



Master Sciences de la Cognition – niveau 2 – mention
recherche

**Comparaison entre catégorisation fonctionnelle et
analyse perceptive de scènes sonores.**

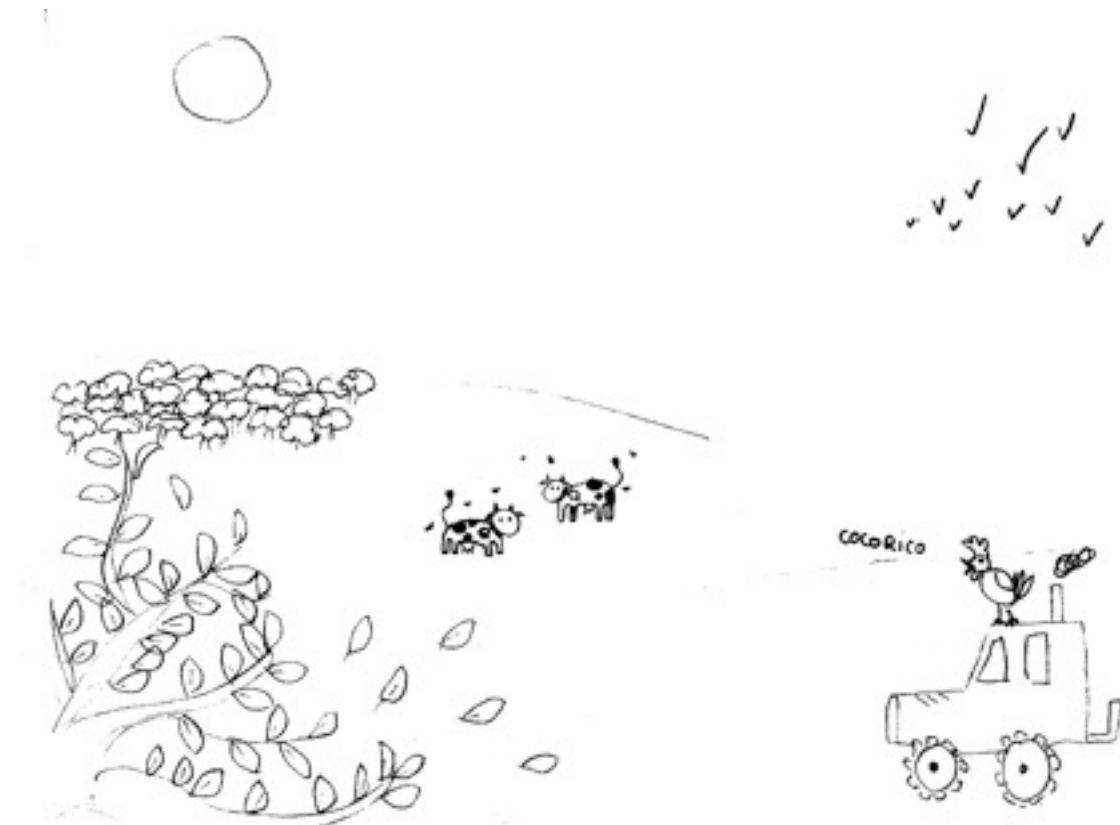
Mémoire de stage

Enseignants : M. Leblanc Benoît & Mme. Sémal Catherine

**Analyse de scènes auditives, sons de l'environnement,
classification de sons, rumeur de fond, événements sonores déterministes
événements sonores transitoires, critères typo-morphologiques,
taxonomie, causalité d'un événement sonore, mémoire sémantique.**

« Soundscape does not exist without a subject or a community. People hear and produce their environment not only through their actions, but also by speaking and writing about their experiences. The cultural study of soundscapes can be compared to the cultural study of music : both need the teller and her listening experience to accompany the sound ».

« 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)



Résumé

L'étude présentée ici s'inscrit dans la continuité du projet initié en M1 avec pour ambition de proposer un processus de validation expérimental du système de classification mis en œuvre précédemment. La finalité du projet reste avant tout de valider un mode de classification des sons permettant d'opérer à des recherches des sons qui seront perceptivement viables dans un contexte donné. Mais d'une façon plus générale la méthodologie expérimentale mise en place va permettre d'aborder les thèmes : de la représentation mentale d'une scène sonore, de la discrimination et de la catégorisation d'éléments sonores aussi bien en mémoire qu'en situation d'écoute, visant ainsi à mettre en avant la dualité de la contextualisation (analyse perceptive) et de la catégorisation (analyse fonctionnelle) dans la perception auditive tout en abordant le rôle de l'émotion dans le traitement cognitif des scènes sonores; et d'une manière plus précise d'évoquer la structuration des représentations mentales des ambiances sonores rurales et d'évaluer leur appréciation qualitative.

1	<u>INTRODUCTION</u>	4
2	<u>METHODE</u>	7
2.1	LE CORPUS EXPERIMENTAL	7
2.1.1	LA REPRESENTATION MENTALE	7
2.1.2	LES ECHANTILLONS SONORES	8
2.1.3	LA CONSTRUCTION DE LA SCENE SONORE	8
2.1.4	LES QUESTIONNAIRES A CHOIX FORCES	9
2.2	DETAILS DU PROTOCOLE EXPERIMENTAL	9
2.2.1	PREMIERE EXPERIENCE : ANALYSE SENSIBLE ET PERCEPTIVE	9
2.2.2	DEUXIEME EXPERIENCE : ANALYSE PERCEPTIVE ET FONCTIONNELLE	10
2.2.3	TROISIEME EXPERIENCE : ANALYSE SENSIBLE ET FONCTIONNELLE	10
2.3	METHODES D'ANALYSE	10
2.3.1	LA TACHE DE REPRESENTATION	10
2.3.2	LES ORGANISATIONS COGNITIVES	11
2.3.3	L'ACOUSTIQUE EN AVAL	11
2.3.4	LES REPONSES : LA VERBALISATION	11
2.3.5	RESPECT DES FORMES LINGUISTIQUES PRODUITES PAR LES SUJETS	12
2.3.6	QUESTIONNEMENT	12
2.3.7	CONCLUSION	12
3	<u>RESULTATS</u>	13
3.1	PREMIERE EXPERIENCE	13
3.1.1	ANALYSE DES REPRESENTATIONS	13
3.1.2	ANALYSE DES ECHANGES VERBAUX	14
3.1.3	LA REPRESENTATION « EMOTIONNELLE »	14
3.2	DEUXIEME EXPERIENCE	15
3.2.1	QUELQUES RESULTATS SUR LA PREPARATION DE LA SCENE ET DES ECHANTILLONS	15
3.2.2	EVALUATION « EMOTIONNELLE » DE LA SCENE SONORE	16
3.2.3	ANALYSE DES REPONSES AU QUESTIONNAIRE	17
3.3	TROISIEME EXPERIENCE	18
3.3.1	ANALYSE DES NOUVELLES REPRESENTATIONS	18
3.3.2	ANALYSE DES ECHANGES VERBAUX	18
3.3.3	LA NOUVELLE REPRESENTATION « EMOTIONNELLE »	19
3.4	COMPARAISON DES DONNEES EMOTIONNELLES	19
4	<u>DISCUSSION</u>	20
4.1	LES LIMITES DE LA PSYCHOACOUSTIQUES	21
4.2	LE CORUPS SONORE	22
4.3	LES REPRESENTATIONS MENTALES	22
4.4	RECONNAISSANCE / IDENTIFICATION ET DEPENDANCE CULTURELLE	23
4.5	PLASTICITE SENSORIELLE VERSUS SELECTIVITE FONCTIONNELLE	24
4.6	LA NECESSITE D'UNE MODELISATION MULTI ECHELLES	26
5	<u>CONCLUSION</u>	26

1 Introduction

Schaeffer (1966) 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, 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".

La genèse du projet provient d'un réel besoin prenant sa source dans l'activité professionnelle des Sound Designers. En effet, la majeure partie du travail d'un designer sonore consiste à trouver le son qui lui conviendra par sa dynamique, par sa texture, par l'émotion et la sensation qu'il va susciter chez l'auditeur. Ce travail est souvent ardu car il s'appuie sur des banques de sons de studio qui comportent plusieurs centaines de giga-octets de sons dont la plupart ne sont pas classés ou disposent d'un nom à rallonge définissant les sources de production des sons ; la seule solution est donc l'écoute systématique de « tranches » de banques de sons pour procéder à une réduction par sélection de sons appropriés ou inappropriés.

Les différents sujets abordés par l'étude réalisée en M1, reposent sur le socle commun de l'audition. Comment les sons sont-ils perçus ? et donc comment les classer ? Nous nous sommes intéressés donc particulièrement aux différentes classifications existantes mais aussi à certains processus physiologiques permettant une sorte de « marquage émotionnel ».

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), "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". Ainsi, reconnaissance et identification s'effectuent par regroupement, sous forme de catégories, des représentations en mémoire d'événements signifiants (Guyot, 1996 ; 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) 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 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 (propriétés pertinentes).

Les catégories s'organisent autour de trois niveaux d'abstraction: 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".

