première percussion sont le plus souvent dédoublés : 1065 et 1120 Hz, 2675 et 2765 Hz, 4690 et 4800 Hz, 7060 et 7130 Hz, 9780 Hz et enfin, 12385 Hz. Les principales fréquences associées aux partiels de la seconde percussion sont les suivantes : 845 et 880 Hz, 2180 Hz, 3830 Hz, 5730 Hz, 7890 Hz, 10280 Hz et 12920 Hz.

Pour les deux sonnettes 10 et 11, l'extinction des partiels élevés est plus rapide, notamment au-dessus de 6 kHz.

### 3.2.12 Signal 12 : Poire en caoutchouc ("pouêt")

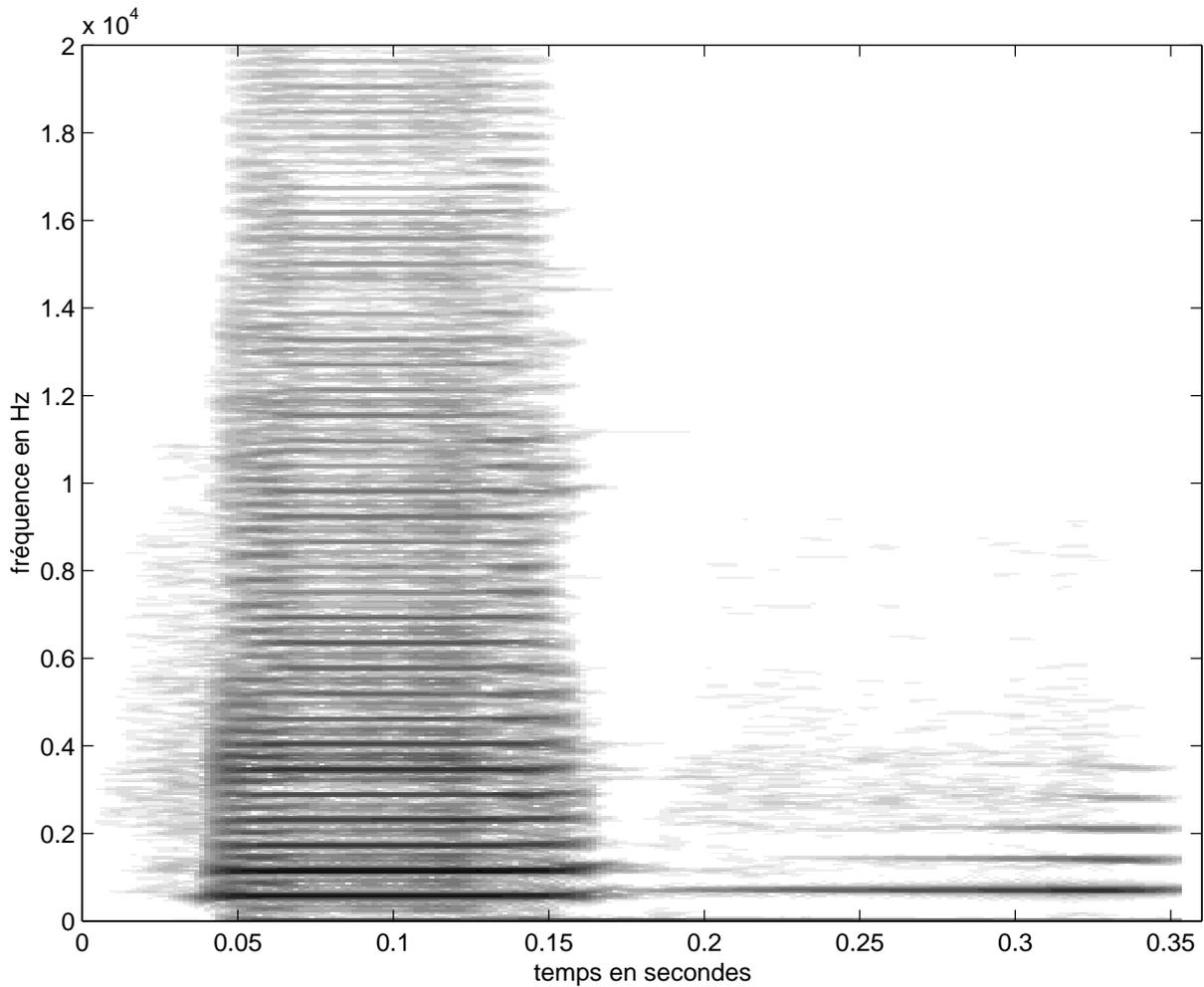


Figure 6-27 : Sonagramme du signal 12  
(Féqh = 44100 Hz, 1024 pts, hanning(1024), 950 pts de recouvrement).

Le spectre du son issu de la poire en caoutchouc ressemble fortement à celui d'un klaxon de voiture. Très riche en harmoniques, il est cependant limité en durée par la quantité d'air disponible dans la poire. Cette durée est de l'ordre de 120 ms. Par ailleurs, le fort débit de l'air expulsé produit un "souffle" qui bruite considérablement le spectre. La fréquence fondamentale est de 578 Hz et l'énergie décroît progressivement avec le rang des harmoniques.

Le relâchement de la poire provoque un appel d'air qui induit un second son, nettement plus faible en intensité et pauvre en harmonique, à une fréquence fondamentale de 700 Hz. Cet intervalle se situe entre une tierce majeure et une tierce mineure.

