

CLAUDE Olivier

La recherche intelligente de sons



**Université de Provence
Département SATIS
IUP Métiers de l'Image et du Son
Mémoire de maîtrise**

2005-2006

UNIVERSITE DE PROVENCE

MEMOIRE DE RECHERCHE

Pour l'obtention du diplôme de Maîtrise

**Discipline Sciences Art et Technique de l'Image et du Son
Option Son**

Présenté et soutenu par

Olivier CLAUDE

Juin 2006

TITRE :

La recherche intelligente de sons

Travail réalisé sous la direction de Rémi Adjiman

REMERCIEMENTS

Je tiens à remercier tout particulièrement Nicolas Misdariis et Patrick Susini, chercheurs à l'IRCAM au département Design Sonore, qui m'ont accueilli dans leurs locaux, dans le cadre d'un stage sur la recherche des sons environnementaux non-musicaux à travers une base de données audio. Ils m'ont suscité la volonté d'écrire ce mémoire sur la recherche intelligente de sons.

Remerciements également à Rémi Adjiman, mon directeur de mémoire, qui a suivi l'avancé de mes travaux et soutenu ma réflexion.

Un grand merci aussi à Matthias Garcia et toute l'équipe informatique de 802point11 qui ont travaillé à la réalisation d'un moteur de recherche suivant le modèle théorique de ce mémoire.

Résumé

Les “bruits” de tous les jours qui nous entourent, les sons de la ville, de nos maisons, de notre Nature, de nos machines, des animaux, les bruits propres à l’homme forment notre environnement sonore quotidien et familier. Contrairement aux sons musicaux que l’on peut classer par tonalité, fréquence, gamme, genre musical, instrument...les sons environnementaux sont plus complexes à organiser, rassembler, décrire, classifier de part les multiples façons de les générer (multitude des sources sonores, multitudes des fréquences écologiques).

Les mots qualificatifs, descriptifs des sons sont très peu employés dans notre quotidien, nous gardons le silence sur la forme, la substance auditive que nous entendons comme représentation d’une cause, d’un sens.

La multiplication des données informatiques nécessite des interfaces de recherches pertinentes navigants dans les bases de données audio. C’est pourquoi, il devient aujourd’hui un besoin dans les structures de post production audiovisuelle, d’établir une théorie pour la description des sons à travers des critères posés. Savoir mettre des mots aux sons et des sons aux mots.

Mots clés : Sons environnementaux non-musicaux, classification, description, base de données audio, moteur de recherche, objet sonore, écoute réduite, perception cognitive, représentation sémantique.

SOMMAIRE

Introduction	11
---------------------------	----

1. Classer les sons

Organisation de la taxonomie.....	12
Taxonomie des sons de la Nature.....	12
Taxonomie des sons des Animaux.....	16
Taxonomie des sons de l'Homme.....	19
Taxonomie des sons Objets/Machine.....	20
Conclusion : classification des sons environnementaux.....	35

2. Description des sons

Introduction.....	36
Les trois écoutes.....	37
Critères de description : Sphère morphologique.....	38
Sphère causale.....	45
Sphère sémantique.....	47

3. Mise en œuvre du système de recherche

Protocole de prise de sons.....	53
Protocole d'intégration et d'indexation des sons.....	57
Conclusion	60
Bibliographie	61

INTRODUCTION

Le projet de recherche intelligente des sons environnementaux non-musicaux est la poursuite du projet ECRINS mené à l'IRCAM par Emmanuel Deruty, il vise à aboutir à un outil de recherche intelligente des sons environnementaux non musicaux par le biais d'une interface de navigation donnant accès à une base de donnée d'échantillons sonores et s'appuyant sur des critères descriptifs des sons. Cet outil est destiné aux professionnels de la production sonore (musique, post-production image, multimédia, design sonore industriel, illustrateur sonore...).

Les « bruits » de tous les jours qui nous entourent, les sons de la ville, de nos maisons, de notre Nature, de nos machines, des animaux, les bruits propres à l'homme forment notre environnement sonore quotidien et familier. Contrairement aux sons musicaux que l'on peut classer par tonalité, fréquence, gamme, genre musical, instrument...les sons environnementaux sont plus complexes à organiser, rassembler, décrire, classifier de part les multiples façons de les générer (multitude des sources sonores, multitudes des fréquences écologiques. On trouve, comme état des lieux des organisations de banques de sons, des systèmes encore indirects par rapport au concept d'écoute réduite et d'objet sonore de Pierre Schaeffer ; au mieux on a une description hiérarchique de la cause de production du son nommé arbitrairement par le producteur du son. Les mots qualificatifs, descriptifs des sons sont très peu employés dans notre quotidien, nous gardons le silence sur la forme, la substance auditive que nous entendons comme représentation d'une cause, d'un sens.

Le projet vise à proposer une recherche des sons en prenant en compte plusieurs aspects de leur nature même : physique, perceptif et sémantique. Ainsi, une recherche peu s'appuyer sur plusieurs attributs du son et répondre à une demande plus précise. L'intérêt demeure là, dans un gain de temps dans le processus de travail de la production sonore, par la limitation d'écoutes de sons inutiles qui correspondraient à une recherche moins descriptive. Dans le support de recherche, la base de donnée informatique apporte une interactivité plus ergonomique et rapide qu'une liste de sons consultables sur papier ou fichier informatique, l'interface graphique donne une réponse et un accès à l'utilisateur à tous les critères de recherche du son. L'organisation des sons à l'intérieur de la base de données se fait par une taxonomie des sons s'appuyant sur la base de l'automate de Quilliam (1969) et de l'étude de la catégorisation d'un corpus de bruits domestiques de Frédérique Guyot, Michèle Castellengo et Benoît Fabre, permettant une hiérarchisation logique des différents types de sons et une description textuelle. Chaque son de cette taxonomie est aussi classé en fonction de sa nature physique, perceptive et sémantique. On associe donc à un son plusieurs attributs qui sont les critères de recherche dans l'interface de la base de donnée. Chaque critère sera associé à un référent, un prototype qui permettra de mettre en relation et de comparer les descriptions de sons.

La tâche n'est pas simple, car elle est encore aujourd'hui très peu explorée, nous tenterons néanmoins, à l'aide de nombreux fondements théoriques et de mon expérience de l'écoute et de la chaîne de production du son, d'établir, tout au long de ce mémoire, une base ouverte à une approche de classement et de description des sons non-musicaux.

1. Classifier les sons

Pour parler des sons de tous les jours on qualifie généralement les phénomènes sonores de « bruits ». Contrairement au terme de « timbre » employé pour décrire la musique et la parole, les sons environnementaux ne sont pas assujettis à un vocabulaire rigoureux prenant en compte la notion d'objet sonore. Pourquoi sont-ils ainsi nommés « bruit » ? Car ils ont la différence de ne pas avoir une hauteur précise et déduite, d'être formés de multiples fréquences irrégulières et non cycliques, d'être complexes en termes schaeffériens.

Pour cette recherche de classification des sons, une écoute de plus de 3000 sons tirés des *Sound Ideas serie 6000*¹ a été préalable. Puis la réflexion s'est orientée vers une taxonomie² où les sons sont classés par rapport à leurs sources dans notre environnement, et de leur relation directe avec notre espace, notre univers, notre culture, notre Nature. L'homme s'est créé au cours de son évolution dans le temps un vocabulaire pour donner des noms à des événements naturels, à des objets, des comportements, des actions, des sentiments... Mettre en relation le vocabulaire courant visuel pour parler d'un son (un son de voiture, un son de machine, un son de la nature, un son de l'homme...) et en déduire en analogie un classement hiérarchique et regroupé de tous les sons qui peuvent exister sur notre planète, et tous les sons qui concernent ce que l'Homme sait faire dans tous les domaines, ainsi que les animaux et la Nature.

Organisation de la taxonomie

Comment regrouper les sons dans différentes catégories ? Lesquelles ? Avec quelle logique ? Dans quel ordre hiérarchique ?

Il existe quatre grandes familles de sources sonores : les sons provenant de la **Nature**, les sons produits par les **Animaux**, les sons produits par l'**Homme** et les sons des **Objets/Machines**.

1.1 Taxonomie des sons de la Nature

Pour les sons concernant la Nature (vent, forêt, ruisseau, océan, l'orage, la pluie...), il a été adopté une classification par élément naturel : l'eau, l'air, la terre, le feu. Ainsi à chaque événement sonore naturel, correspond une appartenance à un des quatre éléments qui constitue la Nature. On peut donc associer à l'**EAU** : la rivière, le lac, l'océan, une chute d'eau, la pluie, le marais... à l'**AIR** : le vent, l'orage (issue d'une turbulence atmosphérique), à la **TERRE** les différentes zones géographiques : la forêt, la jungle, la campagne, le désert, la montagne... et au **FEU** : les feux de forêts.

Les sous ordres³ reprennent une logique qui vient du son à l'image, et des termes employés pour décrire le contenu des événements sonores. Sur le tournage d'un film, on parle de sons seuls et d'ambiances. La prise d'un son seul fige un **événement** sonore court, bref (ex : une pierre qui tombe dans l'eau), au contraire de la prise de son d'**ambiance** qui fige un environnement sonore et qui dure souvent plus d'une minute. (ex : pluie avec de l'orage, vent). Il est utilisé aussi des termes qui décrivent le temps, le moment de la journée (ex : le

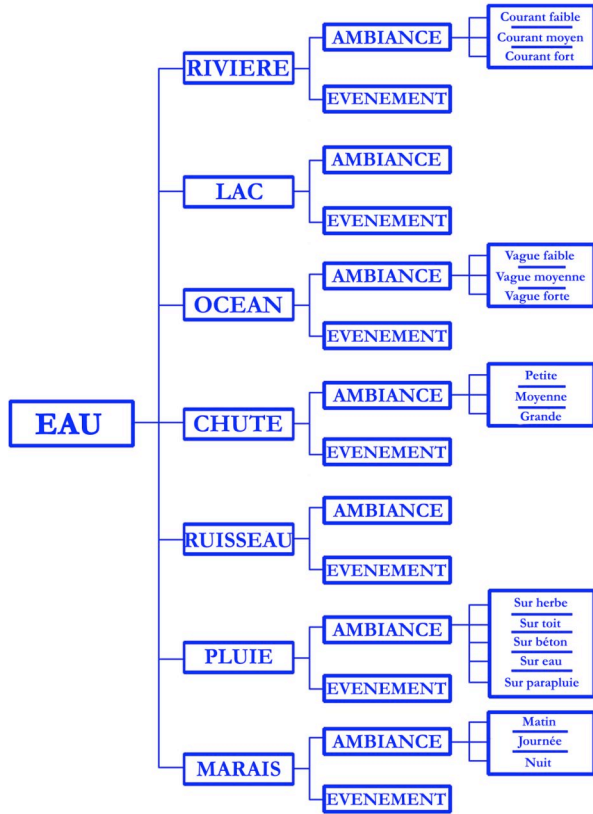
¹ Sonothèque commerciale américaine

² Ou taxinomie, science de la classification

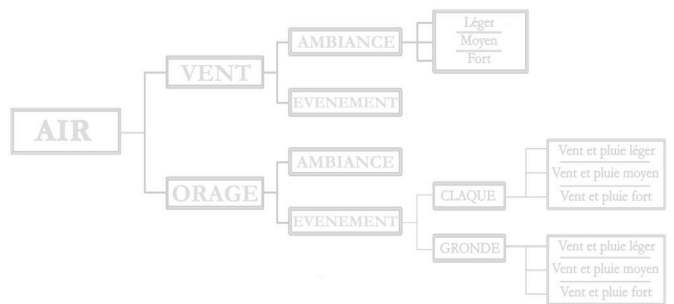
³ Déclinaison hiérarchique de la classification

matin, en journée et la nuit) où la prise de son a été effectuée. On exprime aussi la densité sonore émise par la source (ex : vent léger, moyen , fort).

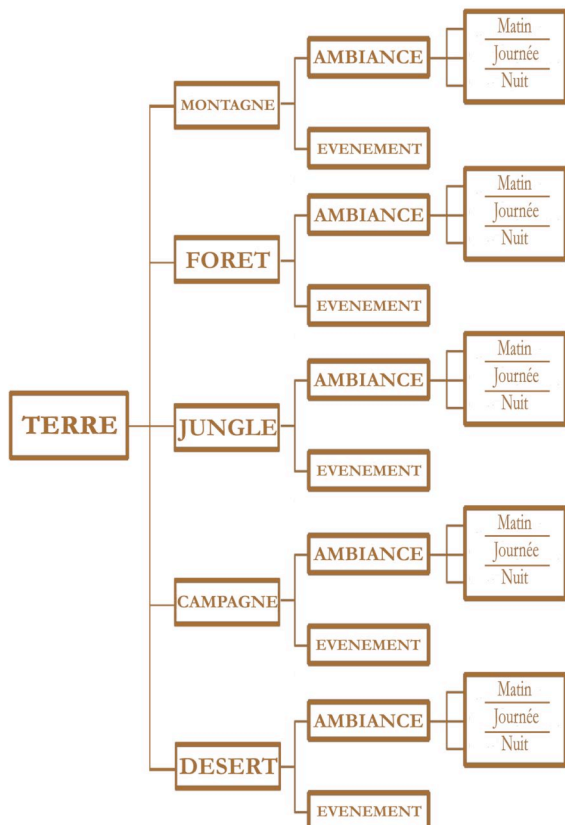
NATURE



Taxonomie des sons de la Nature liés à l'eau



Taxonomie des sons de la Nature liés à l'air



Taxonomie des sons de la Nature liés à la Terre



Taxonomie des sons de la Nature liés au Feu

Il faut considérer cette classification comme une arborescence descendante, c'est à dire informatiquement elle correspond à une hiérarchie de dossiers imbriqués.

On constate que quatre niveaux apparaissent dans la classification, quatre niveaux qui nous permettent d'atteindre le son lui-même, qui nous permettent de le classer suivant cette logique.

- Le **premier niveau** considère la **source** du son, ici nous avons en l'occurrence les quatre éléments qui forment la Nature : l'eau, l'air, la terre, le feu.

- Le **deuxième niveau** prend en compte le **lieu de la source**, il se réfère aux différents caractères des zones naturelles géographiques et climatologiques de la Terre. (ex : désert, montagne, lac, océan...).

- Le **troisième niveau** différencie la durée et l'intention de la prise de son, le son peut être une **ambiance** (plus d'une minute), dans lequel la prise de son prend un plan large sonore sans intention de focalisation ; ou soit un **événement** qui focalise une action sonore brève dans un environnement naturel.

Le **quatrième niveau** relate la densité du son , la dénomination est très variée, elle est fonction du deuxième niveau, c'est à dire qu'elle reste logique, cohérente à l'image sonore collective des sons des éléments naturels, elle doit donner des informations générales sur le contenu du son. C'est l'interaction sonore caractéristique du deuxième niveau qui est fonction du quatrième niveau, c'est à dire la composante sonore qui définit le mieux une ambiance. Pour une rivière ou l'océan nous parlons de courant faible, moyen, fort ou de vagues faible, moyenne , forte, c'est l'intensité du courant ou la force des vagues qui caractérisera les ambiances de rivières ou de l'océan. La **densité** du son dépend du nombre de sources différentes incluses dans une prise de son d'ambiance, elle représente la notion de « masse »⁴ en terme Schaefferien. . Elle définit l'image sonore que nous renvoi des événements comme une pluie forte ou une pluie faible :

- Une pluie forte est formée d'une succession multiple et très rapide de chocs d'eaux, le rythme est rapide et donne une « masse » sonore forte.
- Une pluie faible est rythmée plus lentement, le choc des gouttes d'eaux est plus lent, la « masse » sonore est faible.

Bien entendu, d'autres caractéristiques peuvent être présentes dans l'ambiance (ex : du vent, des mouettes...), j'appellerai ces **composantes sonores de discrètes**⁵, néanmoins les composantes discrètes peuvent apparaître dans le nom des fichiers informatiques ou dans l'ajout d'autres sous niveaux. Dans des ambiances relatives au son de l'élément terre nous utiliserons des termes temporels, car ces ambiances sont fonctions de multiples composantes discrètes qui nous apparaissent comme un tout sonore, nous identifions culturellement les sonorités diurne et nocturne. Exemple d'une composante discrète qui entre en compte dans une structure sonore : la journée nous avons plus d'activité animale que la nuit et les sonorités du jour (oiseaux chanteurs, déplacement banc d'oiseaux ...) sont différentes de celles de la nuit (hibou...).

⁴ Pierre Schaeffer Traité des Objets Musicaux 1966-Critère Morphologique du son, spectre de fréquence complexe

⁵ Prend place dans une structure de sons assemblés pour former une ambiance sonore

Le quatrième niveau correspond au classement final qui dans la lecture progressive de la classification nous donne le renseignement essentiel qui correspond à la représentation sonore. En pratique, il est probable que ce quatrième niveau correspondra au nom du fichier informatique.

Des éléments n'ont volontairement pas été inscrits dans la taxonomie, mais qui ont tout à fait leur place, mais à ma connaissance nous n'avons pas encore de sons de tremblements de terres (Terre), d'éruptions volcaniques ou coulées de magma (Feu), de raz de marées (Eau), de typhons (Air) ; des prises de sons qui demandent une grande prise de risque.

Ici nous cherchons uniquement à créer, à proposer une méthodologie de rangement des sons par catégorie de niveau naturel, de lieu, de durée et de densité. Ce mémoire proposera par la suite d'attribuer une description plus précise de l'objet sonore, restons pour l'instant à l'état de la classification des sons.

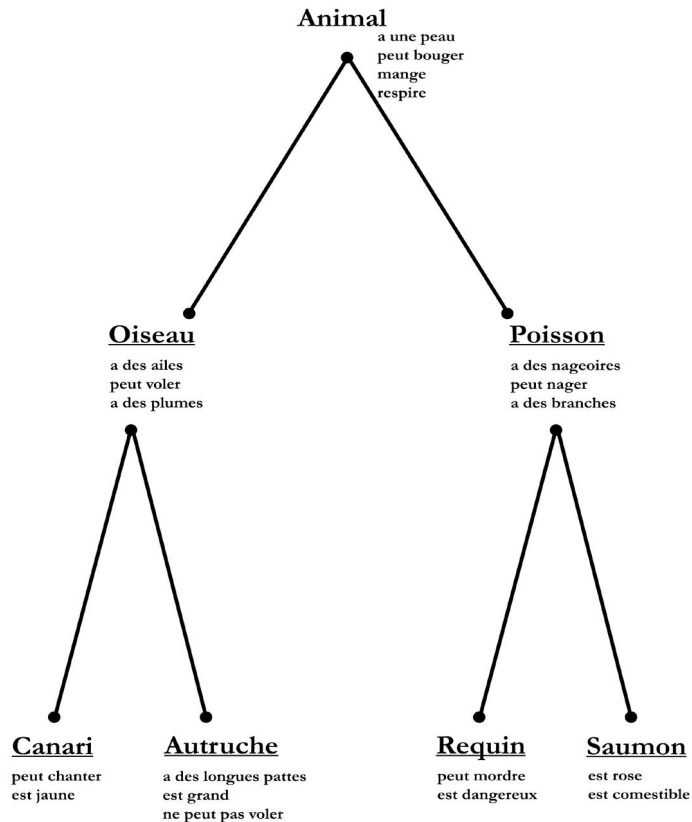
1.2 Taxonomie des sons des Animaux

Nous allons désormais étudier une classification de la deuxième catégorie des sources sonores : les **Animaux**. Notre but est d'aboutir à un rangement des matières sonores efficace et le plus commun possible à nos écoutes. Ma première piste fut d'exploiter les classifications déjà existantes et connues des sciences naturelles, comme carnivore, mammifère arboricole, mammifère herbivore, amphibien, rongeur, oiseau, insecte, pinnipède... Mais le résultat ne fut pas satisfaisant, les termes me semblaient trop scientifiques et pas assez parlants pour l'assimiler à une source sonore. C'est pourquoi j'ai orienté ma réflexion sur comment classer les sources sonores produites par les animaux dans une arborescence où la compréhension serait commune à tous.

