

Qualité sonore de l'environnement urbain : Sémantique et intensité

Valérie Maffiolo, Michèle Castellengo,

Laboratoire d'Acoustique Musicale,
Université Paris 6,
CNRS,
Ministère de la culture,
11, rue de Lourmel - 75015 Paris,
tél. : 01 53 95 43 22,
fax : 01 45 77 16 59,
e-mail : maffiolo@ccr.jussieu.fr

Danièle Dubois,

Langages, Cognitions, Pratiques, Ergonomie,
INaLF - CNRS,
1, rue Maurice Arnoux,
92120 Montrouge,
tél. : 01 46 57 12 86,
fax : 01 42 53 75 86,
e-mail : dubois@idf.ext.jussieu.fr



“Make a quieter world” [1]. Tel est l'objectif de ceux qui, de près ou de loin, se sentent concernés par l'amélioration de la qualité de notre environnement sonore. Habituellement, celle-ci s'exprime en décibels, voire en terme de spectre ou d'émergence des signaux. Depuis quelques années cependant, citoyens, aménageurs et ingénieurs prennent de plus en plus conscience qu'une amélioration de la qualité sonore de l'environnement urbain ne peut se réduire à une “lutte contre le bruit”, qui, en intervenant uniquement sur des intensités, “tend à ramener les niveaux sonores vers un bruit de fond monotone de plus en plus élevé” [2]. Dans une politique de réduction de la gêne liée aux nuisances sonores, les techniques qui corrigent quantitativement le niveau sonore semblent ne plus suffire ; de nouveaux concepts sont ainsi introduits tels ceux “d'ambiances” ou de “paysages” sonores qui imposent de prendre en compte les aspects qualitatifs des phénomènes acoustiques. Le bruit n'est donc plus “simplement” et d'abord un phénomène acoustique, il se définit prioritairement comme l'“effet” d'un événement sonore sur des individus, à la fois “récepteurs” et acteurs dans ce monde. Cette démarche impose alors de recentrer les phénomènes acoustiques sur l'analyse des perceptions et interprétations des individus et usagers, en préalable aux mesures physiques correctrices.

Ce changement d'attitude qui se manifeste dans le monde technique n'est pas sans conséquence sur le plan scientifique : il conduit en effet à une redéfinition de l'objet même des études psychoacoustiques et donc de l'articulation des disciplines qui traditionnellement s'y trouvent concernées : acoustique, psychologie, sociologie... Il ne s'agit plus en effet de considérer l'individu comme “instrument de mesure

du bruit”, instrument particulier qui, en raison de sa subjectivité même, introduirait des distorsions au regard des instruments de mesure physique “objectifs” ; il s'agit plus fondamentalement de réexaminer la question de l'objectivation du bruit, conçu dans sa définition même comme un phénomène subjectif, c'est-à-dire relevant du sujet¹. Cela concerne au premier chef la psychologie scientifique et, au sein de celle-ci, les domaines de la perception et de la mémoire, qui sont actuellement l'enjeu de prises de positions différentes au sein de la psychologie cognitive. En effet, il nous semble important d'identifier dans quelle mesure le “cognitivisme” contemporain est héritier de la tradition psychophysique, et en conséquence d'évaluer dans quelle mesure ces recherches psychologiques répondent, tant sur le plan théorique que sur celui des méthodes, aux questions posées par l'amélioration de la qualité sonore des espaces de vie.

L'approche que développe depuis plusieurs années notre groupe, composée d'acousticiens, de psychologues et de linguistes, s'inscrit dans une démarche concernée en premier lieu par les phénomènes de **signification**. Le bruit est conçu comme effet d'un “stimulus” sur un individu, et inclut nécessairement l'interprétation de ce stimulus en vue d'un comportement adapté. Une telle approche s'appuie donc sur des cadres théoriques et méthodologiques qui permettent l'identification des processus cognitifs d'interprétation des événements du monde physique, à travers à la fois les représentations en mémoire des événements antérieurs et les processus de perception des événements présents. À la différence des perspectives psychophysiques classiques ou cognitivistes contemporaines qui insistent sur “le traitement de l'infor-

¹ On lèvera ici l'ambiguïté du terme subjectif qui s'applique souvent tout à la fois à l'objet d'étude “le sujet”, en opposition à l'objet (du monde physique en particulier) et au mode d'appréhension de cet objet (appréhendé de manière subjective, en opposition avec l'objectivité de la démarche scientifique, dont le modèle est la science physique).

mation" par le "système cognitif humain" [3], sur les mécanismes "élémentaires" de perception de stimuli simples, cette perspective accorde une importance particulière, au rôle des connaissances et représentations préalablement mémorisées, à l'expérience, aux attentes d'un sujet dans la perception².

Celle-ci est donc active, sélective dans la prise d'information à partir d'attentes développées sur la base des connaissances et visées de la perception³.

Une première étude exploratoire de la productivité de cette perspective dans le domaine acoustique a déjà été réalisée par F. Guyot dans sa thèse [6] en partie évoquée dans le n° 7 de la présente revue [7].

