THÈSE de DOCTORAT de L'UNIVERSITÉ DU MAINE

Spécialité :

ACOUSTIQUE

présentée
par

M. Vincent ROUSSARIE

pour obtenir le grade de

DOCTEUR de L'UNIVERSITÉ DU MAINE

Sujet de la thèse :

ANALYSE PERCEPTIVE DE STRUCTURES VIBRANTES SIMULEES PAR MODELE PHYSIQUE

Thèse soutenue le 17 Décembre 1999
devant la commission d'examen
composée de Messieurs

P. AVAN (Rapporteur)	Professeur, Clermont-Ferrand
G. CANEVET (Rapporteur)	Professeur, Marseille
R. CAUSSE (Examinateur)	Chercheur, Paris
A. CHAIGNE (Co-Directeur)	Professeur, Paris
S. McADAMS (Co-Directeur)	Directeur de recherche, Paris
S. SAHRAOUI (Président)	Professeur, Le Mans

Table des matières

Remerciements

Je tiens à remercier tout spécialement Stephen McAdams pour ces trois années passées dans la confiance et l'enthousiasme, pour son aide appuyée dans les moments difficiles et enfin pour son optimisme inépuisable qu'il sait partager sans cesse.

Je remercie Antoine Chaigne pour son soutien technique et moral et pour m'avoir proposé avec S. McAdams ce beau sujet. Merci à l'ENST pour son acceuil.

Je voudrais également remercier M.M. Paul Avan, Georges Canévet, S. Sahraoui ainsi que René Caussé pour me faire l'honneur de bien vouloir participer à mon jury de thèse, d'autant qu'il ne correspond pas nécessairement à la spécialité de chacun.

Merci à l'IRCAM et à Hugues Vinet, Directeur Scientifique, de m'avoir acceuilli au sein d'un institut aussi passionant.

Merci à l'Université du Maine d'avoir accepté un projet qui ne soit ni directement sa spécialité, ni qui s'est déroulé sous son toit, mais qui fait de nous quand même, des fils du LAUM.

Je remercie Vincent Doutaut et Christophe Lambourg pour m'avoir permis de comprendre et de rentrer dans la synthèse par modèle physique tel qu'ils l'ont programmé.

Je tiens à remercier bien sûr tous les membres de l'équipe Perception et Cognition Musicales, en particulier Anne Faure, Caroline Bey, Sophie Donnadieu et Cécile Marin pour un peu de douceur dans ce monde de bruit. Mais aussi Bennett Smith pour ses idées, toutes ses idées, Patrick Susini pour sa bienveillance méridionale, ainsi qu'Olivier Lin'houix, Alain de Cheveigné et

Gérard Bertrand pour leurs conseils et leur aide technique.

Enfin je voudrais remercier ceux qui ont souffert d'être autour de moi, ma petite Sarah, correctrice implacable mais aussi patience et douceur de chaque jour, Florence et César qui m'ont acceuilli et ont été pour moi simultanément mes hôtes, mes conseils et ma famille bienveillante, mes parents imperméables à mes doutes, et mes amis, en particulier Pascal Durrieu, chef latex et caporal zinc.

aux femmes ...

Le phénomène est bien tel qu'il nous apparaît,
mais ce n'est pas la chose en soi.
(E. KANT)

Chapitre 1

Introduction

1.1 Contexte de l'étude

Partons du processus complet de l'audition. Une source est excitée par une impulsion quelconque et rentre en vibration. Les vibrations internes de l'objet vont mettre en vibration les couches limites d'air qui sont à la surface de l'objet propageant ainsi une onde dans l'air sous forme d'une variation de pression. Cette variation de pression se propage ensuite jusqu'au tympan et vient l'exciter à son tour. Celui-ci est relié à une chaîne ossiculaire, le marteau, l'enclume et l'étrier, qui vont transmettre à leurs tours les vibrations à la cochlée au travers de la fenêtre ovale. Les vibrations vont alors être traduites en impulsions nerveuses par les différentes parties de l'oreille interne. L'onde sonore se propage le long de la membrane basilaire délimitant le canal cochléaire situé dans la cochlée elle-même. L'organe de Corti, porté par la membrane basilaire, contient les cellules ciliées internes qui ont pour fonction d'effectuer la transduction mécano-électrique. L'excitation de ces cellules va donner lieu à des potentiels d'action qui vont stimuler les fibres, chacune reliée aux différentes cellules ciliées. Toutes les fibres auditives aboutissent après divers relais au cortex auditif.

Intervient alors le processus d'organisation auditive qui va aboutir à la perception du son. Il peut se décomposer en trois étapes [49]. Dans un premier temps il faut séparer et regrouper les impulsions issues de la même source sonore, ce sont les processus de groupement auditif. Ensuite il faut extraire les propriétés auditives qui vont caractériser la source, c'est le processus d'analyse des données auditives. Enfin, selon les résultats de l'analyse des données, soit l'auditeur aura accès à un lexique auditif correspondant, il reconnaît la source, soit il percevra le son à partir de ses propres associations à des événements sonores et il interprétera alors l'événement.

1.2 Plan de la thèse

Notre thèse étudie les différents processus mis en œuvre dans la figure 1.1 à travers deux attributs essentiels de la perception auditive, la hauteur et le timbre. Nous commencerons donc par rappeler les différentes étapes des études précédentes sur la hauteur et le timbre en général afin de mieux comprendre les méthodes et les questions que posent la psychoacoustique. Les méthodes d'analyses de données employées dans cette thèse seront résumées également. Ensuite nous présenterons la théorie écologique à travers diverses études et expériences sur ce sujet. Pour comparer ces différentes vues du chemin de la perception nous aurons besoins de stimuli "naturels" dont nous puissions néanmoins contrôler les différents paramètres de la source afin de faire l'étude du schéma complet. Pour cela nous présenterons les modèles de synthèse numérique de barres et de plaques par modèle physique. Ensuite nous décrirons les expériences réalisées sur les barres et les plaques afin de remonter les différentes étapes possibles et d'en comparer les résultats, sur quoi nous concluerons.

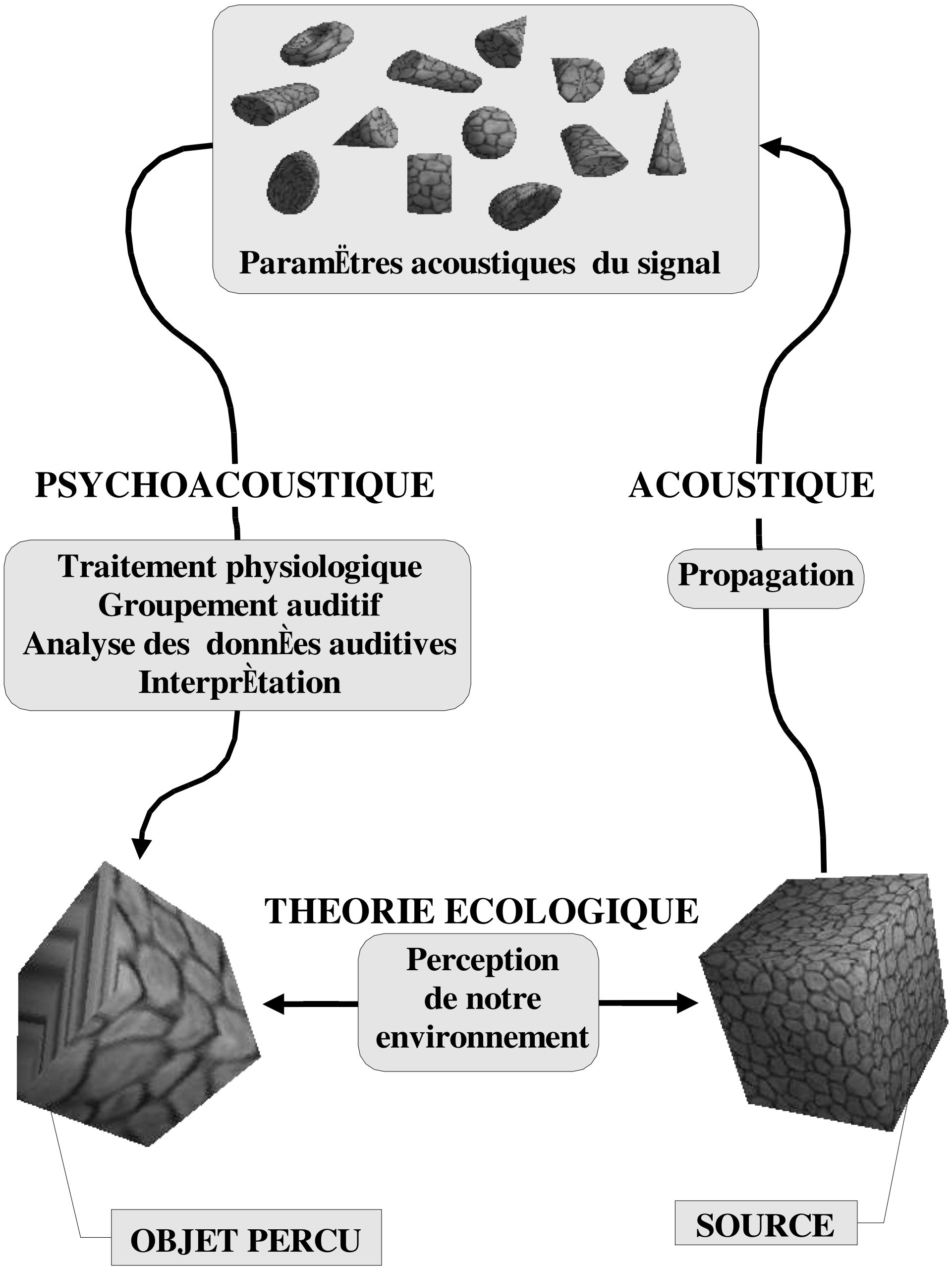
ParamËtres acoustiques du signal
PSYCHOACOUSTIQUE
ACOUSTIQUE
Traitement physiologique
Groupement auditif
Analyse des donnÈes auditives
InterprÈtation
Propagation
THEORIE ECOLOGIQUE
Perception
de notre
environnement
OBJET PERCU
SOURCE