### 3.2.13 Signal 13 : Sifflet à un ton

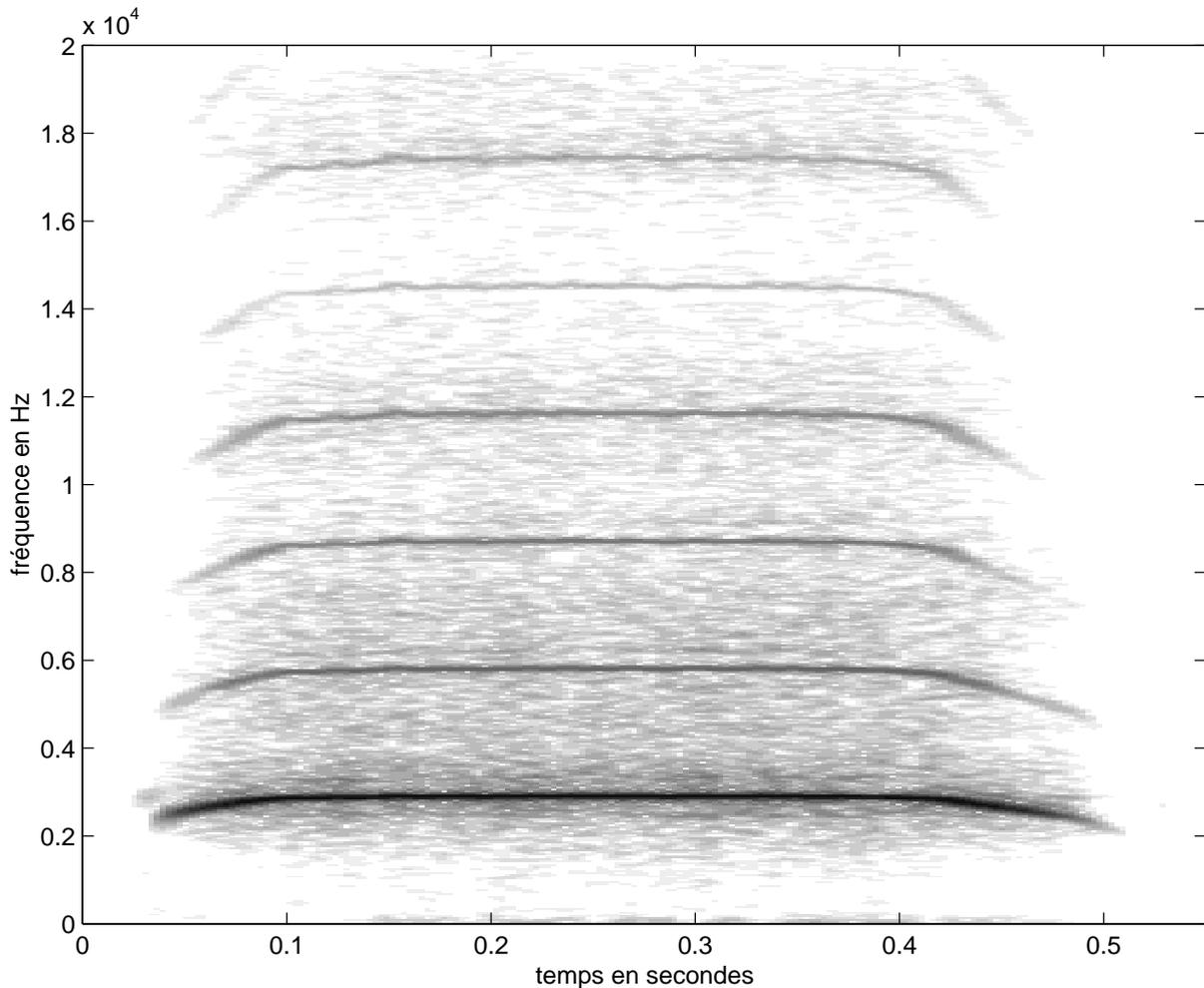


Figure 6-28 : Sonagramme du signal 13  
(Féch = 44100 Hz, 1024 pts, hanning(1024), 800 pts de recouvrement).

Le sifflet à un ton émet une seule note à plusieurs harmoniques et dont la fréquence fondamentale est égale à 2900 Hz en régime stationnaire. Cette dernière est nettement plus intense que les harmoniques qui restent toutefois présents jusqu'au seuil supérieur d'audition. Le début et la fin du coup de sifflet sont marqués par une légère hausse, respectivement une baisse, de la fréquence du signal, entre 2400 et 2900 Hz. Ce son très aigu permet au sifflet d'émerger facilement hors d'un bruit de fond urbain. Il correspond exactement à la description de la "clef forée" par Leipp (1965).

## 3.2.14 Signal 14 : Sifflet à deux tons

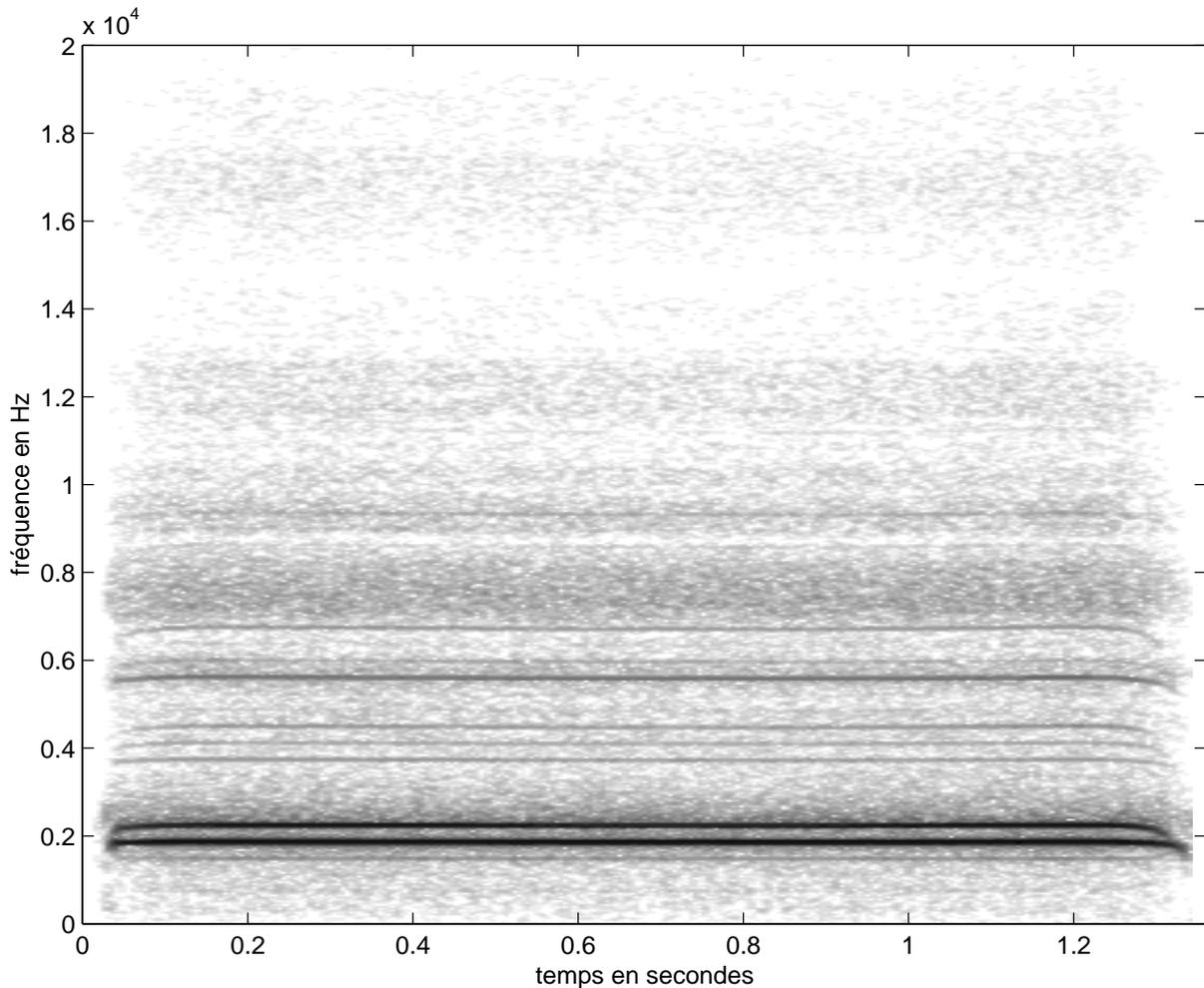


Figure 6-29 : Sonagramme du signal 14  
(Féch = 44100 Hz, 1024 pts, hanning(1024), 800 pts de recouvrement).