Ces deux théories semblent toutes les deux être un bon reflet du fonctionnement cognitif de l'homme, même si toutefois elles omettent le penchant émotionnel qui comme Damasio le suggère va influencer sur le traitement cognitif. Il constate sur certains de ces patients cérébrolésés qu'en fait le raisonnement totalement de sang froid empêche d'attribuer des poids positifs ou négatifs aux différentes options possibles face à un problème. Ce déficit dans l'attribution d'une connotation positive ou négative des événements les empêcherait de prendre une décision effective.

Ces observations cliniques suggèrent que l'état corporel en tant que siège de la perception et de l'expression des émotions se présente au service d'un processus adaptatif de sélection d'informations, qui guide le comportement d'un organisme dans son milieu. Cette proposition correspond à l'hypothèse de Damasio : l'hypothèse des marqueurs somatiques.

Damasio dit que cette hypothèse des marqueurs somatiques fait référence à des états du corps perçus de façon plus ou moins confuse qui connotent automatiquement, positivement, ou négativement, en fonction de l'expérience ultérieure, telle option particulière, et conduit ainsi le sujet à rejeter immédiatement, voir implicitement, certaines options, et à choisir parmi un nombre beaucoup plus limité d'alternatives.

Cette hypothèse suggère que la perception d'un état somatique participe à la sélection d'options pertinentes. Parce qu'on va en rejeter, on va réduire notre choix à 2/3 alternatives possibles, et on va pouvoir justifier notre choix de manière explicite.

Et réciproquement : ressentir une émotion consisterait en la perception de changement de l'état corporel, rapporté à des représentations particulières d'un phénomène donné.

On a vu que la prise de décision était largement pendante de l'état émotionnel. On voit ici à travers cette théorie que l'expérience émotionnelle implique la capacité d'élaborer des représentations de l'évènement et de l'état corporel.

L'étude présentée ici s'inscrit dans la continuité du projet initié en M1 avec pour ambition de proposer un processus de validation expérimental du système de classification mis en œuvre précédemment, se basant essentiellement sur l'analyse perceptive de scènes sonores construites à partir de contraintes sur des critères prédéfinis et s'appuyant sur une étude des corrélations entre la catégorisation fonctionnelle (du point de vue de l'auteur de la scène sonore) et perceptive (du point de vue de l'auditeur naïf). La validation expérimentale du modèle passera donc par la comparaison et l'adéquation de cette catégorisation fonctionnelle et de cette catégorisation perceptive.

La première approche de classification des sons est l'approche taxonomique, il s'agit de

mettre en place une taxonomie des sons. Les mots qualificatifs, et descriptifs des sons sont très peu utilisés dans notre quotidien, nous gardons donc la substance auditive que nous entendons comme la représentation d'une cause ou d'un sens, c'est à cela que l'approche taxonomique s'attèle. Dans la plupart des cas pour couvrir l'ensemble des sons productibles elles sont dissociées en « sous taxonomies » ; c'est-à-dire par exemple une taxonomie des sons musicaux, une taxonomie des sons environnementaux et une taxonomie des sons synthétiques (ou électroniques).

La seconde approche est une approche morphologique, ou plutôt typo-morphologique créée par Pierre Schaeffer. Cette approche considère les sons comme des objets sonores et propose une approche conceptuelle d'un son dans le sens où il entend connaître le son par sa morphologie (sa forme). Selon Schaeffer, la meilleure façon de connaître un son est de le décontextualiser et donc de connaître et caractériser le son hors de sa causalité et donc de la source qui le produit. Tous les sons sont donc potentiellement des objets sonores.

Ces deux approches sont donc fortement antinomiques et font écho à deux courants porteurs dans le paysage sonore international du 20^{ème} siècle : Pierre Schaeffer et le concept d'objet sonore et de son décontextualisé (connaître le son hors de la source qui le produit) faisant référence à une sphère Morphologique ; en opposition à Robert Murray Schäfer et le son environnemental (le son prend sens dans son contexte, dans son environnement et donc aussi de la source qui la crée) faisant donc référence à une sphère Sémantique

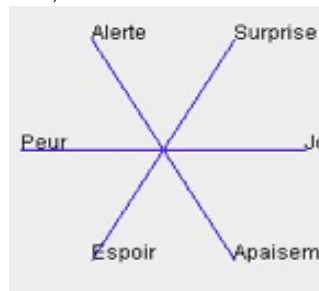
Le modèle de classification construit s'appuie donc sur quatre sphères principales de classification : la sphère sémantique, la sphère morphologique, la sphère causale et la sphère émotionnelle.

SPHERE SEMANTIQUE : recourant à deux types de catégorisation que sont les critères **thématiques** (bruit de vent léger dans les arbres => chant d'oiseaux) et **taxonomiques** (chant d'oiseaux => chant de grive, de rouge-gorge ...).

SPHERE MORPHOLOGIQUE : ces critères se décomposent en groupes : les critères de **matière** (masse, timbre harmonique et grain), les critères de **forme** (dynamique et allure) et les critères de **variation** (profil de masse et profil mélodique).

SPHERE CAUSALE : s'intéressant à la causalité du son et donc à la source qui le produit. Cette sphère va donc s'intéresser à la verbalisation de l'**action**, la dénomination de **matériaux**, de **types d'excitation** et d'**adjectifs timbraux**.

SPHERE EMOTIONNELLE : correspondant à la création d'un diagramme émotionnel en étoile, se rapprochant de la conception des marqueurs somatiques de Damasio permettant alors de situer (connoter) un son en fonction d'émotions basales.



La finalité du projet reste avant tout de valider un mode de classification des sons permettant d'opérer à des recherches des sons qui seront perceptivement viables dans

un contexte donné. La construction des scènes auditives est donc un des éléments clés du projet puisque l'objectif est d'extraire d'une banque de son un ensemble de matériaux sonores pouvant être agencés de façon cohérente entre eux dans le but de créer une scène sonore. Mais d'une façon plus générale la méthodologie expérimentale mise en place va permettre d'aborder les thèmes : de la représentation mentale d'une scène sonore, de la discrimination et de la catégorisation d'éléments sonores aussi bien en mémoire qu'en situation d'écoute, visant ainsi à mettre en avant la dualité de la contextualisation et de la catégorisation dans la perception auditive tout en abordant le rôle de l'émotion dans le traitement cognitif des scènes sonores; et d'une manière plus précise d'évoquer la structuration des représentations mentales des ambiances sonores rurales et d'évaluer leur appréciation qualitative.

2 Méthode

La méthode expérimentale de validation va s'appuyer sur plusieurs points : tout d'abord une étude du matériel et des ressources utilisés, puis une description détaillée des expériences et des conditions d'expérimentations et enfin des pistes pour l'analyse des résultats.

2.1 Le corpus expérimental

La plupart des études sur la qualité des sons environnementaux font appel à des méthodes d'évaluation à la fois quantitatives (mesures de niveau sonore) et qualitatives, dont la mise en œuvre repose sur l'étude des représentations collectives ou des perceptions individuelles. Leur revue critique souligne la non prise en compte de la signification des bruits pour les auditeurs et la nécessité de recourir aux paradigmes de la psychologie cognitive. La première étape de la constitution du corpus expérimental consiste à identifier un contexte rural à partir des représentations verbales et graphiques de l'environnement sonore. L'ensemble «lieu, activité, temps» apparaît alors comme représentatif des ambiances sonores. La seconde étape consiste à recréer un archétype sonore de cette représentation, à partir d'un matériel sonore préexistant classé et connu selon le modèle défini.

2.1.1 La Représentation mentale

Une représentation mentale ou représentation cognitive est l'image qu'un individu se fait d'une situation. Elle est au confluent des sensations et de la mémoire. Dans une situation donnée, les sensations vont susciter l'activation d'informations contenues en mémoire ce qui provoquera les réactions du sujet. Comme toute activité humaine est organisée en vue d'une fin, la notion de représentation est proche de celle d'état mental, et donc du concept d'intentionnalité. Les représentations sont variées : images, mémoire, concepts, émotions... On peut distinguer plusieurs représentations :

- image mentale, reconstitution de la forme physique d'un objet
- mémoire sémantique
- concepts et catégories

L'accès à cette représentation mentale est difficile, et pour y accéder nous allons proposer au sujet de réaliser une représentation graphique de son image mentale d'un paysage sonore rural et de la commenter. Le sujet sera muni d'une feuille A4 pliée en deux, d'un crayon de bois et d'une gomme disposera d'un temps de cinq minutes. Au terme des cinq minutes aura lieu une discussion, orientée par un questionnaire ouvert

connu uniquement de l'expérimentateur. Cet exercice permettra de fournir un corpus de documents à analyser comprenant une image mentale commentée par le sujet lui-même.

2.1.2 Les échantillons sonores

Le corpus d'échantillons sonores utilisé pour l'expérience est une banque de sons préenregistrés proposée par Radio France qui est couramment utilisée par les designers sonores ; elle est scindée en quatre Compact Disques représentant les sons relatifs à la Nature (63 éléments), la Foule (38 éléments), les Transports (52 éléments) et les Activités (68 éléments). L'ensemble de la banque de sons représente donc 221 éléments et un poids de 2,82 Go de données audio (encodées en AIFF).