Fig. 1.1 –

Chapitre 2

Présentation des études classiques de la psychoacoustique

2.1 La hauteur

2.1.1 Introduction

Le système auditif humain extrait des sons qu'il reçoit de la nature une sensation de hauteur si ces sons possèdent certaines régularités spectrales et/ou temporelles. La hauteur perçue bénéficie d'une définition explicite : l'American National Standards Institute présente la hauteur d'un son pur ou complexe comme "that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high". Cette définition basée sur les concepts de "grave" et "aigu", qui restent eux-mêmes à définir, a été jugée insuffisante par la plupart des chercheurs. De plus on distingue plusieurs hauteurs possibles. Il peut s'agir de la hauteur tonale liée à la fondamentale où encore de la hauteur spectrale liée à la répartition énergétique du spectre et ceci à l'intérieur d'un même signal.

2.1.2 Les sons purs

Pour un son pur, c'est-à-dire une variation sinusoïdale de la pression acoustique, on associe la hauteur à la fréquence de l'onde. La correspondance physiologique entre le signal et la hauteur perçue est proposée par Helmholtz en 1846 [30]: il associe la membrane basilaire à un analyseur fréquentiel. L'émission d'un son pur se traduit au niveau de cette membrane par la propagation d'une onde dont le maximum de déplacement est atteint en un point de la membrane qui dépend de la fréquence du son. Ce point est proche de sa base, lorsque la fréquence est élevée, et proche de son apex lorsque la fréquence est basse. On distingue le pic d'excitation et la largeur du domaine de la membrane excitée par le signal qui représente la bande critique associée. La cochlée peut donc être assimilée à un banc de filtres passe-bande, modélisation qui explique bien d'autres phénomènes tels que certains aspects spectraux du masquage. La validité de cette relation entre la fréquence d'un son pur et le maximum du lieu d'excitation de la membrane basilaire a depuis été montrée par von Békésy [4]. D'autres facteurs que la fréquence (niveau, enveloppe, phase) peuvent également avoir un effet sur la hauteur d'un son pur (Demany [15]).

2.1.3 Les sons complexes

Pour un son complexe harmonique la hauteur est associée à la fondamentale, c'est-à-dire la fréquence de répétition de l'onde ou le plus grand commun diviseur des fréquences des composantes. Mais la hauteur perçue peut être aussi assimilée à la fréquence correspondant au lieu d'énergie spectrale maximale. On peut ainsi distinguer, dans le cas où les deux hauteurs peuvent être perçues, une hauteur "tonale" liée à la fondamentale et une hauteur "brute" liée à l'énergie spectrale maximale. Ces deux hauteurs peuvent être variées, indépendamment comme l'ont observé Plomp et Steeneken [57] ainsi que Risset [63, 64]. Dans son expérience Risset a utilisé des sons dont les timbres étaient très différents de celui d'un son pur. Les sons étaient des sons harmoniques auxquels on avait imposé une enveloppe spectrale unimodale. Pour construire la série de sons, Risset faisait baisser la fondamentale vers les basses fréquences et simultanément monter vers les hautes fréquences le sommet de l'enveloppe spectrale. Les auditeurs ont rencontré de grosses difficultés pour ajuster le son inconnu à un son pur. En effet certains auditeurs, à qui l'on présente deux sons ayant la même fondamentale mais une enveloppe spectrale différente, sont gênés dans leurs jugements de hauteur par la différence d'enveloppe spectrale. Le son présentant une énergie spectrale plus élevée en hautes fréquences est jugé plus aigu par ces sujets. Chez d'autres auditeurs, comme les musiciens par exemple, les jugements de la hauteur perçue sont en général la hauteur tonale, donc un jugement de la fondamentale indépendamment de l'enveloppe spectrale. Ces résultats contradictoires entre les sujets entraînent une nouvelle difficulté, la prise en compte de la disparité des jugements entre les sujets. Toutefois ces méthodes de calculs ne permettent pas d'expliquer toutes les hauteurs perçues, notamment dans le cas de la fondamentale absente que nous allons maintenant aborder.

2.1.4 La controverse "temporel-fréquentiel"

Si le rôle de la membrane basilaire comme analyseur fréquentiel n'est pas contesté, l'association directe de la composante fondamentale et de la sensation de hauteur est très largement remise en question. Un phénomène bien connu dans l'histoire de la recherche sur la hauteur est celui de la fondamentale absente. Lorsque nous entendons un son harmonique dans lequel est absente la fondamentale, c'est-à-dire lorsque le signal possède un spectre dans lequel sont présents des multiples entiers de la fondamentale mais pas celle-ci, nous éprouvons tout de même une sensation de hauteur correspondant à la fondamentale. Ce phénomène montré par Seebeck [72] dès 1841 à l'aide d'une sirène de son invention, contredit à priori l'association directe du lieu d'excitation maximale de la membrane basilaire et de la hauteur. Seebeck formule alors l'hypothèse que l'extraction de la hauteur se fait tout naturellement par une mesure de la périodicité du signal. Cette hypothèse contredit le fait que le système auditif fonctionne uniquement comme un analyseur de Fourier, idée fortement défendue par Ohm et Helmholtz. Ces derniers proposent une autre hypothèse pour expliquer la présence d'une hauteur malgré l'absence de fondamentale. Ils suppose que le spectre "vu" de la cochlée est différent de celui du signal original arrivant dans l'oreille. Par non linéarité une combinaison de sons primaires renforcerait la présence de la fondamentale dans la cochlée. Schouten [69] en 1940 a réfuté cette hypothèse grâce à une série d'expériences dont la plus célèbre est celle-ci : Schouten a utilisé un son harmonique composé des raies 1000 Hz, 1200 Hz, 1400 Hz, etc. La hauteur de ce son est perçue alors à 200 Hz. Ensuite il a élevé chacune des composantes du son de 40 Hz, de sorte que les fréquences étaient 1040 Hz, 1240 Hz, 1440Hz, etc. La hauteur perçue selon Helmholtz aurait dû être inchangée, 200 Hz, or les auditeurs ont perçu une hauteur d'environ 205 Hz. Plus tard d'autres expériences, comme celle de Licklider en 1954 [46], ont confirmé ce fait. L'hypothèse d'un codage temporel, en plus d'un filtrage par banc de filtres, a ainsi été prise en compte dans les modèles de calcul de la hauteur. Cette hypothèse déjà évoquée par Wundt en 1880 suppose que la hauteur donne naissance à des impulsions nerveuses dans les cils qui favorisent la détermination de la hauteur par une mesure de la périodicité du signal. Ainsi la modélisation du système auditif ne pouvant se résumer à un analyseur fréquentiel, les modèles ont évolué vers une double interprétation du système auditif : d'une part, la cochlée se comporte comme un banc de filtres passe-bandes où chaque filtre est caractérisé par sa fréquence centrale et une largeur de bande qui dépend de celle-ci, d'autre part, le système auditif extrait la périodicité du signal par un processus de type autocorrélation. C'est une combinaison de ces deux modélisations qui permet de mieux analyser la hauteur perçue dans le cas des sons complexes.