Nous rapportons ici les résultats d'une autre recherche qui repose sur les mêmes présupposés épistémologiques et méthodologiques, à savoir, **partir de la prise en compte de la situation globale dans laquelle les phénomènes acoustiques font sens pour les individus**, avant de l'identifier et de la réduire à des corrélats descriptibles par la science physique.

Les conséquences méthodologiques et techniques de cette approche interviennent dès la phase d'enregistrement et de sélection de "stimuli" sonores pertinents pour l'expérimentation : il convient en effet sur le plan technique, de substituer au concept de fidélité de l'enregistrement et de la reproduction du signal (concept pertinent dans l'espace des sciences physiques) celui de **validité écologique** de l'enregistrement et de la reproduction du son (sciences cognitives), c'est-à-dire la garantie de la conservation de sa signification pour le sujet dans une situation d'écoute spécifiée, qui minimise voire annule la pertinence de variations éventuellement "observables" par des instruments de mesure physique [8].

Une spécificité des représentations mentales est de faire sens comme ensemble complexe de critères pris comme un tout. Il importe que ce tout soit d'abord identifié, avant d'être réduit à de nombreux paramètres physiques dimensionnels. Dans cette perspective, nous avons fait appel aux théories psychologiques de la **catégorisation**. Elles rendent en effet possible l'identification des catégories mentales pertinentes qui sont construites à partir des effets de l'environnement sonore sur les individus et qui permettent la reconnaissance immédiate et efficace des événements (exemples de catégories de bruits : "le-bruit-d'un-verre-qui-casse", "le-bruit-d'une-porte-qui-claque"⁴).

Une des catégories particulièrement pertinentes pour notre propos ici (de même que pour des aménagements de terrain), distingue et contraste les bruits agréables des bruits désagréables. En particulier pour des bruits désagréables, il a été observé que l'effet du bruit ne peut

pas être interprété à partir du repérage de valeurs sur le seul paramètre d'intensité, mais de corrélats d'autres indices pour différents niveaux d'intensité [10]. La question n'est donc plus de traiter l'intensité comme une simple dimension physique isolable dont on mesure l'effet singulier dans l'espace psychologique, mais d'évaluer comment l'intensité ou plus exactement comment différents niveaux d'intensité interagissent avec les autres caractéristiques (physiques et psychologiques) des bruits, comme par exemple les caractéristiques hédoniques des sons, l'identification de la source ou la signification de l'événement sonore en train de se produire.

Ces diverses composantes du bruit, comme effet sonore sur un individu, ne peuvent en effet être identifiées dans une approche qui, au départ, a à la fois éliminé ces caractéristiques exclusivement psychologiques de la définition du bruit, et réduit le bruit à un "stimulus", c'est-à-dire à une dimension isolée a priori et paramétrée uniquement dans l'univers physique. Cette approche héritière de la psychophysique ne peut identifier que les réponses d'un système (Éventuellement humain) à un "bruit mutilé", c'est-à-dire à un phénomène sonore devenu artefact expérimental ; s'il s'agit d'un "instrument-sujet", ce dernier va tenter avec ses ressources cognitives disponibles, d'assigner à ce phénomène sonore une ressemblance à des objets déjà perçus et identifiés, mais par construction nécessairement différents. Il ne s'agit pas alors d'étudier des processus de reconnaissance d'objets sonores à partir des structures préalables de connaissances sur l'environnement sonore, mais de repérer des capacités humaines de traitement d'objets insolites (ce qui constitue deux objets de recherches différents).

Une illustration de cette approche en est donnée ici. On notera que la démarche méthodologique adoptée dans ce travail inclut nécessairement une première étape inductive, afin précisément d'identifier la "composition globale" des différents bruits pris comme "tout", avant leur reproduction et leur manipulation dans un dispositif expérimental qui permettra de reprendre la démarche hypothético-déductive classique des sciences de la nature.

Construire un échantillon d'ambiances sonores urbaines

Écologie des enregistrements

L'étude expérimentale de la qualité sonore de l'environnement urbain que nous avons entreprise nécessite donc un matériau sonore qui, restitué en laboratoire, c'est-à-dire "décontextualisé", n'en reproduit pas moins les qualités significatives des phénomènes sonores perçus dans l'environnement. Il s'avère aussi primordial que les tests d'écoute

² Une telle approche conduit à reprendre les perspectives élaborées par Gibson [4] dans les années 70 et à les reconsidérer dans le cadre des sciences cognitives contemporaines, comme celui de la "cognition située" [5].

³ Dans le jargon des théories cognitives actuelles, influencé par la modélisation informatique, on contraste ainsi les approches qui mettent l'accent sur les processus "bottom-up", ascendant de l'analyse du signal (en périphérie par les récepteurs) "jusqu'à" l'interprétation par les systèmes centraux, aux approches "top-down" ou "descendantes" qui insistent sur le contrôle des récepteurs par les représentations de "haut-niveau", en particulier les représentations sémantiques.

⁴ On notera dès maintenant l'absence de terme simple et la nécessité de recourir à des formes verbales complexes pour désigner ce qui à l'évidence correspond à des catégories mentales simples, aussi simples que des couleurs par exemple pour lesquelles, de manière contrastée il existe des désignations simples (sur ce point, cf. Dubois, [9], par exemple).

mis en place soient reproductibles d'un sujet à l'autre, afin de pouvoir comparer leurs résultats. Cette exigence n'élimine pas pour autant une certaine artificialité de la situation d'écoute en laboratoire qui à l'évidence "mutile" la perception "normale" de l'individu en situation, où, en particulier, l'auditif est "filtré" ou recomposé par le visuel [11, 12].