Le sifflet à deux tons émet, comme son nom l'indique, deux notes qui sont émises simultanément en principe. Leurs fréquences fondamentales sont beaucoup plus basses que celle du sifflet 13 : 1870 Hz pour la plus grave et 2240 Hz pour la plus aiguë. Le rapport entre ces deux fréquences ( $6/5$ ) correspond à un intervalle de tierce mineure. Outre leurs harmoniques, nous observons également d'autres raies avec le même intervalle de 370 Hz. En fait, elles sont dues à un son différentiel résultant des deux notes du sifflet et qui correspond à un artefact de non-linéarité de l'analyse (par saturation), similaire au comportement non linéaire de la membrane basilaire du système auditif.

Ce sifflet se rapproche du sifflet à deux bourdons étudié par Leipp (1965), à la différence près que l'intervalle entre les deux notes correspond à une tierce majeure ( $5/4$ ) et que les harmoniques du son différentiel sont nettement moins marqués. Pour Leipp, il s'agit d'un sifflet "universel" qui émerge de presque tout bruit de fond. Toutefois, nous ne constatons pas une détection en bruit de fond particulièrement performante : -12 dB en circulation et -13 dB en jardin public. Cela correspond à la moyenne sur les quinze signaux (cf. chapitre 4, § 4.2.3).

### 3.2.15 Signal 15 : Sifflet à roulette

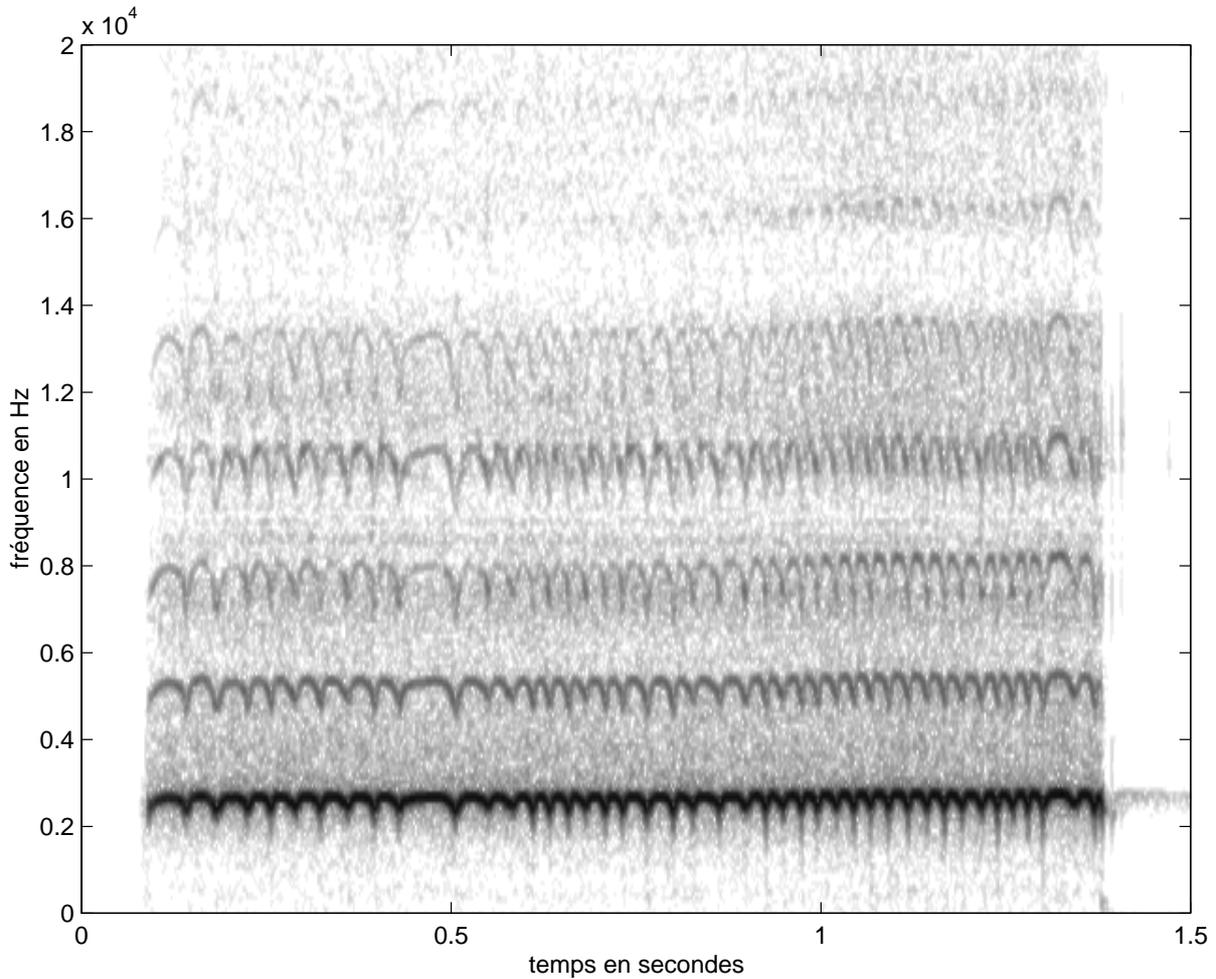


Figure 6-30 : Sonagramme du signal 15  
(Féché = 44100 Hz, 512 pts, hanning(512), 400 pts de recouvrement).

Le sifflet à roulette a comme caractéristique de varier très rapidement autour d'une hauteur donnée. Ces variations sont représentées par une forme en arcs de cercles dont le maximum correspond à une fréquence de 2680 Hz et le minimum à environ 2100 Hz. La période de ces oscillations n'est pas constante, mais varie entre 20 ms et 75 ms ; elle dépend en fait du débit d'air expulsé par la personne qui souffle. Les trois premiers harmoniques sont relativement intenses, ce qui porte ce signal aisément au-dessus de 10 kHz. Ce sifflet correspond exactement à celui décrit par Leipp, qui explique la variation de la fréquence par le passage de la bille devant la fente du sifflet, passage qui modifie ainsi la surface de cette ouverture.

### **3.3 Sémiotique associée aux signaux acoustiques**

Les analyses acoustiques que nous venons de présenter doivent être confrontées aux caractéristiques données par les sujets dans l'expérience de comparaison des signaux "semblables" (chapitres 5, § 5). En effet, une telle démarche permet de vérifier si les descriptions des sujets correspondent aux description physiques des signaux.

#### **3.3.1 Les signaux d'urgence**

Les signaux 1 et 2, faiblement associés à "pompiers" et à "sirène", sont jugés "agressifs" et "stressants" car d'une part, ils ont des variations plus rapides, sont plus aigus et disposent d'intervalles plus grands que le signal 5 et d'autre part, ils ont plus d'attaque et ne sont pas modulés comme les sirènes 6 et 7. Ces dernières ont pourtant un fondamental montant à plus de 1300 Hz, ce qui atténue donc l'importance du critère "aigu" comme propriété liée au caractère "agressif". Les descriptions des sujets sont en grande partie vérifiées, puisque le signal 5 est presque deux fois plus lent que les signaux 1 et 2 et que son intervalle correspond à un ton, tandis que celui de 1 et 2 équivaut à une quarte. Toutefois, la hauteur de la première note est sensiblement la même pour les trois signaux d'avertissement ; le caractère "plus aigu" repose donc avant tout sur la seconde note et rejoint le critère d'intervalle entre les deux notes.