1.2.1 Automate de Quillian (1969)

Quillian⁶ s'est posé en 1969 la question de savoir quel type d'organisation en mémoire sémantique permet à l'être humain de connaître et d'utiliser des notions comme il le fait. Il propose une structure concernant l'appartenance catégorielle et les propriétés cognitives⁷ autour desquelles se forme la mémoire chez l'Homme. Exemple, un canari est un animal, un canari est un oiseau, un canari est jaune, un canari peut chanter, un canari peut voler. L'automate de Quillian est un modèle de la mémoire sémantique humaine.

Automate de QUILLIAN (1969)

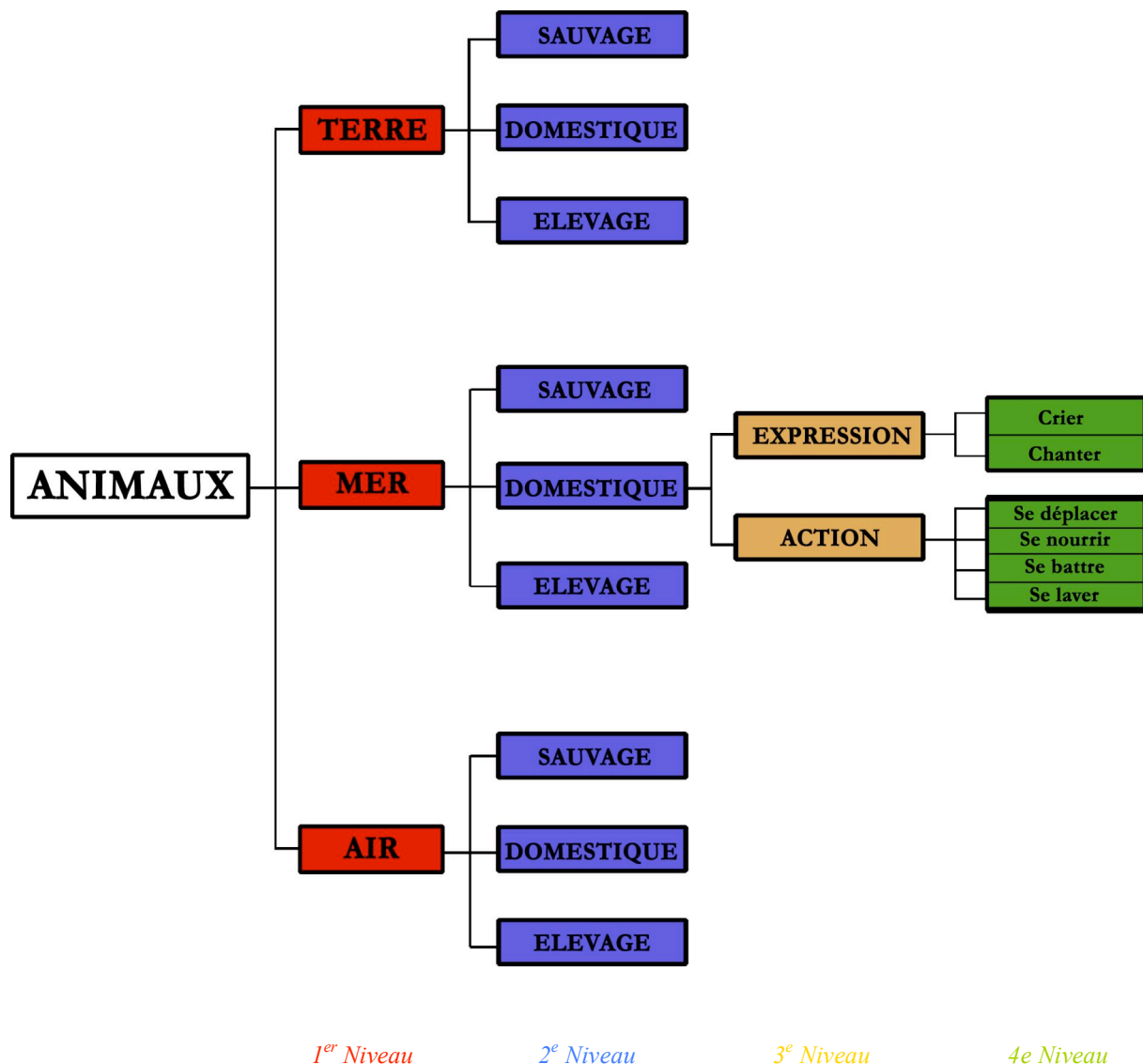


⁶ Quillian R. *Semantic memory* 1969

⁷ intelligente

Cette automate donne beaucoup d'informations sur ce que mémorise l'Homme pour identifier et reconnaître un animal. Ici, *oiseau* et *poisson* sont considérés comme des prototypes⁸, instance centrale de la catégorie qui la représente le mieux. Ce qui organise notre représentation mentale des catégories. Maintenant, si nous transposons cette automate pour l'appliquer à la mémorisation sémantique des sons d'animaux chez l'Homme, nous trouvons deux propriétés responsables de la production du son chez l'animal, ses **expressions** et ses **actions**. C'est à dire, les cris, les chants, se battre, se déplacer, se nourrir, se laver... Dans le schéma suivant, nous trouverons comment l'application de l'automate de Quillian aux propriétés sonores des animaux, prend place dans la taxonomie de la deuxième catégorie des sources sonores :

1.2.2 Taxonomie des sons liés aux animaux



Taxonomie des sons liés aux Animaux

⁸ organise notre représentation mentale des catégories

On constate quatre niveaux dans la classification :

- Le **premier niveau** considère l'espace d'évolution de l'animal, c'est à dire l'eau pour les poissons, l'air pour les oiseaux, la terre pour les animaux terrestre (mammifère, insecte...). Ces dénominations permettent de simplifier et de regrouper l'intégralité des animaux sans passer par la classification de la science naturelle. Si un animal peut évoluer sur terre et dans l'eau (ex : un crocodile) ou évoluer sur terre et dans l'air (ex : un oiseau), nous le classerons en fonction de sa seconde caractéristique de milieu d'évolution, la terre faisant office de frontière commune entre l'eau et l'air.

- Le **deuxième niveau** prend en compte la condition et la place de l'animal dans notre monde, dans notre société, dans notre culture et notre mode de vie. L'animal peut vivre dans la nature librement sans être dompté par l'Homme, il sera donc classé dans la catégorie *Sauvage* (ex : chimpanzé, corbeau, pigeon, loup, grenouille...) . Dans la catégorie *Domestique* sont regroupés les animaux sauvages apprivoisés et qui vivent en contact avec l'Homme (ex : chien, chat, cheval...) . Et enfin, la troisième classe des animaux est celle de ceux exploités par l'Homme pour son alimentation, qui vivent conditionnés dans un lieu d'élevage, nous trouvons dans la catégorie *Elevage* poule, oie, poussin, coq, canard, cochon, vache...

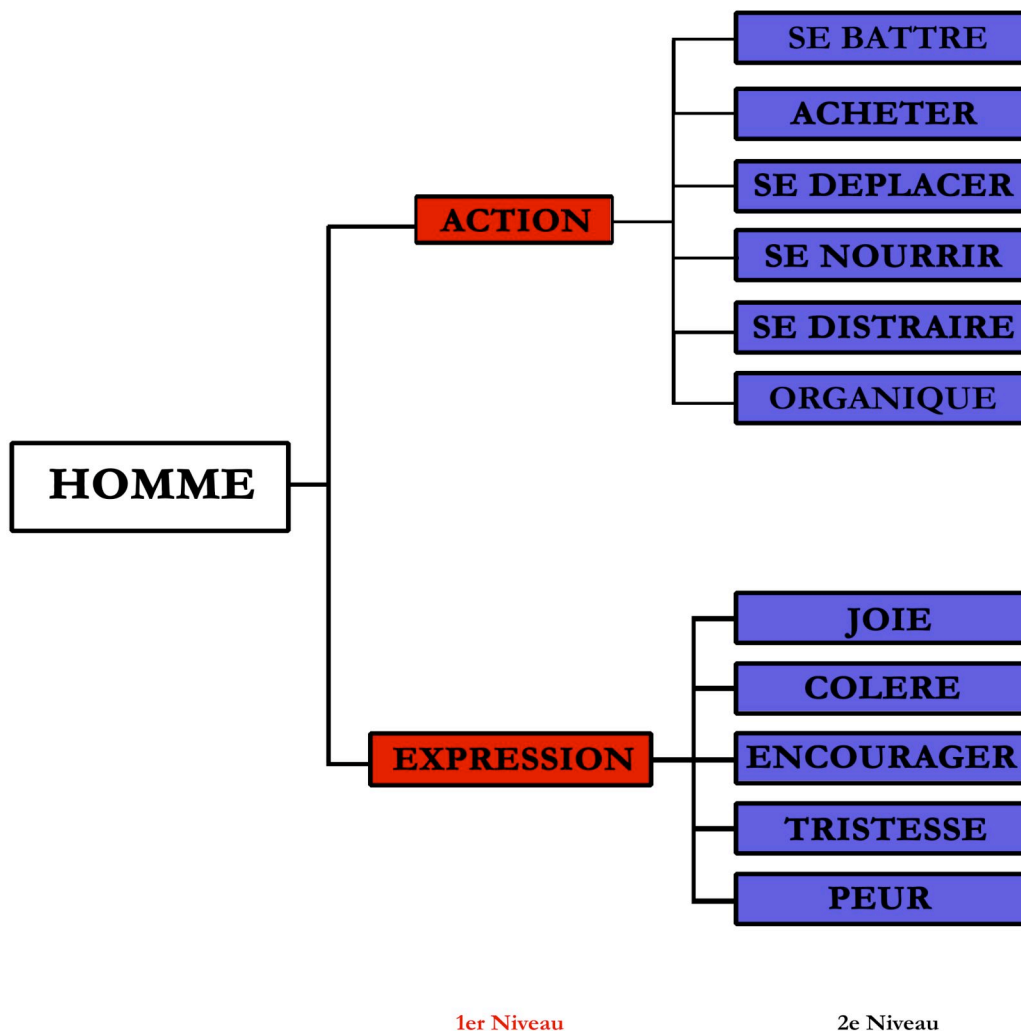
- Le **troisième niveau** reprend la transposition du modèle de Quillian en présentant les deux propriétés d'émission des sons originaires des animaux : leurs différentes *Expressions* et *Actions* .

- Le **quatrième niveau** classe le son par rapport à son intention d' Expression (crier, chanter) et d' Action (se déplacer, se nourrir, se battre, se laver). Ce quatrième niveau, en pratique, peut correspondre au nom du fichier audio.

Cette arborescence me semble la plus représentative et la plus directe pour envisager de classer toutes les sources sonores animales. De plus, la directive que nous avons entreprise de l'automate de Quillian semble être efficace, car par la suite nous l'appliquerons aux sons de l'Homme, ainsi que le concept de densité vu à la classification des sons de la Nature. La voie de la classification entreprise semble être fiable puisqu'elle va rejoindre des notions communes entre les grandes catégories de sources sonores.

1.3 Taxonomie des sons de l'Homme

Auparavant, nous parlions de classer toutes les sources sonores présentes sur notre planète, nous avons vu les sons de la Nature et ceux des animaux. Désormais, la recherche d'une classification s'axe sur les sons dont l'Homme est la source sonore. La tâche est complexe, il se faut pour la traiter émettre des distinctions et définir ce que sont les sons de l'Homme, j'insiste sur le terme « Homme » avec une majuscule car une des premières orientations de la classification qui va se présenter au fur et mesure de cette partie, s'axe sur les sources sonores produites par l'homme, animal doué d'une conscience. Il est placé au centre d'une écoute sémantique⁹, à travers ses actions, ses expressions, et des sons de son corps. Nous devons diriger notre oreille vers l'Homme et considérer la classification avec notre propre conditionnement de réception des impressions des sons, de l'entendre, notre représentation sonore des actions et des expressions, tel le décrit cette arborescence :

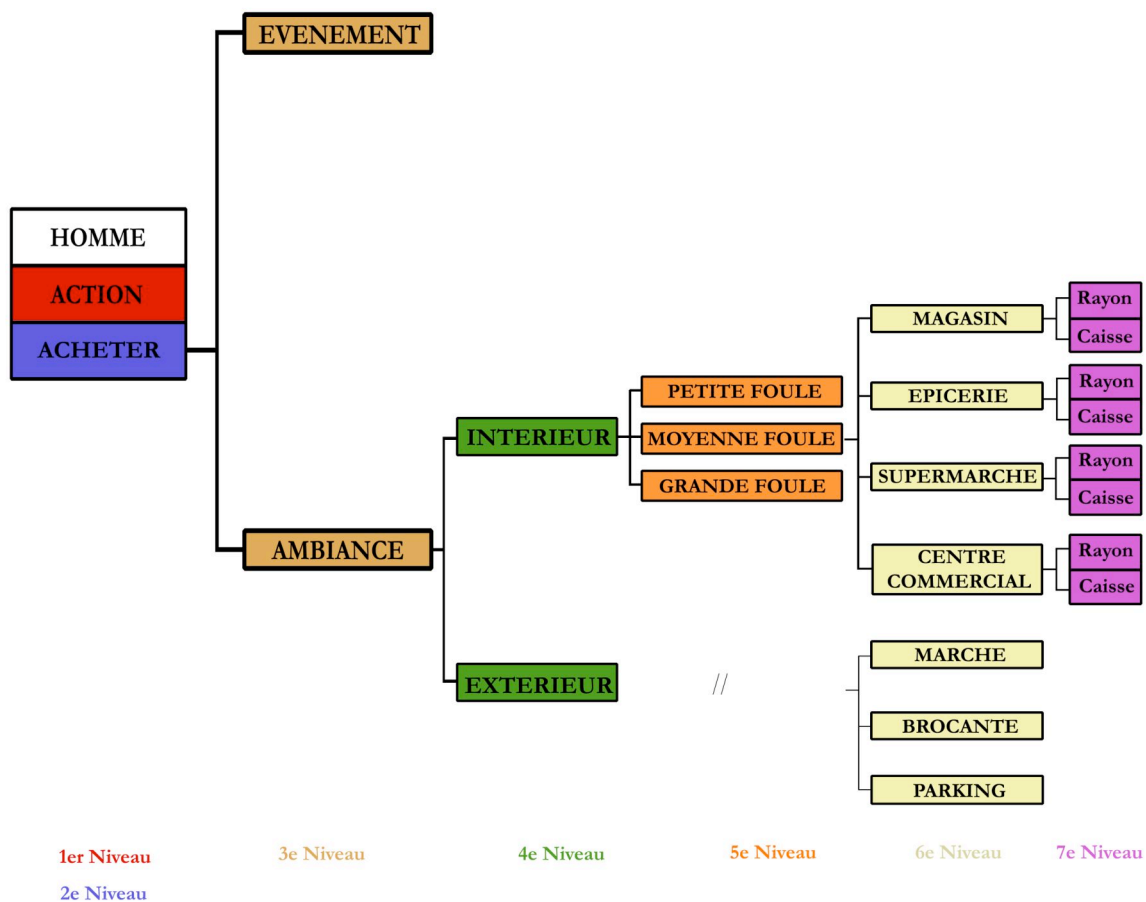


Taxonomie du 1er et 2e niveau des sons de la catégorie Homme

⁹ La classification se base sur le sens du son que produit l'Homme

On retrouve un **premier niveau** similaire à l'intention sémantique vu avec l'automate de Quillian : Action – Expression. Le niveau Action se divise et contient les sons liés à des actions sémantiques de l'Homme, ils sont classés par un verbe à l'infinitif qui définit l'action propre (ex : se battre, se nourrir ...). On trouve au **deuxième niveau**, en plus des verbes d'actions, une branche « organique ». Que devons-nous écouter par son d'action organique de l'Homme ? Il s'agit de tous les sons originaires des réflexes naturels et caractéristiques des sons dont la source est l'interaction entre le système biologique et le corps humain. Exemple, uriner, déféquer, péter, roter, vomir, renifler, se moucher, se reproduire, respirer, tousser, éternuer, battements de cœur. Ces sons peuvent faire sourire, mais ils sont communs à tous les hommes et il faut les distinguer des expressions. Toutes ces actions à l'infinitif sont accompagnées dans leurs définitions du dictionnaire de la notion « avec le bruit ». Un son organique est émit d'une action naturelle de l'Homme, la source est son système biologique. Le niveau Expression classe les sons qui définissent un état d'expression, de sentiment (ex. : Joie, Encouragement, Peur, Colère...), nous ferons par la suite une étude sur la présence de la parole dans les sons de l'Homme. Nous allons à présent approfondir la taxonomie des sons liés aux actions de l'Homme.

1.3.1 Taxonomie Homme-Action
1.3.1.1 Action - Acheter



Taxonomie des sons liés à l'action de l'Homme d'acheter

Nous avons vu précédemment une explication du 1^{er} et 2^e niveau, explorons désormais les étages suivants de la classification.

Ici, nous abordons l'action d'*acheter*, au **3^e niveau** on retrouve le type du son, s'il s'agit d'une ambiance ou d'un événement de l'action concernée, notion abordée à la taxonomie des sons de la Nature (nous reviendrons sur ce niveau par la suite). Au **4^e niveau**, nous classons l'espace dans lequel la prise de son a été effectuée par *Extérieur* ou *Intérieur*, notion utilisée dans le cinéma pour décrire le lieu de tournage d'une scène : enfermé (Intérieur) ou ouvert (Extérieur).

Au **5^e niveau** apparaît la notion de densité qui fait appelle à l'image sonore de masse d'une petite, moyenne ou grande foule.

Au **6^e niveau**, nous classons les sons par lieu (Marché, supermarché, brocante, épicerie, boucherie...), notion très importante dans la considération de nos représentations sonores.

Et au **7^e niveau**, nous trouvons le point d'écoute d'un lieu qui n'est pas applicable à toutes les situations (fonction de la densité, de l'espace et du lieu), mais qui renseigne sur le contenu du son.

Il y a un point à éclaircir sur le 3^e niveau, je pose une séparation entre les sons représentatifs propres à l'Homme lui-même, qu'il met en émergence tous les jours, formant ses repères sonores inconscients, ses environnements sonores où sa présence est acteur de l'ambiance, et les sons que l'Homme met en action avec un objet ou une machine. La classification de la catégorie Homme prend en compte uniquement l'empreinte sonore focalisée sur les actions et les expressions de l'Homme avec comme source ou corps de résonance, son corps physique.

Définition du niveau Ambiance de la catégorie Homme :

Une ambiance de la catégorie Homme se définit par la présence physique et l'empreinte sonore de l'Homme dans un environnement donné. L'empreinte sonore de l'Homme se définit par des sons d'Actions et d'Expressions corporelles, et où la présence du langage est prise en compte sur l'intelligibilité audible de l'entendre de l'Ambiance.

Les évènements sont des composantes discrètes de l'ambiance sur lesquelles nous nous focalisons pour l'isoler du reste de l'ambiance ; souvent ces composantes discrètes sont originaires de sources en provenances de l'interaction de l'Homme avec un objet ou une machine. Nous excluons dans la catégorie Homme ces sons évènements qui ne se rapportent pas ou plus à la représentation d'une ambiance Action ou Expression de l'Homme.

Définition du niveau Evènement de la catégorie Homme :

Un Evènement de la catégorie Homme est un son bref relatif à une action (ex. : un pas) ou une expression (un rire, une parole intelligible ...) de l'Homme.

Par exemple nous serions tentés de classer dans Acheter - Evènement des sons seuls de pièces de monnaies, de billets, de caisse enregistreuse, de caddies... Mais ces sons sont originaires d'objets ou de machines, il se classe dans la catégorie suivante (Objet – Machine).