Pour satisfaire ces exigences à la fois de pertinence et de scientificité, une grande importance doit être accordée à la manière dont nous allons effectuer l'enregistrement et à la reproduction des sons. Notre souci premier est de conserver non seulement la qualité technique mais en plus le caractère écologiquement valide des échantillons sonores ; la signification de l'environnement, perçu à l'écoute de l'enregistrement, doit être la plus proche possible de celle qu'on lui attribue dans le contexte de la vie quotidienne.

La technique d'enregistrement et de restitution a donc fait l'objet d'une étude spécifique, où nous avons travaillé en particulier le concept de "validité écologique" de la restitution en lieu et place de la notion plus traditionnelle de fidélité de l'enregistrement [8] ⁵.

Des ambiances sonores urbaines ont été enregistrées :

- en stéréophonie afin d'obtenir une image sonore proche de notre perception qui rende compte à la fois de la localisation et de l'effet d'espace,
- avec deux microphones électrostatiques cardioïdes (Schoeps MK6), écartés de 60 centimètres avec un angle d'ouverture entre les deux microphones de 100 degrés et recouverts de bonnettes Rycote complètes,
- reliés à un enregistreur numérique R-DAT portable (HHb Portadat) fixé à 48 kHz d'échantillonnage.

Une mesure de niveau sonore a été réalisée pendant toute la durée des enregistrements à l'aide d'un sonomètre CEL-573. Les enregistrements ont duré approximativement 4 à 5 minutes par lieu.

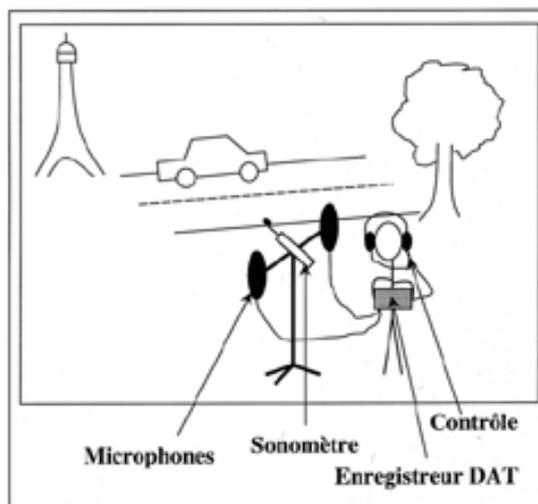


Fig. 1 : procédure d'enregistrement des ambiances sonores urbaines

À l'issue de ces prises de son en extérieur, des séquences sonores ont été isolées, en laboratoire, de chaque ambiance sonore enregistrée. Numérisées sur un PC pentium 200 MHz, elles ont fait l'objet d'un changement de gain, à l'aide du logiciel WaveLab de Steinberg, afin qu'elles soient reproduites aux niveaux sonores qu'elles avaient au moment de l'enregistrement. La reproduction s'est faite en chambre semi-anéchoïque sur deux haut-parleurs (Studer A723) via un lecteur DAT (Tascam DA30-mKII) qui servait de convertisseur numérique-analogique entre la carte son de l'ordinateur et les haut-parleurs. Les auditeurs étaient placés à environ 1,50 m des haut-parleurs.

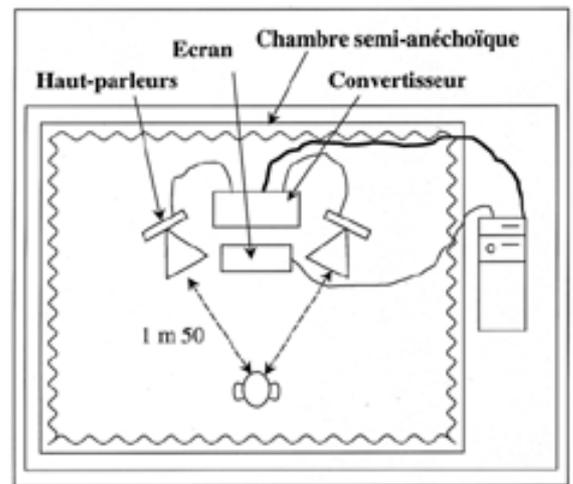


Fig. 2 : Procédure de reproduction du corpus de séquences sonores

De l'ambiance sonore à la scène sonore

La sélection des ambiances sonores urbaines à enregistrer s'est faite à partir d'une liste de lieux obtenus au cours d'entretiens auprès d'usagers de Paris [12]. Les réponses, principalement du type "le bruit des motos à la Bastille le vendredi soir" (description qui intègre un lieu, une activité et un temps), se situent à un niveau de catégorisation pertinent pour les personnes interrogées. Là encore la sélection des lieux à enregistrer et la constitution de l'échantillon lui-même (dans sa composition) constituent une étape non triviale pour la restitution en laboratoire de stimuli et s'inscrit dans l'exigence de validité écologique de l'expérimentation, hors de laquelle les généralisations à partir des résultats obtenus sont trompeuses.