Par ailleurs, le signal 2 paraît plus "poussif" et avec une transition entre les notes moins continue que pour l'avertisseur 1. Les paramètres acoustiques liés à ce critère ne peuvent qu'être contenus dans les différences de transitoires d'attaque des notes. Or, la comparaison des signaux 2 et 4 ne met pas en évidence le même caractère "poussif" pour le second signal, malgré un fonctionnement mécanique similaire. Il semblerait donc que cette sonorité "essoufflée" n'est pas due à l'arrivée retardée des harmoniques de rang élevé, mais plutôt à la zone plus marquée des harmoniques 17 à 34 sur la troisième note de la police.

Enfin, nous avons vu (chapitre 5, § 5.1.3) que le signal 4 est estimé plus "agressif" que l'avertisseur 2, et encore davantage par rapport au signal 3. Ainsi, l'intervalle de tierce majeure de l'ancien signal d'ambulance dérange plus que celui de tierce mineure du nouvel avertisseur. Une autre différence apparaît dans la richesse des spectres, pour lesquels les sujets préfèrent le spectre pauvre en fréquences aiguës du signal 3 et qui lui confère un son plus "rond". L'absence de retard des harmoniques de rang élevé assure un meilleur chevauchement des notes et donc, des transitions moins discontinues. Si l'ancien avertisseur 4 a été estimé plus efficace que le nouveau 3, en ce qui concerne les niveaux seuils de détection (chapitre 4, § 4.2), il s'avère cependant mal identifié au seuil de perception et plus gênant que le récent. Ainsi, un son mesuré moins fort qu'un autre en dBA peut-il être perçu comme plus désagréable.

Les prototypes 3 et 5, associés respectivement aux termes "ambulance" et "pompiers", sont donc en outre estimés moins agressifs que les autres signaux "semblables". De même, le signal 6 est "préféré" à l'avertisseur 1, comme prototype associé aux désignations "police" et "sirène".

#### **3.3.2 Les klaxons**

Les caractéristiques données par les sujets pour les avertisseurs 8 et 9 sont essentiellement de type temporel, liées à la durée et au rythme du signal. Pour leurs interprétations du sens sous-jacent aux deux signaux, les sujets ne se réfèrent plus à la signification qu'eux perçoivent, mais à celle correspondant au comportement du conducteur. Dans ce cas, les sujets répondent clairement en étant extérieurs à l'action qui se déroule. Les possibilités temporelles qui sont

offertes à "l'actionneur" du klaxon priment donc sur le contenu spectral du son en ce qui concerne la signification qui est attribuée au signal sonore. Ainsi, s'établit un code propre à l'aide d'un unique avertisseur et lié à la culture de la collectivité utilisant ce signal sonore.

### 3.3.3 Les sonnettes

Les différences de signification que nous venons de mettre en évidence pour les klaxons ne sont pas valables pour les sonnettes. Cela est dû au fait que le son d'une sonnette forme une unité élémentaire plus riche, composée d'un rythme intrinsèque lié aux deux percussions successives, ainsi que d'une durée d'extinction imposée. Nous retrouvons donc les caractéristiques des signaux d'urgence qui offrent moins de possibilités de variations rythmiques par l'utilisateur. La comparaison entre avertisseurs est généralement effectuée en fonction de la hauteur perçue : la sonnette 11 est jugée plus "douce" car plus grave, mais certainement aussi du fait des discontinuités, temporellement très proches, issues de la double percussion de l'avertisseur 10.

### 3.3.4 Les sifflets

La comparaison des sifflets a conduit à de nombreuses réponses en termes d'identification précise, c'est-à-dire en *noms de noms*, les caractéristiques acoustiques ayant été peu mentionnées. Ainsi, les sujets n'ont plus aussi bien suivi la consigne de comparaison, mais ont plutôt répondu par identification. Ce changement de stratégie peut s'expliquer en relation avec la théorie de la catégorisation prototypique (Rosch, 1978). En effet, si les trois sifflets n'appartiennent pas à une même catégorie subordonnée, mais plutôt au niveau de base ou à un niveau superordonné, la comparaison de leurs caractéristiques significatives devient difficile. Nous avons vu que ce n'est pas le cas des "sonnettes de vélo", qui sont bien catégorisées au niveau subordonné - en *noms de noms*. La stratégie de comparaison n'est donc adoptée que pour des *signaux intracatégoriels au niveau subordonné*.

Néanmoins, le signal 15 est considéré comme plus "strident" et plus "agressif". Or, le signal 13 est celui dont la fréquence fondamentale est la plus élevée. Le caractère "strident" serait donc plutôt lié à l'évolution temporelle régulière et rapide du spectre. Les discontinuités dues au passage de la bille devant la fente du sifflet ont une fréquence allant de 13 à 50 Hz et se rapprochent du phénomène de double percussion de la sonnette 10 et de celui de *raucité* (cf. § 3.2.10 ; Zwicker, 1982).

### 3.3.5 Les signaux "incongrus"

Quatre signaux ont conduit à des interprétations par les sujets différentes des significations attendues pour une utilisation donnée de l'avertisseur. Ainsi, le signal 4 de l'ambulance est-il identifié comme "klaxon" à son niveau seuil de reconnaissance dans un fond sonore urbain. Si ce décalage par rapport à l'utilisation d'urgence qui en est faite disparaît dès +9 dB d'émergence perceptive - voire même moins -, il n'en demeure pas moins que l'efficacité de ce signal en est affectée : la distance maximale nécessaire entre un individu et le véhicule se signalant est en effet réduite pour une identification correcte du danger et donc du choix du comportement à adopter. En outre, le caractère "agressif" de l'avertisseur 4 rend ce dernier moins acceptable pour les sujets que le signal 3.

Par ailleurs, la sonnette 11 pose le problème de son caractère atypique, ce qui conduit à une identification parfois ambiguë, même si la présence du contexte urbain permet de lever

partiellement cette ambiguïté. Pourtant, sa sonorité est préférée à la sonnette 10, "classique" dans la culture française. Ainsi, ce signal est suffisamment signifiant pour ne pas être mal interprété et ne pas conduire à d'éventuelles situations périlleuses, mais nécessite cependant encore un apprentissage de la part des citadins appartenant au même groupe culturel que les sujets des expériences - nous retrouvons l'importance de la représentativité des sujets.

Enfin, le sifflet 14 est parfois encore identifié comme un "klaxon" à +18 dB d'émergence perceptive. Cette identification non résistante à la contextualisation du signal le rend inacceptable pour des situations d'urgence. Si Leipp (1965) considère ce sifflet comme "universel" et d'une "efficacité extraordinaire" pour être perçu dans les bruits de fond les plus divers, son identification est particulièrement ambiguë et nécessite par conséquent une utilisation pour laquelle il existe une attente de la part d'un groupe restreint de personnes - activités sportives, scoutisme, etc.

### 3.4 Propriétés acoustiques et signification des signaux

Les descriptions des sujets ont permis de rapprocher les critères de préférences et de jugements qualitatifs de divers paramètres acoustiques décrivant les signaux d'avertissement. Nous allons présenter ces différentes propriétés en fonction du sens attribué par les sujets et indépendamment du contexte d'écoute.