2.1.5 Remarque sur les sons inharmoniques

Certains des sons utilisés dans nos expériences sont inharmoniques comme les sons de plaques. Il devient donc impossible de parler de fondamentale, on s'intéresse alors aux méthodes de recherche du lieu d'énergie fréquentielle maximale d'une part, et de mesure de périodicité dans la fonction d'autocorrélation, d'autre part. Les pics de la fonction d'autocorrélation correspondent aux périodicités du signal, le premier d'entre eux étant pris comme premier partiel. Là aussi, du fait de la présence de fréquences dispersées dans le spectre et non-harmoniques, il est possible que les sujets soient sensibles à plusieurs hauteurs. C'est une perception analytique de la hauteur que nous décrirons plus loin. La perception de sons complexes harmoniques peut elle-même être modifiée par des variations de l'enveloppe spectrale : c'est la sensation de hauteur brute [15]. Dans le cas des sons inharmoniques, une interprétation possible de la hauteur perçue est que la hauteur spectrale perçue correpond à un intervalle de temps entre deux pics d'amplitude du signal brut. Ce qui implique comme hypothèse que le système auditif mesure la périodicité du signal. Une autre interprétation, vraie dans certain cas, mais qui ne peut être généralisée, est fondée sur une représentation fréquentielle de l'échelle de hauteur, suppose que l'auditeur cherche à ajuster les fréquences qu'il perçoit sur une fréquence fondamentale dont les harmoniques seraient proches des fréquences pures perçues [15].

2.2 Le timbre

Nous parlions déjà du timbre d'un son avant même d'en avoir une définition ; cet usage répondait en effet à un certain nombre d'interrogations sur la perception et l'interprétation auditive de notre environnement. Aussi a-t-on choisi de présenter ici cet attribut auditif complexe sous forme de questions, sur sa définition et ses fonctions, mais aussi sur ses liens avec d'autres attributs auditifs et enfin le caractériser par les études précédemment menées.

2.2.1 La non définition du timbre

Quelle définition pour le timbre?

Dans le cadre des recherches sur le timbre de nombreux auteurs se sont appliqués à trouver une *définition* en identifiant dans un premier temps les fonctions essentielles de cet attribut auditif perçu par tous. Grey [28] prend l'exemple de l'enfant à qui l'on fait écouter une mélodie jouée sur un instrument pour lui apprendre le nom de cet instrument. L'enfant, lorsqu'il entendra par la suite le même instrument avec une autre mélodie, en extraira l'information sous-jacente (le timbre) et, si la comparaison est un succès, l'enfant reconnaîtra l'instrument. En revanche si le timbre lui est inconnu, soit il *enregistrera* le nouveau timbre, soit il cherchera à le comparer en associant plusieurs timbres. Timbre : "Caractéristique de la tonalité d'un instrument ou d'une voix humaine, autre que la hauteur ou la sonie, permettant son identification", cette définition donnée par Helmholtz [30] en 1877 est reprise depuis par l'ANSI [1] et de nombreux auteurs tels que Risset et Wessel [65], Houtsma [32] ou encore Plomp [56]. Elle reste la plus répandue dans la littérature concernée, mais c'est aussi une définition par la négative, le timbre n'est ni la sonie, ni la hauteur et rien de plus, constate Plomp. D'où les difficultés rencontrées pour améliorer la connaissance du timbre. Comment trouver une liste de mots capables de décrire toutes les sensations auditives pour tous les sujets ? Comment calculer les paramètres acoustiques pertinents pour les décrire ? Quelle interaction existe-t-il avec la sonie et la hauteur ?

Peut-on dissocier le timbre de la hauteur ou de la sonie ?

Le timbre n'est ni la sonie, ni la hauteur mais n'est pas pour autant indépendant de chacun d'eux. Plomp [56] décrit l'enveloppe spectrale comme un attribut essentiel du timbre et pose ainsi la question de l'intersection avec la hauteur. Le timbre d'un stimulus varie avec sa hauteur et son niveau sonore. Pour cela Plomp souligne que même un son pur possède son propre timbre différent de sa hauteur. Un sinus basse fréquence sonne extrêmement plat, terne, tandis qu'un sinus haute fréquence sonne pointu, tranchant. La même voyelle prononcée par un homme ou une femme sonne différemment ; la place des harmoniques, déterminée par la fréquence fondamentale, sera déterminante dans la perception du timbre avec des enveloppes spectrales équivalentes mais de hauteurs différentes. Krumhansl [37] fait la même remarque et se pose la question de savoir si nous pouvons admettre que les différences dans les distributions d'énergies spectrales sont complètement découplées des mécanismes de perception de la hauteur. Fletcher [22] pense aussi que le timbre dépend principalement du comportement des harmoniques mais note à son tour que de fortes variations dans la hauteur et la sonie des sons produiront de larges variations dans la perception du timbre. Bregman et al. [5] en 1990 reviennent à leur tour sur la définition de l'ANSI et de sa référence à la hauteur. En effet que penser du timbre d'une pelle qui racle un tas de gravas dont le son n'a pas de hauteur ? La présence de la hauteur dans la définition pose donc un sérieux problème, d'abord parce que celle-ci est mal définie pour un grand nombre de sons, ensuite parce que le timbre n'est pas indépendant de la hauteur quand elle existe. Notre étude qui met en jeu à la fois le timbre et la hauteur devra donc être prudente quant aux interactions possibles notamment dans l'analyse des réponses des sujets.

Comment décrire alors le timbre ?

La définition va être largement remise en cause. L'ANSI [1] la complète en 1960 par une note dans laquelle il est précisé que : "le timbre dépend d'abord du spectre du stimulus mais aussi de sa forme d'onde, de la pression et des caractéristiques temporelles du signal". La définition reste néanmoins floue et ne décrit toujours pas le timbre précisément, ni dans ses fonctions auditives, ni dans sa représentation analytique. Roederer en 1975 [66] décrit la perception du timbre comme la première étape du mécanisme de reconnaissance de la tonalité de la source. Il présente le timbre comme un élément acoustique par lequel nous pourrons reconnaître un instrument de musique dans une situation qui nous est inconnue. Le cerveau extrait des informations qu'il compare avec des références en mémoire. Dès lors, notre mémoire enregistre-t-elle un certain nombre de sons qu'elle compare systématiquement ou enregistre-t-elle des invariants sonores qui caractérisent directement la source, quel que soit le mode d'excitation ? Indépendamment du processus d'extraction des propriétés acoustiques du timbre, il manque encore un outil, s'il existe, capable de le mesurer.

Vers une approche multidimensionnelle

D'ores et déjà il apparaît difficile de pouvoir associer une seule valeur au timbre d'un son. Rasch et Plomp [61] parviennent à la conclusion que les sons ne peuvent pas être représentés sur une simple échelle au sens du timbre. Le timbre est donc un attribut multidimensionnel de la perception des sons. Schouten [70] fait ce même constat d'insuffisance et définit le timbre comme un terme vague qui permet de couvrir toutes les propriétés acoustiques invariantes permettant de reconnaître un violon dans la reconnaissance d'un instrument de musique. Il propose pour combler cette insuffisance une série de paramètres analytiques dont la combinaison pourrait expliquer le timbre : le rang tonal, l'enveloppe spectrale, l'enveloppe temporelle, le glissement de fréquence et l'attaque. Mais ces paramètres acoustiques doivent décrire les dimensions saillantes dont les noms restent à trouver. Bregman et al. [5] conseillent donc de définir ces dimensions de sorte que, d'une part les définitions soient dans une voie psychologique simple, d'autre part que ces attributs aient une définition physique claire. Ils concluent prudemment et proposent de ne pas trop parler de timbre tant que ses dimensions ne sont pas clarifiées. La nécessité de trouver une décomposition du timbre en plusieurs attributs est néanmoins admise. En effet l'identification directe de ces attributs à travers les auditeurs pose une première difficulté de sémantique : sommes-nous en effet capables de décrire notre univers sonore et surtout sommes-nous capables de tous le décrire de la même façon ? A cette délicate question les travaux de Faure et McAdams [21] ont partiellement répondu. Leurs expériences ont principalement consisté à demander aux sujets de décrire des différences entre timbres, d'évaluer les timbres sur des échelles sémantiques et de reconnaître des instruments à partir d'une description donnée. Leurs conclusions sont doubles : d'une part il semble très difficile de décrire avec des mots nos perceptions sonores et il n'apparaît pas de vocabulaire spécifique du timbre musical, d'autre part il semble très facile de reconnaître un son, dans un ensemble de douze sons, à partir d'une des descriptions produites par d'autres personnes. Plusieurs types de "portraits" verbaux ont été testés lors de cette reconnaissance. Les portraits décrivant un objet ou une action ayant pu produire le son semblent beaucoup plus faciles à reconnaître que les descriptions de qualité des sons entendus. Alors qu'il nous est difficile de parler de nos perceptions sonores, l'association qu'un autre auditeur a pu faire entre un son entendu et une description produite est facile à comprendre. La communication du contenu de nos perceptions sonores est donc possible même en l'absence d'un vocabulaire spécifique. De ce résultat très important on peut tirer la conséquence immédiate qu'il faut être prudent quant au vocabulaire que l'on soumet aux sujets puisqu'ils n'y associent pas le même sens. La détermination des différents attributs du timbre ne se fera donc pas par une description méthodique directe des sons par les sujets. Il faut donc trouver un moyen de les isoler par d'autres méthodes et à ce titre les expériences de dissemblance, pour lesquelles le sujet se contente d'estimer la différence entre deux sons, conviennent. Nous allons donc devoir étudier le timbre, de façon analytique. Il s'agit de l'approche multidimensionnelle dont la formulation mathématique sera présentée au paragraphe 4.1.