Une fois la bande sonore constituée, notre objectif était d'extraire des séquences sonores de ces enregistrements. Étant donnée l'hétérogénéité d'une ambiance sonore urbaine, même d'une durée de 4 minutes, quelle séquence sonore doit-on extraire pour la donner à écouter ? De quelle durée doit-elle être pour restituer un "objet" sonore unitaire signifiant et reconnaissable comme tel par le sujet ?

À l'écoute de la bande sonore, il s'est avéré que ces ambiances sonores urbaines pouvaient être décomposées

⁵ Cette étude a entre autres montré que les systèmes stéréophoniques utilisés en intérieur pour des sources fixes (type ORTF) ne semblaient pas convenir pour des ambiances de circulation dans lesquelles les sources sont en mouvement.

en plusieurs fragments de l'environnement sonore porteur de sens et d'identité que nous avons désignés comme **"scènes sonores"**. À chaque séquence isolée de l'enregistrement correspond donc une scène sonore. La durée des séquences n'a donc pas été normalisée sur une base temporelle arbitrairement "égale" vis-à-vis de la variable temporelle "abstraite", mais sur la base d'une **unité de temps signifiante en relation avec le "contenu" ou la "nature" des phénomènes acoustiques**, et ce afin de restituer, pour le sujet, une écoute de chaque scène sonore dans leur intégralité ; les scènes sonores ont ainsi des durées qui s'échelonnent entre 15 et 20 secondes.

Du point de vue du contenu, ces scènes sonores sont par hypothèse de deux types. À partir tant des remarques des aménageurs, que de nos hypothèses sur la représentation psychologique du bruit, nous avons distingué celles qui se caractérisent par un ensemble d'événements sonores discriminables, imputables à des sources identifiables que nous pouvons appeler **"événementielles"**, et celles composées d'un bruit de fond continu avec peu voire aucun événement sonore spécifique, que nous pouvons appeler **"amorphes"**. L'hypothèse est de différencier la représentation en mémoire et les traitements psychologiques de

ces deux types de scènes sonores urbaines à savoir :

- si la scène est composée d'événements sonores discriminables, alors elle est en priorité traitée selon ses propriétés sémantiques,
- si la scène est composée d'un bruit de fond, d'où n'émerge aucun événement faisant sens, alors elle est traitée, comme les stimuli isolés et paramétrés des expériences psychophysiques, selon des paramètres physiques comme par exemple l'intensité.

L'objectif est de tester si cette hypothèse de traitement différentiel, en fonction précisément de la possibilité d'affecter des significations, est essentielle et pertinente par rapport aux données expérimentales et aux observations de terrain [2].

Basés sur cette différenciation, nous avons mis en place un corpus expérimental de 16 séquences sonores, 8 séquences événementielles (1, 2, 5, 6, 9, 10, 13, 14) et 8 séquences amorphes (3, 4, 7, 8, 11, 12, 15, 16), la numérotation correspondant à l'ordre d'enregistrement. Cet ordre n'a aucune incidence sur la passation du test puisque toutes les séquences sont présentées en même temps à l'auditeur.



Par la suite, notre travail a porté sur ces 16 séquences d'ambiances sonores urbaines afin de tester si :

- l'intensité est subordonnée à la présence ou à l'absence de sémantique,
- dans le traitement psychologique de scènes événementielles, l'intensité est en conjonction avec d'autres critères ;
- dans le traitement psychologique de scènes amorphes, l'intensité est davantage prise en compte comme critère de traitement.

Des catégories de scènes sonores urbaines

Catégories sémantiques et intensité

La méthode utilisée pour l'étude de la qualité sonore de l'environnement urbain est une méthode de catégorisation libre. Elle a été conçue et validée dans la logique de notre approche qui est de ne pas imposer au sujet de contrainte a priori, qui lui permette d'identifier les paramètres de construction du corpus de sons et donc d'orienter ses activités de traitement comme réponses à ces paramètres (cf. Tenin cité dans Dubois, [13]).

23 sujets, sans déficience auditive, ont participé à un test de catégorisation libre, sans contrainte sur le nombre de catégories ni sur celui des éléments par catégorie. Afin de pallier la difficulté de transfert en laboratoire de la situation d'écoute naturellement contextualisée dans un site urbain, c'est une consigne verbale qui précise aux sujets la nature des bruits qu'ils vont entendre et donc leur suggère le contexte d'écoute dans lequel ils doivent s'imaginer être pour identifier ces bruits. En effet, un tel acte de décontextualisation constitue un artefact par rapport à une situation écologiquement valide, mais nous supposons que la consigne contribue à restituer efficacement un contexte et donc un traitement similaire à celui d'une écoute "située".

La consigne était donc : "Vous allez entendre 16 très courts extraits d'ambiances urbaines. Il vous est demandé de les écouter sans attention particulière en vous imaginant piéton dans la rue. Votre tâche est de grouper ces séquences en catégories selon l'intensité perçue. Vous êtes libre de faire autant de catégories que vous voulez, avec autant d'extraits dans chaque catégorie"⁶.