Nous allons aborder d'autres exemples qui permettront d'illustrer cette distinction entre les sons propres à l'Homme et ceux qui proviennent d'une interactivité avec des objets ou des machines :

- Ex 1 : Si on veut classer le son d'un homme qui coupe du bois, on peut penser le classer de la façon suivante : Homme – Action – Evènement - Travailler – Couper du bois. L'exemple est démonstratif de la différence à assimiler ; l'Homme ne laisse pas

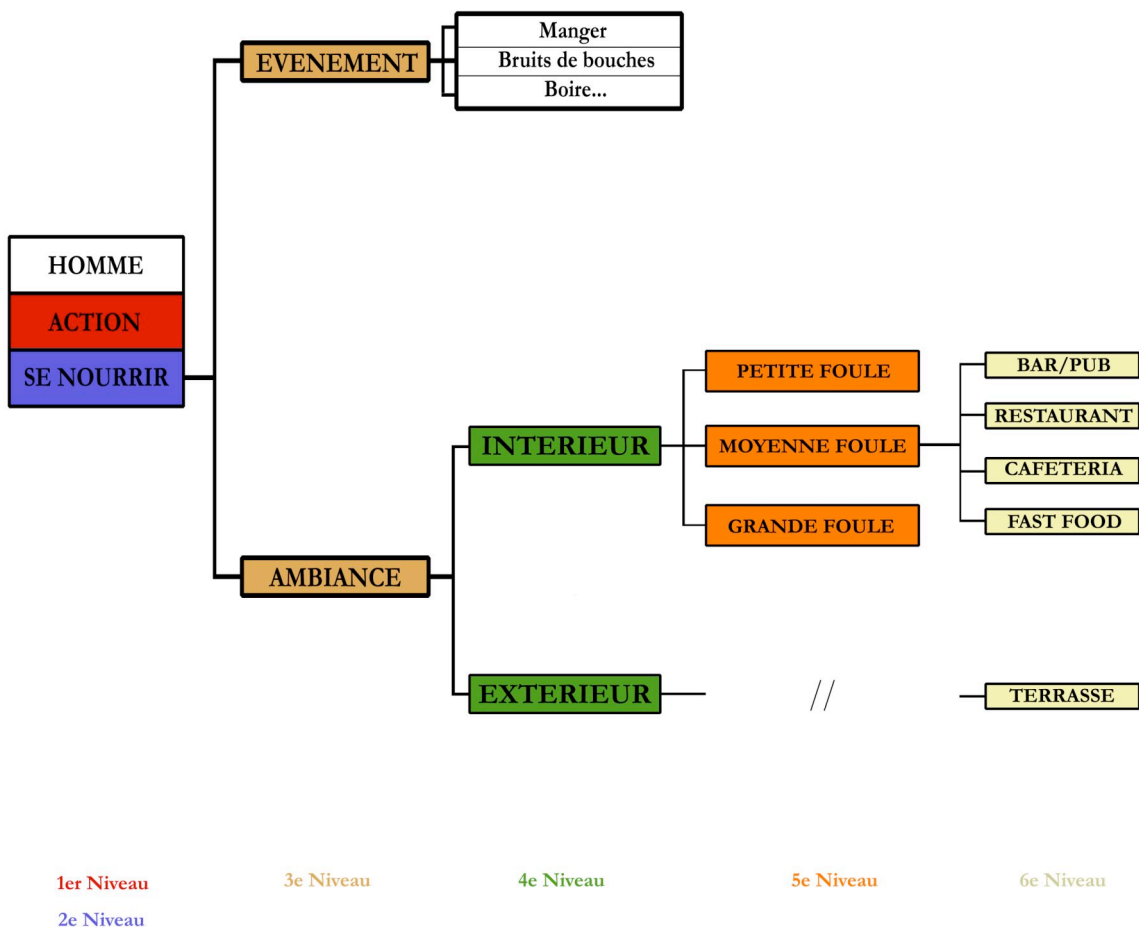
d’empreinte sonore, il n’émet aucune source sonore de part lui-même ; l’empreinte sonore caractéristique est celle du choc entre la hache et le bois, c’est l’objet hache qui focalise et qui définit l’image sonore de cette action. Ce son se classera dans la catégorie Objet-Machine que nous étudierons par la suite.

Dans une Ambiance ou un Evènement de la catégorie Homme, s’il n’y a pas la présence de sons vocaux, de sons corporels à travers une Action et une Expression, le son entre dans la catégorie Objet-Machine.

- Ex 2 : Considérons un son classé de la façon suivante : Homme – Action – se déplacer – Ambiance – Train – Wagon ; si dans l’écoute réduite on entend des passagers, le contrôleur, le vendeur de boissons... le son se classe dans la catégorie Homme. Par contre si on entend uniquement le son caractéristique du train, c’est à dire les roulements du wagon et que l’on n’entend pas la présence des passagers, qu’ils sont muets ; alors le son se classe dans la catégorie Objet-Machine. On parlera de sons d’ambiances ou événements muets de l’Homme.

Cette distinction est primordiale à assimiler, elle régit la logique de la taxonomie. Je continuerai à la justifier avec les classes de sons suivantes de l’Homme.

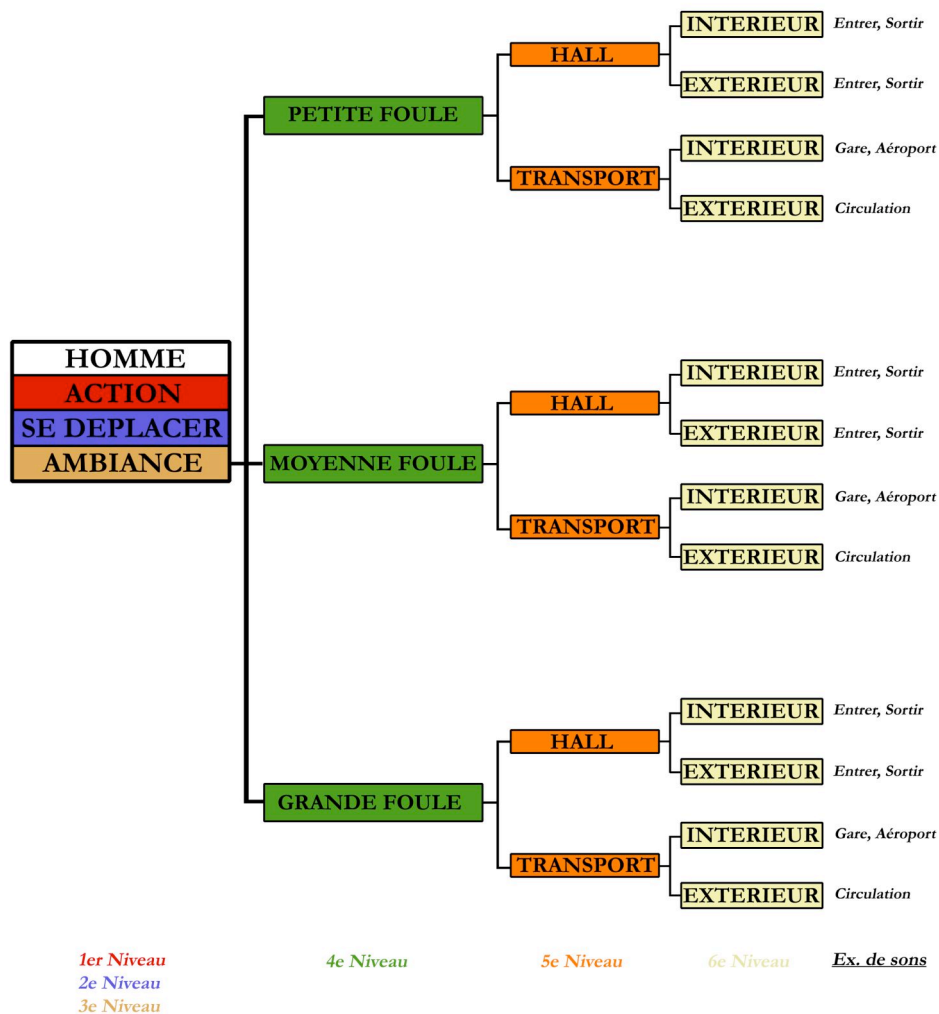
1.3.1.2 Action – Se nourrir



Taxonomie des sons liés à l’action de l’Homme de se nourrir

Ici, dans le classement des sons liés à l'action de se nourrir, on retrouve bien le **3^e niveau** des types de sons « Ambiance » ou « Evènement ». La classe Evènement corrobore la définition vu précédemment, on trouve uniquement des sons corporels de l'Homme (Sons de la bouche, mâcher, boire...), nous ne trouverons pas de sons seuls de verres ou d'assiettes ici (C .f catégorie Objet-Machine). Au **4^e niveau**, on retrouve la notion d'espace (Intérieur ou Extérieur). Au **5^e niveau**, nous retrouvons l'étage de la densité et la notion de masse de l'image sonore de l'ambiance. Et le **6^e niveau** donne place aux déclinaisons de lieux du secteur d'activité (Restaurant, Cafétéria, cantine scolaire, restaurant, fast-food, Bar/Pub...).

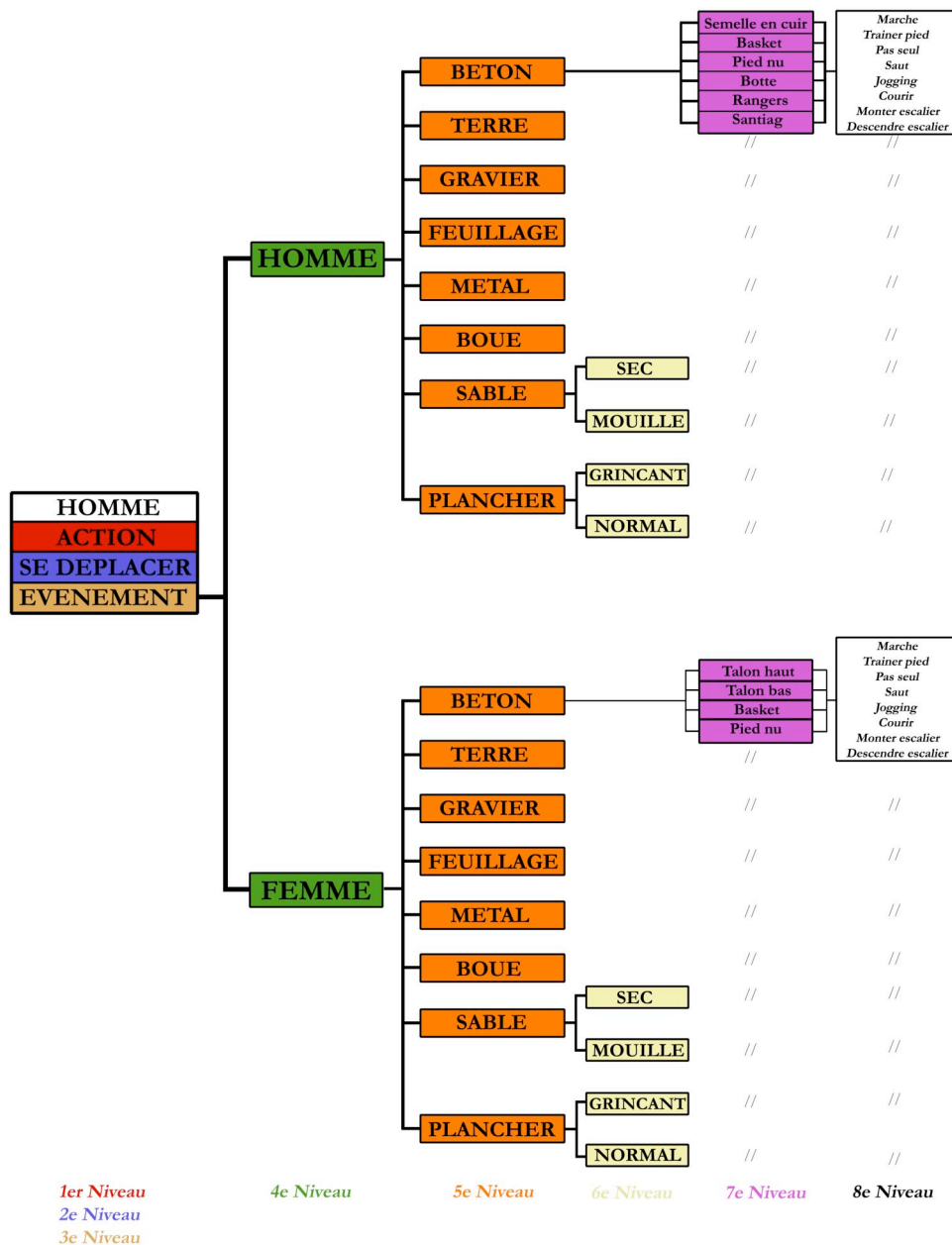
1.3.1.3 Action – Se déplacer - Ambiance



Taxonomie des sons d'ambiance liés à l'action de l'Homme de se déplacer

La classification suit sa logique, l'arborescence précédente classe les sons Homme – Action – Se déplacer – Ambiance ; au 4^e niveau la notion de classe se réfère à la densité, au 5^e niveau nous avons deux branches qui se divisent, l'une est le point d'écoute du lieu de transport (Hall de gares, aéroport...) et le second prend en compte le type de transport. Au 6^e niveau on différencie si l'espace du son est à l'extérieur ou l'intérieur. Nous classons dans cette partie de l'arborescence des sons du type ambiance de circulation (voiture, centre ville, entrée et sortie de gares de personnes...) ; j'insiste sur le fait que des ambiances où l'Homme est muet du type circulation (voiture, camion...) et où il n'y a aucune sonorité d'origine humaine, se classeront dans la catégorie Objet – Machine.

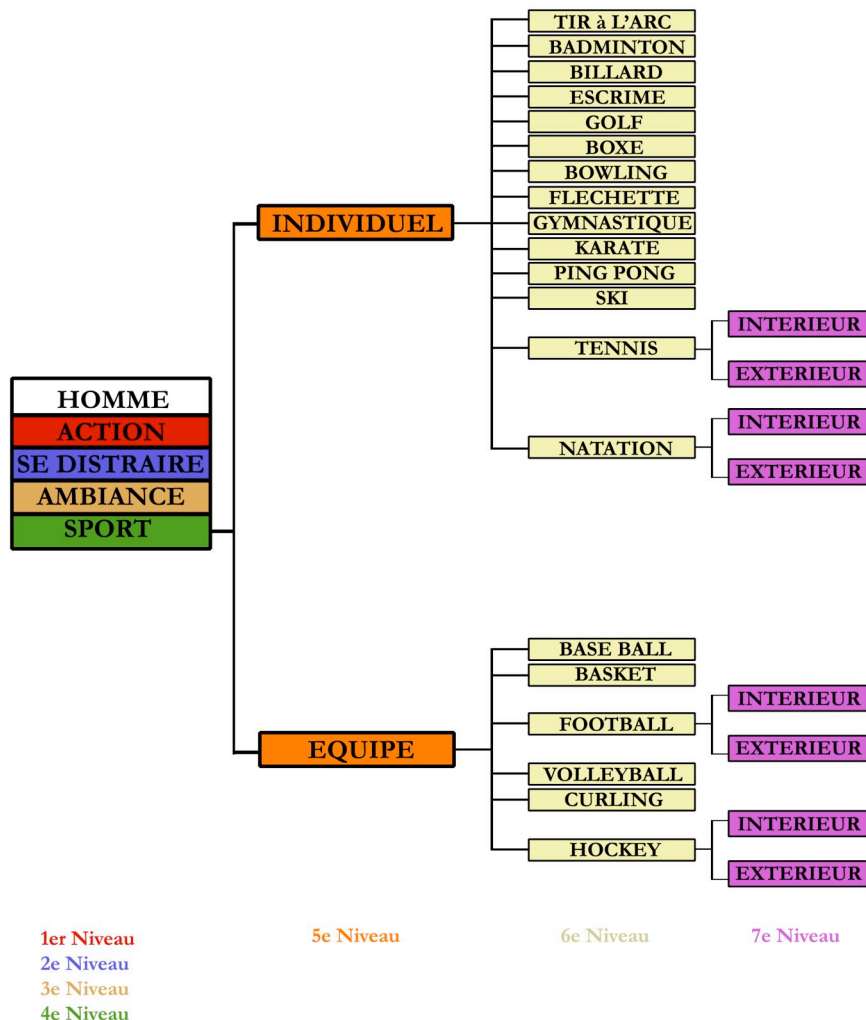
1.3.1.4 Action – Se déplacer – Evènement



Taxonomie des sons évènements liés à l'action de l'Homme de se déplacer

La déclinaison de Homme – Action – Se déplacer – Evènement place au **4^e niveau** le genre humain masculin ou féminin, c'est une caractéristique que notre perception cognitive différencie, nous différencions les sons de pas d'un homme à ceux d'une femme, comme les aveugles peuvent différencier la frappe d'applaudissement d'une femme à celle d'un homme. Au **5^e niveau**, nous avons la matière du support de déplacement (Béton, terre, gravier...). Au **6^e niveau** est précisé l'état du sable et du plancher, respectivement sec, mouillé, grinçant, normal. Cette notion d'état de la matière est un renseignement supplémentaire que notre représentation sonore assimile très bien et sait l'exprimer. Au **7^e niveau**, nous avons le type de chaussure que l'homme ou la femme porte. Et le **8^e niveau** définit le rythme des pas (traîner les pieds, marcher, courir, monter et descendre escalier...).

1.3.1.5 Action – Se distraire – Ambiance - Sport



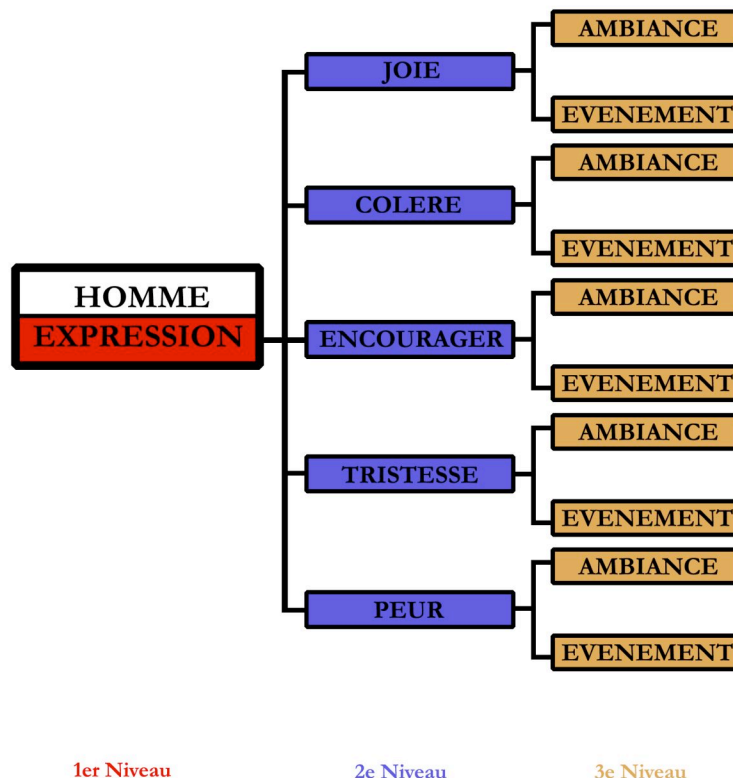
Taxonomie des sons liés à l'action de l'Homme de se distraire par le sport

Dans cette catégorie Action – Se distraire – Ambiance – Sport , on trouve au **5^e niveau** la notion de densité, en utilisant les termes de sport d'équipe ou individuel, ces termes rappels à notre image la masse sonore du son. Au **6^e niveau** nous avons le nom du sport. Et en **7^e niveau** nous trouvons l'espace où se joue le sport : Intérieur ou Extérieur.

Nous venons d'étudier la taxonomie des sons d'actions qui sont liés à l'Homme, nous avons défini des critères d'intégrations à la classification. Nous poursuivons l'arborescence en étudiant la deuxième branche du 1^e niveau de la catégorie de source Homme : Expression.