#### 3.4.1 Les informations temporelles

Des résultats du paragraphe 3.3, il apparaît clairement que les structures temporelles ont une importance capitale pour l'identification des signaux d'avertissement. En effet, d'une part, la *rapidité des variations* entre notes est directement corrélée à l'aspect "agressif" d'un signal. Ainsi, les lents changements de l'avertisseur des pompiers 5 rend ce dernier moins "stressant". D'autre part, la *modulation de la hauteur* de note est plus appréciée par les sujets que les brusques changements de hauteur.

Un autre aspect temporel concerne la *répétition régulière et rapide de discontinuités* - double percussion de la sonnette, roulement de la bille d'un sifflet... Le rythme qui en découle contribue fortement à caractériser un signal sonore qui émergera sans difficulté d'un fond sonore, puisque sa nature à la fois impulsive - raies verticales sur le sonagramme - et continue - partiels horizontaux - se distingue nettement de la structure de tout fond sonore urbain. Toutefois, cette répétition rapide - périodes inférieures à 70 ms - conduit également aux caractères "strident" et "agressif" du signal.

#### 3.4.2 Les transitoires d'attaque

Le *retard des harmoniques de rang élevé* semble être une des propriétés signifiantes d'un avertisseur mécanique qui différencie ce dernier d'un avertisseur électronique, pour lequel le régime stationnaire est établi dès l'attaque. Ce retard rend les transitions entre notes discontinues, alors que le passage est perçu comme continu pour les signaux électroniques, ce qui leur confère un caractère moins "agressif".

Le caractère "poussif" du signal mécanique de la police est donné par une attaque privilégiant les harmoniques 17 à 34, soit la zone fréquentielle s'étendant de 8 à 14 kHz.

### 3.4.3 Les informations spectrales

Si le caractère "agressif" d'un signal est souvent justifié par la hauteur "aiguë" des notes, la mise en parallèle des descriptions des sujets avec les propriétés acoustiques met plutôt les *intervalles de hauteurs* en exergue plutôt que la hauteur absolue (cf. § 3.3.1 et 3.3.4).

Par ailleurs, un avertisseur est jugé moins "agressif" s'il est *pauvre en harmoniques aigus*, c'est-à-dire si l'énergie est concentrée sur les premiers harmoniques. Le son du signal paraît alors plus "rond". Ce critère va toutefois à l'encontre de la psychoacoustique qui considère que les centres de gravité spectrale - la brillance - du fond sonore et du signal doivent être aussi éloignés que possible l'un de l'autre afin de faciliter la détection (cf. chapitre 1, § 3.2.2). Néanmoins, au-dessus de 2 kHz, les fonds sonores urbains sont fortement atténués, alors qu'un signal "rond" et moins "agressif" comme 3 ou 5 disposent d'harmoniques intenses jusqu'à 7 kHz, ce qui leur permet tout de même d'émerger aisément des contextes urbains.

## **4. Conclusion**

La comparaison des résultats issus des deux expériences a permis d'obtenir des signaux prototypiques par rapport à des critères significatifs pour les sujets. Ces prototypes correspondent aux meilleurs signaux en contexte urbain, du point de vue de leur efficacité pour répondre à une situation de danger imminent. Ils résultent d'un apprentissage par une collectivité d'une culture donnée.

Par ailleurs, la mise en relation des réponses verbales des sujets avec les propriétés acoustiques des signaux a permis de déterminer des caractéristiques significatives des signaux d'avertissement. Ainsi, la rapidité des variations de hauteur de notes, les répétitions rapides de discontinuités et une forte richesse en harmoniques de rang élevé - au-dessus de 7-8 kHz - contribuent à rendre un signal plus "agressif". Au contraire, une modulation lente de hauteur - inférieure à 4 Hz - et de petits intervalles consonants rendent les signaux moins "stressants". Une faible richesse en harmoniques de rang élevé est signe de son "rond", c'est-à-dire plus doux.

La modification de ces paramètres permet d'optimiser l'efficacité des signaux d'avertissement prototypiques et de minimiser la gêne qui peut leur être associée. Ces changements ne doivent pas être trop importants sous peine de perdre la signification du signal et de nécessiter un nouvel apprentissage, long et risqué, par les usagers de la ville.

Pour accéder à une stratégie de comparaison de la part des sujets, il faut toutefois considérer des stimuli appartenant à une même catégorie de niveau subordonné. Des comparaisons intercatégorielles, notamment entre prototypes de différentes catégories, paraissent plus difficiles à obtenir. Une étude spécifique serait nécessaire pour analyser ces changements de conduites.

# Conclusion générale

Quand on voit ce qu'on voit  
et qu'on entend ce qu'on entend,  
on a bien raison de penser ce qu'on pense  
Coluche

## 1. Les résultats

Tout au long de ce mémoire, nous avons insisté sur le fait que l'étude de la perception des sons de l'environnement nécessite la prise en compte de la signification de ces sons. De ce fait, d'une part, le sujet doit être placé au centre de la problématique pour appréhender le sens qu'il attribue aux signaux et d'autre part, ses réponses doivent correspondre à celles qu'il aurait produites en situation réelle. Les travaux sur les signaux sonores en contexte urbain exigent par conséquent l'*intégration du contexte* de situation dans les stimuli testés en laboratoire, ainsi que la *validité écologique du corpus sonore* considéré et des *consignes* données. Or, les approches des études antérieures, souvent physicalistes, ne tiennent pas compte de ces conditions de validité. La première exigence, la prise en compte du contexte, si elle paraît aisée à mettre en oeuvre, rend en fait la mesure des paramètres acoustiques particulièrement délicate, du fait de la complexification des stimuli. La seconde condition, généralisation de la première, suppose quant à elle la conservation du sens contenu dans les stimuli lors du travail d'abstraction du monde réel - l'enregistrement et la restitution - et de la réalisation de scénarios pertinents - le montage. La troisième nécessité enfin, celle des consignes, doit permettre d'orienter les sujets vers une stratégie de réponses qui corresponde à la problématique de recherche. Par ailleurs, les méthodes d'analyse ont parfois conduit certaines recherches à des résultats incomplets du fait des contraintes sous-jacentes aux méthodes. Ainsi, des analyses dimensionnelles présupposent le caractère dimensionnel de la perception des signaux sonores (Dubois, 1996). En revanche, le concept de catégorisation prototypique (Rosch, 1978) et la consigne d'identification libre que nous avons adoptées n'imposent aucune contrainte au sujet et respectent les formes linguistiques produites par le sujet.

La mise en application de cette méthodologie a apporté une quantité inespérée d'informations, mais a cependant nécessité plusieurs expériences préliminaires pour créer des échantillons sonores écologiquement valides. Nous avons ainsi pu élaborer une chaîne complète d'enregistrement et de restitution stéréophonique, munie d'un système de montage audionumérique, qui conserve la signification intrinsèque des signaux en situation réelle. Les résultats issus des tests d'écoute conduisent quant à eux, à plusieurs conclusions au niveau linguistique, sémiotique et acoustique.