Dans un second temps, il était demandé aux sujets de qualifier verbalement les catégories qu'ils avaient formées. La durée de l'expérience était en moyenne d'une heure.

Nous n'avons pas choisi un groupe particulier d'individus pour nos tests d'écoute ; nous avons supposé que l'expertise de tout un chacun sur la perception des ambiances de rue était équivalente.

À cette méthode de questionnement est associée une méthode d'analyse statistique des résultats dite d'analyse arborée déjà évaluée dans sa pertinence par rapport aux phénomènes psychologiques de catégorisation [14], qui

propose une représentation des résultats sous forme de catégories correspondant à une moyenne des catégories formées par les sujets. L'analyse de Barthélemy et Guénoche [15] donne une représentation qui permet d'interpréter les distances en termes de similarités et de dissimilarités.

Un arbre se caractérise par des feuilles (1 à 16) qui sont les séquences étudiées ; des nœuds (A à G) qui sont des prototypes virtuels des différentes catégories que le programme a créées, chaque nœud représentant ainsi une entité qui partage avec un groupe d'objets les caractères communs aux membres de ce groupe ; des branches qui sont des segments avec une longueur variable assimilable au degré de typicalité ou ressemblance entre les exemplaires [16].

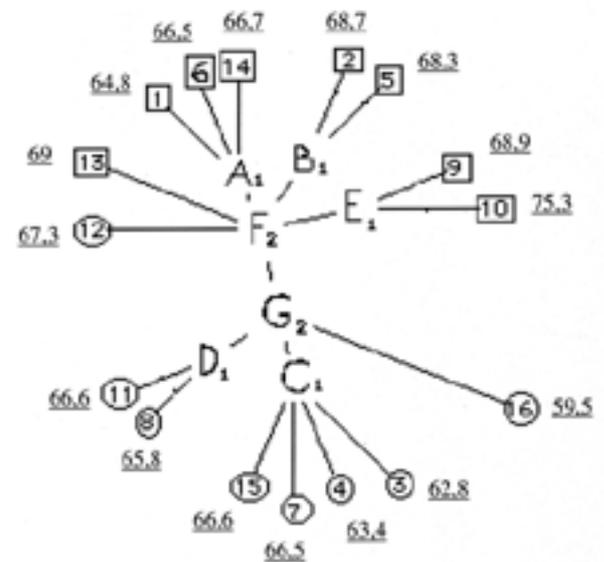


Fig. 3 : Arbre obtenu par l'algorithme d'analyse de Barthélemy et Guénoche. Les nombres dans un carré sont les séquences événementielles, ceux dans un rond les séquences amorphes. Les nombres soulignés sont les niveaux sonores moyens (Leq) en dB (A).

La formation d'un arbre s'opère en une succession d'itérations (l'indice 1 de A₁ représentant la 1ère itération). La pertinence des catégories formées décroît avec le numéro de l'itération.

De par leurs numéros d'itération, les catégories de A à E sont donc les plus pertinentes. Toutefois, l'arbre peut être divisé en deux grandes catégories, l'une regroupée sous le nœud F2 et l'autre sous le nœud G2.

D'abord, lorsqu'on associe chaque séquence à son niveau sonore mesuré au moment de l'enregistrement, on constate que les catégories les plus pertinentes ne regroupent pas des séquences de même niveau sonore moyen (que celui-

⁶ Conscientes du fait que la consigne sur l'intensité oriente les traitements cognitifs vers une dimension physique des phénomènes sonores, mais dont l'isolement n'a pas nécessairement de pertinence d'un point de vue psychologique, cette expérience a été mise en place parallèlement à une autre expérience dont la consigne était de grouper les séquences selon un critère plus "naturel" de traitement à savoir l'agrément. Seuls les résultats du test sur l'intensité sont présentés ici.

ci soit mesuré en dB SPL ou en dB (A)). Pourtant la catégorie G2 regroupe les séquences de plus faible niveau moyen (de 59,5 à 66,6 dB (A)). À l'intérieur de cette catégorie, on peut remarquer que la séquence 16 est isolée des autres, ce qui correspond au fait que son niveau est de loin le plus bas. La catégorie F2 contient, quant à elle, les séquences de niveau moyen plus fort. La question est maintenant de savoir si on peut parler d'intensité moyenne pour des ambiances dont l'essence même est d'être composées de diverses sources sonores ayant des intensités sonores différentes. D'un point de vue psychologique, cela est pertinent dans la mesure où l'intensité est associée à une source. Par exemple, si une moto est présentée à différentes intensités, le regroupement se fera sur la reconnaissance de la (source) moto et non pas sur les différentes intensités perçues.

Ensuite, le groupement des séquences 9 et 10, de par son numéro d'itération, peut être considéré comme robuste et pertinent en tant que catégorie pour les sujets interrogés. Pourtant les séquences sont loin d'avoir le même niveau sonore moyen (68,9 dB (A) pour la séquence 9 et 75,3 dB (A) pour la séquence 10). Cet exemple implique que **des critères autres que l'intensité entrent en jeu dans le processus de catégorisation**.