1.3.2 Expression

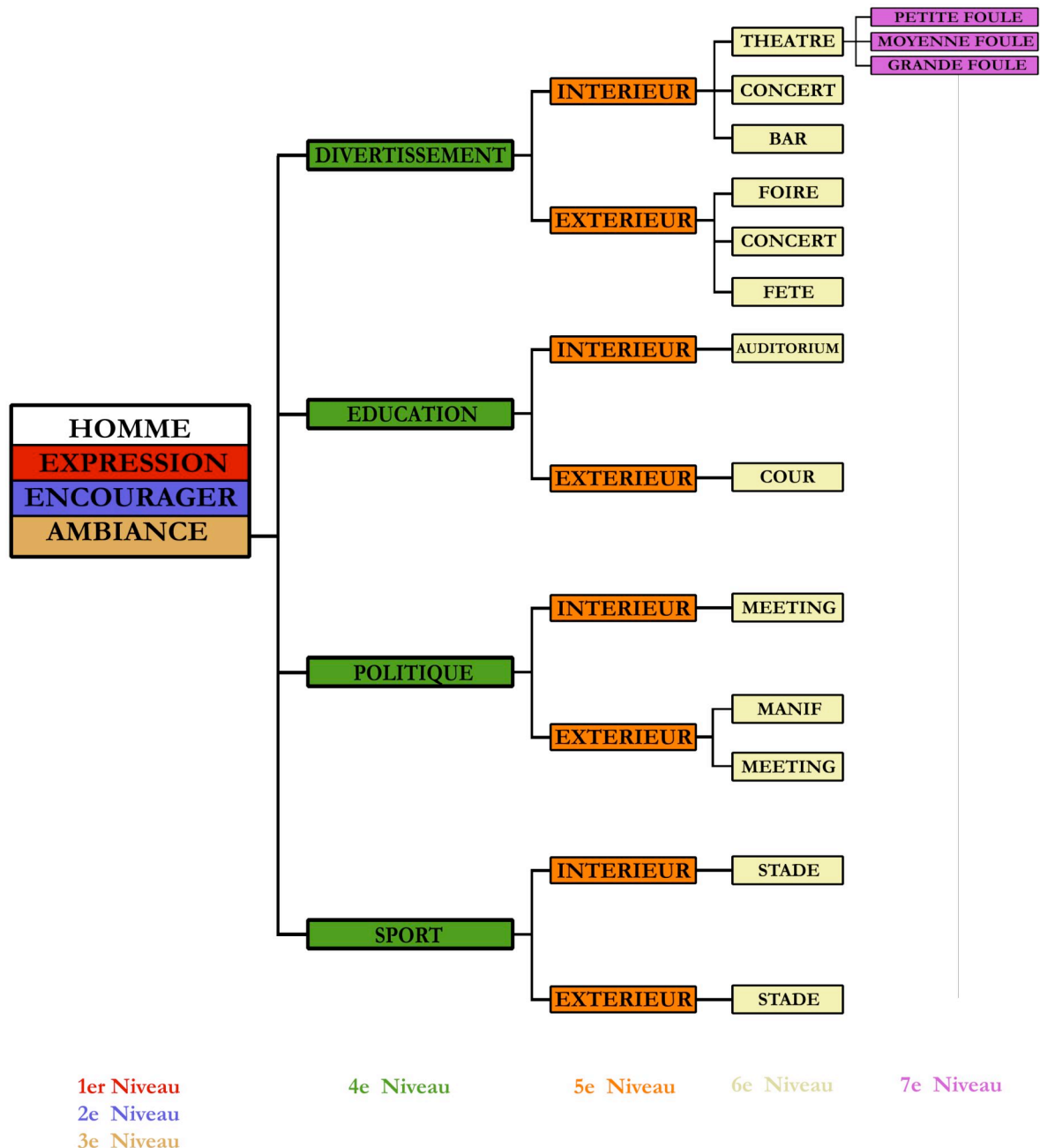
La classe « Expression » de 1^{er} niveau, organise les sons en fonction de leurs intentions sémantiques. Elle considère aussi le contexte de l'expression ; la principale source sonore de cette catégorie est la corde vocale de l'Homme, elle émet des sonorités caractéristiques à des états, des sentiments et à la parole. Elle traite la parole intelligible dans l'événement et la parole de l'entendre, dans la structure, dans l'ambiance du son. Je rappelle que nous sommes dans l'étude d'une classification des sons non-musicaux, c'est pourquoi les sons d'Expression ne sont pas classés en fonction de repères musicaux, si un chant est présent dans une ambiance ou un événement, il est uniquement lié à son intention sémantique et à son contexte d'expression.



Niveau 1, 2 & 3 de la Taxonomie des sons de l'Homme liés à l'expression

Le 2^e niveau est l'état sémantique du son, c'est l'expression de sentiments ; ce sont des sentiments naturels propres à l'Homme (ex : la joie, la colère, la tristesse, l'encouragement...) Ils sont à travers leurs expressions et leurs intentions, reconnaissables dans leurs sonorités par tous. Au 3^e niveau on retrouve la notion de durée de la prise de son même ; les « événements » sont des sons du type rire seul, applaudissements seuls, des slogans seuls, des expressions liées à la parole et à un contexte donné... que nous classons.

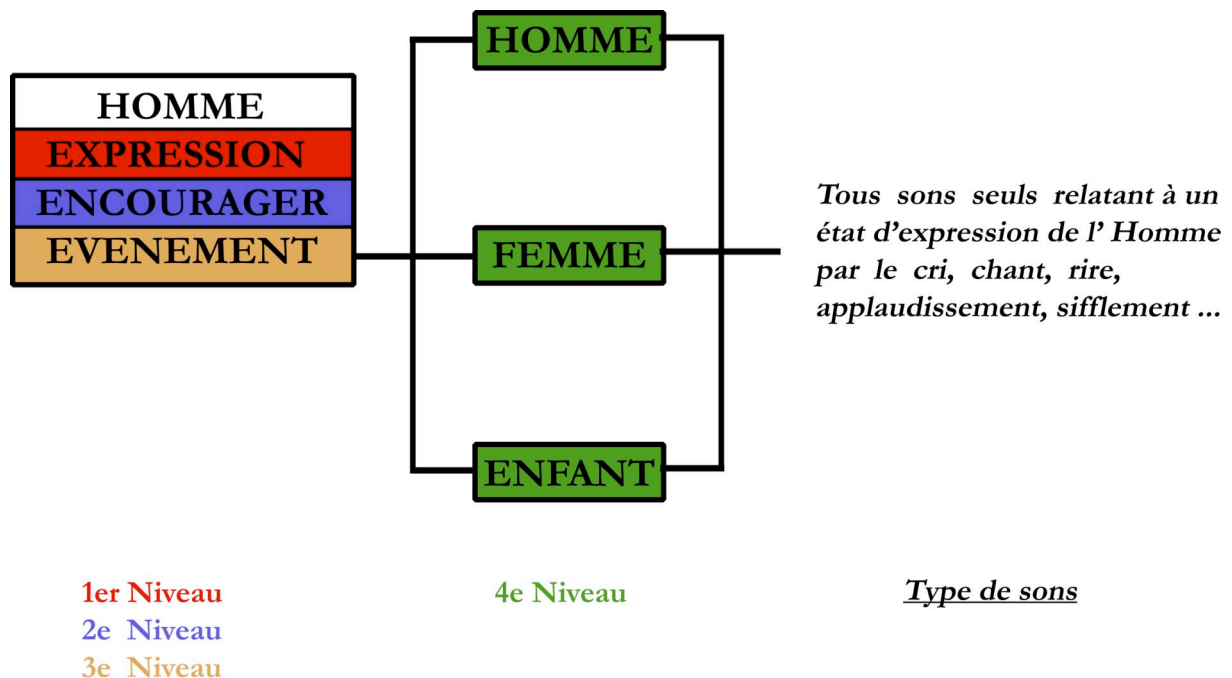
1.3.2.1 Expression – Ambiance



Taxonomie des sons d'ambiance liés à l'expression d'encouragement de l'Homme

Nous n'étudierons que l'expression de l'encouragement pour expliquer la déclinaison de l'arborescence suivant le 3^e niveau. Le 4^e niveau peut être assimilé à la « coopérative » des différentes instances de contextes sociaux, plaçant l'intention sémantique du son dans une case contextuelle générale. Le 5^e niveau définit l'espace du lieu du contexte : *Intérieur* ou *Extérieur*. Le 6^e niveau est le lieu d'expression du son. Le 7^e niveau renseigne sur la densité du son, il pose une jauge définissable par trois adjectifs de grandeurs : *petite - moyenne - grande foule*. Le fait d'avoir la notion d'espace définie au 5^e niveau me permet d'aménager le classement des lieux en amont, ces 5^e et 6^e niveaux nous engagent sur le renseignement de l'image acoustique du son.

1.3.2.2 Expression - Evènement



Taxonomie des sons évènements de l'Homme liés à l'expression

L'arborescence des expressions d'évènements est générale, elle ne s'applique pas uniquement à l'encouragement mais représente toutes les intentions d'expressions. Le 4^e niveau ; nous n'utiliserons pas de classement par gamme ou hauteur, néanmoins la distinction d'une expression si nous voulons la définir des autres sources de même niveau (le 4^e) se réfère à l'image sonore du timbre du son, c'est à dire s'il est émit par un homme, une femme ou un enfant. Et comme le définit le schéma, les exemples de sons qui se classent ici sont des sons seuls, ils peuvent être isolés d'une acoustique ou pas et ils relatent d'un état d'expression de l'homme, de la femme ou de l'enfant. Ex : cri de joie, de peur, chant, slogan, rire, applaudissement, pleurer, siffler, hurler...

Pour conclure cette partie sur la taxonomie des sons liés à l'Homme, rappelons les philosophies fondamentales qui l'organise ; il faut considérer un son d'action muet d'Homme comme un son Objet – Machine. Il faut assimiler que la source des sons de l'Homme n'est uniquement originaire de lui-même, en terme de présence physique ; son système biologique, ses réflexes naturels, son expression par la voix et la résonance de ses sons par son corps sont les notions fondamentales qui permettent de distinguer les sons de l'Homme et les sons de la catégorie Objet – Machine. Nous allons, par ailleurs, poursuivre cette étude dans le but de clarifier la taxonomie Objet – Machine. Néanmoins, l'ordre des niveaux reste flexible, la mise en pratique permettra de confirmer les bases vues précédemment et d'ajuster la classification des sons liés à l'Homme, au fur et à mesure de l'alimentation de la base de données de sons.

1.4 Taxonomie des sons Objet / Machine

Notre environnement sonore a beaucoup évolué au cours des deux derniers siècles, avec l'avènement de l'électricité et de la Révolution Industrielle. Notre quotidien sonore, aujourd'hui, n'a plus rien à entendre avec la société du Moyen Age, notre temps moderne c'est accaparé des sons industriels, objets et machines qui régissent désormais, au 21^e siècle, notre mode de vie et notre quotidien de l'entendre. C'est en grande partie ces nouvelles sources sonores que nous appelons par ignorance « Bruit ». Pour la classification des sons Objet / machine, je vais prêter attention à l'identification de la source et à l'analyse de la qualité du son, me permettant d'aboutir à une taxonomie des sons en relation avec l'image sonore du timbre des « bruits ». Lier source sémantique et notions psychophysiques¹⁰ pour corréler au maximum un classement efficace et accessible à tous.

J'ai exploré plusieurs pistes avant la proposition de taxonomie que nous allons étudier. Je suis parti des définitions des mots « Objet » et « Machine ».

- « **Objet** » : Chose solide considérée comme un tout, fabriquée par l'Homme et destinée à un certain usage.

- « **Machine** » : Appareil ou ensemble d'appareils capable d'effectuer un certain travail ou de remplir une certaine fonction, soit sous la conduite d'un opérateur, soit d'une manière autonome.

En vue de ces définitions, j'étais parti sur l'orientation d'une classification en fonction de l'utilisation que l'Homme a des objets et des machines. (Ex. : Transport – Véhicule - Se déplacer – Usage – Armée, Travaux, particulier, pompier ...). Une voie sémantique vite écartée, en raison d'un trop grand nombre de niveaux de classifications et d'une structure trop peu directionnelle et efficace pour distinguer les sons . Une autre voie psychophysique en rapport au type de fonctionnement de l'objet ou machine (Mécanique, Hydraulique, Electrique, Manuel...) . Un registre trop unilatéral qui ne peut, à lui seul être un critère suffisant de classification.

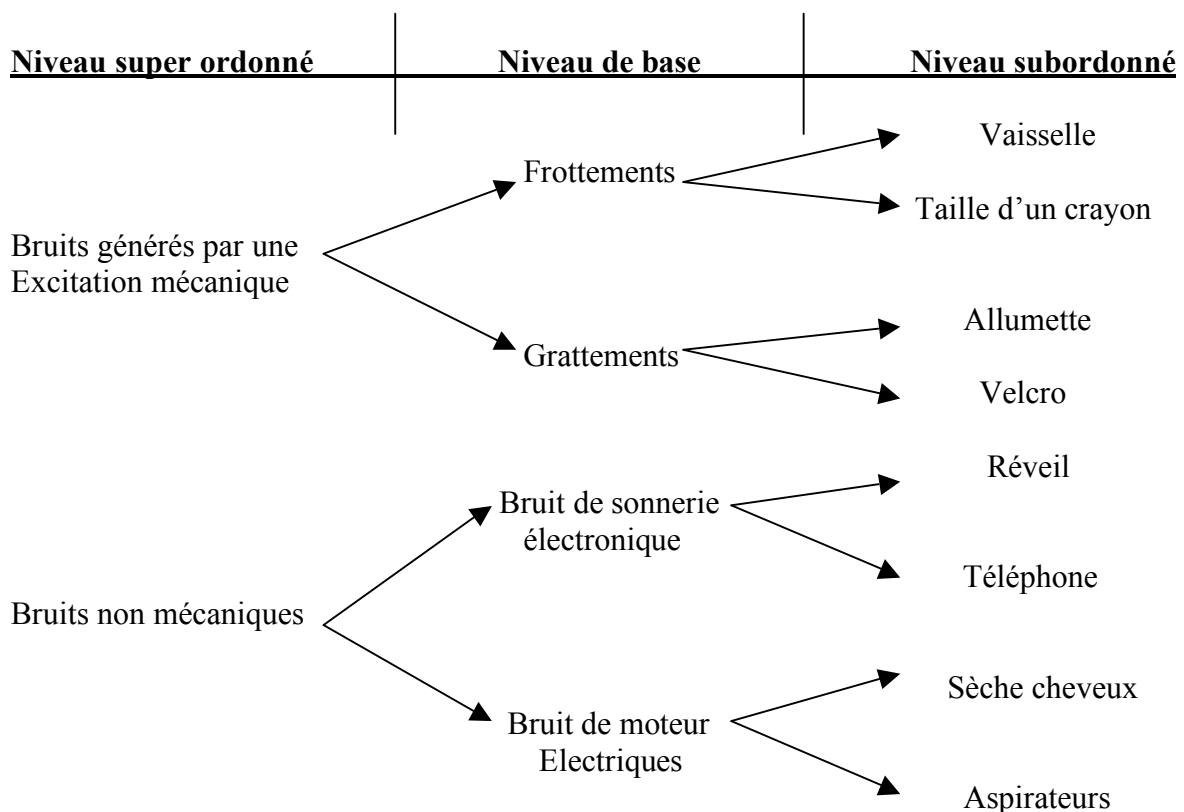
Intéressons-nous à l'étude de la catégorisation d'un corpus de bruits domestiques menée par Frédérique Guyot, Michèle Castellengo et Benoît Fabre¹¹. Cette étude est partie de l'accent mis par Ballas (1993) sur l'importance du contexte dans l'identification d'un son. Il a montré que des événements sonores impulsifs sont difficilement identifiables à l'écoute hors contexte, malgré une fréquence écologique élevée (Probabilité d'apparition d'un son dans un contexte donné), alors que les sons électroniques, type alarme, sont facilement identifiables. C'est le rôle important des situations et des contextes d'écoute, ainsi que du degré d'expertise de chaque individu, qui a amené Ballas à conclure qu'il fallait établir un schéma général de l'organisation perceptive des sons. Il propose de se référer pour cela au concept de catégorisation prototypique de Rosch¹² (1975), concept bien connu des psychologues cognitivistes mais totalement ignoré des « psychoacousticiens ». Il s'agit de présenter une notion commune caractéristique (prototype) qui représente un groupe de catégorisation. L'étude de catégorisation d'un corpus de bruits domestiques est partie d'une écoute de différents bruits courants de ce contexte (brossage de dents, grattement d'une allumette, sèche

¹⁰ étude du timbre

¹¹ Chercheurs CNRS-LAM Laboratoire d'Acoustique Musical

¹² Eleanor Rosch-*Cognitive representations of Semantic Categories* 1975 -Les objets basiques sont les catégories les plus inclusives dont les membres possèdent des nombres significatifs d'attributs en commun, ont des programmes moteurs semblables, ont des formes similaires et peuvent être identifiés à partir des formes moyennes des membres de la classe.

cheveux, taille d'un crayon, cliquetis de clés...) qui font partis de notre univers sonore et qui sont entendus quotidiennement par une majorité de gens. Il a été supposé que l'expertise de tout un chacun sur la perception des bruits était équivalente et qu'il n'y avait pas lieu de choisir une catégorie précise d'individus pour cette expérience, car nous sommes fort de constater que même les musiciens ayant pourtant une expérience d'écoute et de qualification des sons musicaux, ont les mêmes difficultés à décrire les bruits que les sujets non musiciens. En regard des résultats, on constate deux types de catégories, une qui est décrite par le nom ou la nature de la source qui a produit le son (porte, vaisselle, moteurs, sonnerie). Et l'autre qui est définie par l'action ou le mouvement qui a produit le son (grattement, froissement...) Ces deux modalités de catégorisation correspondent à deux niveaux d'abstraction différents. Un niveau pour lequel la source a été identifié (porte). Il est, dans ce cas, difficile de savoir si les bruits sont regroupés parce qu'ils sont perceptivement proches ou tout simplement parce qu'ils sont engendrés par des sources similaires. L'autre niveau, moins affiné que le premier, est caractérisé par l'action qui génère des bruits. Un son de « bruit de porte qui grince » a été catégorisé soit comme « bruit de porte », soit comme « grincement », soit comme « bruit mécanique ». A l'inverse nous ne pouvons parler de geste ou de mouvement de la source pour les bruits électroniques, tels que les alarmes et les bruits de moteurs électriques. Ces deux niveaux de catégorisation s'expliquent par l'identification des sons en fonction de leur nature. On identifie parfaitement un frottement de deux assiettes l'une sur l'autre à un bruit de vaisselle, mais on n'identifie que le mouvement à l'écoute de la taille d'un crayon. Cette typicalité peut se justifier soit par un fréquence écologique plus élevée de certains bruits, soit par des indices acoustiques plus pertinents (couleurs spectrales, harmonicité, structure rythmique...). Ainsi de cette étude a été suggéré une organisation de la classification en trois niveaux comme sur le schéma suivant :



Exemple d'organisation hiérarchique de catégories perceptives de bruits domestiques

Nous pouvons, déduire de ce schéma trois niveaux de hiérarchisation. Le **Niveau super ordonné**¹³ différencie les sources sonores, les sons *mécaniques* sont ceux qui sont excités par un mouvement ou une action de l'Homme : tout objet qui émet un son uniquement sous la manipulation de l'Homme. Et un second niveau super ordonné, les sons non mécaniques ; ils sont autonomes, leur timbre ne dépend pas de leur excitation, ils sont émis par une énergie autre que celle de l'Homme (Hydraulique, électrique, pneumatique, électronique...). Ensuite, le **Niveau de base**¹⁴ classe les sons par famille de source sonore, les termes employés se réfèrent à l'image sonore ou au timbre de celles-ci ; pour les sons *mécaniques* nous les classons suivant leur type d'excitation (action et bruit associé) :

- Claquement
- Couinement
- Craquement
- Craquètement
- Crépitement
- Crissement
- Déchirement
- Frottement
- Froissement
- Grésillement
- Grattement
- Grincement
- Pétillement
- Raclement
- Roulement
- Tapement

Liste établie après une recherche approfondie dans *Le Petit Robert*. Une source excitée de différentes manières peut créer des bruits tout à fait différents. Par exemple, une porte fermée lentement est un grincement, alors que cette même porte fermée violemment générera un claquement. C'est le son en tant qu'objet sonore que nous classons et n'ont pas les sources sonores, ce qui explique la présence possible d'une même source dans différentes catégories de la classification.

Pour les sons *non mécaniques*, le niveau de base les regroupe sous leur nom de famille sonore, exemple bruit de sonnerie électronique, Bruit de moteur électrique, Bruit électrique... Et enfin, le **Niveau subordonné**¹⁵ correspond à la source sonore elle-même : l'objet ou la machine concerné.