2.2.2 L'approche multidimensionnelle

L'approche multidimensionnelle admise, restait à caractériser ce qui dans les signaux correspond aux dimensions perçues. La difficulté qu'ont les auditeurs à juger sur un grand nombre d'attributs restreint l'espace perceptif des sujets à un espace à deux où trois dimensions, par exemple lors d'études générales telles que celles menées par Plomp [56]. S'inspirant des expériences de Plomp [56] et de Miller et Carterette [54], Grey [28], en 1977, propose aux sujets d'évaluer les différences perçues dans toutes les paires de sons possibles dans un groupe de seize sons. Les sons sont ceux de seize instruments de musique égalisés en sonie, en hauteur et en durée. Les trois dimensions rendues par l'analyse multidimensionnelle représentent l'enveloppe spectrale, le flux spectral et l'attaque (degré de présence des transitoires). Chacune de ces dimensions n'ayant pas le même rôle dans l'espace de timbre, l'attaque et le flux spectral caractérisent davantage la reconnaissance de l'instrument. Les expériences de Clark et al. [12] ont montré que supprimer l'attaque diminue significativement les performances de reconnaissance d'instruments de musique. L'enveloppe spectrale représentée par le centre de gravité spectral décrivant la brillance des sons est un indice qualitatif qui apparaît comme une contribution majeure de la perception du timbre. Il restait donc à voir dans quelle mesure ces dimensions étaient effectivement orthogonales et dans quelle mesure des inversions de coordonnées sur une seule dimension pouvaient affecter une des fonctions du timbre.

Aussi Grey et Gordon [29] ont repris l'expérience avec toujours seize sons d'instruments de musique mais huit d'entre eux ont été modifiés par paires, de sorte que l'enveloppe spectrale globale de l'un des signaux soit échangée avec celle de l'autre. L'une des observations les plus intéressantes de cette étude est que l'inversion dans les signaux proposés aux sujets se retrouve dans la configuration de l'espace perceptif donnée par l'analyse. Ceci confirme que les trois dimensions représentent bien les attributs avec lesquels les sujets ont jugé les stimuli et que ces attributs ont des correspondances directes et indépendantes avec les paramètres analytiques du signal. La similarité proprement dite des différents instruments s'est trouvée modifiée par ces inversions provoquant des confusions et des inversions dans la tâche de dissemblance. Ces éléments posent alors une double question : quels sont les rôles relatifs des attributs entre eux dans la fonction de reconnaissance et quel est le rôle de la mémoire qui ignore à priori les stimuli hybrides ? Wedin et Goude [81] remarquent la saillance de certains attributs tels que l'attaque ou les transitoires. Ainsi des sons avec une attaque similaire mais des spectres différents sonneront plus éloignés si on retire les transitoires alors que des sons avec des spectres identiques mais des attaques différentes sonneront plus proches si on retire les transitoires. Ils soulignent ainsi le rôle des transitoires dans les tâches de comparaison. Krumhansl, Wessel et Winsberg (cités dans [37]) remarquent que la rapidité de l'attaque et les fluctuations spectrales influencent les jugements de similarité. Iverson et Krumhansl [34] prolongent ces études dans leurs expériences sur les attributs dynamiques du timbre dans les sons musicaux. Ils utilisent une expérience de dissemblance sur trois groupes de sons d'instruments de musique (16 stimuli) : les sons entiers, les sons sans l'attaque et les sons avec seulement l'attaque. Ils constatent que premièrement l'absence d'attaque ou l'attaque seule affectent peu les jugements de similarités et deuxièmement qu'il paraît nécessaire de faire appel à plusieurs attributs acoustiques pour décrire une même dimension perceptive du timbre. Ils rappellent ainsi que l'étude analytique des signaux ignore à priori les paramètres acoustiques pertinents pour décrire les dimensions psychologiques du timbre.

L'approche multidimensionnelle, désormais admise, a été retenue dans la plupart des études menées depuis. Nous donnons ci-dessous un rappel des définitions des principaux paramètres analytiques utilisés dont ceux pertinents dans notre étude.

2.2.3 Les paramètres analytiques

Le centre de gravité spectral (CGS)

Le centre de gravité spectral est la moyenne des fréquences du spectre (f_i) pondérées par leurs amplitudes (A_i):

$$CGS = \frac{\sum_i A_i f_i}{\sum_i A_i} \tag{2.1}$$

Ce paramètre correspond à la mesure de la moyenne au cours de la durée d'un son des centres de gravité instantanés calculés à l'intérieur de fenêtres temporelles glissantes de 2 ms. Comme nous l'avons vu le CGS est défini comme un attribut majeur du timbre. On le retrouve dans de nombreuses études (McAdams et al. [51], Krumhansl [37], Krimphoff et al. [36], Grey et Gordon [29]). Grey et Gordon ont trouvé une forte corrélation entre une dimension qu'ils ont interprétée comme la brillance du son, et le centre de gravité spectral. Ce paramètre qui leur a permis d'expliquer 88 % de la variance des données perceptives sera essentiel dans notre étude. Nous verrons plus loin que c'est en effet un paramètre analytique puissant pour décrire l'attribut brillance du timbre.

La déviation spectrale

Un autre paramètre, lui aussi mis en évidence dans des études sur les intruments de musique [36], est la déviation spectrale. Elle représente la déviation des harmoniques par rapport à l'enveloppe spectrale globale. La déviation est calculée en prenant l'écart entre chaque niveau de fréquence et la moyenne des amplitudes pondérées (dBA) des trois partiels adjacents :

$$DEV = \sum_{k=1}^{0} \left(L(k) - \frac{L(k-1) + L(k) + L(k+1)}{3} \right) \tag{2.2}$$

où L est le niveau en dB et k le nombre d'harmoniques.

La variation spectrale

La variation spectrale est calculée à partir des corrélations entre deux spectres consécutifs. On calcule une série de spectres par fenêtres glissantes, puis on estime la moyenne des coefficients de corrélations [36]. La variation spectrale est ainsi estimée par :

$$VS = \frac{1}{M}\sum_{p=1}^{M}|r_{p,p-1}| \text{ avec } r_{p,p-1} = \frac{\sum_{k=1}^{N} A_k(t_{p-1}).A_k(t_p)}{\sqrt{\sum_{k=1}^{N} A_k^2(t_{p-1})}\sqrt{\sum_{k=1}^{N} A_k^2(t_p)}} \tag{2.3}$$

L'attaque

L'attaque peut se mesurer de plusieurs façons ; elle dépend bien sûr des stimuli que l'on veut mesurer. Dans le cas général des sons percussifs c'est le temps nécessaire à l'enveloppe temporelle pour atteindre son maximum à partir du seuil preceptif. Pour un son de barre frappée le temps d'attaque est de moins de 20 ms. L'attaque peut être très importante, notamment dans la dissemblance des sons d'instruments de musique. C'est ce que décrivent les travaux de Krumhansl [37] entre autres.