Cette banque de sons a donc été importée dans la base de donnée puis classée selon les différentes sphères d'analyse et les sons ont été évalués émotionnellement par trois réalisateurs sonores différents une fois par semaine pendant un mois à des horaires différents. Les valeurs présentes dans la base de donnée représentent une moyenne de ces évaluations.

La construction de la scène sonore va donc utiliser ces matériaux comme support qui sont définis de façon précise dans la modélisation mise en place.

2.1.3 La construction de la scène sonore

La construction de la scène sonore va s'appuyer sur des requêtes sur la base de donnée visant à mettre en avant des attributs thématiques, morphologiques, et émotionnels. La scène sera donc construite sur des matériaux ayant entre eux une forte cohérence taxonomique comme ici une ambiance rurale avec des sons connus mais pouvant offrir une difficulté de représentation (vent dans les arbres, cris d'animaux divers, bruits d'eau, voire de pas en suggérant un déplacement...).

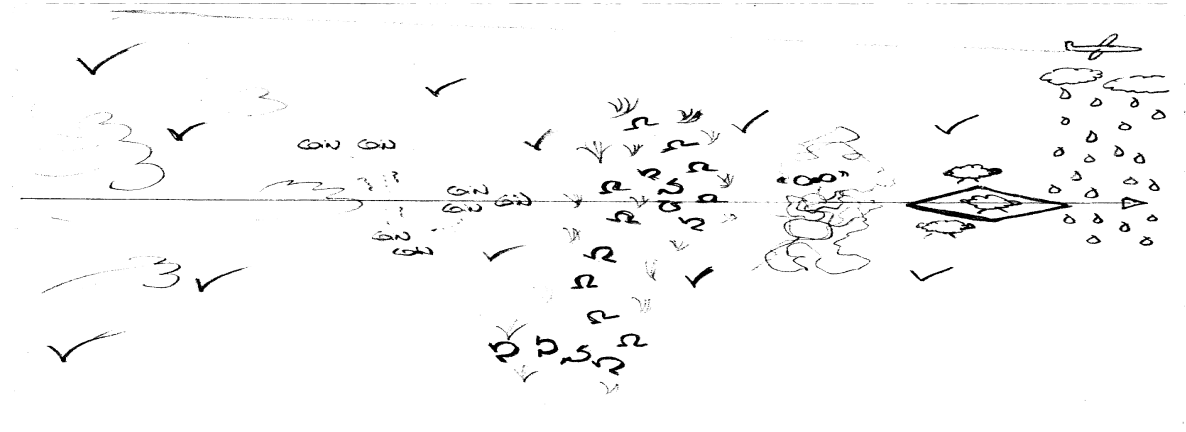
Dans l'objectif d'une scène réaliste, il faudra s'attacher tout particulièrement à identifier de façon précise les éléments constitutifs d'une scène : **la rumeur de fond** (fond sonore ou texture : comme un bruit de vent ou de vague), les **événements déterministes** (événements de premier plan courts et accrocheurs (souvent sinusoïdaux): comme des chants d'oiseau) et les **événements transitoires** (événements de premier plan, courts, de type bruité : comme des sons de passage de voitures). Ces trois éléments sont des caractéristiques évaluées et représentées dans la base de données, ils ont donc en commun un critère sémantique fort et identifiable, et ont cependant des critères morphologiques et émotionnels propres.

D'un point de vue plus créatif, l'art sonore permet de se libérer de l'emprise du verbal et du langage pour établir une communication entre les imaginaires des réalisateurs et auditeurs. L'écriture, impropre pour transposer les idées et évocations suscitées par l'écoute, est néanmoins utile dans la pratique créative. L'écriture de la musique (texte, graphisme ou enregistrement) peut ainsi avoir deux utilités :

- ⇒ descriptive : pour enregistrer une suite temporelle sur un support matériel.
- ⇒ prescriptive : pour générer de la musique à partir de l'utilisation d'un langage écrit approprié.

Le rôle de l'écriture est d'apporter une représentation de l'information sémantique qui permette de produire la musique de manière répétable à partir d'une description. Nous allons donc choisir comme procédé d'écriture de la scène sonore le même procédé que celui qui sera utilisé par les sujets : la représentation mentale ; la composition par dessin est une technique utilisée par certains compositeurs de musique

électroacoustique pour créer des partitions d'oeuvres, ou par certains designers sonores pour une écriture utilisant le son comme matière première.



La « partition » : le temps est représenté puis archétypes graphiques représentant des caractéristiques d'un son (ex la taille va représenter la masse, une texture par une densité...)

2.1.4 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 prédéfinies... 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. Les réponses des sujets à ces questionnaires après l'écoute de l'extrait sonore permettront de valider donc l'adéquation du modèle du sujet avec celui du chercheur.

2.2 Détails du protocole expérimental

2.2.1 Première Expérience : Analyse sensible et perceptive

Cette expérience va porter sur un exercice de représentation graphique de la représentation mentale du sujet d'une scène sonore. La discussion qui lui sera posée est la suivante : « Pouvez-vous dessiner votre représentation mentale d'un paysage sonore rural ? ». Après cinq minutes de dessin le sujet sera d'abord interrogé sur la représentation émotionnelle qu'il se fait d'une telle image mentale, en lui demandant de noter de zéro à six la sensation que lui procure cette image selon les six émotions basales étudiées; puis il sera interrogé sur la signification de sa représentation : il lui sera demandé de verbaliser l'ensemble des éléments présents en répondant à la question « Qu'avez vous représenté ? », puis la discussion sera orientée successivement vers les notions de représentations abstraites et concrètes (s'il y a lieu), l'importance des échelles représentées (la différence entre la taille des éléments), de la disposition spatiale de l'élément sur la feuille et enfin vers la notion d'archétype de représentation

(ou de représentation prototypique) avec une sémantique, une fonction, une cause et peut-être une forme ?

1ère expérience :

- 1) Représentation graphique (5 min maximum, feuille blanche A4, crayon de papier, gomme)
- 2) Evaluation émotionnelle.
- 3) Discussion (orientée par un questionnaire ouvert inconnu du sujet, prise de note, (5min maximum).

2.2.2 Deuxième Expérience : Analyse perceptive et fonctionnelle

Cette deuxième expérience va permettre l'analyse perceptive via l'identification et la reconnaissance des sons, les groupements de sons (flux auditifs), localisation des sons dans l'espace, et l'estimation des distances ; et l'analyse fonctionnelle par la catégorisation sur la valeur fonctionnelle d'un son précis pris en contexte dans une ambiance : appartenance d'un son aux catégories : rumeur de fond, événement transitoire, et événement déterministe. Enfin le sujet sera interrogé sur la validité de cette ambiance : une ambiance réelle ou construite ?

2ème expérience : Ecoute puis réponse à un questionnaire fermé visant à la reconnaissance et à la catégorisation d'événements sonores (temps imparti de 15min).

2.2.3 Troisième Expérience : Analyse sensible et fonctionnelle

Cette nouvelle expérience de dessin va permettre une étude comparative des dessins et de l'influence de l'expérience de catégorisation sur la nouvelle représentation. Induisant l'influence de l'apprentissage de connaissances analytiques sur la représentation et la perception d'une scène auditive. (=> vers une représentation sensible et fonctionnelle ?)

3ème expérience : Représentation graphique (même conditions que précédemment)

2.3 Méthodes d'analyse

2.3.1 La tâche de représentation

Dans cette tâche l'analyse va porter sur les choix stratégiques de représentations qu'a fait le sujet. L'analyse de la représentation graphique reposera sur les items représentés et par exemple leur niveau d'abstraction (ou de symbolisation), la place relative qu'ils occupent dans la représentation en concurrence avec les autres items, leur disposition, et la mise en place de relation entre eux.

L'analyse de la représentation sera donc fortement orientée par l'échange qui suivra.

- 1) « *Pouvez vous dessiner votre représentation mentale d'un paysage sonore rural ?* »
- 2) « *Qu'avez vous représenté ?* » : verbalisation des items représentés
- 3) « *Quel est l'importance de cet élément dans la scène ?* » au niveau :
 - de sa sémantique,
 - de ses propriétés sonores (masse (intensité du son), timbre ; fixe/en mouvement)
 - de sa fonction (rumeur de fond, événement déterministe ou transitoire)

Dans cet échange, les sujets n'ont aucune contrainte langagière, mais doivent s'exprimer, soit oralement, soit par écrit. 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.

2.3.2 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.

2.3.3 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.

2.3.4 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 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, - qui imposent aux sujets des a priori lexicaux.

2.3.5 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). À 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.

2.3.6 Questionnement

- => Procédure de questionnement sur la catégorisation perceptive et fonctionnelle d'un auditeur naif.
- => Analyse des corrélations et dissemblances entre catégories fonctionnelles et catégories perceptives.
- => Comparaison des représentations avant et après l'expérience 2.
- => Recueil de données sur le plan émotionnel : comparaison entre les différentes évaluations.