Le type de séquences groupées sous les nœuds F2 et G2 permet de constater que le nœud F2 regroupe presque uniquement des séquences événementielles, c'est-à-dire des séquences comportant des événements, sources et activités sonores discriminables (excepté la séquence 12 dont le rattachement "tardif" (dernière itération) aux autres catégories a peu de pertinence). Le nœud G2, quant à lui,

regroupe uniquement des séquences amorphes, c'est-à-dire des séquences ne comportant aucun événement sonore spécifique. La figure 4 représente les sonagrammes (représentations temps/fréquence) des séquences 2 et 15, des crissements de freins et un démarrage de moto pour la première, un bruit de circulation continu et monotone pour la deuxième.

Nous avons déjà signalé que les séquences 9 et 10 sont regroupées bien que leurs niveaux moyens diffèrent de 6 dB (A). Il s'agit en fait de deux séquences événementielles contenant des sources sonores similaires (bruit de marteau-piqueur et bus qui passe). Ceci met en évidence l'importance de l'identification des sources sonores et le fait que l'auditeur soit tenté de regrouper des séquences composées de sources équivalentes sur la base sémantique.

Le caractère (événementiel ou amorphe) ainsi que le contenu sonore (identification de sources) des séquences apparaissent avoir une grande influence sur l'évaluation de l'intensité perçue par les sujets. Il semblerait qu'il soit pertinent, voire fondamental, de prendre en compte, dans l'étude de la perception d'ambiances sonores urbaines, le nombre d'événements sonores discriminables ainsi que leur identification.

Données Verbales : des signaux et des sources

Afin d'interpréter plus précisément les catégories formées, une analyse linguistique des verbalisations données par les sujets, étayée dans le cadre théorique de la linguistique et de la psycholinguistique [17, 18], a été menée.

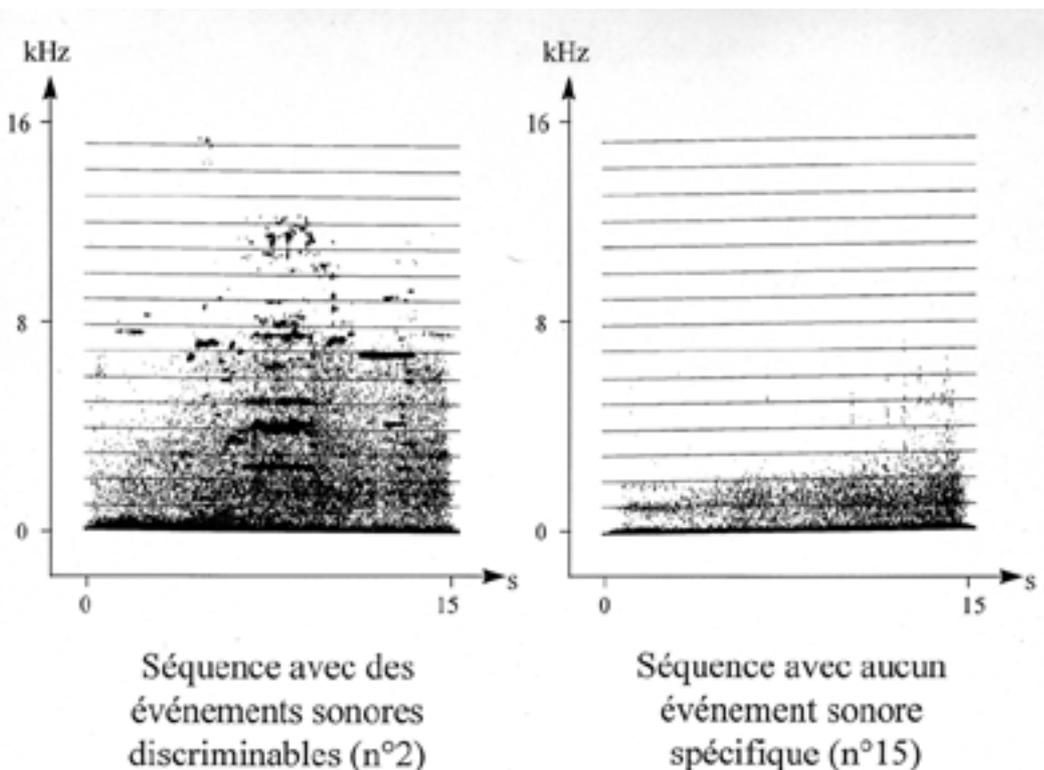


Fig. 4 : Sonagrammes d'une séquence événementielle et d'une séquence amorphe.