Le temps de décroissance

Comme c'est souvent le cas dans la représentation de petites vibrations naturelles, l'enveloppe temporelle des sons percussifs est souvent assimilée à une exponentielle décroissante. Aussi la décroissance de l'enveloppe temporelle peut être approchée par une enveloppe du type $A\ exp(-\alpha\ t)$. Il est alors pratique de caractériser son amortissement dans le temps à travers la constante α que nous appellerons **le facteur d'amortissement**. Plus celui-ci est élevé plus le son se dissipe rapidement et inversement.

La rugosité

Dans leur étude sur la tension musicale Pressnitzer et al. [58] utilisent la rugosité comme paramètre acoustique qu'ils définissent à partir de la proximité ou non de fréquences à l'intérieur d'une bande autour de la fréquence moyenne. En réalité la rugosité s'avère être un paramètre complexe de la perception et plusieurs modèles ont été définis pour essayer de le décrire. Nous donnons ici le modèle spectral. Pour chaque fréquence un coefficient g est calculé (Plomp et Levelt, 1965) tel que pour un éloignement de 120 % de la bande critique g vaut 0 et pour un éloignement de 25 %, g vaut 1. Les bandes critiques sont calculées suivant le modèle de Hutchinson et Knopoff [33]. Puis les coefficients g sont sommés sur les fréquences et normalisés par leurs amplitudes. L'indice D de rugosité est ensuite calculé comme suit :

$$D = \frac{\sum_{ij} A_i A_j g}{\sum_i A_i^2} \tag{2.4}$$

Toutefois ce modèle échoue dans les prédictions de rugosité d'intervalles de sons complexes ainsi que dans la description des effets de phase. Cet attribut n'intervenant pas dans notre étude on se reportera à la thèse de Pressnitzer pour une étude approfondie.

2.2.4 Conclusion

Ce chapitre nous a permis de décrire les différents processus mis en œuvre dans les étapes classiques de la psychoacoustique. Elles consistent principalement à déduire directement du signal acoustique et de son traitement les relations entre les attributs perceptifs et les stimuli proposés à l'expérience. Une partie de la compréhension du timbre consistant à établir des paramètres capables de décrire le timbre. Quelles que soient les techniques utilisées dans le traitement du signal, y compris les filtres simulant les transducteurs physiologiques, à aucun moment la source même qui a servi à créer le signal n'est prise en compte. De plus les signaux utilisés sont en général construits artificiellement sur des modèles simples et éloignés des signaux naturels. C'est ce que conteste la théorie écologique que nous allons maintenant présenter.

Chapitre 3

La théorie écologique

3.1 Présentation de la théorie

Une théorie importante dans la recherche expérimentale sur la perception des sources est la théorie écologique proposée par Gibson [24, 25], qui se développa au début dans la recherche sur la perception visuelle. Selon cette théorie, toutes les données nécessaires à la perception sont contenues dans notre environnement, il suffit de les extraire. Cette théorie de la perception, que l'on appelle aussi perception directe, analyse la perception comme un processus d'interaction entre l'animal et son environnement et non pas seulement comme un phénomène interne à l'animal. Cette théorie s'oppose à d'autres théories dites "indirectes" par lesquelles la perception implique l'intervention de phénomènes de mémoire, de représentations et de calculs effectués sur les représentations. Gibson prétend au contraire que les stimuli que nous recevons sont extraordinairement riches et fournissent une telle précision sur les spécificités de notre environnement, que nous n'avons qu'à lire ou détecter cette information et non pas à la transformer.

3.1.1 La perception indirecte

La perception indirecte est en général associée à la théorie de l'information appauvrie. Nous sommes confrontés chaque jour à un grand nombre d'informations sous forme de stimuli divers, sons, images, odeurs. Certaines parmi ces informations vont aboutir à une perception extrêmement riche par rapport à l'information reçue. Typiquement, c'est le cas de la photographie qui est une représentation de l'espace à deux dimensions et dont nous allons systématiquement reconstituer la dimension manquante. Ainsi s'ajoute à la réception de l'information un certain nombre de processus qui vont constituer le traitement de l'information. La simple observation de l'information reçue serait inadéquate pour décrire ce que nous en ressentons puisque s'ajoutent un grand nombre de traitements annexes inférant au cognitif de l'individu (Michaels et Carello [53]). De plus l'information, lorsqu'elle est échantillonnée et donc séparée des informations qui la précèdent ou la succèdent, oblige le receveur à déduire les informations manquantes. Dans ce sens nous pouvons citer l'exemple de la photographie d'une voiture qui ne donne pas d'information a priori sur le sens du mouvement. C'est une autre photographie prise un instant plus tôt ou plus tard qui permettra de déduire le sens de la voiture. Les processus physiologiques n'étant eux-mêmes pas continus, nous recevons les stimuli sous forme d'une suite d'échantillons, lesquels peuvent eux-mêmes être déformés par nos propres capteurs (déformations dues à la rétine, distorsion du son). C'est ensuite cette succession d'échantillons qui sera traitée et deviendra *in fine* un événement perçu (Michaels et Carello [53]). "Il doit exister un processus intégré qui transforme une succession d'instantanés éphémères et discontinus pris par la rétine en un monde stable et perceptible" résume Neisser en 1967 [55]. La question selon une telle théorie est alors claire : comment une information appauvrie, incomplète, peut-elle être arrangée, organisée, structurée de telle façon que nous recevions une perception acceptable de notre environnement ?

3.1.2 La perception directe

L'approche écologique rejette en particulier les études dites "de laboratoire" dans lesquelles les stimuli sont des reproductions artificielles de stimuli naturels, d'où le terme écologique. Le système de perception lui-même serait sensible aux aspects de l'environnement lesquels ont une importance biologique pour l'auditeur ou ont acquis une importance comportementale à travers l'expérience (McAdams et Bigand [49]). Ainsi l'affirmation selon laquelle la reconnaissance d'un objet dans l'environnement est perçue directement, sans traitement, semble éliminer toute la question consistant à savoir comment des organismes, possédant des systèmes auditifs stimulés par des vibrations sonores, peuvent prendre conscience de l'identité d'une source sonore ou comment ces sources acquièrent une identité et une signification pour ces auditeurs (McAdams et Bigand [49]). C'est donc à travers une description des structures-mêmes du monde physique que cette théorie propose de chercher, mettant de coté d'une certaine façon les adeptes du traitement de l'information. Pour cette raison Gibson [24] commençe l'élaboration de sa théorie par une reformulation de la notion même de stimulus. Celui-ci est en général décrit par des données physiques primaires telles que la fréquence et le niveau pour un son. Cette réduction du stimulus à des descripteurs physiques primaires ignore en réalité la véritable information contenue dans le stimulus. Voilà pourquoi Gibson insiste sur le mot "information" en tant que structure qui spécifie l'environnement pour un animal (Gibson [24]). La notion d'échantillonnage est elle-même rejetée, le stimulus est rattaché à un événement particulier. Il ne s'agit pas de décrire une suite d'événements mais de détecter l'information dans son ensemble, en particulier le temps qui lui est rattaché. Prenons l'exemple d'une pomme qui tombe d'un arbre, l'événement n'est pas placé dans une succession d'instants mais dans des événements emboîtés. Le temps de la chute de la pomme est emboîté dans la vie de l'arbre qui lui-même est emboîté dans la reproduction des espèces, etc (Michaels et Carello [53]). Il importe de savoir si on cherche à percevoir l'événement "pomme qui tombe" ou l'existence de l'arbre ou autres. L'événement lui-même contient toutes les informations nécessaires ; c'est ensuite que nous sélectionnons ce que nous cherchons à en extraire. A cela s'ajoute la notion primordiale de contexte. Une grenouille qui tombe du ciel dans une rue de Chicago est un événement très proche de la pomme qui tombe mais n'aura pas les mêmes conséquences perceptivement puisque le contexte est radicalement différent (McAdams [50]). Ce qui montre encore une fois que c'est l'ensemble de l'évènement qu'il faut considérer (présence d'un pommier dans un cas et passage d'une tornade sur le lac Michigan dans l'autre). La figure 3.1 représente de manière caricaturale les deux approches de la question.