Dans un premier temps, les différences de lexiques utilisés par les sujets, en fonction du niveau de contextualisation des signaux, nous apportent des informations utiles à la méthode d'analyse des données verbales. En effet, nous avons distingué deux niveaux de signification : des noms de sources interprétées - "pompiers", "vélo"... - et des propriétés liées au signal sonore - "sifflement", "pouêt"... Une des difficultés consiste d'une part, à savoir dans quelle classe intégrer des désignations comme "klaxon", "sirène" ou "sonnette" (cf. chapitre 4, § 5.2.4) ; d'autre part, les *noms de noms*, considérés comme des sources interprétées, peuvent être composés de deux types de signifiés : des *sources de source*, telles que "sirène de

pompier", "sonnette de vélo", "alarme de voiture", "sifflet d'enfant"... et des *propriétés de source*, tels que "coup de klaxon", "pin pon de police", "sonnerie de vélo"... Par conséquent, il conviendrait *a priori* de distinguer quatre niveaux de signification qui seraient mieux distingués et de ce fait, plus stables :

- les noms de *sources de sources*, désignations les plus précises en termes de signification liée au signal sonore,
- les noms de *sources interprétées*, désignations significantes, à un niveau supérieur à la classe précédente, en termes de classification prototypique (Rosch, 1978 ; chapitre 1, § 4.3),
- les *propriétés physiques de sources*, descriptions en termes physicalistes ou onomatopéiques de sources significantes,
- les *propriétés physiques du son*, descriptions physicalistes ou onomatopéiques du signal sonore, sans signification explicite.

Dans un second temps, les verbalisations des sujets permettent de valider l'importance du contexte d'écoute (cf. chapitre 6, § 2.3) et de déterminer des signaux d'avertissement prototypiques en fonction des contextes d'écoute et de l'émergence perceptive des signaux (cf. chapitre 6, § 2.3, chapitre 5, § 4 et chapitre 4, § 6). Ainsi, hors contexte, certains signaux ont un caractère ambigu, tandis que les contraintes contextuelles influencent l'identification des signaux et leurs seuils de perception et donc, leur efficacité : d'une part, la mise en contexte améliore l'univocité de certains signaux, c'est-à-dire leur prototypicité ; d'autre part, si le type de contexte - circulation ou jardin public - influence les niveaux seuils des signaux et leur identification à ces niveaux seuils, ce n'est plus valable au-dessus du seuil.

Dans un troisième temps, les données verbales peuvent être mises en relation avec des propriétés acoustiques des signaux. Ce rapprochement a été possible grâce notamment à une consigne de comparaison pour un corpus de signaux semblables et aisément identifiables (cf. chapitre 5, § 5). Ainsi, même si la relation entre le signal acoustique et le signe qui lui est associé est arbitraire, l'apprentissage de cette relation par une population d'une culture donnée a pour conséquence que *certaines propriétés acoustiques revêtent un caractère signifiant*, lié à l'usage de la collectivité.

## 2. Les perspectives

Les résultats de nos expériences ont soulevé un certain nombre d'interrogations théoriques. En effet, si la consigne d'identification des signaux ne pose pas de problème aux sujets, celle de comparaison n'est pas adoptée pour certains groupes de signaux et est remplacée par une stratégie d'identification. D'une part, cette constatation renforce l'hypothèse selon laquelle reconnaissance et identification sont les fonctions premières de la perception auditive et que la qualification n'est possible que consécutivement à la reconnaissance. Nos expériences ne donnent pas suffisamment d'éléments de réponse pour vérifier cette hypothèse, pour laquelle plusieurs tests de comparaison inter- et intra-catégorielles sont nécessaires. D'autre part, ces différences de stratégies permettent de préciser les niveaux hiérarchiques des signaux et leurs connexions en termes de catégorisation prototypique.

Par ailleurs, les propriétés acoustiques significantes mises en évidence à l'issue de nos expériences n'ont pas pu être validées. Pour ce faire, il s'agit de constituer un nouveau corpus sonore constitué à la fois de signaux d'avertissement existants et d'autres modifiés à partir des propriétés significantes obtenues, puis de tester l'identification, la prototypicité, et les jugements de valeurs donnés par les sujets. Les modifications des paramètres acoustiques significants

doivent donc permettre de créer de nouveaux signaux d'avertissement qui soient à la fois plus efficaces et moins gênants pour une population donnée. Cette possibilité est d'un intérêt considérable pour le design sonore des signaux d'avertissement. Elle sous-entend cependant la création de nouveaux avertisseurs par l'intermédiaire de faibles modifications sur les prototypes existants. En effet, la mise en circulation de nouveaux signaux d'avertissement exige des temps d'apprentissage d'autant plus longs que les signaux sont différents de ceux déjà connus, puisque trop de différences conduisent à la perte des propriétés acoustiques signifiantes des autres avertisseurs.

Enfin, les expériences menées sur la validité écologique des échantillons sonores ont été développées jusqu'à obtention d'une chaîne complète d'enregistrement et de restitution sonore respectant le critère de validité écologique. Les contraintes temporelles et les restrictions matérielles n'ont pas permis d'étendre ce concept à d'autres systèmes de prise de son tels que la tête artificielle ou la configuration de prise de son compacte M-S, ou encore à d'autres systèmes de reproduction multicanaux tels que les formats Surround ou Ambisonic. Ces derniers ont comme avantage de permettre l'étude de la localisation - autre fonction première de la perception auditive - des signaux d'avertissement dans tout le plan horizontal des sujets, y compris le demi-plan arrière. Par ailleurs, l'ajout d'un support visuel devrait permettre d'étudier des *scénarios sonores* plus complexes, tels que des déplacements à travers des zones urbaines, afin d'aboutir à des outils de simulation d'ambiances sonores comme aides à la décision pour les aménageurs urbains.