De cette hypothèse de classement de sons domestiques proposée par Frédérique Guyot, Michèle Castellengo et Benoît Fabre, je suis parvenu à une classification qui considère tous les sons objets – machines qu'ils soient domestiques, industriels, quotidiens etc... En différenciant les sons *mécaniques* (excitation de l'Objet par l'action de l'Homme) et les sons *non-mécaniques* (son autonome, Machine, la source est excitée par une énergie). Ces deux niveaux de classes permettent une organisation hiérarchique des connaissances image sonore des sons Objet/Machine, elles différencient leurs origines sonores et le niveau de base représente les catégories proutypiques du son, c'est une catégorie typique, l'appartenance d'un son en fonction de sa ressemblance avec un élément commun, un prototype ; même s'il n'a jamais été vu auparavant, les dénominations sont telles que celles employées pour les bruits *non mécanique* réunissent le maximum de données sémantiques pour faire une analogie

¹³ Les éléments superordonnés ont peu de propriétés clés .

¹⁴ Le niveau de base est le plus informatif et celui qui possède le plus de propriétés clés.

¹⁵ Le niveau subordonnés à peu de propriétés clés, ayant beaucoup de traits en commun avec le niveau de base.

de classement par une personne lambda (audiophile ou pas), la plus proche de la perception réelle, qu'une oreille rencontre le plus fréquemment.

OBJETS / MACHINES

BRUITS GENERES PAR UNE EXCITATION MECANIQUE		
FROTTEMENT	Arme	sortir/remettre arme (pistolet, couteau...) étui
	Outil	Tournevis, Lime métal, Feuille ponce
	Papeterie	lettre, carton, livre, ciseaux, taille crayon
	Véhicule	mettre/enlever ceinture
CLAQUEMENT	Arme	Charger, recharger arme Ouverture/fermeture couteau
	Bruit Domestique	Cliquetis de clés, cadenas, O/f poubelle, fermer fenêtre déplier parapluie, mettre une lettre dans boîte à lettre
	Eclairage	Interrupteur lumière
	Nourriture	Battre les œufs, Couper légumes
	Outil	Agrafeuse, fermer boîte à outil
	Papeterie	fermer livre, classeur, serviette
	Véhicule	fermeture portière, fermeture coffre
TAPEMENT	Machine	Machine à écrire, calculatrice
	Outil	Coup de marteau,
	Papeterie	Coup de tampon
COUINEMENT	Bruit Domestique	Ouvrir et fermer rideaux vénitien
	Outil	Serrer étau
	Papeterie	Ciseaux
RACLEMENT	Nourriture	Râper légume, Broyer café avec moulin à café
	Outil	Rabot, Racloir, Scie
	Papeterie	Dérouler scotch
	Véhicule	enfoncer, retirer clé dans serrure ou contact
PETILLEMENT	Nourriture	Boisson soda, gazeuse, champagne versement dans verre
ROULEMENT	Bruit Domestique	Dérouler papier toilette Tirer rideau à l'horizontale
	Papeterie	Déchirer lettre, Déchirer carton, Déchirer livre
CRAQUEMENT	Bruit Domestique	Déchirer papier toilette
	Nourriture	ouvrir boîte de conserve avec un ouvre boîte

OBJETS / MACHINES

BRUITS NON-MECANIQUE			
BRUITS DE MOTEURS A EXPLOSION	VEHICULE	Evènements	Démarrage moteur Ralenti moteur Arrêt moteur Passage
		Ambiance	Ambiance circulation
	ENERGIE	Groupe électrogène à essence	
BRUITS DE MOTEURS A REACTION	AVION	Evènements	Passager Décollage Pilote Vol Exterieur Atterrissage
		Ambiance	Ambiance passage d'avion
	OUTIL	Perceuse électrique Foreuse électrique Meuleuse électrique Scie circulaire Tourneuse à bois, métal électrique	
BRUITS DE MOTEURS ELECTRIQUE	DOMESTIQUE	Rasoir électrique Sèche cheveux Ventilateur Tondeuse à gazon Taille haie Four électrique	Aspirateur Congélateur Centrifugeuse Cafetière Frigo Micro onde
	VEHICULE	Train, métro, Tramway, funiculaire	
	INFORMATIQUE	Unité centrale ordinateur	
BRUITS DE MOTEURS A AIR COMPRI ME	MACHINE OUTIL	Marteau piqueur	
BRUIT EXPLOSION	ARME	Pistolet, fusil, mitraillette, canon, grenade, mortier lance rocket, mine	
	VEHICULE	Explosion Véhicule	
BRUIT DE SONNERIE ELECTRONIQ	DOMESTIQUE	Réveil, sonnette	
	VOITURE SIGNAL	Signal d'alarme	
	COMMUNICT	Téléphone, Fax, beep répondeur	
BRUIT EAU	DOMESTIQUE	Douche, baignoire, faire couler eau du robinet Tirer la chasse d'eau	
	VILLE	fontaine, écoulement eaux d'égout	
	VOITURE	Lavage voiture, lave vitre	
BRUIT MUSICAUX	VEHICULE	Klaxon	
	DOMESTIQUE	Sonnerie horloge	
BRUIT ELECTRIQUE	ECLAIRAGE	Bruit d'un néon, lampadaire	
	VOITURE	Ouverture et fermeture centralisé	
BRUIT DE DEMOLITION	TRAVAUX	Déversement gravas	
BRUIT DE CONSTRUCTION	TRAVAUX	Epandeur, camion ciment	
BRUIT DE CAROSSERIE	VOITURE	Crash voiture, choc voiture	

Cette classification n'est pas exhaustive, encore de nombreuses catégories et de sons peuvent s'y greffer, elle a la prétention de poser une base de classification des sons Objets/Machines. On peut constater qu'il y a des ajustements de niveaux de classification, le schéma général étant :

1^{er} Niveau : Type d'excitation Mécanique ou non mécanique

2^e Niveau : Excitation ou Famille représentative

3^e Niveau : Domaine d'application sociale du son

4^e Niveau : L'objet sonore

Le 3^e Niveau ajusté est une sous catégorie du niveau de base, il permet d'identifier les différents corpus de sons existants dans des domaines d'applications ou de lieux.

Certes, cette organisation des sons Objets/Machine implique une écoute réduite en fonction du prototype déterminé ; mais celle-ci c'est basé le plus possible sur le regard du réel, sur la communion à apprécier, à percevoir ou à parler d'un son.

CONCLUSION : Classification des Sons environnementaux

En conclusion, la taxonomie des sons non musicaux que je viens de vous soumettre est basée sur l'écoute sémantique, sur l'utilisation des mots utilisés en cinéma pour définir un son dans son contexte ou hors contexte. L'arborescence use de termes qui sont des attributs parmi tant d'autres du son comme différencier un son d'ambiance d'un son événement ou un son extérieur d'un son intérieur, ce sont des notions qui nous permettent d'être le plus juste dans l'organisation et la place des sons à l'intérieur de cette taxonomie. Elle est une synthèse de travaux antérieurs et de ma perception cognitive des sons en tant qu'entité ou objet sonore en terme schaefferien. Elle me semble pouvoir convenir à tous types d'utilisations : post production, design sonore, acoustique ... Revenons à notre objectif de créer une base de données des sons, cette classification devient alors notre organisation principale de la sonothèque. Une fois qu'elle est bien assimilée, elle s'avère dans un premier temps efficace dans l'accès au son désiré, mais nous allons aller plus loin désormais dans la description même du son, nous sommes partis de quelques critères de l'écoute réduite pour notre classification, maintenant dans sa recherche nous allons procéder à l'attribution de critères physiques, sémantiques et perceptifs qui seront associés dans la base de données, aux sons en tant qu'attributs de descriptions pour une recherche intelligente des sons.

2. DESCRIPTION DES SONS

Auparavant, nous avons développé une classification des sons axée principalement sur la représentation que nous pouvons avoir des sons (perception cognitive, sémantique et fonctionnel) dans notre culture et notre société. Une classification organisée hiérarchiquement autour de catégories significatives. Il reste une phase importante de notre étude, leurs descriptions ; le son est un phénomène vibratoire qui prend place dans notre espace temps, il défile et se transforme rapidement. Il est très difficile de les qualifier, car nous ne percevons pas que l'objet sonore, mais les structures ¹⁶qui nous permettent de les identifier. Depuis notre naissance, notre sens de l'ouïe n'a cessé de se développer, de s'éveiller, dans une époque, dans une civilisation, dans une culture. Des objets aux structures, des structures au langage il y a donc une chaîne continue, d'autant plus non identifiable qu'elle nous est absolument familière, spontanée et que nous y sommes entièrement conditionnées. La notion d'entendre est complètement innée, alors que toute une éducation est à entreprendre sur l'écoute des sons.

C'est pourquoi, il nous faut pouvoir fixer cet événement sonore sur un support, le rendre intemporel et infini, et aujourd'hui, le magnétophone, l'enregistreur numérique sont des machines, des laboratoires à observer les sons, à les isoler, à redécouvrir, à réécouter les sons, d'une oreille indépendante détachée de son environnement sonore. Le support d'enregistrement devient alors le moyen d'analyser, de comprendre, de déduire, d'émettre des hypothèses, des problématiques sur les caractéristiques, les critères morphologiques qui sont le fondement propre du son.

Le premier homme à tirer les conséquences théoriques de ce nouveau statut du son est Pierre SCHAEFFER ¹⁷, il tenta de généraliser au delà de la musique un système de classification et de description des sons nommé : Solfège de l'Objet Sonore. Proposant ainsi la science d'écoute réduite de l'objet sonore. La révolution schaefferienne ouvre la voie à la compréhension de ce que nous entendons, par des mots qualificatifs, descriptifs ; ce qui aujourd'hui n'intéresse que peu de monde, car ce n'est pas un besoin vital. Mais pourtant ce que nous entendons détient une substance, l'objet sonore est d'une richesse que peu veulent ou ne peuvent entendre par manque d'intérêt, mais la description du sonore est le « dictionnaire » ou le « solfège » de ce qui nous transmet un sens, nous produit une émotion, nous raconte une cause. Pourtant, aujourd'hui les multitudes de métiers que regroupe le domaine du sonore nécessitent le besoin de parler des sons. Il faut établir un vocabulaire, un langage, pouvoir mettre des mots aux sons, sans analogie abstraite, ni incohérente. Notamment, dans le milieu du cinéma, entre les relations réalisateur, monteur image, monteur son et mixeur, chez les illustrateurs sonores, les bruiteurs, les architectes, dans l'éducation... Il est nécessaire de pouvoir parler justement et objectivement de la structure du sonore dans l'intention de donner tel ou tel ressenti à l'auditeur. Proposer un langage commun du sonore, une morphologie commune, approfondir les critères schaefferiens en développant d'autres notions de descriptions pour que le mot « bruit » que nous employons à tort et à travers ne devienne plus une facilité de qualification, ni une péjoration des sons environnementaux non musicaux.

¹⁶ Ensemble des caractéristiques du son

¹⁷ Pierre Schaeffer (1910-1915), Polytechnicien, poète, philosophe, il consacra la plus importante partie de sa recherche intellectuelle à la nature et aux formes de la musique.

Cette partie vise à développer les attributs physiques, perceptifs et sémantique de l'objet sonore, elle est une synthèse des travaux de Pierre SCHAEFFER, de Emmanuel DERUTY (Projet Ecrins IRCAM), de Romain LEBLANC (Recherche à Structure horizontale), et de Anne FAURE (Thèse : Des sons aux mots, comment parle-t-on du timbre musicale-Sciences cognitives). Je tente ainsi de chercher une amélioration et une orientation du langage sonore.

2.1 Les trois écoutes

Selon la finalité attribuée à un système de description, il est nécessaire de s'interroger quant aux multiples possibilités d'appréhension du son, puisque celles-ci conditionnent nécessairement les critères de description.

Pierre SCHAEFFER définit d'abord les deux écoutes qui traversent le son et le considèrent comme un indice ou un message : l'écoute **causale** et l'écoute **sémantique**.

L'écoute **causale** s'intéresse, à travers le son, à tous les indices susceptibles de renseigner l'auditeur sur sa cause, c'est à dire l'objet, le phénomène qui produit le son, où il se trouve, comment il se comporte, se déplace, etc. Chacun, en fonction de son champ de travail, de compétence et de sa curiosité, développe des écoutes causales spécialisées et raffinées (musique, sons de moteurs, cris d'animaux, etc.), mais seulement sur des domaines très précis, dont chacun ne peut généraliser les acquis à d'autres univers sonores.

Deuxièmement, SCHAEFFER nomme « **sémantique** » l'écoute qui s'intéresse au message que porte le son, à sa signification.

Enfin, l'écoute **réduite**, telle que l'auteur du *Traité des objets musicaux* la pose, est celle qui fait volontairement et artificiellement abstraction de la cause et du sens, pour s'intéresser au son pour lui-même, dans ses qualités sensibles non seulement de hauteur et de rythme, mais aussi de grain, matière, forme et masse.

Par rapport aux autres écoutes, l'écoute réduite est celle qui prend le son, qu'il soit verbal, musical ou réaliste, comme objet d'observation en lui-même, au lieu de le traverser en visant en lui autre chose. Activité forcément volontaire et culturelle, elle s'oppose ainsi à deux écoutes plus quotidiennes et spontanées.

On peut dire que nous pratiquons tous une écoute réduite inconsciente, mais surtout pour fournir aux autres écoutes des éléments d'interprétation et de déduction. Cette écoute réduite spontanée se passe de mots, et ne franchit donc pas un certain niveau de finesse et de développement. En revanche, le repérage des jeux de hauteur et de rythme dans une musique relève bien de l'écoute réduite. Seulement, il ne concerne qu'une mince partie des qualités perceptives du son, même musical, et il abandonne les sons non toniques et les sons non régulièrement pulsés¹⁸ (donc la majorité des sons) à leur prétendue « confusion ». L'écoute réduite telle que SCHAEFFER la définit et la préconise est au contraire une démarche non spontanée et pratiquée collectivement, faisant entrer en jeu une méthode spécifique, un échange de langage et la nomination. Sa pratique fait assez vite surgir les « formes élémentaires » que chacun connaît sans avoir des mots pour les nommer.

L'écoute réduite n'oblige pas à refouler, encore moins à nier nos associations figuratives et affectives. Il s'agit simplement de les placer, temporairement, en dehors d'un champ de nomination et d'observation.

¹⁸ Les sons environnementaux

Tout comme sur le plan visuel, si nous voulons décrire une orange en termes de forme, couleur, texture, nous ne sommes pas obligés d'oublier que c'est le fruit nommé orange ni de faire comme si nous ne le savions pas. Nous le soumettons simplement à un questionnement descriptif qui met entre parenthèses la nature de cet objet, les associations auxquelles il donne lieu, les plaisirs qu'il procure, etc.

Isolement des unités sonores significantes

Un son ne peut pas toujours être saisi dans sa globalité, c'est à dire que l'on ne peut pas percevoir d'un seul coup tous ses traits caractéristiques, aussi bien d'un point de vue temporel que structurel.

Or, une des particularités du son sur le texte est qu'il ne propose pas de découpage préalable-comme des mots dans une phrase ou les lettres dans un mot, il n'existe pas de répertoire fini de signes (alphabet) qui assemblés, permettront de tout définir. Le son ne préconise aucun découpage, c'est pourquoi, décrire le son nécessite avant tout une phase de « localisation » qui consiste à dégager des unités significantes, à repérer des objets.

SCHAEFFER, qui lui-même a été confronté à ce problème de l'isolation des unités sonores significantes sous-jacent à toute tentative de description et de classement du son, a distingué deux types de morphologies des objets sonores. Alors que la morphologie dite *interne* va s'intéresser à la structure interne des objets, la morphologie *externe* sera l'étude d'un objet sonore « constitués d'éléments distincts, dont les formes se détachent les unes des autres ». Lorsqu'il s'intéresse à la morphologie *externe* de ses objets sonores, SCHAEFFER pose le problème du dégagement des « éléments distincts » et distingue les objets dits « composés », constitués de plusieurs éléments distincts et simultanés, et les objets « composites », constitués de plusieurs éléments distincts et successifs.

Ceci revient, en quelques sortes, à faire deux types de découpages du son : l'un est temporel, l'autre est structurel. Le découpage temporel du son est une segmentation du contenu sonore selon l'axe du temps, en fonction de critères de segmentation bien précis, souvent morphologiques. Cette discrétisation temporelle du son permet donc d'extraire des unités élémentaires de leur contexte d'existence à l'intérieur de l'extrait sonore total.

Le découpage structurel est une décomposition qui ne considère plus des éléments sonores successifs, mais simultanés, superposés (le son est « vu » comme un objet « composé »). C'est, en quelque sorte, une analyse des différentes « couches sonores » qui composent le son.

2.2 Critères de description

Nous allons désormais étudier une série de critères descriptifs du son qui seront intégrés dans le système de recherche intelligente de sons. Le système de description des sons environnementaux est de type pluridimensionnel. Trois Dimensions, appelées Sphères, déterminent l'ensemble des critères de description : La **Sphère Morphologique** (Critères Physique), la **Sphère Causale** et la **Sphère Sémantique**.

a) Sphère Morphologique

Dans une perspective d'écoute réduite du son, définie comme étant celle qui fait volontairement et artificiellement abstraction de la cause et du sens, elle s'intéresse au son pour lui-même, dans ses qualités sensibles. Cette sphère recouvre les descriptions d'une perception technicienne du sonore : identification et qualification des attributs perçus (écoute réduite schaefferienne). Cette description se situe à mi chemin entre la psycho-acoustique et l'analyse du signal. Les critères descriptifs physique du son sont les plus objectifs et admis

comme acquis commun. Rappelons sa définition : Le son est une onde produite par la vibration mécanique d'un support fluide ou solide et propagé grâce à l'élasticité du milieu environnant sous formes d'ondes longitudinales. Il se propage selon des lois bien particulières qui au passage, touche ce qu'on appelle l'oreille où elle donne matière à des sensations auditives, non sans toucher aussi d'autres parties du corps, où elle déclenche des chocs, des co-vibrations, etc., plus diffus et non réifiées. Les critères de la sphère morphologique reprennent des données proposées par Pierre Schaeffer comme le critère d'évolution dynamique, le profil mélodique et le profil de masse, le registre (Site et Calibre) ; nous rajoutons aux critères du Traité des Objets Musicaux adaptés et développés dans une nouvelle optique celle de l'analyse de sons environnementaux et non-musicaux, d'autres critères physique du son, comme sa Durée et les effets acoustiques présents dans le son (Réverbération, Echo, effet Doppler). Ces critères deviennent nécessaires dans une perspective de recherche des sons dans les bases de donnée audio de la postproduction audiovisuelle.

a.1.La durée

Les sons peuvent être classés selon leur durée réelle (durée de l'échantillon en secondes). A chaque son sera donc attribuée une icône de durée, indiquant à quelle plage temporelle il appartient :

0-1s 1-5s 5-10s 10-30s 30-60s >60s

Durée délimitée ou non délimitée

En 1992, un groupe de compositeurs de musique électroacoustique du Laboratoire Musique et Informatique de Marseille (M.I.M) a entamé un travail de recherche dont l'ambition était de prolonger et de compléter le Traité des Objets Musicaux en abordant le problème de la signification musicale.

Pierre Schaeffer a établi un système de classification sur des critères typo morphologiques, en faisant abstraction des considérations de sens. Mais paradoxalement, les compositeurs de musique électroacoustique choisissent les fragments sonores qu'ils vont utiliser, plutôt en fonction des sensations et impressions que l'écoute suscite chez eux que sur des critères de forme.

Face à cette contradiction, le M.I.M ¹⁹a proposé de réintroduire la signification dans la description des éléments sonores, et de définir quelque chose comme des Objets Sémiotiques. A cette première problématique s'ajoute une seconde, celle du temps, dont la spécificité de son utilisation dans les musiques électroacoustiques n'avait pas fait l'objet d'une réflexion suivie. Du croisement de ces deux problèmes, signification musicale et temporalité, est né un outil d'analyse musicale : l'Unité Sémiotique Temporelle (U.S.T).

La première étape de la recherche consista à repérer, dans le répertoire discographique, quelques configurations sonores bien typées sur le plan de la signification temporelle. Les figures qui avaient en commun une même signification furent regroupées en catégories et chaque catégorie fut désignée par une appellation métaphorique, un qualificatif qui fait référence à quelque chose d'extra musical : contracté-étendu, qui avance, en suspension, évocation du contenu sémantique, etc.