Dans un cas (information indirecte), l'auditeur analyse une succession de stimuli qu'il traite et décode pour en extraire l'événement. Dans l'autre cas

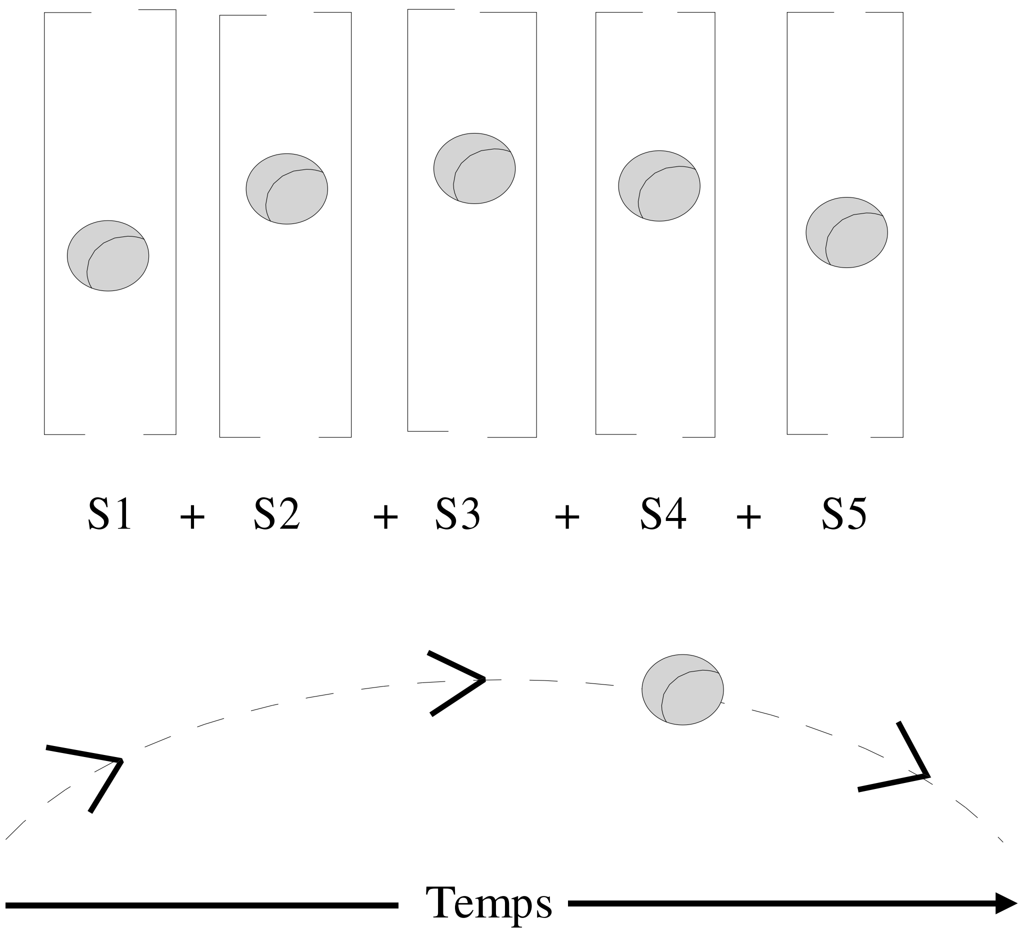

FIG. 3.1 – *Représentation traditionnelle et écologique des évènements et stimuli ; la partie supérieure représente la vue traditionnelle avec décomposition de l'évènement en une suite d'instants, chacun décrit par son propre stimulus S_i ; pour percevoir l'événement il faut réassembler les stimuli ensemble et reconstituer la dynamique. Dans la partie inférieure l'information recouvre le temps et ce dans la même durée que l'événement. La tâche de celui qui perçoit est de détecter l'événement spécifié par l'information (d'après Michaels et Carello [53]).*

(information directe) l'information n'est pas partitionnée, l'événement doit seulement être détecté, dans et à travers son contexte, et non pas analysé.

3.2 L'approche écologique et le timbre

Les études sur le timbre cherchent en général à déterminer quels sont les paramètres acoustiques calculés sur le signal pertinents pour l'analyse de la perception des stimuli. Ces études sont menées la plupart du temps à partir de sons artificiels offrant une meilleure lisibilité de ceux-ci et surtout un contrôle analytique plus facile. Seulement ces stimuli, quoi que de mieux en mieux simulés, le sont à partir de modèles spectraux ou mécaniques ne permettant pas de révéler toutes les composantes acoustiques du son. La solution pour mieux explorer cette question passe évidemment par l'affinement des modèles physiques afin de chercher des corrélats perceptifs dans les paramètres acoustiques et mécaniques simultanément. Parmi les études menées sur ce sujet on distinguera les études à partir de sons naturels (parfois restitués artificiellement) et celles menées à partir de sons de synthèse ; les conclusions du point de vue de cette théorie étant fortement dépendantes de l'origine des sons.

3.2.1 L'approche stimuli naturels

Les expériences de Lakatos et al. [42] cherchent à'évaluer les capacités de l'auditeur à discriminer la géométrie d'un objet mécanique simple, une barre, uniquement sur les bases des composantes acoustiques du son de l'objet. Les stimuli sont des barres frappées enregistrées, douze barres en acier et seize barres en bois. Les barres pour chaque matériau sont toutes de même longueur, seules les sections changent. Ainsi les barres en bois (sapin norvégien) sont de quatre largeurs et quatre hauteurs différentes soient seize sections différentes et les barres en acier laminé de trois largeurs et quatre hauteurs différentes soient douze sections différentes. Pour chaque matériau les stimuli sont présentés par paire et simultanément les géométries exactes des sections des deux barres sont affichées sur l'écran de l'ordinateur dans les deux ordres de présentation possible. L'auditeur doit alors choisir quel ordre correspond au stimulus qu'il vient d'entendre. L'auditeur a donc une chance sur deux de se tromper. Les performances des sujets sont ensuite transformées en dissemblances qui permettent de traiter les résultats par l'analyse multidimensionnelle. Les stimuli, sont souvent confondus, sont traités comme similaires et inversement. L'analyse multidimensionnelle des performances des sujets a révélé un espace perceptif à deux dimensions pour l'expérience sur les barres en acier et à une dimension pour l'expérience sur les barres en bois. La première dimension révéla de fortes corrélations avec le rapport hauteur/largeur et ce pour les deux matériaux. La deuxième dimension trouvée pour l'expérience avec les barres en métal est fortement corrélée avec le centre de gravité spectral. Les performances des sujets augmentent quand les ratios géométriques des deux barres présentées sont éloignés.

Pour chaque barre l'expérimentateur a également calculé les répartitions des différentes ondes présentes, (flexion et torsion) en fonction des caractéristiques physiques et géométriques de la barre. Il a aussi trouvé de fortes corrélations avec la deuxième dimension de l'espace perceptif. Selon Lakatos et al. [42] ce résultat suggère que "les auditeurs peuvent coder les propriétés auditives de sources de sons par extraction de certains invariants physiques caractéristiques de leur forme géométrique et de leur comportement acoustique." Ils proposent ainsi de chercher dans la matière, à travers ce qui peut influer sur les propriétés auditives des sons qui la caractérisent, des invariants que notre système auditif pourrait reconnaître. Dans cette expérience, les paramètres choisis pour décrire les dimensions perceptives données par l'analyse MDS sont directement liés à la source, ratios géométriques des sections et ratios fréquentiels des modes propres des barres. Cette expérience permet ainsi d'aborder le problème de la reconnaissance, et donc une fonction essentielle du timbre, en observant comment l'auditeur lui-même utilise les propriétés spécifiques des sources qui lui sont proposées. Dans la tâche ellemême, l'auditeur doit intégrer la géométrie des barres et ce qu'il en perçoit, rendant originale cette méthode expérimentale.

Sur ce type de processus Repp [62] s'est intéressé aux caractéristiques auditives d'un applaudissement. Vingt sujets ont été enregistrés en train d'applaudir normalement, comme ils le feraient à la fin d'un spectacle. La tâche pour ces mêmes sujets était ensuite de reconnaître les auteurs des applaudissements à l'écoute des 200 sons ainsi construits. Les sujets se connaissaient tous entre eux. Les performances des auditeurs ont montré de faibles capacités à reconnaître les sujets (10 % de succès pour les applaudissements des autres auditeurs et 50 % pour eux-mêmes), en revanche leurs réponses présentent d'autres caractéristiques intéressantes. Car si les distinctions de sexe n'ont pas été reconnues, en revanche le choix du sexe dans les réponses obéit aux mêmes caractéristiques pour tous les auditeurs. Le choix découle de stéréotypes : un applaudissement perçu comme lent et fort est associé à un homme, un applaudissement rapide et faible à une femme. Repp montre ainsi que quelle que soit l'information que l'on cherche à détecter dans la source, il faut se méfier de son corrélat perceptif qui lui-même obéit à d'autres règles psychologiques. Par ailleurs, les façons d'applaudir, c'est-à-dire les positions relatives des mains au moment du "clap", sont en revanche bien détectées (en particulier pour les applaudissements d'un seul sujet) montrant ainsi que l'information acoustique issue des mains et de leur position a été perçue. Son étude a montré, bien que ce ne soit pas le but recherché, que les femmes et les hommes applaudissent de la même façon, ce qui n'est pas l'idée qu'en ont les auditeurs dans leurs jugements.