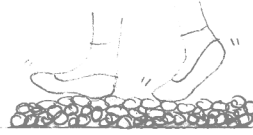
2.3.7 Conclusion

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. Deux points sont essentiels : dans un premier 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; dans un second 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. 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 perceptives et fonctionnelles des signaux.

3 Résultats

3.1 Première Expérience

3.1.1 Analyse des représentations



Parmi les 12 représentations recueillies, nous retrouvons un ensemble de 75 items représentés soit (6,25 items par représentation) et parmi eux 37 % d'*hapax* (28 représentations), représentations qui n'ont été produites qu'une seule fois. Par exemple, une seule personne a dessiné « une personne faisant des ricochets sur l'eau ».

Les représentations sont partagées entre représentations temporelles (4 représentations), et représentations égocentrées (dessine un sujet au milieu de la scène avec les items spatialisés ou représentation d'une vue « à la première personne » - 8 représentations). Dans le cas d'une représentation égocentrée, les représentations prennent en compte l'espace et le déplacement des sons autour de l'auditeur (cela étant intrinsèque à ce type de représentation) ; alors que dans le cas des représentations temporelles aucune ne prend en compte la spatialisation des sons. La dynamique et l'évolution du son n'est pas représentée non plus. Dans les deux cas la taille de l'item représentée semble être corrélée à son importance dans la scène sonore (intuitivement cette notion d'importance est représentative de la notion de masse ou de calibre).

Parmi les 75 items produites au moins deux fois, les plus fréquentes sont : les « chants d'oiseaux » (12 occurrences), le "vent" (8 occurrences), les « meuglements d'une vache(5), les "braiement de moutons" (4), les « bourdonnements d'un insecte » (3).

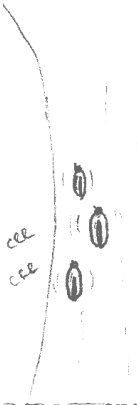
Les représentations les plus fréquentes permettent d'identifier celles qui correspondent à :

- des **propriétés physiques du son** : **15%** "craquements, bruissements, bruitages, sifflement, bourdonnement..." et onomatopées ("ding ding, coïn coïn, bzzz..."). Ces propriétés sont généralement représentées par des dessins abstraits parfois accompagnés d'onomatopées inscrites-

- des **représentation de sources interprétées** : **85%** au niveau sémantique, le son est décrit à travers l'identification de la source et donc sa représentation . On peut distinguer deux types de sources interprétées : le nom de la source ou du "statut" de la source ("moutons, canards, insectes, cloche..."), ou les "**représentations de phénomènes sonores associés**" ("moutons qui braient, caquètements de canards, vol d'un insecte, ") représentent **33%** des désignations les plus fréquentes. Du point de vue de la représentation la deuxième alternative se rapproche des représentations des propriétés physiques du son (dessins abstraits accompagnés d'onomatopées), même si parfois les sujets ont recours à la métonymie accompagnée d'un indice (en représentant par exemple le bec d'un canard accompagné d'une onomatopée - 15 occurrences soit un peu plus d'une fois par dessin en moyenne).

Les noms de sources interprétées 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.

3.1.2 Analyse des échanges verbaux

Une première analyse sémantique des désignations les plus fréquentes permet de regrouper ces dernières en deux classes (les mêmes que précédemment) :

- des propriétés physiques du son, plus ou moins lexicalisées : "craquements, bruissements, bruitages, sifflement, bourdonnement..." et onomatopées ("ding ding, coin coin, bzzz..."),

- 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 ("moutons, canards, insectes, cloche...") et le nom de l'objet sonore associé à une source plus globale ("braiement, caquètement..."). 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.

La quasi-totalité des productions verbales des sujets correspond à des sources interprétées (68 des 75 productions). Parmi elles, les noms de noms représentent près de la moitié des désignations (32 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 (7 productions) et même dans ce cas peu fréquent, les propriétés sont associées à des sources interprétées.

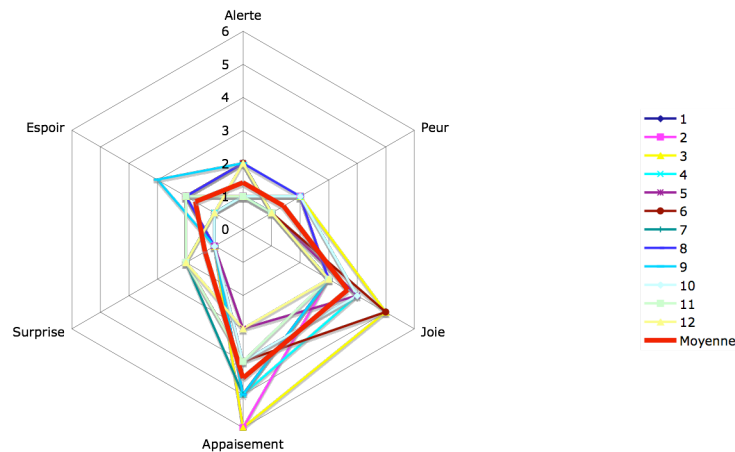
L'identification a priori de signaux déterministes nous fournit ainsi des prototypes de productions langagières signifiantes. Nous remarquons que tous les types de signaux déterministes testés (essentiellement des chants d'oiseaux) sont, non seulement cités par les sujets en tant que signifiés, mais représentent aussi la majorité des signifiés produits. Par ailleurs, les sons ruraux sont avant tout associés à des sons jugés de façon positive, tels chants d'oiseaux et d'animaux, avant d'être considérés parfois comme plus désagréable, tel le hennissement d'un cheval ou le bourdonnement d'un insecte .

Enfin, les termes employés par les sujets correspondent généralement à des effets sur les sujets (« reposant, apaisant, calme, fraîcheur... »).

3.1.3 La représentation « émotionnelle »

	1	2	3	4	5	6	7	8	9	10	11	12	Moyenne
Alerte	1	1	1	1	2	2	1	2	2	1	1	2	1,42
Peur	1	2	2	2	1	1	1	2	1	2	1	1	1,42
Joie	4	3	5	4	4	5	3	3	3	4	3	3	3,67
Appaisement	4	6	6	5	3	4	5	5	5	4	4	3	4,50
Surprise	1	1	1	1	1	2	2	1	1	1	2	2	1,33
Espoir	1	2	2	1	2	2	1	2	3	1	2	1	1,67

Représentation émotionnelle de l'image mentale des sujets



3.2 Deuxième Expérience

3.2.1 Quelques résultats sur la préparation de la scène et des échantillons

La construction de la scène sonore a été réalisée à l'aide de d'un programme de classification de sons. Les requêtes ont essentiellement porté sur la classification taxonomique des sons (sons d'éléments naturels, d'animaux...) et leurs propriétés fonctionnelles (rumeur de fond, événements déterministes, événements transitoires). Les sons ont été sélectionnés ensuite en fonction de leur morphologie et de leur représentation émotionnelle. Un ensemble de huit sons a été sélectionné pour confrontation avec la classification des sujets. Voici les résultats fournis par la base de donnée concernant les sons sélectionnés.

3.2.1.1 La classification taxonomique

	Vent	Abeille	Mésange	Canard	Moutons	Cheval	Pluie	ULM
1er niveau	Eléments	Animaux	Animaux	Animaux	Animaux	Animaux	Eléments	Objets / Machines
2eme niveau	Air	Air	Air	Air	Terre	Terre	Eau	Machines
3ème niveau	Ambiance	Sauvage	Sauvage	Sauvage	Elevage	Elevage	Ambiance	Machine
4ème niveau	Vent	Se déplacer	Chanter	Crier	Crier	Crier	Sur Herbe	Moteur à explosion
				Se déplacer	Se déplacer	Se déplacer		
5ème niveau	Léger							

3.2.1.2 La classification morphologique

	Vent	Abeille	Mésange	Canard	Moutons	Cheval	Pluie	ULM
Durée	>60s	30-60s	1-5s	10-30s	30-60s	30-60s	>60s	>60s
Timbre	complexe	complexe	tonique	tonique + complexe	tonique + complexe	tonique + complexe	complexe	complexe
masse (calibre)	large	fin	fin	Moyen	moyen	moyen	large	moyen
profil dynamique	aléatoire	aléatoire	delta	Aléatoire	delta	delta	dilaté	neutre
profil mélodique	aléatoire	aléatoire	torculus	Aléatoire	aléatoire	aléatoire	neutre	clivis

3.2.1.3 La classification fonctionnelle

Rumeur de fond : Vent, Pluie, ULM.