Dans un deuxième temps, il a fallu baliser ces catégories, déterminer et hiérarchiser leurs traits caractéristiques respectifs. Pour rédiger cette typologie des U.S.T, les musiciens

¹⁹ www.labo-mim.org

chercheurs du M.I.M ont établi un certain nombre de caractéristiques morphologiques et sémantiques.

L'une de ces caractéristiques concerne la durée et permet de distinguer les U.S.T « non délimitées dans le temps » des U.S.T « délimités dans le temps ».

Les premières, « **non délimitées dans le temps** » sont perçues comme un processus continu, qui pourrait durer éternellement, sans début ni fin probable. La sémantique qu'elles véhiculent ne dépend nullement du moment à partir duquel on les écoute.

A l'opposé, les U.S.T « **délimitées dans le temps** » s'inscrivent dans un moment précis du temps, elles peuvent être perçues comme une forme finie dont il est possible de préciser le moment précis de début et de fin.

Ce critère de durée délimitée ou non délimitée sera associé dans la représentation du profil dynamique que nous allons désormais aborder.

a.2. Le profil dynamique

Un mémoire de recherche concernant la classification des sons non figuratifs ²⁰a été réalisé par Gurwal Loïc, élève de l'Ecole Nationale Supérieure Louis Lumière, pour le GRM.

Ce mémoire expose les différentes étapes de travail aboutissant à cinq grandes catégories, appelées figures élémentaires, permettant de classer les sons non figuratifs, c'est à dire les sons que l'on ne « reconnaît » pas, et dont on ne peut donc déterminer la provenance. Ce travail est basé sur l'utilisation du **profil dynamique**, désignant l'évolution de l'amplitude d'un signal dans le temps, comme critère de base à l'élaboration des cinq figures élémentaires. Le choix de ce critère ayant été fait, suite à des expériences consistant à répertorier les termes généralement utilisés pour décrire les sons non figuratifs, et à différentes observations allant dans ce sens.

En répertoriant les critères développés par le M.I.M et par Pierre Schaeffer, six critères en rapport direct avec le profil dynamique, et cités lors des expériences sont retenus :

La durée : critère faisant référence au caractère délimité ou non délimité dans le temps, développé par le M.I.M, que nous avons évoqué précédemment.

Le mouvement : critère correspondant à la sensation de mobilité ressentie lors de l'écoute du son.

Le son des vagues donne dans l'ensemble une impression de « surplace » bien qu'à chaque cycle, on ait une impression de mouvement (formation et éclatement de la vague).

La direction : désignant l'évolution dans le même sens d'une variable ou d'un ensemble cohérent de variables, avec prévisibilité de cette évolution. Un coup de grosse caisse est un son avec direction, la résonance de la frappe s'éteignant progressivement.

Le profil mélodique : correspondant à un déplacement dans le champ des hauteurs de tout le son. C'est le son lui-même qui bouge, au lieu d'être « sculpté » (égalisation) par une évolution interne (profil de masse).

L'accélération : une impression d'accélération dans le son peut être obtenue grâce à une évolution de rythme (tempo de plus en plus rapide) mais aussi par la variation d'un paramètre morphologique (de plus en plus aigu. L'accélération peut être positive (e plus en plus vite), négative (de moins en moins vite) ou variée.

²⁰ c'est à dire les sons que l'on ne « reconnaît » pas, et dont on ne peut déterminer la provenance

L'attaque : elle correspond à la première phase temporelle du son : la phase d'établissement, et représente, dans le cas des sons de percussion-résonance, le moment déterminant du point de vue de l'évolution dynamique.

Quatre formes d'attaques parmi les sept proposées par Pierre Schaeffer seront retenues : abrupte, molle, plate, douce.

Ces critères sont alors utilisés pour décrire les sons d'un corpus pré-établi, afin de faire apparaître d'éventuels regroupements entre ces sons.

Ces familles permettant de classer les sons non figuratifs peuvent bien sûr servir à classer les sons figuratifs. C'est pourquoi les résultats de ce travail seront utilisés dans le cadre de notre méthode en ce qui concerne la description de type morphologique, et plus particulièrement la description du profil dynamique des sons.

Finalement, six familles ou figures élémentaires apparaissent clairement. A chacune de celles-ci sont associés un nom inspiré de la conception dynamique des sons de François Bayle et des Esquisses Pédagogiques de Paul Klee, ainsi qu'une représentation de type iconique.

Les figures élémentaires obtenues sont :

1. La figure **solide** : durée délimitée, avec déplacement, direction vers le bas, profil négatif ou stable, accélération négative, forme d'attaque abrupte.

Elle correspond essentiellement aux phénomènes dits de « percussion-résonance », c'est à dire où l'impact énergétique est donné en bloc au début du son et où la suite du son n'est qu'un retour progressif au silence.

Ex. : Un coup de gong.

Représentation iconique :



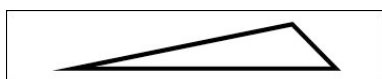
Attaque rapide et dégression vers le silence.

2. La figure **flèche** : durée délimitée, avec déplacement, direction vers le haut, profil positif ou stable, accélération positive, forme d'attaque douce ou molle.

C'est en quelque sorte l'inverse du solide, il y a cette fois accroissement d'énergie.

Ex. : Un coup de cymbale à l'envers.

Représentation iconique :

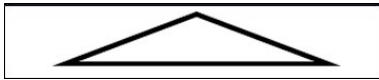


Dynamique progressive

3. La figure **parabole** : durée délimitée, avec déplacement, direction variée (vers le haut puis vers le bas), tout profil possible, toute accélération possible, forme d'attaque douce. Du silence, on passe à un maximum et on retourne au silence.

Ex. : le passage d'une voiture.

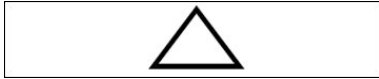
Représentation iconique :



Attaque douce

4. La figure **disque** : durée délimitée, sans déplacement, sans direction, tout profil possible, sans accélération, forme d'attaque abrupte. C'est une impulsion.
Ex. : un bref coup de klaxon.

Représentation iconique :



Impulsion

5. La figure **fluide** : durée non délimitée, forme d'attaque plate (il est difficile d'être plus précis pour cette catégorie). C'est un son qui s'écoule continûment, globalement sans évolution dynamique.
Ex. : un ruisseau qui coule.

Représentation iconique :



Continu

Par ailleurs, une sixième icône, se situant entre la figure solide et la figure flèche, est ajoutée pour caractériser les sons pour lesquels l'impact énergétique n'est pas donné en bloc au début du son mais entretenu pendant une durée plus ou moins longue avant le retour progressif du son au silence.

6. Représentation iconique :



Attaque rapide maintenue et dégression rapide

Seules les représentations de type iconiques seront utilisées puisqu'elles s'avèrent être beaucoup plus significatives que leurs homologues textuels et par ces icônes, nous indiquerons plus particulièrement le caractère délimité ou non délimité du son ainsi que son évolution dynamique générale.

a.3. La masse

La **masse**, déjà évoquée dans la typologie élémentaire schaefferienne, correspond au « mode d'occupation du champ des hauteurs par le son ». On retrouve alors les deux grands cas posés dans la typologie : la masse tonique, lorsque le son a une hauteur définie et repérable (cas des sons musicaux ex. : sonnerie téléphone, sirène, klaxon...), et la masse

complexe, lorsque ce n'est pas le cas. Ce critère sera utilisé dans la recherche intelligente de son. Il peut arriver que la masse d'un son comporte aussi bien des parties toniques que des parties complexes (ex. : Ambiance de circulation avec sirène de police...). Trois types de classement descripteur de la **masse** se présentent :

- Tonique**
- Complexe**
- Tonique + Complexe**

a.4. Le profil mélodique

Le critère de **profil mélodique** s'applique au profil dessiné par un son évoluant dans la tessiture, c'est à dire quand l'ensemble de sa masse se promène dans le champ des hauteurs et y ondoie, pour dessiner des profils caractéristiques, que ces variations dans le registre des fréquences soient discontinues, scalaires (cas de la mélodie traditionnelle) ou continues, serpentines, en sons glissés. Ex. de sons : *Moteur électrique variable, roue de vélo à différentes vitesses, moteur à réaction...*

Dans notre système de recherche des sons, nous intégrerons un champ **profil mélodique**.

a.5. Le Calibre

Le **calibre** d'un son par rapport à sa masse complexe est son encombrement du champ des hauteurs, en fait, sa minceur ou son épaisseur en masse. Il ne concerne que les sons à masse complexe. Ce sont les variations continues en calibre (amincissements, épaississements) qui caractérisent les variations naturelles d'un son de vague. Lorsqu'on joue sur un son complexe dont la masse est un « paquet » d'une certaine épaisseur (son nodal), du filtre pour le sculpter, c'est sur le calibre des sons que l'on intervient.

Se présente alors, dans le système de recherche au champ **calibre** quatre options :

- Calibre fin**
- Calibre large**
- Calibre croissant**
- Calibre décroissant**

Calibre fin : minceur.

Calibre large : épaisseur.

Calibre croissant : le calibre passe progressivement de fin à large .

Calibre décroissant : le calibre passe progressivement de large à fin.

a.6. Le registre (site)

Le site d'un son (qu'il soit tonique ou complexe) est sa situation dans le champ de perception des hauteurs. Pour parler du **registre** fréquentiel du son et de son évolution dans le temps, on considérera qu'il est :

- Suraigu**
- Aigu**
- Médium aigu**
- Médium grave**
- Grave**

- Surgrave
- Aigu à Médium
- Aigu à Grave
- Médium à Aigu
- Médium à Grave
- Grave à Aigu
- Grave à Médium

a.7.Effet

L'environnement dans lequel le son est émis ou encore la manière dont une source sonore se déplace par rapport à un auditeur sont autant de paramètres qui influent sur la perception du son.

Nous avons la **réverbération** qui est le résultat de la réflexion d'une onde sur les parois du milieu qui l'entoure (sol, murs, plafonds, objets...).

Nous trouvons aussi l'effet d'**écho** qui est la réflexion d'une onde qui atteint l'auditeur au moins 50 ms après l'onde direct. Il y a alors une sensation de répétition. Dans la mythologie grecque, Echo est une nymphe des sources et des forêts du mont Hélicon (plus grande montagne de Béotie en Grèce) et la personnification de l'écho acoustique. Elle fut élevée par les nymphes et instruite par les Muses. Elle était très prolixe et inventait de nombreuses histoires : c'est ainsi qu'elle distrayait Héra (femme de Zeus) pour favoriser les amours de Zeus. Mais la déesse finit par s'apercevoir de ce manège et pour punir Echo, elle lui ôta la parole, l'obligeant à répéter ce que disait son interlocuteur. Elle tomba amoureuse de Narcisse, mais fut méprisée par celui-ci. Le cœur brisé, elle s'enfuit dans une grotte solitaire et se laissa dépérir. La légende dit qu'elle s'est tellement amaigrie qu'elle a totalement disparu et que seule sa voix lui reste. Ainsi est né le phénomène de l'écho.

Nous intégrons aussi un troisième effet acoustique naturel, l'effet **doppler**. Cet effet est le décalage entre la fréquence de l'onde émise et de l'onde reçue lorsque l'émetteur et le récepteur sont en mouvement l'un par rapport à l'autre ; il apparaît aussi lorsque l'onde se réfléchit sur un objet en mouvement par rapport à l'émetteur ou au récepteur. Ceci explique que la hauteur du son du moteur de voiture, ou d'une sirène d'un véhicule d'urgence, est différent selon que l'on est dedans (l'émetteur est immobile par rapport au récepteur), que le véhicule se rapproche du récepteur (le son devient plus aigu) ou qu'il s'éloigne (le son devient plus grave).

a.8 La spatialisation

Il s'agit de l'évolution du son dans l'espace, de sa position spatiale. On s'intéressera uniquement à la position spatiale du son sur le plan horizontal. Celle-ci est déterminée par deux paramètres : la distance et la position latérale. Pour chacun de ces paramètres, la position du son peut être fixe ou en mouvement. Donc deux paramètres à prendre en compte pour la recherche en fonction de la position spatiale, la position fixe et le mouvement :

Fixe

- A droite
- A gauche
- Proche
- Eloigné

Mouvement

- De éloigné à proche
- De proche à éloigné
- De gauche à droite
- De droite à gauche

Les mouvements pouvant être multiples, il est possible d'associer les positions fixes avec les mouvements. Ex. : A droite + De éloigné à proche ou Proche + De gauche à droite.

Nous pouvons aussi dans une optique d'intégrer des sons surround (4.0) intégrer les notions de l'avant à l'arrière, de arrière à l'avant.

b) Sphère Causale

La sphère causale est le domaine de l'identification des sources réelles ou supposées. Les **critères causaux** sont ceux qui résultent d'une écoute causale. Approcher le son de manière causale, c'est donc s'intéresser, se reporter à la cause réelle, supposée ou imaginée de ce son. Dans notre cas, il s'agira toujours de la cause puisque c'est précisément aux sons dont la cause réelle est immédiatement identifiable que nous nous intéressons.

Pour Michel Chion, la cause réelle d'un son peut désigner un corps sonore ou un complexe causal. Or, comme disait Schaeffer, « les mêmes causes ne produisent pas toujours les mêmes effets ». En effet, un même objet peut être à l'origine d'une multitude de sons différents. Un morceau de carton, par exemple, peut produire des sons très divers selon qu'on le gratte, qu'on le froisse, qu'on le déchire ou qu'on le découpe.

C'est pourquoi il est souvent nécessaire, pour désigner la cause réelle d'un son, de s'intéresser à l'**événement** qui en est à l'origine (« **complexe causal** »), plutôt qu'au « **corps sonore** », ce dernier restant trop ambigu, comme nous venons de le voir avec l'exemple du carton. C'est bien à un événement que l'on fait référence lorsqu'on parle d'une « porte qui claque » ou d'une « pierre qui roule » car en désignant que l'objet « porte » ou « pierre », les informations sont insuffisantes à définir le son résultant.

La description de la cause d'un son est plus aisée que la description du son lui-même, et cette approche est souvent utilisée. En effet, celle-ci permet de faire appel à des archétypes sonores qui parlent d'eux-mêmes et grâce auxquels il est possible de saisir le son instantanément. En revanche, ce mode de description peut s'avérer insuffisant pour différencier des sons très proche car il ne permet pas un niveau de détail de description très élevé. Mais cumulé aux différents critères de description proposés, il trouve indéniablement sa place dans le système de recherche. Les illustrateurs sonores, pour qui la recherche de sons est capitale, ne se préoccupent pas de la source de production des sons mais plutôt des sons eux-mêmes. Robert L. Mott, illustrateur sonore, explique que dans ce métier, le but est d'atteindre les attentes de l'auditeur. L'origine du son n'a que peu d'importance par rapport aux attentes de l'auditeur concernant la sonorité elle-même. Souvent, dans les films, les sons directs, naturels ne correspondent pas aux attentes de tout le monde. Le monteur son, remplace alors ces sons ou les modifie afin d'obtenir des sons convenant davantage à l'image, au scénario. Le choix des sons se fait par rapport à leurs qualités sonores et non par rapport à leur mode de production.

Dans notre système de recherche trois composantes de la description causale du son seront mises en place : l'**action** qui est à l'origine du son, le **matériau** qui caractérise la qualité du son et le **type d'excitation** de « l'objet sonore » qui permet de différencier des sons émis d'une même source.

b.1 L'action

Ce champ de description regroupe les verbes d'action à l'infinitif qui sont la cause d'un son. Le mode infinitif étant impersonnel (ne variant pas en personne), en considérant qu'il regroupe les diverses formes que peut prendre le verbe selon le mode (indicatif, etc.), le temps ou le nombre (singulier, pluriel), il permet de réduire considérablement la liste des

actions possibles. Une liste des verbes d'actions correspondant à la cause d'un son a été établie :

- Claquer
- Couper
- Déchirer
- Effleurer
- Froisser
- Frotter
- Glisser
- Gratter
- Percuter
- Pincer
- Racler
- Rouler
- Souffler
- Tomber
- Taper
- Tirer

Cette liste regroupe les actions qui d'après leur définition dans le Petit Robert produisent un bruit.

b.2 Le matériau

La description par le type de matériau impliqué dans la cause d'un son, donne une indication primordiale sur l'origine de la matière qui est à la source des vibrations acoustiques. Elle rejoint notre conditionnement des sons et nous réfère à la couleur du son que tel ou tel type de matériau nous fait penser. Ce champ d'attribut de recherche regroupe différents type de corps de matière : fluide, liquide, solide, synthétique, organique, végétal, tous impliqués dans la structure du corps de l'objet sonore.

- Air
- Bois
- Cristal
- Liquide
- Métal
- Papier
- Peau
- Plastique
- Terre
- Tissu
- Végétal
- Verre

b.3 Le type d'excitation

Le critère de type d'excitation déjà vu dans la première partie, pour la classification des sons Objet/Machine Mécanique sera réutilisé comme attribut de recherche, il sera intégré dans la recherche par mots clé et dans notre système de nomination des sons pour les sons

mécaniques. Ce critère de recherche permet de différencier les sons d'une même source. Ex. : une porte qui claque ne fait pas le même son qu'une porte qui grince. Dans le moteur de recherche nous intégrons cet attribut en plus de sa participation dans la nomination des sons mécaniques dans les critères de la sphère causale.

- Claquement**
- Couinement**
- Craquement**
- Crépitement**
- Crissement**
- Déchirement**
- Frottement**
- Grésillement**
- Grattement**
- Grincement**
- Pétilllement**
- Raclement**
- Roulement**
- Tapement**

Le type d'excitation est l'action qui génère le son.

c) Sphère Sémantique

La sphère sémantique, du sens, des significations, décrit les sensations, émotions et attributions de sens par l'homme à l'objet entendu. La reconnaissance et la mémorisation d'un son fonctionnent souvent de manière associative. Une des difficultés majeures rencontrée lors d'un processus de classification – recherche d'un son donné est la difficulté de mémoriser un son en soi, c'est à dire de construire une image mentale qui permette de le reconnaître et de s'en souvenir facilement. Le son du piano est associé à notre connaissance de l'objet et notre expérience d'écoute, mais pour les sons ne faisant pas partie du monde causal, qu'on n'associe pas directement à une cause précise, le processus de mémorisation est relativement long, plusieurs écoutes sont nécessaires avant de pouvoir écouter un son dans notre tête sans l'écouter. D'une manière générale, le sens attribué ou imposé à un son peut avoir des caractéristiques tout à fait subjectives (angoisse, chaleur) ou des caractéristiques qui font penser à des attributs morphologiques ou physiques mais qui relèvent du sens attribué (rapide, aigu...).

c.1 Adjectifs de perception

Il existe déjà un certain nombre d'adjectifs utilisés habituellement pour décrire des caractéristiques de notre perception. Des translations de sens sont souvent utilisées dans la description des sons étant donné leur efficacité subjective. Par exemple les concepts de « rugueux » et « lisse » ont été utilisés par Schaeffer dans sa morphologie. Il existe par ailleurs un nombre considérable d'adjectifs qui s'appliquent à la description des impressions causées par les sons, souvent extraits d'autres contextes.