Le plus intéressant dans les conclusions de Repp est sans doute qu'il n'exclut jamais l'analyse spectrale, au contraire. Les 200 sons ont été soumis à un programme de comparaison des spectres par inter-corrélations avec les spectres moyens pour chaque auditeur. L'ordinateur a su reconnaître à 90 % les auditeurs à partir des corrélations de spectre et ce malgré les différentes positions des mains lors des applaudissements qui modifient considérablement les spectres. L'auteur suggère que les auditeurs ont une "clap signature" et en déduit que son étude "conforte l'hypothèse que le son émanant d'une source naturelle transporte des informations perceptibles sur les états modifiés des mécanismes de la source" (traduction personnelle, pp 1100 de l'article [62]).

Ainsi les auditeurs peuvent reconnaître une signature parmi d'autres contenues dans les sources et ceci en répondant aux questions perceptives que peuvent se poser les auditeurs, parfois de façon sélective. Sur ce sujet Freed [23] propose aux auditeurs d'évaluer sur une échelle monodimensionnelle la dureté relative de maillets frappant un jeu de différentes casseroles en cuivre. Freed veut analyser les capacités de l'auditeur à faire une évaluation perceptive de l'excitateur indépendamment du résonateur. Les maillets présentent six duretés différentes (métal, bois, bois recouvert de toile, bois recouvert de feutre, caoutchouc, caoutchouc recouvert de feutre). Les performances des sujets augmentent avec la dureté des maillets mais surtout indépendamment du résonateur (casserole). Freed a calculé quatre paramètres acoustiques sur l'attaque des signaux (325 ms) qui se sont trouvés bien corrélés avec les réponses des sujets et qui lui ont permis d'évaluer une fonction de prédiction pertinente dans la perception de la dureté des maillets. Toutefois les spectres, et donc les paramètres acoustiques calculés, étaient fortement modifiés par les différentes casseroles sans pour autant que le sujet en soit gêné. L'auditeur extrait donc des invariants que nous ne pouvons voir dans le spectre. Il renforce ainsi l'idée que des indices invariants sont contenus dans la source en s'appuyant là encore sur des sons naturels. Ces invariants permettent à l'auditeur d'isoler dans son écoute le résonateur et le spectre (McAdams [49]).

3.2.2 L'approche stimuli artificiels

Un grand nombre d'études ont été menées pour tenter de caractériser l'influence de tel ou tel paramètre acoustique à partir de signaux synthétisés avec une somme de sinusoïdes (méthode additive). Toutes ces études cherchent en général à mesurer notre échelle perceptive par rapport à un paramètre acoustique donné et ne se soucient donc pas de l'identifier. Mais pour des tâches plus compliquées de notre système auditif, il faut reconnaître de nouveaux paramètres du signal capables de décrire les résultats des sujets (dimensions perceptives données par l'analyse multidimensionnelle). Lutfi et Oh [48] ont cherché des corrélats entre certaines caractéristiques physiques classiques et la reconnaissance des matériaux. Pour cela ils ont choisi un objet physique simple, une barre encastrée, dont on connaît les solutions des équations d'ondes associées. Le modèle calcule les composantes fréquentielles et temporelles. Les solutions de ces équations d'ondes dépendent de caractéristiques mécaniques simples telles que la densité ou le module d'Young. Lutfi et Oh font varier ces différents paramètres pour construire ensuite leurs sons par calcul additif. Les sons sont considérés comme des sommes de sinusoïdes exponentiellement amorties. Le but est de juger notre sensibilité aux variations des paramètres physiques introduits dans les équations d'ondes utilisées pour la synthèse. Les différents paramètres mécaniques (la densité ρ et le module d'Young) et acoustiques (fréquence, amplitude et constante temporelle), correspondant aux différents matériaux, varient autour de leurs valeurs initiales. Les sons sont présentés au sujet par paires. La tâche du sujet est de reconnaître quel est le stimulus cible (deux cibles étaient proposées, verre puis fer). Le paramètre acoustique le plus saillant pour la tâche des sujets est la fréquence. Lutfi et Oh en déduisent que la reconnaissance des matériaux est pauvre, en grande partie parce que nous ne sommes pas capables d'utiliser d'autres attributs auditifs que celui induit par la fréquence. Ces résultats sont peu différents de ceux des nombreuses études utilisant des sons non naturels qui dans le passé ont montré des limites dans l'aptitude de l'auditeur à traiter l'information en deux dimensions (ou plus) à la fois.

Le problème majeur de cette étude réside dans la modélisation, impliquant que les sons eux-mêmes sont éloignés de sons naturels. Les stimuli ne comportent donc pas toutes les composantes du signal. De plus, et c'est de là sans doute que viennent les fortes corrélations trouvées uniquement avec la fréquence, les sons n'ont que trois composantes fréquentielles (dont une présente pendant moins de 20 ms), permettant à l'auditeur de se concentrer sur la composante principale et de ne réaliser les tâches que sur cette seule composante. Le cas barre encastrée-libre n'est pas non plus le cas optimal pour évaluer le matériau. Malgré une approche originale dans laquelle les différents jeux de paramètres (acoustiques et mécaniques) ont été pris en compte, Lutfi et Oh s'éloignent d'une certaine façon de l'approche écologique du sujet en proposant des sons trop éloignés de la réalité. Il ne s'agit pas pour autant de rejeter cette méthode. Elle a ses avantages et elle permet en particulier de chercher dans le signal les éléments auxquels nous sommes sensibles et de les mesurer. C'est donc une approche complémentaire mais insuffisante pour estimer la pertinence de l'approche gibsonienne.

3.3 Conclusion

Les études que nous venons de présenter dans ce chapitre montrent que la théorie écologique peut effectivement être utile à la connaissance des processus auditifs courants puisqu'elle met en valeur des phénomènes tels que la reconnaissance d'un objet (existence d'un invariant) malgré des effets de contexte (changements d'excitateurs, taille de l'objet) difficilement maîtrisables au niveau du signal, lequel contient néanmoins toutes les informations décodées ensuite par le processus auditif. Sommes-nous capables de détecter ces invariants physiques sans les suivre à travers le processus physique de la propagation et le processus physiologique du traitement dans l'oreille interne ? La seule façon d'apporter une réponse est d'utiliser des stimuli construits sur un modèle physique suffisamment précis pour que les sons contiennent toutes les informations d'un son naturel. Dès lors que le sujet reconnaît les sons comme des sons naturels, il écoutera et accomplira la tâche expérimentale selon les mêmes processus qu'avec de vrais sons. Cette méthode permettrait donc d'isoler dans la physique de la source, que l'on peut désormais mieux contrôler, les paramètres jugés pertinents pour suivre tel ou tel attribut auditif utilisé par l'auditeur dans l'expérience. L'analyse spectrotemporelle permettrait de suivre parallèlement les modifications du signal, imposées par la physique de la source, et leurs rôles dans le décodage de l'auditeur. On pourra ainsi se conformer aux recommandations de Lutfi et Oh et à celles de Freed.

Maintenant que nous avons présenté, d'une part, les interrogations que posent les études sur le timbre et la hauteur en psychoacoustique et, d'autre part, les voies proposées par la théorie écologique, nous allons nous attacher à décrire les méthodes d'analyse de la psychoacoustique qui constituent les outils de base pour analyser les expériences menées sur ce sujet.

Chapitre 4

Analyses de données

4.1 Analyses multidimensionnelles

4.1.1 Introduction

Le problème de la représentation mentale d'objets sonores complexes comme le timbre est à l'origine de nombreuses études, notamment dans le choix des expériences et des analyses les plus pertinentes pour déterminer les structures sous-jacentes. L'analyse perceptive du timbre peut se faire par le jugement d'objets sonores en les comparant entre eux, les données rendues par les sujets étant des données de proximité. A partir de celles-ci, qui reflètent le degré de similarité des objets entre eux, deux méthodes principales décrivent la structure sous-jacente à la représentation mentale de l'ensemble des objets présentés. L'une part du postulat que la structure mentale serait continue et multidimensionnelle se basant ainsi sur une représentation géométrique de la structure mentale, tandis que l'autre s'appuie sur une représentation catégorielle des structures mentales. En réalité ces deux méthodes ne s'opposent pas, elles sont simplement plus ou moins adaptées selon les ensembles de stimuli et peuvent être complémentaires ou même combinées comme c'est le cas pour les modèles hybrides. Nous nous attacherons dans cette partie à décrire les modèles utilisés dans nos analyses. Pour une description exhaustive de ces méthodes de représentation mentale des objets sonores on se reportera à la thèse de Donnadieu [17].