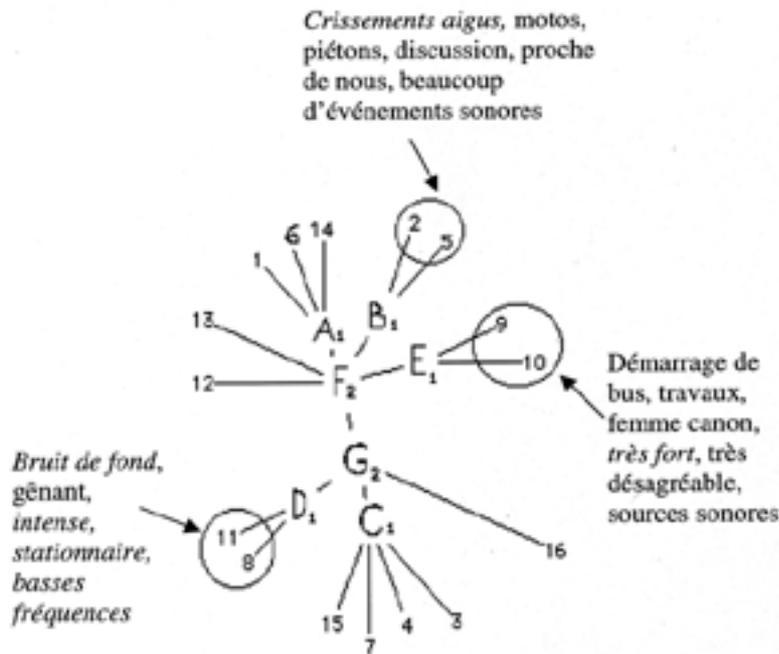


Fig. 5 : Arbre avec quelques verbalisations données par les sujets

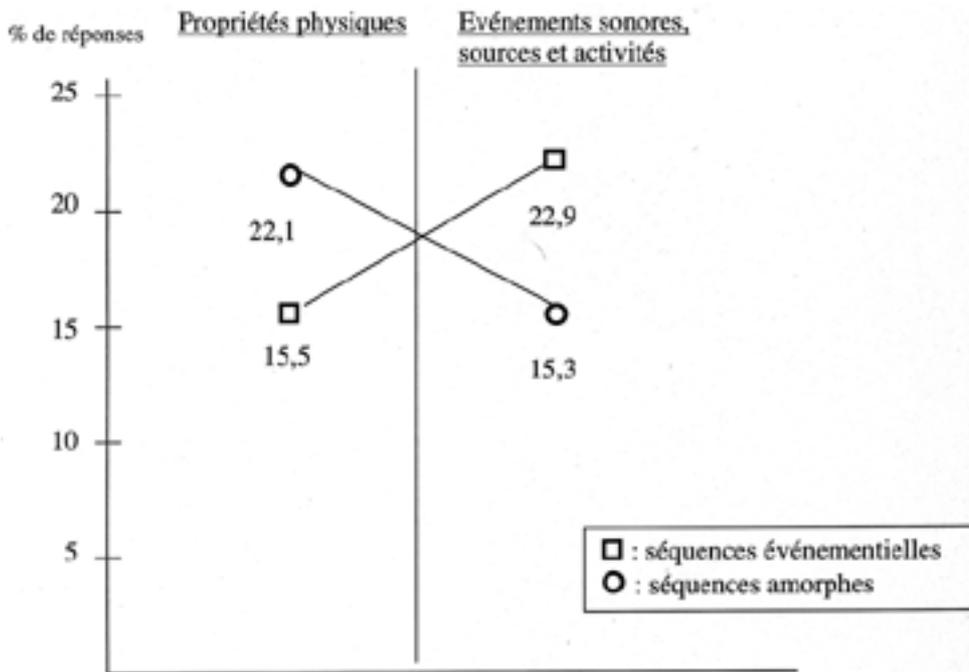


Fig. 6 : Distribution des données verbales

Elle consiste à relever les unités verbales employées et à calculer le pourcentage de chaque descripteur pour chaque catégorie. Deux principaux types de données verbales sont apparus : les descripteurs se rapportant aux paramètres physiques et aux signaux sonores (en italique sur la figure 5) et les descripteurs se rapportant aux événements, sources sonores et activités significatifs.

La figure 6 montre la distribution de ces deux types de données verbales en relation avec le caractère événementiel ou amorphe des séquences sonores.

Les séquences événementielles sont plus fréquemment décrites par des événements sonores significatifs que les séquences amorphes. Inversement, les séquences amorphes sont plus fréquemment décrites par des propriétés physiques du son comme signal, que les séquences événementielles. En effet, pour les sujets, aucune source sonore n'émerge dans les séquences amorphes, donc ces stimuli seraient mentalement traités comme des bruits privés de sens. Ceci explique pourquoi ils sont principalement décrits par des propriétés physiques des signaux, dont l'intensité.

Conclusion

L'étude de la qualité sonore de l'environnement urbain que nous avons menée semble démontrer la nécessité de prendre en compte les phénomènes de signification des sons donnés à écouter aux auditeurs. En outre, la mesure dans l'espace physique doit être rapportée au qualitatif, en particulier décrite par sa valeur sémantique dans l'espace psychologique. Dans notre travail, il apparaît une nette influence du contenu sonore et du caractère évènementiel ou amorphe des séquences dans l'évaluation de l'intensité perçue.

Certaines séquences sont groupées du fait de l'identification par les sujets, dans le contenu sonore de ces séquences, de sources sonores identiques, éventuellement variables en intensité. Même si ce contenu est différent, il semble que le simple fait que les séquences comportent des événements émergents tende à donner à ces séquences un "air de famille" ; leur intensité sera perçue différemment de l'intensité de séquences amorphes de même niveau moyen.