Il a donc été établie une liste d'adjectifs pouvant au mieux parler des sons environnementaux non-musicaux, à partir de plusieurs travaux, ceux réalisés lors du projet ECRINS et de la thèse « Des sons aux mots, comment parle-t-on du timbre musical » (2000) de Anne FAURE, en

psychologie cognitive. Il s'agit de comprendre, à travers l'étude du timbre musical comment s'organisent les liens entre nos perceptions et les concepts qu'on y associe, révélés par des verbalisations. Effectivement, après étude des résultats des expériences effectuées auprès de musiciens. Expériences portant sur l'analyse perceptive, sur le ressenti et les adjectifs employés pour parler des sons musicaux, suite à plusieurs écoutes, des portraits verbaux ont pu être établis. De ces portraits verbaux musicaux, certains semblent être utilisés pour parler des sons environnementaux non musicaux. Une liste d'adjectifs serait la suivante :

Aigu, Agréable, Agressif, Attaque lente, Attaque rapide, Bouché, Brillant, Chaleureux, Clair, Constant, Continu, Diffus, Doux, Eloigné, Envoûtant, Etouffé, Faible, Feutré, Fort, Franc, Froid, Grave, Intérieur, Lisse, Long, Métallique, Nasillard, Net, Précis, Proche, Pur, Résonnant, Riche, Rond, Rugueux, Sec, Sourd

Aigu : notion de hauteur tonale

Agréable : sensation de bien-être à l'écoute et notion de belle esthétique du son

Agressif : sensation de désagrément, de gêne (ex. : fraise du dentiste)

Attaque lente / rapide : Dimension perceptive du son lié au temps d'attaque

Bouché : Absence d'aigu et de médium aigu

Brillant : Riche en harmonique aigu

Chaleureux : provoque une sensation de confiance

Clair : qualité agréable qui oscille entre un caractère plutôt aigu ou plus proche de la notion de brillance (luminosité)

Constant : son d'une seule fréquence, qui n'est pas interrompu dans le temps

Continu : son qui n'est pas interrompu dans le temps et qui garde la même variation dans l'espace

Diffus : Notion de netteté imprécise du son, l'extinction d'un son noyait dans son acoustique environnante.

Doux : Entre la notion d'agréable et de velouté

Eloigné : notion de position spatiale

Envoûtant : sentiment d'être englobé et transporté par le son (appréciation personnelle)

Etouffé : absence d'harmonique aigu et médium

Faible : notion de faible amplitude, souvent associé à la pauvreté et de l'intérêt du son

Feutré : comportement d'une matière dont l'effet sonore est d'amortir les réflexions du son

Fort : notion de forte amplitude, souvent associé à la puissance et la persistance d'un son

Franc : Attaque rapide, son sec et net

Froid : manque de définition dans l'esthétique et dans les harmoniques

Grave : présence d'harmoniques graves

Intérieur : son très présent, au premier plan sonore et absence presque de réflexions

Lisse : grain faible

Long : notion de résonance

Métallique : notion entre brillant et froid

Nasillard : approche similaire à la sensation perceptive de parler du nez

Net : attaque rapide

Précis : son directif, brillant, attaque rapide, intelligible

Proche : notion de position spatiale

Pur : une seule fréquence

Résonnant : Notion se rapprochant de l'effet d'écho et de réverbération

Riche : Son riche en harmonique, son opposé à un son synthétique, son « naturel »

Rond : son doux, voluptueux et présence d'harmonique dans les bas médium-grave

Rugueux : son au grain épais
Sec : notion temporel, attaque rapide
Sourd : notion de grave et de faible intensité

Cette liste sera accompagnée dans notre moteur de recherche de sons prototypes qui définiront leurs références. Elle est le résultat d'expériences menées par Anne FAURE et Emmanuel DERUTY.

c.2 Onomatopée

L'onomatopée consiste à former des mots imitant les sons (ex. :Boum, Ding etc.). Les bases de données d'échantillon commercialisées actuellement utilisent ce type de description qui fonctionne bien mais pour un nombre assez limité de sons seulement. De plus, certains sons peuvent être dits universels, comme l'abolement de chien, le miaulement du chat, quelques bruits organiques et naturels, la pluie sur le gravier et le caillou, mais en revanche leur écoute est culturelle et est dépendante de la culture proprement dite de chaque pays. Notre système de recherche intelligente des sons s'axe uniquement sur le Français. Le Français exprime bien le grave et l'aigu par l'utilisation des voyelles fermées et des voyelles ouvertes : voir la différence entre « clic » (cliquètement) et « tac » (dans claquement), « tic » et « tac », « plic » et « ploc ». Ses nasales exprime bien la résonance et ses diverses qualités, plus aiguës ou plus grave : le « ding », tintement clair, ou le « dong » plus grave. Ce type de recherche paraît pertinent car même s'il ne représente qu'un groupe limité de sons permet une identification du timbre harmonique de ceux-ci. On distingue entre une voyelle ouverte comme « a » et une voyelle plus fermée comme « o » un timbre harmonique différent. Nous utiliserons le critère de recherche par onomatopée en suivant le glossaire utilisé dans la bande dessinée, utilisé pour exprimer les sons avec un dessin.

Aboiements : OUAH! OUAH! - WHAF! - WAF! - WHOUAF!

Acclamations : HOURRA!.

Ane : HI-HAN - HI-HAN

Applaudissements : CLAP-CLAP! - KLAP - KALP!

Avertisseur : (automobile) TUT! - TUUUT! ; T-U-U-U-U-T!

(Vieux klaxon) COIN - COIN - POUËT - POUËT

(Voiture de pompiers) PIN-PON - PIN-PON.

Appel : HEP! - HELP! - HE!

(Entre les dents) PSSSSST - KSSSS

(Au loin) OHË ! – OOHËËË !

Bal : (orchestre) FLON-FLON

Bébé: AREU-AREU

Bêlement : (mouton) BÊÊÊÊÊÊÊ

Boire : GLOU-GLOU - GLOUP-GLOUP

(goulûment) GOUL-GOUL.

Chat : (miaulement) MIAOU

(lapant du lait) LAP-LAP

(ronronnement) RON-RON

Chute : (d'un corps sur le sol) POUF! - PAF! - PATATRAS - SPALF - BAM!

(D'un corps dans l'eau) PLOUF! - POF! - SPLOUF! - SPLAF! - GLOUP

(sur une surface métallique) DING - BING - DONG - BONG

Choc : (violent) BING! - CRAC! - CRAAAC! - KRAAAK! - CRASH! - BANG!

(Sur objet métallique creux) DOIING.

Clapotis : (eau, vague...) CLAP-CLAP.

Claquement : (fouet...) CLAC - CLAAAC - CLAK - SPLATCH
 (Bois sur bois) CLAP
 (Des dents) CLAC-CLAC - AGLAGLA - AGLE AGLE

Cliquetis : CLIC-CLIC

Cloche : (d'église) DING-DONG
 (grosse cloche, bourdon...) DONG-DONG - BONG-BONG
 (dans le lointain) DIN-DIN - TIN-TIN.

Coassement : (grenouille) COÂÂÂÄ-COÂÂÂÄ.

Cochon: (grognement) GROIN-GROIN.

Colère : (grognement) ARGN
 (rentrée) GREEEU.

Coq : COCORICÔÔÔÔÔ.

Coup : (sec) PAF! - CRAC! - CHLAC! - BANG!
 (De cravache) STICK!
 (De poing) BING! - SPLAF! - TCHAC! - POING - THAN
 (Coup de feu, fusil, revolver) BANG! - PAN! - PAW! - PIAW!
 (mitrailleuse) TATAATATA, TACATA-CATA , TA-TAC-TAC
 (canon) BAOUM , BOUM.

Court circuit : (électrique) SCHLAF! , SCRATCH!.

Cri : (de douleur) OUAÏE! , AÏE! , WHOUAAAH!
 (De joie) HIPIPIHOURRA!
 (D'étonnement) OOOH!
 (De frayeur) AAAAH! , AIEEEE! , IIIIII!
 (Petit cri d'étonnement) HEEEEEE?!?

Crissement : CRIIII, CRIIIS-CRIII
 (De chaussures neuves) CRITCH-CRITCH, CRATCH
 (De pneus) CRIIII-III.

Déchirement : (papier, étoffe...) CRAAATCH, SCRAAATCH

Déclat : CLIC, CLAC, FLIP, PLIP

Écrasement : (d'une manière molle) SPLAF! , POF! , POUF! , PLOUF!
 (D'une manière solide) CRAC! CRAAAC! , CRASH! , CRAAAK!

Effort violent : HAN! , IIIAAARRRR!

Éternuement : HAT... HAT... HATCHOUM

Étoffe : (froissement) FROU-FROU

Étranglement : (gargouillements) GARGL, ARGL

Explosion : BAOUM, BOUM, BOOOOM , BOMBADABOUM!

Fermeture : (rideau de fer) ZIP
 (Fermeture éclair) ZIIIIIP

Frapper : (à une porte) TOC-TOC
 (Du poing, à une porte) PAN-PAN
 (Avec un objet creux) CLONC, BONG, GONG
 (Sur du métal) KONNNNNG
 (Sur du métal creux) DOIING
 (Du pied avec impatience) TAF-TAF, TAP-TAP.

Fuite : (d'un liquide, de gaz, ...) PCHHHHHHH-PCHIII.

Galop : (cheval) TAGADA-TAGADA, TABADA-TABADA.

Gifle : BAF! , SPLAF! , PAF, SPAF! , BING!

Grattement : SCRATCH-SCRATCH.

Grésillement : (électricité) ZZZZZZZZZZ, ZZZIIZZZZI.
(moteur) CRIIITCH.

Grincement : (porte métallique) CRIIII-IIIIII.

Grogement : GREU, GREEEUUUU, GRRRR
(D'animale sauvage) : GRRR, GROAAR.

Horloge : TIC-TAC
(sonnerie) DONG-DONG.

Hélicoptère : (vol) FLAP-FLAP

Hennissement : (cheval) HIIIIII.

Hurlement : (de douleur) WHOUAH! , WHOOOAH! , HOULA!

Liquide : (projection de) SPLATCH
(coulant) GLOU-GLOU ; GOUL-GOUL ; COUL-COUL
(bouillonnant) PSCHII
(gazeux) PSCHIIIIIT

Machine : (à coudre) TIKETIKETIC, TIC-TIC,TIC-TIC
(D'usine) BAM-BAM, BONG-BONG
(À écrire) TAC-TAC-TAC,TACATACATA

Manger : (aliments liquides, soupe...) SLURP ; SLUUUUURP
(Aliments solides) SCRUNCH ; SCROTCH ;CRATCH-CRATCH
(De bon appétit) MIAM-MIAM

Marcher : (sur un vieux plancher) CRAAAC-CRAAAAC
(Sur un sol mouillé) FLIP-FLAP
(Dans la boue) PLATCH, FLOC, SPLATCH
(D'un éclopé) CLOP-CLOP
(Sur les pavés) CLAP-CLAP , TAP-TAP.

Moteur : (vieille voiture) TEUF-TEUF , PEUF-PEUF
(De course) : VROOOOM-VRAOUM.

Objet volant : (légers) VRRRRRRR , PFFFFFFT , PFIIIIT.

Oiseau : (battement d'ailes) FLAP-FLAP-FLAP, FLIP-FLIP
(cris) CUI-CUI.

Orage : (foudre) SCRAAAATCH
(Tonnerre) BAOUMBADABOUM

Raclement : (métal contre bois) CRAAAAASH
(Métal contre métal) CRIIIISH.

Reniflement : SNIF-SNIF

RESSAC : (mer contre quai, rocher...) CLAP-CLAP-CLAP

Rire : AHAHAHAH ! OH ! OH !AH ! AH ! AH ! HI ! HI !
(Gros rire) WHA HA ! HA !

Ronflement : (dormeur) ZZZZZZZZZZZZ, RRRRRRRR

Roulement : (objet lourd sur le sol) BADONG-BADONG.

Scie : ZZZZZZ

Sifflet : TRIT, TRIIIIT, TRIII, TRRRRRRIT

Son sourd : (déménagements...) BAOUM - BADABOUM – RAMDAM

Sonnerie : (téléphone) DRIIING, DRIIIIN
(trompette) TARATATA, TSOIN-TSOIN
(Électrique, alarme...) DRIIN, TUUUUT, UUUUT

Sonnette : (porte d'entrée) DILING-DILING
(clochette) DRELIN-DRELIN

Sirène : (bateau, usine...) TUUUUUT

Souffler : (après un effort) OUF, HOUF

Vent : (dans les arbres) WHOU-OU-OU-OU...

Verre : (choc sur un) DIIING

(Sur un verre de cristal) DIIIIIIIIIIING

(Se brisant) CRASH , GLING , GLIIING

Violon : IIIIII

(Interprétation déplorable) CRIN-CRIN

Volets : (fermeture) CLAP

Vibrations : DOIING

L'idéal est de pouvoir cumuler les possibilités de description (morphologique, causale, sémantique), les critères, et qu'ils soient proposés à l'utilisateur afin qu'il ait le choix de décrire un son donné selon tel ou tel critère, suivant ce qui convient le mieux au son pour une application donnée. Mais dans tous les cas, rappelons que ces différentes appréhensions possibles du son sont complémentaires et n'ont donc vraiment d'intérêt que si elles sont combinées entre elles.

3. MISE EN ŒUVRE DU SYSTEME DE RECHERCHE

3.1 Protocole de prise de sons pour la base de donnée audio

Nous allons aborder les différents types de prise de sons adaptées à l'enregistrement de sons environnementaux non musicaux : mono, stéréo, multicanal. La base de données intégrera ces différents protocoles qui seront fonction de la matière enregistré. Nous sommes dans l'optique d'optimiser la qualité de prise de sons afin que notre base de données soit alimentée de sons broadcast. Nous privilégierons les sons non compressés, c'est à dire que nous excluons le format Mp3 et nous éviterons le plus possible les sons enregistrés sur support Mini Disc qui applique une compression à l'enregistrement. Il faut privilégier la qualité du registre fréquentiel des fichiers sonores, ce qui implique un support d'enregistrement professionnel (Nagra, DAT, Enregistreur numérique sur compact flash ou disque dur), la résolution doit être au minimum de 16 bits et 44.1 kHz , pour un support DAT elle est de 16 bits et 48 kHz , et les nouveaux support numériques (enregistreurs sur disque dur ou carte compact flash) proposent aujourd'hui des résolutions allant jusqu'à 24 bits et 192 kHz. Néanmoins, les fréquences d'échantillonnage supérieur à 48 kHz sont peu intéressantes car elle n'ont que peu d'influence sur la qualité des sons environnementaux, elles sont une application plutôt dédié à la musique et présente une augmentation du poids du fichier sonore. Par contre, une résolution de 24 bits propose une aisance à la prise de son sur la gestion de la dynamique du son, on obtient une dynamique beaucoup plus grande et nous évitons des problèmes d'écrêtage du son. Quant au choix des micros, il est clair que ceux ci doivent être obligatoirement statiques pour privilégier leur sensibilité et leur courbes de réponse large en spectre fréquentiel. Il existe une série assez grande de micros adaptés à la prise de sons environnementaux, pour ne citer que les meilleurs et les plus favorables à une qualité professionnel nous avons les marques Schoeps et Neumann. Il me semble nécessaire d'insister sur l'exigence d'un choix professionnelle de matériel d'enregistrement car celui ci influe énormément sur la qualité d'exploitation du son, néanmoins d'autres paramètres rentrent en compte dans la qualité du son, moins technique : l'analyse à écouter la qualité d'un son à son enregistrement.

A) Prise de son Monophonique

La prise de son monophonique s'applique aux sons que nous avons dénommé Evènement (son plus ou moins court). Elle est une prise de son de proximité, car nous cherchons à obtenir la prise la plus précise possible dans le timbrage et la dynamique. Pour cela, il faut privilégier les directivités cardioïde et hyper cardioïde qui permettent d'être les plus précises sur la captation de l'évènement sonore. L'utilité d'avoir un son monophonique est de pouvoir exploiter facilement sa spatialisation dans la procédure de mixage. Par exemple lors d'une séance de bruitage, les prises sont effectuées en mono, souvent avec plusieurs micros (pas, déplacement...), disposés pour obtenir un son dans son espace (ex. : des pas qui arrivent de loin et qui se rapprochent de plus en plus, l'ingénieur du son peu ainsi mixer à la prise avec plusieurs micros cette effet de rapprochement) . Bien sur, la prise de proximité n'est pas exclusive à la monophonie, elle peut dans certain cas s'avérer intéressante d'être en stéréophonie mais en prise d'intensité (XY). En résumé, la prise monophonique est a privilégier pour l'enregistrement d'un évènement sonore précis isolé de son environnement, elle permet d'obtenir une source sonore au premier plan, elle est un « zoom focal » du son ,

voir même une « macro » du son.

B) Prise de son Stéréophonique

La prise de son stéréophonique intervient dans l'enregistrement d'ambiance sonore environnemental et dans le rendu à la prise de la spatialisation, de la position dans l'espace, de l'évolution spatiale d'un événement sonore. Elle recrée notre écoute du monde par nos deux oreilles similaires. Les techniques de prise de son stéréo entrent dans deux catégories principales : les techniques coïncidentes et les techniques non-coïncidentes. Avec les techniques coïncidentes, les micros sont placés extrêmement près les uns des autres, tandis qu'avec les techniques non-coïncidentes, ils sont plus espacés. Les méthodes coïncidentes sont considérées comme étant très précises, elles reproduisent notre perception stéréophonique. Pour élargir le champ stéréo on éloigne légèrement les micros coïncidents afin qu'un décalage temporel soit créé.

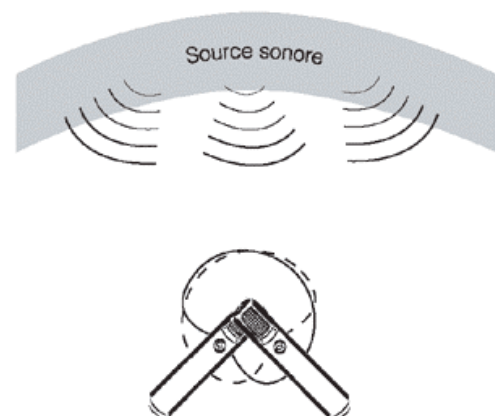
Dans le cadre de la prise de son d'événements sonores destinés à être intégrés dans une sonothèque nous privilégierons les techniques coïncidentes et non coïncidentes, qui permettent d'élargir le placement virtuel de nos oreilles ; nous nous recherchons au maximum la fidélité du son.

Prise de son d'intensité X-Y

La technique X-Y de prise de son emploie un couple apparié de micros qui "s'imbriquent" autant que leurs corps le permettent. Comme illustré, on place un couple de micros à directivité cardioïde aussi proches que possible l'un de l'autre de manière à ce que leurs capsules forment un angle. Le micro de gauche capture le signal venant de la droite et vice versa. Si 90 degrés est l'angle qui est le plus utilisé, la fourchette de fonctionnement se situe environ entre 60 et 130 degrés.

Généralement, la distance par rapport à la source sonore combinée avec l'envergure stéréo voulue permet de déterminer l'angle approprié.

L'utilisation d'une directivité cardioïde signifie que la configuration X-Y dans son ensemble rejette les signaux qui viennent de l'arrière. Ce rejet de l'arrière présente plusieurs avantages. Cette réduction du désordre acoustique est également un atout lorsqu'il s'agit de convertir l'enregistrement stéréo en enregistrement mono. Si on augmente la distance entre la paire de micros coïncidents et la source sonore, la séparation stéréo diminue et davantage de réflexions sont capturées. La technique X-Y, qui utilise des cardioïdes, génère une image stéréo précise présentant des réflexions acoustiques minimales, bien que la séparation ne soit pas aussi significative qu'avec d'autres techniques de prise de son stéréo.

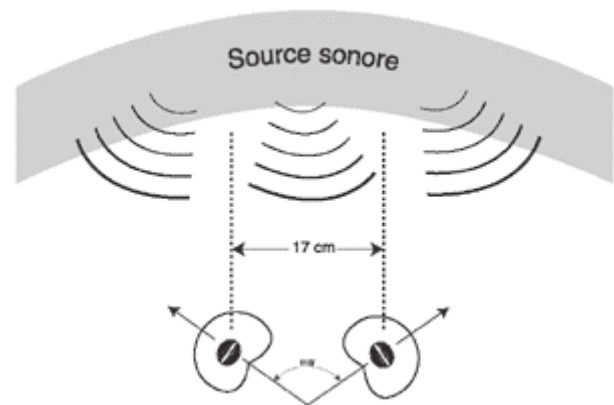


La prise de son X-Y emploie un couple apparié de cardioïdes coïncidents.

Prise de son ORTF

Conçue par l'ancienne agence de diffusion nationale française, l'Office de Radio Télévision Française, la technique ORTF vise à imiter l'emplacement des oreilles de la tête d'un adulte. Deux capsules cardioïdes sont placées à 17cm l'une de l'autre et forment un angle de 110 degrés. La technique ORTF peut produire une image et une profondeur large, mais l'utilisation de cardioïdes a pour effet de capturer beaucoup moins de réflexions réverbérées.