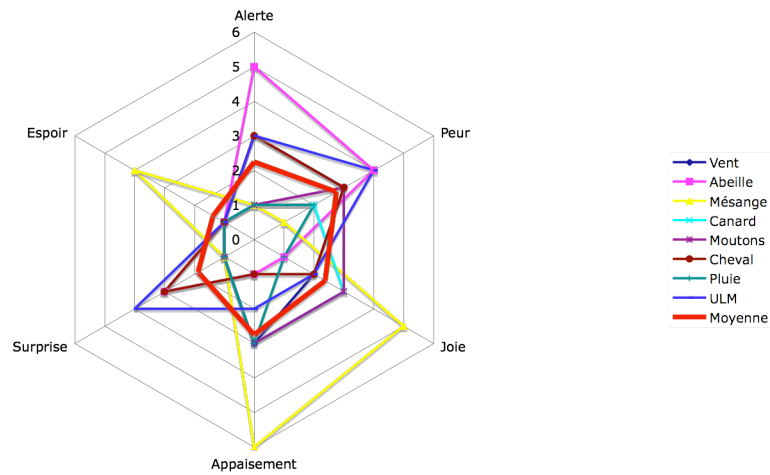
Événements transitoires : Canards, Moutons, Cheval.

Événements déterministes : Mésange, Abeille.

3.2.1.4 La classification émotionnelle

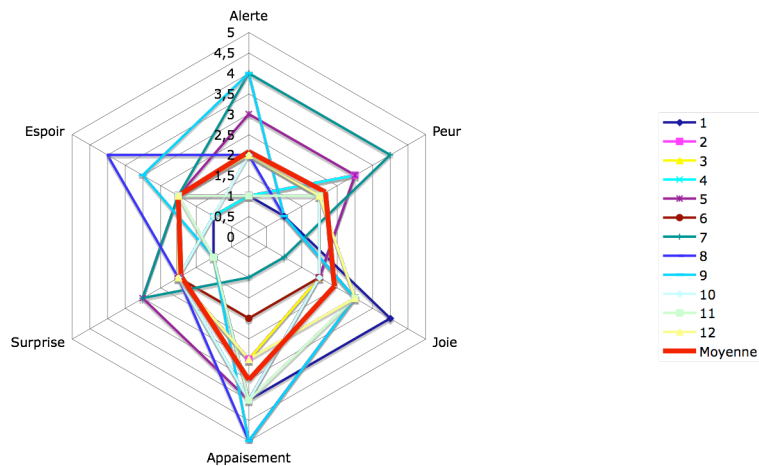
Les valeurs fournies ont été arrondies automatiquement par le logiciel

Représentation émotionnelle de la Base de données



3.2.2 Evaluation « émotionnelle » de la scène sonore

Représentation émotionnelle de la scène après écoute

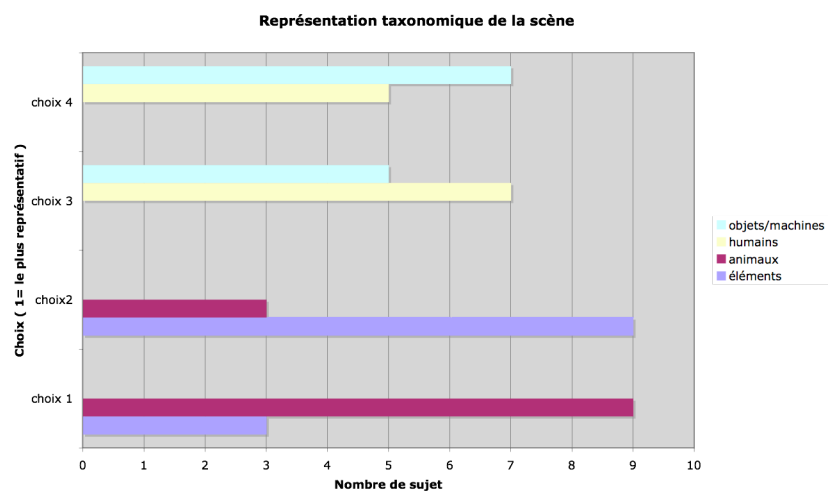


	1	2	3	4	5	6	7	8	9	10	11	12	Moyenne
Alerte	1	1	2	1	3	2	4	2	4	2	1	2	2,08
Peur	1	3	2	3	3	2	4	1	1	2	2	2	2,17
Joie	4	2	2	2	2	2	1	3	3	2	3	3	2,42
Appaisement	4	3	3	4	4	2	1	5	5	4	4	3	3,50
Surprise	1	2	2	2	3	2	3	2	1	2	1	2	1,92
Espoir	1	2	2	1	2	2	2	4	3	1	2	2	2,00

3.2.3 Analyse des réponses au questionnaire

Cette deuxième expérience va permettre l'analyse perceptive via l'identification et la reconnaissance des sons, les groupements de sons (flux auditifs), localisation des sons dans l'espace, et l'estimation des distances ; et l'analyse fonctionnelle par la catégorisation sur la valeur fonctionnelle d'un son précis pris en contexte dans une ambiance : appartenance d'un son aux catégories : rumeur de fond, événement transitoire, et événement déterministe.

L'étude des réponses au questionnaire révèle un bon phénomène de reconnaissance du point de vue du premier niveau de la taxonomie (d'un point de vue sémantique), la majorité des sujets (9 sur 12) ont reconnu la scène sonore comme étant principalement composée de sons d'animaux, les 3 sujets restant ayant identifié la scène sonore comme composée essentiellement de sons relatifs aux éléments ;



D'un point de vue plus général, l'identification taxonomique est précise aux premier et second niveaux, cependant au-delà les représentations des sujets ne sont plus corrélés aux résultats de la base de données.

L'analyse statistique des réponses au questionnaire, a permis de mettre en avant des éléments de corrélations entre les catégories fonctionnelles et leur représentation perceptive (du point de vue de la taxonomie, ou de leur morphologie).

Dans les réponses des sujets au questionnaire, certaines corrélations sont observées entre catégories fonctionnelles et taxonomiques, ainsi les sons relatifs aux éléments ont été identifiés comme étant des rumeurs de fond, les sons d'animaux sont eux représentés comme des éléments transitoires, mis à part les chants d'oiseau considérés comme déterministes.

De la même façon, certains aspects morphologiques sont fortement corrélés à la catégorisation fonctionnelle ; de cette façon on remarque que les rumeurs de fond seraient caractérisées par une durée relative (>30s), à un calibre de masse large ; en revanche les tests employés révèlent l'indépendance du timbre. Les événements transitoires eux sont corrélés à une durée moyenne (entre 5s et 30s), à un calibre de masse (moyen), mais aussi à la notion d'une source en mouvement; de la même façon on note l'indépendance du timbre. Enfin les événements déterministes sont corrélés à une durée courte (<5s), à un calibre de masse fin, et à un timbre tonique. Les critères de profil mélodique, et de profil dynamique ne semblent être liés à aucune catégorie fonctionnelle ; ces critères sont de plus difficiles d'appréciation pour un néophyte et

leur caractérisation chez un utilisateur chevronné passe plutôt par une représentation graphique que par une verbalisation.

A la dernière question du questionnaire : « Pensez-vous que cette ambiance a été enregistrée ou a été fabriquée à partir de sons préenregistrés ? », la quasi-totalité des sujets (9 sur 12) a conclu à une ambiance enregistrée en contexte.

3.3 Troisième Expérience



3.3.1 Analyse des nouvelles représentations

Parmi les 12 représentations recueillies, nous retrouvons un ensemble de 69 items représentés (soit 6 items de moins que lors de la première expérience) et parmi eux seulement 7 % d'*hapax* (5 représentations), représentations qui n'ont été produites qu'une seule fois.

La répartition du nombre de sujets ayant opté cette fois pour une représentation temporelle est de 7 (3 représentations de plus), contre 5 représentations égocentrées.

Les techniques de représentation n'ont que peu évolué ; dans le cadre d'un référentiel temporel certains sujets (2), ont dessiné un point d'ancrage (prenant la forme d'une oreille) et ont dessiné le déplacement du son autour de ce point. De plus la notion de dynamique à cette fois été représenté par 3 sujets par la représentation successive du même item mais à des dimensions différentes.

Parmi les 75 items produits au moins deux fois, les plus fréquents sont : les « chants d'oiseaux » (18 occurrences, représentation de plusieurs types d'oiseaux pour 4 sujets), le « vent » (12 occurrences), les « bourdonnements d'un insecte » (12), les « braiement de moutons » (11), les « meuglements d'une vache » (3) ; on note aussi, dans l'ensemble des scènes, l'apparition de l'homme ou de son action à travers l'utilisation d'une machine.

Les représentations les plus fréquentes permettent d'identifier celles qui correspondent à :

- des **propriétés physiques du son : 32%**
- des **représentation de sources interprétées : 68%** parmi lesquelles les "**représentations de phénomènes sonores associés**" représentent **42%** des désignations les plus fréquentes. Les nouvelles représentations vont donc plus s'attacher à la représentation de propriétés physiques de sons, et avoir a priori un niveau de représentation des sources interprétées plus précis.

3.3.2 Analyse des échanges verbaux

La majorité des productions verbales des sujets correspond à des sources interprétées (50 des 69 items représentés). Parmi elles, les noms de noms représentent maintenant la quasi totalité des désignations (44 productions), ces désignations, plus précises, correspondent donc à des représentations mentales bien définies chez l'individu. Seulement 5 sujets ont produit des expressions en termes de propriétés physiques du son (19 productions) , dans ce cas les propriétés sont associées à des sources interprétées.