Du point de vue linguistique, les descriptions verbales des catégories formées sont souvent liées à la description des séquences (événements particuliers discriminables). Lorsque les catégories regroupent des séquences amorphes, ne comprenant aucun événement sonore spécifique, les descriptions verbales des catégories se réfèrent alors à des propriétés physiques du son comme signal. Ce qui illustre le fait que le système cognitif humain, à la différence de l'instrument de mesure physique, est capable de différents types de traitements. Ce qui d'un point de vue physique ne peut être traité que de manière identique, peut de fait, pour le sujet humain, être traité soit comme un événement en y assignant des significations, soit comme un signal, de manière un peu similaire à un "instrument"⁷.

La spécificité de l'humain est de faire sens, d'être un producteur de sens ; c'est la signification donnée par l'humain qui conditionne le traitement des paramètres de l'environnement sonore décrits du point de vue de la physique. Le point essentiel de l'approche que nous venons de présenter, à la fois théorique et méthodologique, est de considérer l'analyse **sémantique** comme incontournable. Elle est préalable à tout travail qui se voudrait être non seulement "économiquement pertinent" mais scientifiquement valide, pour ce qui est, évidemment des phénomènes humains.

Cette recherche a bénéficié d'un soutien du ministère de l'environnement.

Remerciements

Nous remercions tout particulièrement Sophie David (SILEX – CNRS, Université de Lille 3) et Coriandre Vilain pour leur collaboration à ces travaux de recherche ainsi que tous les sujets pour leur participation bénévole aux tests d'écoute.

Références bibliographiques

- [1] Berry B, Standards for a quieter world ; some acoustical reflections from the UK National Physical Laboratory, *Internoise 97*, Budapest, Hungary, PP. 3-16, 1997.
- [2] Gualezzi JP, *Le bruit dans la ville. Avis et rapports du Conseil Economique et Social*, Les éditions des Journaux Officiels, 1998.
- [3] McAdams S, *Audition : physiologie, perception et cognition. Traité de psychologie expérimentale*. Presses universitaires de France, 1, PP. 283-344, 1994.
- [4] Gibson J, *The ecological approach to visual perception*, Lawrence Erlbaum Ass., Hillsdale, New Jersey, 1986.
- [5] Norman DA, *Things that make us smart*. Reading, M.A. : Addison-Wesley, 1993.
- [6] Guyot F, Castellengo M, Vogel C, Maffiolo V, Une méthode d'étude de la qualité acoustique des sons réels complexes. *Acoustique et Techniques*, n° 7, PP. 23-26, 1996.
- [7] Castellengo M, Guyot F, Viollon S, Perceptive characterisation of the acoustical quality of real complex sounds – Validation with synthesis, *Forum Acusticum (Anvers)*, *Acustica/Acta Acustica 82 Suppl.1*, S78, 1996.
- [8] Vogel C, Maffiolo V, Polack JD, Castellengo M, Validation subjective de la prise de son en extérieur. *Congrès français d'acoustique 97*, Marseille, France, S. f. d'acoustique, PP. 307-310, 1997.
- [9] Dubois D, *Catégorisation et cognition : de la perception au discours*, Kimé, Paris, France, 1997.
- [10] Guski R, *Psychological determinants of train noise annoyance*. Euro-Noise, München, PP. 573-576, 1998.
- [11] Cage J, M : *Writings '67-72*, Wesleyan University Press, Middletown, Connecticut, 1973.
- [12] Maffiolo V, *Méthodes d'approche de l'environnement sonore urbain*. Mairie de Paris, Direction de la protection de l'environnement, SPAAS, Paris, 1997.
- [13] Dubois D, *Sémantique et Cognition - Catégories, prototypes, typicalité, Sciences du langage*, CNRS Éditions, 1993.
- [14] Barthélemy JP, *Similitude, arbres et typicalité In Sémantique et Cognition - Catégories, prototypes, typicalité*, D. Dubois, Sciences du langage, CNRS Éditions, 1993.
- [15] Barthélemy JP, Guénoche A, *Les arbres et les représentations des proximités*, Masson, Paris, 1988.
- [16] Rosch E, Lloyd B, *Cognition and categorization*, Lawrence Erlbaum Ass., Hillsdale, New Jersey, 1978.
- [17] David S, Souchart M, *Analyse de discours et traitement automatique de données textuelles : le logiciel Termino, Recherches Documentaires, Cahiers de Linguistique sociale, Coll. "Bilans et perspectives"*, Bulot, Th., Dubois, D. et Ch. Batime Ed., PP. 61-76, 1995.
- [18] David S, Dubois D, Rouby C, Schaal B, *L'expression des odeurs en français : Analyse lexicale et représentation cognitive*, Intellectica, 1997.
- [19] Maffiolo V, Vogel C, Polack JD, Castellengo M, Dubois D, David S, *Ambiances sonores représentatives d'une ville : le cas de Paris. Congrès français d'acoustique 97*, Marseille, France, S. F. d'acoustique, PP. 303-306, 1997.
- [20] Maffiolo V, Dubois D, David S, Castellengo M, Polack JD, *Loudness and pleasantness in structuration of urban soundscapes. Internoise 98*, Christchurch, New Zealand, 1998.

⁷ Ce que ne peut pas faire l'instrument physique, mais ce vers quoi tend éventuellement la modélisation informatique de reconnaissance des sons et bruits en intelligence artificielle.