# Références bibliographiques

- Amphoux, P., (1991), *Aux écoutes de la ville. La qualité sonore des espaces publics européens - Méthode d'analyse comparative - Enquête sur trois villes suisses*, Rapport de recherche, Institut de Recherche sur l'Environnement Construit. Département d'Architecture. Ecole Polytechnique Fédérale de Lausanne., n°94, Lausanne.
- Aures, W., (1985b), Berechnungsverfahren für den sensorischen Wohlklang beliebiger Schallsignale, *Acustica*, **59** (2), 130-141.
- Aures, W., (1985a), Der sensorische Wohlklang als Funktion psychoakustischer Empfindungsgrößen, *Acustica*, **58** (5), 282-290.
- Ballas, J. A., (1993), Interpreting the Language of Informational Sound, *Journal of the Washington Academy of Sciences*, **83** (3), 143-160.
- Bataille, J.-P., (1992) *Cloches, Sonneries et Alarmes*, collection Sonothèque, disque Auvidis, n° A6183.
- Bermond, R., (1996), *Détermination et traitement des modes propres dans une salle de faible volume*, Université de Paris 6, Paris, (mémoire de Maîtrise).
- Binnington, J. D., McCombe, A. W., Harris, M., (1993), Warning signal detection and the acoustic environment of the motorcyclist, *British Journal of Audiology*, **27**, 415-422.
- Bismark, G. v., (1974b), Sharpness as an attribute of the Timbre of Steady Sounds, *Acustica*, **30**, 159-172.
- Bismark, G. v., (1974a), Timbre of Steady Sounds: A Factorial Investigation of its Verbal Attributes, *Acustica*, **30**, 146-159.
- Bonnet, C., (1986), *Manuel pratique de psychophysique*, Collection U - Série Psychologie, Armand Colin Editeur, Paris.
- Botte, M.-C., et al., (1989), *Psychoacoustique et perception auditive*, Série Audition, Editions INSERM / SFA / CNET - Editions Médicales Internationales, Paris.
- Botte, M.-C., Chocholle, R., (1960), *Le bruit, Que sais-je ?*, vol. n°855, 4ème édition refondue de 1984, Presses Universitaires de France, Paris.
- Canévet, G., (1994), *Eléments de psychoacoustique*, Cours du D.E.A. de Mécanique de l'Université Aix-Marseille II, Laboratoire de Mécanique et d'Acoustique, Marseille.
- Castellengo, M., (1994), La perception auditive des sons musicaux, in *Psychologie de la musique*, Zenatti, A. (dir.), Paris, Presses Universitaires de France, p. 55-86.
- Cohen, G., (1999), *Etude de paramètres physiques et cognitifs influant sur le temps de reconnaissance et le temps d'identification des bruits de l'environnement*, Université de Paris 6, Paris, (rapport de D.E.A.).

- David, S., (1997), Représentations sensorielles et marques de la personne : contrastes entre olfaction et audition, in *Catégorisation et cognition : de la perception au discours*, Dubois, D. (dir.), Paris, Editions KIME, p. 211-242.
- David, S., et al., (1997), L'expression des odeurs en français : analyse morpho-syntaxique et représentation cognitive, *Intellectica*.
- Delorme, A., (1982), *Psychologie de la perception*, Editions Etudes Vivantes, Montréal.
- Dubois, D., (1993a), Comment l'homme communique-t-il ?, in *L'homme cognitif*, 4ème édition de 1998, Paris, Presses Universitaires de France, p. 199-301.
- Dubois, D., (1996), Matériel et consignes : un type de questionnement social dans la recherche expérimentale en psycholinguistique, in J. Richard-Zappella : Le questionnement social, *Les cahiers de Linguistique Sociale*, n° special **89-98**.
- Dubois, D., (1993b), *Sémantique et cognition - Catégories, prototypes, typicalité*, Sciences du langage, CNRS Editions, Paris.
- Dubois, D., Resche-Rigon, P., Tenin, A., (1997), Des couleurs et des formes : catégories perceptives ou constructions cognitives, in *Catégorisation et cognition : de la perception au discours*, Dubois, D. (dir.), Paris, Editions KIME, p. 17-40.
- Ebata, M., Baba, H., (1997), *Detection of the ambulance siren while driving*, Inter-Noise 97, Budapest, Hongrie, 2, p. 1009-1014.
- Edworthy, J., Loxley, S., Dennis, I., (1991), Improving Auditory Warning Design: Relationship between Warning Sound Parameters and Perceived Urgency, *Human Factors*, **33** (2), 205-231.
- Edworthy, J., Stanton, N., (1995), A user-centered approach to the design and evaluation of auditory warning signals: 1. Methodology, *Ergonomics*, **38** (11), 2262-2280.
- Eldred, K. M., Sharp, B. H., (1972), *Are present horns, whistles and sirens necessary for communications ?*, Wyle Laboratories, n° 720640.
- Gaver, W. W., (1993), What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception, *Ecological Psychology*, **5** (1), 1-29.
- Gibson, J. J., (1979), *The Ecological Approach to Visual Perception*, réédition de 1986, Lawrence Erlbaum Associates, Hillsdale, New Jersey.
- Gibson, J. J., (1966), *The senses considered as perceptual systems*, Houghton-Mifflin, Boston.
- Greimas, A. J., (1970), *Sémantique structurale - Recherche de méthode*, Formes sémiotiques, édition de 1986, Presses Universitaires de France, Paris.
- Guillaume, P., (1937), *La psychologie de la forme*, Collection Champs, édition de 1979, Flammarion, Paris.
- Guyot, F., (1996), *Etude de la perception sonore en termes de reconnaissance et d'appréciation qualitative : une approche par la catégorisation.*, Université du Maine, Le Mans, 269 p. (Thèse de doctorat).
- Haas, E. C., Casali, J. G., (1995), Perceived urgency of and response time to multi-tone and frequency-modulated warning signals in broadband noise, *Ergonomics*, **38** (11), 2313-2326.

- Harnad, S., (1987), Preface, in *Categorical perception - The groundwork of cognition*, édition de 1990, Harnad, S. (dir.), Cambridge, Cambridge University Press, p. 599.
- Hassall, J. R., Zaveri, K., (1988), *Acoustic Noise Measurements*, 5ème édition, Brüel & Kjaer, Naerum (DK).
- Hugonnet, C., Walder, P., (1995), *Théorie et pratique de la prise de son stéréophonique*, première, Editions Eyrolles, Paris.
- Jakobson, R., (1942), *Six leçons sur le son et le sens*, Arguments, édition de 1976, Les Editions de Minuit.
- Kryter, K. D., (1970), *The Effects of Noise on Man*, Academic Press, New York.
- Lass, N. J., et al., (1982), Listeners' Identification of Environmental Sounds, *Perceptual and Motor Skills*, **55**, 75-78.
- Lass, N. J., et al., (1983), Listeners' Identification of Human-Imitated Animal Sounds, *Perceptual and Motor Skills*, **57**, 995-998.
- Lax, R., (1993), *Journal D*, Pendo-Verlag, Zurich.
- Leipp, E., (1971), *Acoustique et Musique*, 4ème édition de 1984, Masson, Paris.
- Leipp, E., (1965), Le problème de la perception des signaux acoustiques par effet de contraste. Les signaux d'avertissement, les sifflets., *Annales des télécommunications*, **20** (5-6), 103-118.
- Maffiolo, V., (1999), *Contribution à la caractérisation sémantique et acoustique de la qualité de l'environnement sonore urbain*, Université du Maine, Le Mans, (Thèse de doctorat).
- Maffiolo, V., Dubois, D., Castellengo, M., (1998), Qualité sonore de l'environnement urbain : sémantique et intensité, *Acoustique & Techniques*, **16**, 14-21.
- McAdams, S., (1994), La reconnaissance de sources et d'événements sonores, in *Penser les sons*, McAdams, S., Bigand, E. (dir.), Paris, Presses Universitaires de France, p. 157-213.
- Møller, H., et al., (1996), Binaural Technique: Do We Need Individual Recordings?, *J. Audio Eng. Soc.*, **44** (June), 451-469.
- Musil, R., (1930), *Der Mann ohne Eigenschaften*, vol. Tome 1, édition de 1978, Rohwolt, Reinbeck (Hambourg).
- Musil, R., (1930), *L'homme sans qualités*, Le don des langues, vol. Tome 1, 1956, Editions du Seuil, Paris.
- Patterson, R. D., (1982), *Guidelines for Auditory Warning Systems on Civil Aircraft*, Civil Aviation Authority, n°CAA Paper 82017, London.
- Robinson, G. S., Casali, J. G., (1995), Audibility of reverse alarms under hearing protectors for normal and hearing-impaired listeners, *Ergonomics*, **38** (11), 2281-2299.
- Rosch, E., (1978), Principles of Categorization, in *Cognition and Categorization*, Rosch, E., Lloyd, B. B. (dir.), Hillsdale, New Jersey, Lawrence Erlbaum Associates, p. 27-48.
- Rossi, M., (1986), *Traité d'électricité - Electroacoustique*, vol. 21, 1ère, Presses Polytechniques Romandes, Lausanne.
- Saussure, F. d., (1916), *Cours de linguistique générale*, édition critique de 1975, Payot, Paris.