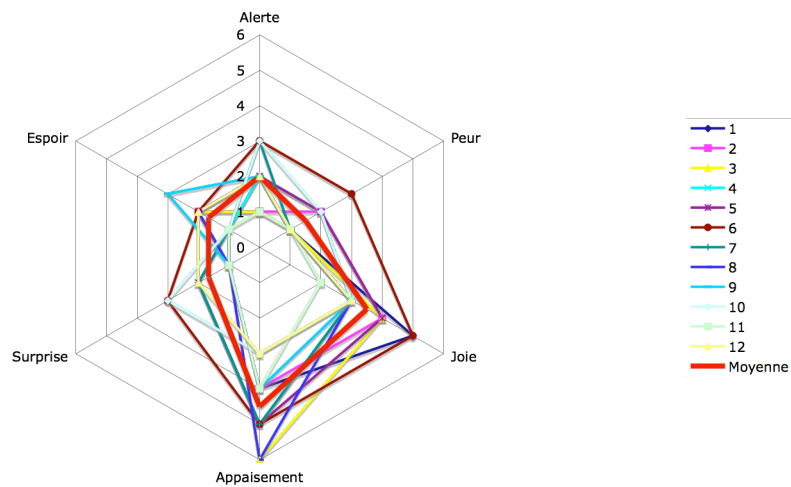
D'une façon générale la totalité des sujets a repris les expressions de catégorisation fonctionnelle dans son discours avec plus ou moins d'hésitation. La notion de rumeur de fond (la plus intuitive a priori a été retenue et utilisée à bon escient par la totalité des sujets) ; du point de vue des termes « événement déterministe » et « événement



transitoire » leur utilisation a été beaucoup plus rare (évoqués par seulement 5 sujets) mais de façon relativement pertinente.

3.3.3 La nouvelle représentation « émotionnelle »

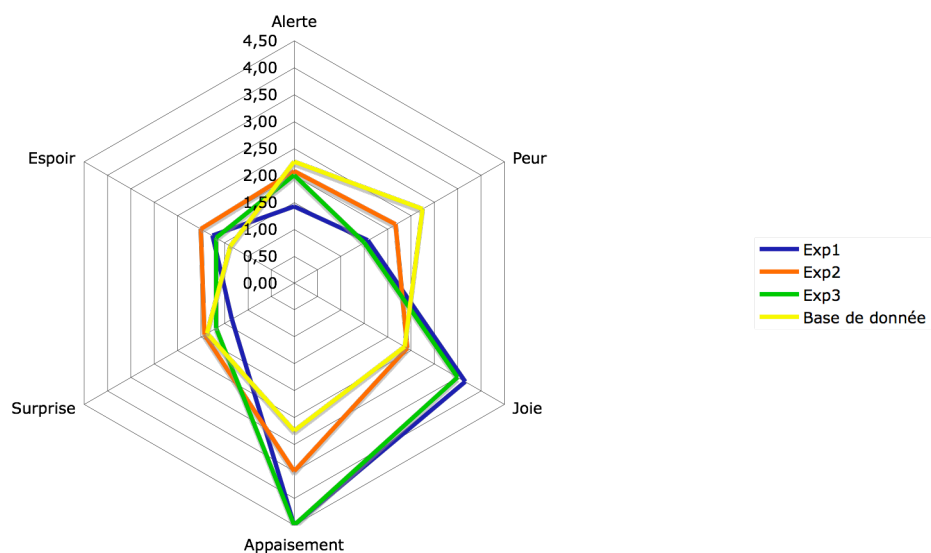
Représentation mentale émotionnelle (2)



3.4 Comparaison des données émotionnelles

	Exp1	Exp2	Exp3	Base de donnée
Alerte	1,42	2,08	2,00	2,25
Peur	1,58	2,17	1,50	2,75
Joie	3,67	2,42	3,50	2,37
Appaisement	4,50	3,50	4,50	2,75
Surprise	1,33	1,92	1,67	1,87
Espoir	1,75	2,00	1,67	1,37

Comparaison des moyennes des représentations émotionnelles



	Exp1	Exp2	Exp3
Alerte	NON	OUI	OUI
Peur	NON	OUI	NON
Joie	NON	OUI	NON
Appaisement	NON	OUI	NON
Surprise	NON	OUI	OUI
Espoir	OUI	NON	OUI

Comparaison à la Base de donnée

	Exp1
Alerte	NON
Peur	OUI
Joie	OUI
Appaisement	OUI
Surprise	OUI
Espoir	OUI

Comparaison de l'expérience 1 à l'expérience 3

Du point de vue des données recueillies sur les représentations émotionnelles, la comparaison des différentes données obtenues permet d'observer une forte correspondance entre les données des expériences 1 et 3 (la comparaison des moyennes à un risque de 5%, suggère que les échantillons diffèrent uniquement pour le sentiment d'Alerte), et une correspondance moindre mais non négligeable entre les valeurs de l'expérience 2 et les valeurs de la base de donnée (à 5% les valeurs diffèrent uniquement pour les sentiments de Peur et d'Espoir). De plus les écart-types des séries de données 1 et 3 sont inférieurs aux écart-types des séries de données de l'expérience 2 suggérant ainsi une plus grande précision de l'expression des sentiments dans le cadre d'une représentation mentale que dans un contexte imposé de catégorisation.

4 Discussion

On peut distinguer trois formes de mémoire sensorielle, dans le domaine auditif comme dans le domaine visuel. Il existe d'abord une mémoire à très court terme, dont l'empan se mesure en dixièmes ou centièmes de seconde. Celle-ci est en réalité une composante de la perception elle-même, de sorte qu'on la qualifie parfois de "pré-perceptive" : tout objet auditif étant étendu dans le temps), n'importe quelle "perception" auditive fait intervenir une intégration temporelle et donc une mémoire.

Une seconde forme de mémoire auditive, dite "à court terme", la mémoire à court terme strictement auditive (non symbolique) semble être très largement automatique et insensible à l'attention. La constitution d'une trace mnésique de hauteur tonale est légèrement dépendante de l'attention (Demany, Montandon, et Semal, 2004), mais une fois la trace constituée sa rétention n'apparaît plus du tout dépendre de l'activité mentale du sujet. La rétention d'une trace de hauteur tonale est automatique au point que l'on peut même retenir plusieurs secondes une hauteur tonale qui n'a pas été perçue consciemment. Le mécanisme sur lequel repose cette surprenante détection automatique de changement est susceptible de jouer un rôle important dans l'analyse des scènes auditives. Toutefois, la rétention à court terme de la hauteur tonale d'un son suivi d'autres sons, de même que la rétention de séquences mélodiques, semble reposer sur l'existence de traces mnésiques représentant la hauteur indépendamment du timbre ou de la sonie (Semal, Demany, Ueda et Hallé, 1996).

Il existe enfin une mémoire strictement auditive à long terme (à peu près permanente). Dans le cadre des recherches concernant la mémoire sémantique, le concept d'amorçage sémantique est particulièrement intéressant dans le cadre de notre problématique. Un concept activé en mémoire engendrera une activation vers les concepts voisins dans le réseau (« fleur » propagera de l'activation vers « rose », « tulipe », « coquelicot », « rouge »...). L'ensemble des concepts sera alors pré-activé, et donc plus facile à traiter. Cette facilitation de traitement s'exprime dans les effets

d'amorçage sémantique ou un mot précédé d'un autre mot sémantiquement relié sera reconnu plus facilement. Les théories de la propagation de l'activation sont restées l'explication dominante, même si certaines de ces théories diffèrent énormément (Collins & Loftus, 1975 ; Anderson, 1983), elles partagent des principes essentiels :

- Récupérer un item en mémoire correspond à l'activation de sa représentation interne
- L'activation se diffuse à l'intérieur d'un réseau de traces / nœuds / concepts interconnectés.
- L'accumulation de l'activation du niveau d'un concept facilite sa récupération ultérieure.

L'ensemble des données recueillies ont permis d'étudier les relations entre les différentes sensibilités de perception auditive : à travers les représentations mentales des sujets, l'analyse de leur discours et de leur réponse au questionnaire.

4.1 Les limites de la psychoacoustiques

Les techniques de la psychoacoustique perdent un peu de leur intérêt dans le cadre de l'étude de critères sémantiques des scènes auditives. La démarche proposée reste très scientifique et va se confronter à des problématiques d'apprentissage, de mémoire : de psycholinguistique et sémio acoustique; ces deux dernières approches peuvent donc revêtir un intérêt tout particulier pour notre étude.

L'approche psychoacoustique, 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 béhaviorisme 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 se fonde sur des modèles associationnistes qui, à l'instar de la physique atomique du XIXème 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 hypothèse 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 stationnaire. 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.

4.2 Le corpus sonore

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.

Comment peut-on transposer des conditions naturelles et quotidiennes à une situation d'expérimentation, a priori artificielle ? En effet, l'expérimentation 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 expérimentation, 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. 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.