La distance spécifiée pour la technique ORTF assure la cohérence de phase des longueurs d'onde inférieures à environ 500 Hz. Les décalages temporels ou incohérences de phase au-dessus de cette fréquence contribuent généralement à donner une impression de séparation stéré associée à la perception d'une qualité de son ouverte ou aérienne. La technique ORTF offre en outre une compatibilité monophonique suffisante.



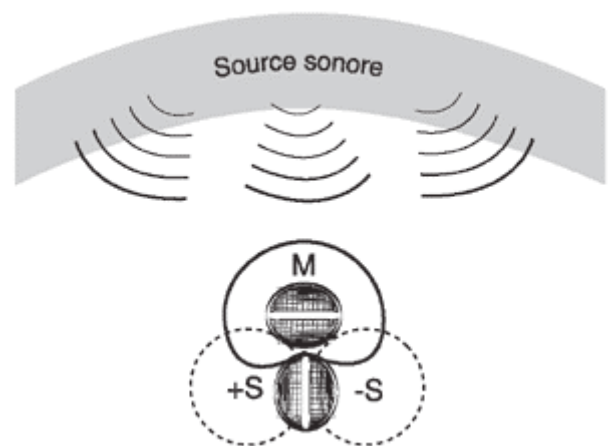
La technique de l'ORTF consiste à placer un couple apparié de microphones selon une configuration similaire à celle des oreilles humaines.

Prise de son M-S (Middle-Side)

La technique M-S utilise un traitement spécial pour capturer des images stéréo très précises tout en assurant une excellente compatibilité mono. Un micro "middle" ou "central" en français (en général un cardioïde) est placé en face du centre de la source sonore et capture les sons principaux. Un micro à directivité en 8 ("side" ou "latéral" en français) est placé sur le même axe vertical, ses lobes en face des côtés droit et gauche, ce qui lui permet de capturer les informations des extrémités droite et gauche grâce au rejet latéral propre à la directivité en 8.

Cette configuration ne devient stéréo que lorsque les signaux ont été traités à l'aide d'une matrice d'encodage M-S. L'encodeur combine les signaux centraux et latéraux pour créer un côté du signal stéréo et retranche le signal latéral du signal du milieu pour créer l'autre côté. Le résultat est une traduction très précise du champ d'écoute stéréo. La présence d'un contrôle de balance M/S sur l'encodeur permet également de contrôler le rapport signal central/signal latéral, et donc l'ampleur perçue du champ stéréo.

La technique M-S est l'unique technique de prise de son stéréo ne se basant pas explicitement sur un couple apparié de micros. Mais il est impératif d'employer des micros de haute qualité pour obtenir une intégrité acoustique globale ainsi que pour garantir l'équilibre entre les micros bidirectionnels. De plus cette technique offre une grande souplesse puisque le micro central ne doit pas nécessairement être un cardioïde. Si l'on souhaite davantage de présence



D'un point de vue électronique, la technique M-S dérive un signal stéréo à partir d'un micro central coïncident à directivité en 8.

ou de réflexions de l'arrière, un micro omnidirectionnel peut être utilisé en tant que micro "central". La technique M-S offre également une excellente compatibilité mono puisque la recombinaison des deux signaux latéraux déphasés entraîne leur annulation pour ne laisser que le signal du milieu. Ce procédé réduit simultanément les réflexions latérales pouvant provoquer la confusion dans une conversion mono.

En résumé, ces types de prises de sons stéréophoniques offrent une précision de la restitution de l'espace stéréo et une compatibilité monophonique qui peut s'avérer utile laissant de nombreuses possibilités de choix de mixage.

C) Prise de son Multicanal

Dans notre système de recherche, nous devons prendre en compte l'intégration de fichiers audio multicanaux, notamment issu de prises de son surround (4.0). Nous sommes dans une époque où les supports et les systèmes de diffusion développent le son multicanal (DVD, HomeTheatre, SACD), le souci de recréer une meilleure localisation du son dans l'espace, de produire une sensation d'enveloppement agréable du son devient une priorité dans le cinéma, la musique etc. Auparavant, le multicanal ne prenait forme qu'à l'étape de mixage, désormais, la chaîne de production du son s'adapte à ce système, le montage son et la prise de son intègrent cette nouvelle possibilité.

Ce paragraphe ne vise pas à être une théorie de la prise de son multicanal, domaine encore en développement et en expérimentation, mais une approche de la prise de son multicanal. Les avantages qu'elle apporte sont divers, reproduction de l'espace sonore d'une ambiance de lieu ou d'action, captation des réflexions du son, fidélité de réverbération. Il existe de nombreux systèmes de captation multicanal plus ou moins coûteux (holophonie, ambisonie), les moins coûteux et les plus simples à mettre en place sont les techniques que nous allons aborder.

Le Decca-Tree

Le Decca-Tree est un ensemble composé d'un triplet microphonique espacé entre eux de 1.5m à 2m. Les microphones sont omnidirectionnels. L'utilisation de trois capsules pour la reproduction frontale génère trois images fantômes qui tendent à rendre plus confuse la précision. Deux microphones placés à l'arrière servent à nourrir les enceintes surround afin d'assurer une compatibilité en 5.1. La distance séparant les microphones surround des principaux, permet d'ajuster l'image subjective de la prise de son. Ce système donne de bons résultats quant aux critères de spatialisation et du renforcement de l'image centrale.

La Tétraphonie

La tétraphonie est un système d'enregistrement avec deux couples ORTF qui donne des résultats intéressants car il génère un codage frontal et arrière de la localisation. Cette captation est assez souple car la distance entre les deux couples a peu d'influence. L'interaction entre les deux couples contribue à recréer une spatialisation agréable. C'est une méthode de captation qui ne prend nullement en compte la localisation latérale, mais qui devant la difficulté de perception de la direction des images fantômes dans cette zone, est certainement le système d'enregistrement multicanal donnant les meilleurs résultats en regard de sa facilité de mise en œuvre. C'est dans les situations où deux masses sonores seront réparties l'une à l'avant, l'autre à l'arrière que ce système donnera de meilleures performances (enregistrement d'un orchestre et d'un orgue dans une église par exemple).

Bien sur, il existe d'autres techniques qui peuvent être efficaces, comme l'utilisation de deux couples M-S (avant-arrière) ou la combinaison d'un couple M-S frontal et d'un couple ORTF arrière... Ces systèmes permettent de pouvoir gérer l'élargissement ou le rétrécissement de l'espace sonore frontal et arrière. La base de donnée intégrera des sons multicanaux à long terme, elle l'intègre déjà dans le moteur de recherche avec le critère morphologique de la spatialisation où la recherche descriptive prend en compte les notions de avant – arrière.

Conclusion

Voici donc un aperçu, des types de prises de son qui semblent les plus adaptés pour être intégrés dans la base de donnée audio. Néanmoins, ces techniques doivent dépendre de l'écoute du preneur de son qui se doit de réfléchir pour déterminer le bon compromis entre le type de dispositif, la bonne distance à la source, la position des micros en fonction des sons à enregistrer et du rendu souhaité. Le preneur de son doit comprendre ce qui l'enregistre, l'intention qu'il veut donner à sa prise.

3.2 Protocole d'intégration et d'indexation des sons dans la base de données audio

A) Protocole de conformation du son

Cette sous partie vise à structurer une méthode de travail pour l'optimisation de la qualité d'exploitation des sons intégrés à la base de donnée audio. Nous avons vu précédemment, la première étape, celle de la prise de son ; la deuxième étape consiste à transférer les supports d'enregistrements en fichiers informatiques à l'extension .wav ou .aiff. Le transfert s'effectue dans un logiciel d'édition (Pro Tools, Soundforge, Nuendo...), pour le transfert d'un support analogique vers un fichier informatique, nous parlerons de numérisation, d'une quantification de 16 bits et d'une fréquence d'échantillonnage de 44.1kHz (Stéréo ou Mono). Quant au transfert de supports d'enregistrement numérique (DAT, enregistreur sur compact flash ou disque dur), la conversion analogique-numérique s'étant établie à la prise, nous parlerons de transferts de données audio, pour un DAT la résolution du son est de 16 bits et 48 kHz, pour des supports sur disque dur et carte compact flash la résolution du transfert dépend de la résolution établie à l'enregistrement de la prise, elle est de 16 ou 24 bits et de 44.1 kHz à 192 kHz.

Ensuite, intervient une étape de dérushing. Il s'agit alors de sélectionner les sections audio du fichier en fonction de sa qualité et de sa pertinence pour en déterminer les points de début et de fin, on peut éventuellement monter le son (pour éliminer des parties indésirables du son) en établissant des raccords ajustés à un fondu croisé adapté à une continuité sonore du son. On applique ensuite un léger fondu d'entrée et de sortie d'une durée d'environ 200 ms aux points respectifs du fichier audio. Le son peut nécessiter un nettoyage de bruits indésirables comme des bruits de perche ou des infra basse dû au vent, généralement ils sont éliminés à la prise à l'aide d'un filtre coupe bas ajustable de 80 à 120 Hz, mais certains enregistreurs n'en sont pas équipés, on peut donc appliquer ce type de filtrage au moment de la conformation du son. Il faut aussi appliquer un limiteur de crête à 0 dB en sortie pour appliquer une légère atténuation du son dans ses pics de dynamique, son temps de relâche est très court, d'environ 10 ms pour éviter la sensation de « pompage » du son. Enfin, la dernière

étape de conformation consiste à normaliser le son à 0 dB, on maximise les niveaux. Le processus de normalisation des niveaux est relativement simple. Le logiciel repère le niveau maximal existant dans le fichier audio, estime sa différence avec le niveau numérique maximal possible (full scale) et amplifie en conséquence tout le fichier, de façon à aligner le niveau le plus élevée sur le niveau maximal possible. Néanmoins, les bruits de fonds sont amplifiés, mais sont traités lors de l'étape de mixage.

Récapitulons les manipulations du fichier audio pour sa conformation à intégrer la base de donnée :

- Dérushage, et éventuellement raccord des parties du son entre elles par fondu enchaîné
- Fondu d'entrée et de sortie d'environ 200 ms
- Filtre coupe bas de 80 à 120 Hz
- Limiteur réglé à 0 dB en sortie et un release court d'environ 10 ms
- Normaliser à 0 dB

Une fois ce processus de conformation effectué, il s'agit d'exporter notre fichier audio avec un nom de fichier correspondant à l'étude de la 1^{er} partie de ce mémoire.

B) Nom et indexation des fichiers audio

Pour la nomination des sons nous avons étudié auparavant la classification des sources, notre fichier audio sera classé parmi sa catégorie correspondante ; il est inutile de rappeler la hiérarchie de catégorie des sources dans le nom du fichier. Le nom de fichier doit être une description générale du son, le nom de l'objet ou de la machine, le contexte du son, les actions...

Ex. : - Applaudissements Théâtre
- Cri de femme
- Femme pleure
- Marché
- Coups de poing
- Activité calme sur chantier
- Ville-piétons et passage voiture
- Machinerie chantier
- Rasoir
- Alarme Voiture
- Main sur porte en métal
- Chant d'oiseau au bord d'une crypte

Pour la catégorie des sons Nature description de l'ambiance ou de l'événement EAU, AIR, TERRE ou FEU

Pour la catégorie des sons Animaux description de l'Action ou de l'Expression

Pour la catégorie des sons Homme description de l'Acteur(femme, enfant, homme, bébé, foule), de l'Action ou de l'Expression.

Pour la catégorie des sons Objets-Machines description de l'objet ou de la machine, de l'action de l'objet ou de la machine, de son contexte ou environnement sonore (Domestique, travaux...)

Le fichier audio est ensuite exporté et sauvegardé dans un répertoire qui est le « container » des sons à indexer dans la base de données ; l'indexation s'effectue à partir de l'interface de l'outil informatique. Lors de cette indexation, nous classons le son dans l'arborescence de sa catégorie. Puis, après une écoute morphologique, causale et sémantique on détermine les critères descriptifs vus dans la deuxième partie de ce mémoire, correspondants au son. Le

fichier audio est ainsi indexé dans la base de données, il est classé et nommé selon notre modèle et contient nos attributs de descriptions.

C) Le moteur de recherche

Le moteur de recherche est programmé en langage de programmation PHP et SQL, langage adapté à la construction d'une base de données. Le moteur de recherche est donc un outil informatique qui présente dans son interface, une recherche par mots clés qui correspond à la nomination du son et à son classement, cumulé avec la possibilité de préciser les critères de description du son (morphologique, causal, sémantique). Il propose ainsi des réponses en fonction de la recherche désirée ; plus la recherche est précise plus les réponses seront précises. On obtient donc un compte rendu avant même l'écoute du son, résumant l'appartenance de sa catégorie de sources sonores et l'ensemble des critères descriptifs qui le compose. Nous avons donc, avant même d'avoir écouté le fichier audio, une idée de ce qui l'est. Nous arrivons donc à l'objectif de départ pouvoir obtenir un son souhaité à travers une recherche complète de ces caractéristiques. Une fois le son désiré obtenu, il sera à travers l'outil informatique de recherche intelligente des sons, extrait vers un répertoire « container » qui regroupe tous les sons répondants aux attentes de l'utilisateur.

Ensuite, il ne reste plus qu'à l'utilisateur de l'exploiter dans son logiciel d'édition. Le moteur de recherche doit être utilisé dans de nombreux domaines du son : cinéma, audiovisuel, design sonore, radio, musique, multimédia, éducation, architecture, design sonore industriel... Il demeure un outil de base pour la réalisation et la création sonore.

CONCLUSION

Ce mémoire présente un intérêt dans la façon de classer les sons environnementaux par source et de les décrire ; il propose de mieux comprendre et d'appréhender la réalité de notre environnement sonore. La classification des sons par catégorie propose une hiérarchie et une organisation des sources sonores de notre planète. Elle est une taxonomie des sons qui englobe la plupart des sons environnementaux, néanmoins elle n'est qu'une base, qui reste ouverte à des améliorations issues d'une mise en pratique qui permettra de valider l'efficacité d'une telle classification. A travers une bonne organisation des sons, les critères de description proposés sont eux aussi assujettis à une mise en pratique pour que leur pertinence soient validées. Le solfège traditionnel ne permet pas une description comme nous avons pu l'étudier, nous nous sommes très vite rendu compte que les seuls paramètres de hauteur, de durée, d'intensité et de timbre sont bien largement insuffisants pour décrire les sons non-musicaux. La description mise en œuvre se concentre sur les trois écoutes du son, elle prétend pouvoir aux travers de critères morphologiques, causaux et sémantiques permettre à n'importe quel type d'utilisateur (des plus avertis au moins avertis au son) de travailler plus rapidement dans son approche de la recherche d'un son.

Le contenu de ce mémoire est une synthèse, un approfondissement, une recherche issue de nombreux travaux et d'une longue réflexion sur comment définir et parler de ce que nous avons trop longtemps appelé « bruit ». J'espère que d'autres critères de description pourront s'ajouter à ceux cités, que la mise en œuvre d'un tel outil informatique de recherche puisse être pertinente dans ses propositions et auprès des utilisateurs. De plus, nous sommes dans une actualité technologique où la description des médias multimédias devient un besoin pour le web. Notamment, le standard MPEG-7 qui se propose de standardiser la représentation des contenus de données multimédias, dispersées dans de nombreuses bases de données, pour la recherche et le filtrage d'information. Beaucoup d'applications peuvent bénéficier du MPEG-7, telles que la recherche dans les bases de données audiovisuelles ou la sélection de programmes diffusés. Avec la classification et la description étudiée avec ce mémoire, il peut y avoir des éléments correspondants aux besoins du MPEG-7.

L'avenir de l'outil informatique serait dans le développement d'outils logiciels ayant la capacité de calculer et d'analyser automatiquement les critères de description. Ce qui permettrait d'avoir un logiciel intelligent qui classerait et attribuerait en autonomie, tous les éléments de définition du son environnemental non-musical. Ainsi, ce pour lequel nous nous étions peu préoccupés, « le bruit », aura désormais une identité, une place dans une hiérarchie donnée et obtiendra une nouvelle considération.

Bibliographie

LIVRES

CHION, Michel, *Le Son*. Paris, Armand Collin (2004).

GUYOT Frédérique, CASTELLENGO, Michèle, FABRE, Benoît, *Etude de la catégorisation d'un corpus de bruits domestiques*. Chapitre 2, Paris, PUF (1998).

SCHAEFFER, Pierre, *Traité des objets musicaux*, Paris, Seuil (1966).

THESES ET MEMOIRES

DECAUX, Sébastien, *Banques de sons*. Sous la direction de Rémi Adjiman. Mémoire de Maîtrise : Image et Son : Université de Provence : 1996.

DERUTY, Emmanuel, *Projet ECRINS*. Sous la direction de Daniel Teruggi. Mémoire de DESS : Acoustique architecturale et Urbaine : Université Pierre et Marie Curie, Paris 6 : 2001.

FAURE, Anne, *Des sons aux mots, comment parle-t-on du timbre musical ?*. Sous la direction de M. Stephen McAdams. Paris : EHESS, 2000. Thèse de doctorat : Sciences Cognitives : Université de Paris 5 : 2000.

HOUIX, Olivier, *Catégorisation auditive des sources sonores*. Sous la direction de René Caussé et Stephen McAdams. Thèse de doctorat : Acoustique : Université du Maine : 2003.

Netographie

Comparisonic Corporation, *Find Sound*, Adresse : <http://www.findsounds.com/types.html>
Dernière mise à jour juin 2006. Moteur de recherche de sons en ligne.

Audible Magic Corp., Los Gatos, California, USA, *Muscle Fish*, Adresse : <http://www.musclefish.com/>
Dernière mise à jour 2006. Groupe d'ingénieurs travaillant sur la recherche de sons musicaux, développement de logiciels.

Sound Ideas Corp., Ontario, Canada, *Sound ideas*, Adresse : <http://www.sound-ideas.com/>
Dernière mise à jour 2006. Vente en ligne de sonothèques.

<http://www.univers-son.com>

Dernière mise à jour juin 2006. Moteur de recherche de liens sur le son (logiciels, hardware, microphones...)

Banque de données

Sound Ideas Serie 6000 General Sound Effects Library [banque de données sur CD].
Version 1. Sound Ideas Corp., Ontario, Canada. 1992.

Les "bruits" de tous les jours qui nous entourent, les sons de la ville, de nos maisons, de notre Nature, de nos machines, des animaux, les bruits propres à l'homme forment notre environnement sonore quotidien et familier. Contrairement aux sons musicaux que l'on peut classer par tonalité, fréquence, gamme, genre musical, instrument...les sons environnementaux sont plus complexes à organiser, rassembler, décrire, classifier de part les multiples façons de les générer (multitude des sources sonores, multitudes des fréquences écologiques).

Les mots qualificatifs, descriptifs des sons sont très peu employés dans notre quotidien, nous gardons le silence sur la forme, la substance auditive que nous entendons comme représentation d'une cause d'un sens.

La multiplication des données informatiques nécessite des interfaces de recherches pertinentes navigants dans des bases de données audio. C'est pourquoi, il devient aujourd'hui un besoin dans les structures de post production audiovisuelle, d'établir une théorie pour la description des sons à travers des critères posés. Savoir mettre des mots aux sons et des sons aux mots.

Mots-clés : Sons environnementaux non-musicaux, classification, description, base de données audio, moteur de recherche, objet sonore, écoute réduite, perception cognitive, représentation sémantique.

The "Noise" of every day, which surrounds us, the sounds of the city, of our houses, of Nature, of our machines, of animals, human's sounds make our environment vibratie with sound. Contrary to musical sounds which can be classified by tonality, by frequency, by scale, by musical instrument...the environmental sounds are more complex to organize, to describe, to grade because there are many ways to generate them (many roots of sounds, many ecological frequencies). Descriptive words used to speak about sounds are not employed in our life. We remain silent about the auditive structure heard as representation of a cause, of a sense. The multiplication of computer data needs a good sound browser which can search effectively in a sound data base. For this reason, there is a need in audiovisual postproduction to establish a theory about the description of sounds with interesting criteria. To know how to describe sounds with words.

Key words : Environmental sounds, classification, description, sound data base, sound browser, sound object, reduce listening, cognitive perception, semantic representaion.