- Schaeffer, P., (1966), *Traité des objets musicaux*, Collection Pierres Vives, Editions du Seuil, Paris.
- Streicher, R., Dooley, W., (1985), Basic Stereo Microphone Perspectives - A Review, *J. Audio Eng. Soc.*, **33** (July/August), 548-556.
- Streri, A., (1993), Comment l'homme perçoit-il le monde ?, in *L'homme cognitif*, 4ème édition de 1998, Paris, Presses Universitaires de France, p. 91-197.
- Szeto, A. Y. J., Valerio, N. C., Novak, R. E., (1991), Audible pedestrian traffic signals: Part 1. Prevalence and impact, *Journal of Rehabilitation Research and Development*, **28** (2), 57-64.
- Szeto, A. Y. J., Valerio, N. C., Novak, R. E., (1991), Audible pedestrian traffic signals: Part 2. Analysis of sounds emitted, *Journal of Rehabilitation Research and Development*, **28** (2), 65-70.
- Szeto, A. Y. J., Valerio, N. C., Novak, R. E., (1991), Audible pedestrian traffic signals: Part 3. Detectability, *Journal of Rehabilitation Research and Development*, **28** (2), 71-78.
- Theile, G., (1980), *Untersuchung zur Wahrnehmung der Richtung und Entfernung von Phantomschallquellen bei 2-Kanal-Stereofonie*, Institut für Rundfunktechnik IRT, n°24/80, Francfort.
- Vanderveer, N. J., (1979), *Ecological Acoustics: Human Perception of Environmental Sounds*, Dissertation, Dissertation Abstract International, n°4543B (University Microfilms n° 8004002).
- Vogel, C., (1995), *Etude psychoacoustique de l'influence des niveaux de bruits et vibrations sur le confort des passagers dans les avions ATR*, Aérospatiale, (rapport de D.E.A.).
- Vogel, C., et al., (1997b), *Validation subjective de la prise de son en extérieur*, 4ème Congrès Français d'Acoustique, Marseille, 14-18 avril 1997, vol. 1, p. 307-310.
- Vogel, C., et al., (1997a), *Sound Characterization of Urban environments: an Approach Based on Ecological Validity*, 102ème Convention de l'Audio Engineering Society, Munich, 22-25 mars 1997.
- Weil-Barais, A., (1993), Repères théoriques, in *L'homme cognitif*, 4ème édition de 1998, Paris, Presses Universitaires de France, p. 33-64.
- Zwaga, H., Easterby, R., (1984), Developing effective symbols for public information, in *Information design*, Zwaga, H., Easterby, R. (dir.), Chichester, Wiley.
- Zwicker, E., (1982), *Psychoakustik*, Hochschultext, Springer Verlag, Berlin.
- Zwicker, E., Feldtkeller, R., (1981), *Psychoacoustique - L'oreille, récepteur d'information*, Collection Technique et Scientifique des Télécommunications, Masson / CNET - ENST, Paris.

## Bibliographie complémentaire

- Aragon, L., (1926), *Le paysan de Paris*, Collection Folio, édition renouvelée en 1953, Editions Gallimard, Paris.
- Augoyard, J.-F., Torgue, H., (1995), *A l'écoute de l'environnement : Répertoire des effets sonores*. Marseille, Editions Parenthèses.
- Ballas, J. A., Mullins, T., (1991), Effects of context on the identification of everyday sounds, *Human performance*, **4** (3), 199-219.
- Barbaras, R., (1994), *La perception : Essai sur le sensible*, Collection Optiques. Philosophie, première, Hâtier, Paris.
- Benjamin, W., (1939), Paris, capitale du XIXème siècle. Exposé, in *Ecrits français*, édition de 1991, Paris, Editions Gallimard, p. 289-309.
- Beranek, L., (1996), *Concert and Opera Halls - How They Sound*, Acoustical Society of America, Woodbury, NY.
- Botte, M. C., McAdams, S., Drake, C., (1994), La perception des sons et de la musique, in *Aspects neurobiologiques et cognitifs de la perception*, p. 55-99.
- Boulez, P., (1989), *Le pays fertile - Paul Klee*, nrf, Editions Gallimard, Paris.
- Chemetov, P., (1996), *20000 mots pour la ville*, Flammarion, Paris.
- Dubois, D., Fleury, D., Mazet, C., (1993), Représentations catégorielles : perception et/ou action ? - Contribution à partir d'une analyse des situations routières, in *Représentations pour l'action*, Weill-Fassin, A., et al. (dir.), Paris?, Octares Editions, p. 79-93.
- Dubois, D., Rouby, C., (1997), Une approche de l'olfaction : du linguistique au neuronal, *Intellectica*, **24** (1), 9-20.
- Dubois, D., Sprenger-Charolles, L., (1988), Perception/interprétation du langage écrit : contexte et identification des mots au cours de la lecture, *Intellectica*, **5** (1), 113-146.
- Edworthy, J., Hellier, E., Hards, R., (1995), The semantic associations of acoustic parameters commonly used in the design of auditory information and warning signals, *Ergonomics*, **38** (11), 2341-2361.
- Guyot, F., Castellengo, M., Fabre, B., (1997), Etude de la catégorisation d'un corpus de bruits domestiques, in *Catégorisation et cognition : de la perception au discours*, Dubois, D. (dir.), Paris, Editions KIME, p. 211-242.
- ISO 7731, (1986), Danger signals for work places - Auditory danger signals.
- ISO 8201, (1987), *Acoustics - Audible emergency evacuation signal*.
- Leipp, E., (1968), *Physique, technique et perception du bruit*, V° Colloque du COMAPI, Strasbourg, 16 mai 1968, 11 p.

- Merleau-Ponty, M., (1945), *Phénoménologie de la perception*, Collection Tel, Editions Gallimard, Paris.
- Nathanail, C., (1999), *Influence des informations visuelles sur la perception auditive. Conséquences sur la caractérisation de la qualité acoustique des salles*, Université du Maine, Le Mans, (Thèse de Doctorat).
- Russolo, L., (1975), *L'art des bruits*, Collection Avant-Gardes, L'Age d'Homme.
- Schaeffer, P., (1952), *A la recherche d'une musique concrète*, Collection Pierres Vives, Editions du Seuil, Paris.
- Vogel, C., et al., (1999), *Perception and Meaning of Warning Signals in Two Urban Contexts*, Forum Acusticum 1999, Berlin, 14-19 mars 1999.