De ce point de vue la sélection du corpus sonore, celle-ci est faite par le biais de requêtes sur la base de données qui a été « entraînée » préalablement et assure par le biais des nombreux critères de sélection des sons. De plus les deux approches proposées dans le cadre des expériences sont deux logiques représentationnelles résolument différentes.

4.3 Les Représentations mentales

Une représentation mentale ou représentation cognitive est l'image qu'un individu se fait d'une situation. Elle est au confluent des sensations et de la mémoire.

Dans le cas de l'audition, puisque les événements sonores se succèdent dans le temps : l'élaboration d'une représentation mentale s'avère indispensable pour percevoir leur structure, c'est à dire pour établir des relations entre des événements séparés par plusieurs minutes ou même plusieurs heures.

Lorsque l'information immédiatement disponible au niveau sensoriel se révèle insuffisante, le système perceptif analyse la situation en considérant la connaissance acquise sur le monde sonore. Les connaissances interagissent avec les données sensorielles actuelles dans l'interprétation des stimulations auditives.

De plus, tout apprentissage consiste en l'acquisition de " cartes cognitives " ou " cartes mentales", c'est-à-dire l'acquisition d'informations sur des séquences ordonnées d'événements se produisant de façon régulière dans l'environnement. Frédéric C Bartlett (1886-1969) a été un des premiers à comprendre que la récupération des informations en mémoire est un phénomène complexe : des processus cognitifs ont réorganisé les souvenirs pour les rendre signifiants.

« La récupération (...) est une reconstruction inventive ou une construction fondée sur notre attitude à l'égard d'une masse entière active de réactions et d'expériences passées organisées... » Une des critiques majeures des connexionnistes est que la psychologie cognitive peut difficilement étudier scientifiquement les représentations internes.

Chez l'homme, en mémoire sémantique, le rôle des images mentales est déterminant car elles permettent une facilitation du processus de mémorisation pour les mots concrets.

Pour les mots abstraits où la représentation imagée est plus difficile, il semble que le système acoustique donne de meilleurs résultats. On parle de double codage lorsqu'un mot peut être codé dans le système visuel et acoustique : sa mémorisation en sera facilitée.

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 allons voir 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.

4.4 Reconnaissance / Identification et dépendance culturelle

Lass *et al.* (1982) ont étudié l'identification de quarante sons : vingt sons d'animaux, douze sons mécaniques, quatre sons d'instruments de musique et quatre sons humains. 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.

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.

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 acoustiques, topologiques et socioculturelles 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.

Cette expérience a permis à Amphoux 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. Enfin, Amphoux 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).

Dans le cadre de notre méthodologie, l'itération des exercices de représentation et de commentaire de ces représentations nous donne des indications, sur l'effet de plasticité du traitement de l'information induite par un apprentissage, en l'occurrence la réponse à un questionnaire à choix fermé permettant de caractériser les sons.

4.5 Plasticité sensorielle versus Sélectivité fonctionnelle

Les recherches des bases neurologiques de l'apprentissage et de la mémoire ont souvent ignoré l'adaptation corticale des cartes sensorielles. Ceci est dû au fait que durant les quatre dernières décennies, la neurophysiologie sensorielle a mis en évidence un nombre considérable de relations précises entre certains paramètres du stimulus et la réponse des neurones corticaux, spécialement dans des préparations anesthésiées et paralysées. Ce succès de la physiologie sensorielle dans l'étude de la représentation de l'environnement par les cortex sensoriels primaires a mené à une surévaluation de la stabilité des représentations corticales. Ce même succès a donc renforcé l'idée que l'étude de l'apprentissage et la mémoire commencent là où celle de l'analyse sensorielle s'achève. La plasticité en tant que propriété intrinsèque des aires sensorielles corticales semblait incompatible avec la nécessité des organismes à obtenir une image stable du monde où ils vivent.

Plasticité développementale. Dès les années 60 cependant, il avait été démontré que le cortex visuel a un fort potentiel de plasticité pendant une période précoce de la vie d'un animal. Par ex. la distribution corticale de l'orientation préférée est modifiée par l'expérience précoce chez l'animal jeune. (Cooper and Blakemore, 1971).

Plasticité adulte. Dans les années 80, la recherche des sites neuronaux de la plasticité induite par un apprentissage associatif dans des systèmes sensoriels des mammifères adultes (cortex auditif et somatosensoriel et, dans une moindre mesure, visuel) a mis en évidence l'implication des niveaux primaires d'intégration des systèmes sensoriels.

Au cours des 10 dernières années de nombreux laboratoires ont mis en évidence des phénomènes de plasticité dans les systèmes sensoriel d'animaux adultes, et ce dans toutes les modalités (voir pour revue Calford 2002 ; Gilbert, 1993; Kaas, 1991; Weinberger, 1995; Edeline, 1999, 2003). Bien qu'initialement, peu d'études concernaient réellement des situations d'apprentissage comportemental, depuis la fin des années 90 plusieurs laboratoires ont montré que le Champ Récepteur (ChR) des

neurones du système auditif se modifiait sélectivement après apprentissage. Ces modifications de ChR, observées pendant le temps limité d'un enregistrement unitaire, ont souvent été présentées comme les premières manifestations des réorganisations de cartes sensorielles décrites dans les cortex somesthésique (Recanzone et al. 1992a,b,c,d) et auditif (Recanzone et al., 1993). Bien que de nombreux résultats soient venus étayer ces données au cours des dernières années, plusieurs questions fondamentales se posent encore:

Y a-t-il vraiment un lien de causalité entre les modifications de ChR et les réorganisations de topographie ? En effet, la plasticité des ChR thalamocorticaux s'observe rapidement lors de protocoles relativement simples. Par contre, les modifications de cartographie corticale sont détectées après plusieurs mois de pratique dans des entraînements perceptifs où le sujet est quotidiennement poussé à ces limites perceptives. On peut légitimement se poser la question de savoir si la mobilisation du réseau thalamo-cortical est la même dans ces deux situations.

Quelle est la signification fonctionnelle de modifications se produisant dès le versant sensoriel lors d'un apprentissage ? Peut-on réellement dissocier Mémoire et Perception ou ces deux fonctions sont-elles nécessairement imbriquées l'une dans l'autre ? Cela est même trivial si on définit la perception comme une interprétation des sensations basée sur l'expérience passée... et ce qu'il en reste en mémoire ?

Le phénomène de plasticité et de réorganisations corticales est étudié pour différents systèmes sensoriels chez l'humain, et la plasticité corticale auditive a attiré récemment plus d'attention (Rauschecker, 1999). Pour la perception auditive, un exemple de cette capacité du cerveau est l'adaptation à traiter des nouveaux stimuli sonores lorsqu'une personne malentendante ou sourde reçoit des signaux grâce à un implant cochléaire (Giraud et al., 2001).

Un autre exemple est lié à la perception de la musique et concerne l'auditeur sain qu'il soit musicien ou nonmusicien. D'autres études portent sur d'autres aspects que les caractéristiques acoustiques des sons musicaux, et analysent la perception des structures musicales de plus haut niveau d'organisation, notamment la perception de la 'syntaxe musicale' : quels événements musicaux sont utilisés ensemble et comment ils sont organisés dans le temps. Un ensemble d'études neurophysiologiques (utilisant des méthodes EEG, MEG et IRMf) et d'études comportementales a mis en évidence que même les auditeurs non musiciens sont sensibles aux structures musicales entendues et cela de façon similaire que les auditeurs musiciens (Bigand et al., 1999 ; Besson et al., 1995; Regnault et al., 2001; Maess et al., 2001). Cette sensibilité de l'auditeur aux structures musicales représente un exemple de la capacité cognitive à apprendre implicitement les régularités sous-jacentes à notre environnement (Seger, 1994). Ces processus d'apprentissage implicite mènent à des connaissances de l'auditeur qui influencent par la suite la perception et dont des corrélats neurophysiologiques peuvent être observés. Afin de mieux comprendre les processus impliqués, des modèles connexionnistes (des réseaux de neurones artificiels) ont été appliqués à la perception musicale (Leman, 1995; Tillmann et al., 2000). Les modèles connexionnistes sont basés sur des caractéristiques d'apprentissages du système neuronal, par exemple, les cartes d'auto-organisation (Kohonen, 1995) sont inspirées par les représentations topologiques du cerveau. Ces cartes parviennent à simuler l'apprentissage des connaissances musicales et leurs influences sur la perception. Les simulations permettent également de générer des nouvelles hypothèses pour des études comportementales et neurophysiologiques.

4.6 La nécessité d'une modélisation multi échelle

La perception de notre milieu sonore dépend de l'échelle à laquelle nous le regardons et restituons, dans la plupart des études scientifiques, les sons sont étudiés à « petites » échelles : soit des « petits » sons , soit des propriétés microscopiques à l'intérieur du son lui-même (le timbre, la dynamique ...). Dans le cas des scènes sonores la plupart des études vont s'attacher à l'étude de l'identification et la reconnaissance des sons composant l'ensemble de la scène ; l'intérêt de ces études n'est absolument pas mis en cause, mais le but est de suggérer une approche multi échelle comportant une approche à une échelle « microscopique » permettant l'analyse fine d'objets sonores (description des notion d'attaque, de transitoire, de timbre...), une approche à l'échelle de l'objet sonore (description morphologique et taxonomique du son) et une approche « macroscopique » (description du son en tant que signifiant et signifié). Les approches combinées permettront ainsi de modéliser des processus de choix entre les différents aspects de la perception : les sens fournissent une variété de représentations et liens entre les différents éléments perçus de manière à faire un tri et orienter notre perception vers une vision globale de l'ensemble.

L'application d'un modèle de classement va jouer sur la sémantisation des informations sonores. Le modèle appliqué pourra alors faire émerger, à force d'utilisation, la cohérence entre les informations ressenties et les informations classées dans la Base de données ne devraient donc que s'améliorer avec l'expérience du logiciel et donc du modèle de Base de donnée développé. Cette hypothèse pourrait s'accorder avec l'implémentation de ce type de modélisation à l'aide de cartes d'auto-organisation.

La question est donc maintenant de comprendre comment ces représentations mentales peuvent-elles se retrouver au niveau neuronal et donc essayer de tenter de comprendre quelles formes elles pourraient prendre. Un des champs d'investigation possible pour cette question serait aussi l'étude de sujets synesthésiques présentant une synesthésie dont le point de départ serait une source sonore.

5 Conclusion

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 avons pu aborder 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, rejoignant ainsi une représentation prototypique des sons en mémoire comportant un versant perceptif et sensible (du point de vue de la morphologie) et un versant fonctionnel (taxonomie, et signification du son).

Si les représentations physicalistes sont essentielles et nécessaires à la description des sons perçus en contexte, 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.

L'analyse des données émotionnelles ne permet pas de conclure, elles nous guident cependant à soutenir les hypothèses neurobiologiques et neuropsychologiques de modulation du traitement cognitif par les émotions et de modulation des émotions par le contexte d'expérimentation proposé ; il paraît cependant difficile d'établir un lien entre perception d'un son et l'émotion provoquée.

Dans de futures recherches, il pourrait donc paraître intéressant de développer des modèles informatiques simulant un apprentissage perceptuel des configurations et catégorisations morphologiques des objets sonores. Ces catégorisations sont basées sur les classifications typo- morphologiques proposées par Pierre Schaeffer. Car il semblerait que seulement une approche dans ce sens pourra s'occuper de la perception auditive de façon plus large et générale, en tenant compte non seulement de la musique tonale mais aussi de la musique contemporaine, de la perception du langage et de la perception des scènes auditives environnantes.

Dans le cadre du sujet posé, il paraît clair que les sons ont une forme physique, une forme symbolique et qui plus est une forme qui va évoluer selon le contexte, ainsi que la mémorisation et la sémantisation de ces informations. Le sujet est évidemment beaucoup plus large et complexe que cette infime partie qu'on a abordée dans ce texte. Mais l'interrogation qui émerge ici est la suivante : dans quelle direction pouvons-nous bien mener nos recherches sur la modélisation informatique de la perception auditive et musicale ? Question sine qua non pour le développement de l'intelligence artificielle dans ce domaine. Certains aspects ici effleurés sont visiblement encore ignorés – ou peut-être regardés avec une certaine négligence – par certaines approches. Notamment les approches purement computationnelles (symboliques) dans la modélisation des traitements du langage. Ces approches, dans l'idée de combler l'écart entre la nécessité d'un niveau sémantique (représentationnel) et la contrainte exigeant que ce niveau soit, après tout, physique, introduisent les symboles à la fois pourvus de signification et physiques. Mais comment, dans cette stratégie, les symboles acquièrent-ils leur signification dans un apprentissage perceptuel ? Sans les contraintes de la programmation de l'ordinateur et la prédéfinition de l'ensemble des stimuli, les symboles n'acquièrent pas de signification. Or, la modélisation mixte et à plusieurs niveaux peut, comme on a essayé de démontrer, combler ce déficit. Les recherches de Leman, Carreras et Mattusch ont déjà produit des résultats que nous permettent d'affirmer que la modélisation informatique de la perception auditive et musicale doit se faire dans ce sens.

6 Bibliographie

- 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.
- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22, 261-295
- Bartlett, F. C. (1932; reprint 1964). *Remembering: A study in experimental and social psychology*. London: Cambridge University Press.
- Besson, M. and Faïta, F., (1995) An event-related potential (ERP) study of musical expectancy: Comparison of musicians with nonmusicians, *Journal of Experimental Psychology: Human Perception and Performance*, 21, 1278-1296.
- Bigand, E., Madurell, F., Tillmann, B. and Pineau, M., (1999) Effect of global structure and temporal organization on chord processing, *Journal of Experimental Psychology: Human Perception and Performance*, 25, 184-197.
- Bonnet, C., (1986), *Manuel pratique de psychophysique*, Collection U - Série Psychologie, Armand Colin Editeur, Paris.
- Collins, A.M. & Loftus, E.F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- Damasio, A. R., & Blanc, M. (2001). *L'erreur de Descartes*: Paris: Odile Jacob.
- 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.
- Demany L., Montandon G., Semal C., (2004). Pitch perception and retention: two cumulative benefits of selective attention. *Perception and Psychophysics*, 66, 609-617.
- 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-Fassina, A., et al. (dir.), Paris, Octares Editions, p. 79-93.

- Edeline J.M., (2003) The thalamo-cortical auditory receptive fields: regulation by the states of vigilance, learning and the neuromodulatory systems. *Exp. Brain Res.*, 153, 554- 572.
- 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.
- Gaver, W. W., (1993), What in the World Do We Hear ?: An Ecological Approach to Auditory Event Perception, *Ecological Psychology*, 5 (1), 1-29.
- Giraud, A-L, Truy, E., and Frackowiak, R. (2001). Imaging plasticity in cochlear implant patients, *Audiology Neurootol.*, 6, 381-393.
- 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).
- Kaas, J. H. (1991). *Ann. Rev. Neurosci.*, 14, 137-167.
- Kohonen, T. (1995). *Self-Organizing Maps*. Springer: Berlin.
- Lass, N. J., et al., (1982), Listeners' Identification of Environmental Sounds, Perceptual and Motor Skills, 55, 75-78.
- Leman, M. (1995). *Music and Schema Theory*. Springer: Berlin.
- Leman, M. & Carreras, F, (1997) « Schema and Gestalt :Testing the Hypothesis of Psychoneural Isomorphism by Computer Simulation », *Music, Gestalt, and Computing*, Springer-Verlag, Berlin.
- Maess, B., Koelsch, S., Gunter, T. and Friederici, A.D., (2001) 'Musical syntax' is processed in the Broca's area: An MEG-study., *Nature Neuroscience*, 4, 540-545.
- Mattusch, U., (1997), « Emulating Gestalt Mechanisms by Combining Symbolic and Subsymbolic Information Processing Procedures », *Music, Gestalt, and Computing*, Springer-Verlag, Berlin.
- Recanzone, G. H. et al. (1992b) *J. Neurophysiol.*, 67, 1057-1070.
- Recanzone, G. H., et al. (1992c) . *J. Neurophysiol.*, 67, 1031-1056.
- Recanzone, G. H., et al. (1992d) . *J. Neurophysiol.*, 67, 1071-1091.
- Recanzone, G. H., Schreiner, C. & Merzenich, M. M. (1993) *J. Neurosci.*, 13, 87-103.

- Regnault, P., Bigand, E. and Besson, M., (2001), Event-related brain potentials show top-down and bottom-up modulations of musical expectations, *Journal of Cognitive Neuroscience*, 13 241-255.
- 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.
- Schaeffer, P., (1966), *Traité des objets musicaux*, Collection Pierres Vives, Editions du Seuil, Paris.
- Semal C., Demany L., Ueda K., Hallé P.A., (1996). Speech versus nonspeech in pitch memory. *Journal of the Acoustical Society of America*, 100, 1132-1140.
- Seger, C. A. (1994). Implicit learning. *Psychological Bulletin*, 115, 163-169.
- Theile, G., (1980), *Untersuchung zur Wahrnehmung der Richtung und Entfernung von Phantomschallquellen bei 2-Kanal-Stereofonie*, Institut für Rundfunktechnik IRT, n°24/80, Francfort.
- Tillmann, B., Bharucha, J.J. and Bigand, E., (2000) Implicit learning of tonality: A Self-Organizing approach, *Psychological Review*, 107, 885-913
- Vanderveer, N. J., (1979), *Ecological Acoustics: Human Perception of Environmental Sounds*, Dissertation, Dissertation Abstract International, n°4543B (University Microfilms n° 8004002).
- Weil-Barais, A., (1993), *Repères théoriques*, in *L'homme cognitif*, 4ème édition de 1998, Paris, Presses Universitaires de France, p. 33-64.
- Weinberger, N. M. (1995) *Ann. Rev. Neurosci.*, 18, 129-158.
- Zwicker, E., (1982), *Psychoakustik*, Hochschultext, Springer Verlag, Berlin.