Les modèles spatiaux

De nombreux auteurs (Torgerson [80], Shepard [73, 74], Carroll [7], Kruskal [38, 39], Young [88], Winsberg et Carroll [82]) se sont penchés sur le problème de la représentation spatiale des données de proximité. Ils postulent que les similarités entre les objets rendues par les sujets sont reliées à la distance psychologique entre les objets, représentée par la distance euclidienne ou semi-euclidienne dans un espace géométrique. Cette approche suggère que l'espace subjectif est continu et homogène et que ses axes représentent les différents attributs pertinents pour décrire le modèle.

Les modèles non spatiaux

Les modèles non spatiaux (Johnson [35], Carroll [7], Shepard et Arabie [76], Barthélemy et Guénoche [3]) décrivent les similarités entres les stimuli en terme d'"appariement" des traits distinctifs entre les objets ou par la somme de propriétés partagées par les stimuli. La similarité des objets est ainsi représentée par l'appartenance de ceux-ci à une même classe d'objets ou au contraire à deux classes d'objets distinctes, si les objets sont jugés dissemblants. Certains auteurs s'appuient sur la faiblesse des données expérimentales par rapport aux axiomes métriques constituant les modèles géométriques.

Les modèles hybrides

Toutefois les modèles hybrides basés sur les deux types de modèles donnent de très bons résultats ainsi que l'ont montré les études de Carroll [7] et McAdams et al. [51]. Les objets sont ainsi représentés dans un espace géométrique dans lequel on a pris en compte les propriétés discrètes spécifiques à chaque objet ou encore l'appartenance de groupes de stimuli ou de sujets à des classes distinctes. Ces modèles sont présentés à la fin du chapitre.

Parmi les différentes tâches possibles pour recueillir les données nécessaires, on s'intéressera essentiellement à deux méthodes expérimentales utilisées pour notre étude : les expériences de jugements de similarités offrant des données de proximité directes, et les jugements d'identification appartenant aux données de proximité dérivées.

4.1.2 Expériences de dissemblances

Les données de proximité recueillies peuvent être obtenues à partir de différents types de jugements. On distingue les jugements directs, les jugements ordonnés (Stevens [78]), les jugements de triade (Torgerson [79], Burton et Nerlove [6], Roskam [67]) et les jugements de tétrade (Torgerson [79]). On ne s'intéressera ici qu'aux seuls jugements directs utilisés dans nos expériences. Il s'agit de la "méthode de comparaison par paires" (David [14], Grenn et Rao [27]) : l'auditeur doit juger le degré de dissemblance de chaque paire d'un ensemble de stimuli. Le terme d'expérience de dissemblance est en général employé pour ce genre de protocole. Dans ce type d'études, l'auditeur n'a normalement pas d'autre tâche que celle d'évaluer la dissemblance entre les paires d'objets. En particulier aucune consigne ne lui est donnée de porter attention à tel ou tel attribut pouvant caractériser les stimuli, sauf par exemple dans les cas de comparaisons entre pays, qui peuvent être jugés selon un grand nombre de critères. Les données sont ensuite recueillies et rangées dans une matrice qui peut être triangulaire si l'on ne s'est intéressé qu'aux seules proximités δ_{ij} ou complète si l'on a recueilli également les proximités symétriques δ_{ji}. Dans ce cas, étant donné que dans les valeurs rendues par les sujets on aura en général δ_{ij} différent de δ_{ji}, ces différences seront traitées comme du bruit dans les données qui sera ignoré en moyennant par exemple les proximité δ_{ij} et δ_{ji}. C'est cette matrice dite de dissemblance qui servira ensuite d'entrée aux programmes d'analyse multidimensionnelle.

4.1.3 Tâche d'identification

Dans une tâche d'identification (Shepard [75], Wish et Carroll [85]), les stimuli sont présentés un par un à l'auditeur qui doit les identifier. Les proximités sont alors assimilées au nombre de fois qu'un sujet a identifié le stimulus i comme étant le stimulus j. La matrice de confusion est ensuite traitée comme une matrice de dissemblance : plus on confond les choses, plus elles se ressemblent.

4.1.4 Méthodes d'analyses multidimensionnelles

L'objectif

Ce type d'analyse a pour but de construire un espace géométrique dans lequel est représenté un ensemble de points décrivant les stimuli proposés pendant l'expérience. Les distances entre ces points assimilés aux objets sonores représentent les similarités perceptives rendues par les sujets lors de l'expérience. Ainsi plus les points représentant les stimuli seront éloignés dans l'espace, plus ils seront perçus dissemblables et inversement. Cette technique présente ainsi plusieurs avantages :

- la tâche ne présente pas de difficulté particulière pour les sujets,
- il n'est pas nécessaire de faire des hypothèses sur le nombre de dimensions ou sur le type de modèle adapté, ceux-ci étant déterminés par l'analyse,
- le nombre de dimensions, peu élevé en général, permet de visualiser facilement l'espace rendu par les sujets et donc facilite son interprétation subjective.

Les difficultés auxquelles on peut être confronté avec ce type d'analyse diminuent avec l'affinement des modèles. Ainsi nous noterons comme principales difficultés :

- la confiance dans les indices statistiques qui évoluent avec les modèles,
- la détermination des attributs perceptifs correspondant aux dimensions de l'espace rendu, la pertinence de ces interprétations s'améliorant aussi avec l'affinement des modèles,
- l'évaluation de la fiabilité du modèle par rapport aux bruits introduits par les sujets indépendamment et entre eux (cohérence inter-sujets). De nouveaux modèles prenant en compte des classes de sujets ont permis là aussi de limiter les difficultés (CLASCAL [84]).

Le postulat supplémentaire, complètant cette analyse, est que le nombre de dimensions rendu par les sujets correspond au nombre d'attributs psychologiques utilisés par ceux-ci pour construire leurs jugements. Les axes, quant à eux, sont à priori libres en rotation. Il existe toutefois des modèles avec pondération par sujet ou par classe de sujets, qui permettent d'orienter les axes selon la variation entre les sujets ou les classes (Wish et Carroll [85], Winsberg et de Soete [84]).

La démarche calculatoire

Pour une description plus complète des méthodes de calcul on se reportera à l'ouvrage de Kruskal et Wish [41] ainsi qu'aux articles de Young [86, 87], Young et al. [89], Winsberg et De Soete [84] et Winsberg et Carroll [83].

Mathématiquement le modèle n'est pas limité en dimensions. Toutefois l'espace perceptif rendu peut avoir un grand nombre de dimensions. D'une part, les sujets ne peuvent pas faire leurs jugements sur un trop grand nombre de critères à la fois. D'autre part, la précision du modèle diminue avec le nombre de paramètres à calculer donc avec le nombre de dimensions de l'espace rendu par l'analyse. Les différentes étapes qui constituent la méthode algorithmique sont expliquées ci-dessous et le schéma global de calcul est représenté figure 4.1.

Première étape

La première étape consiste à établir une configuration initiale des points à représenter, celle-ci pouvant être déterminée de façon aléatoire ou suivant une méthode métrique ou quasi-métrique. Dans le cas d'une configuration métrique, les dissemblances sont prises comme des distances euclidiennes pour ensuite être converties en produit scalaire. Pour cette conversion on s'assurera de trois conditions :

- la minimalité : $\hat{d}_{ij} \geq \hat{d}_{ii} = 0$,
- la symétrie : $\hat{d}_{ij} = \hat{d}_{ji}$,
- l'inégalité triangulaire est respectée ; pour cela on fait subir à chaque δ_{ijk} une transformation monotone : $\hat{d}_{ijk} = a + b\delta_{ijk}$ ou non linéaire $\hat{d}_{ijk} = a + b\delta_{ijk}^n$ telle que soit respectée l'inégalité triangulaire $\hat{d}_{ijk} \geq \hat{d}_{jlk} + \hat{d}_{ilk}$ où δ_{ijk} représente la dissemblance entre les stimuli i et j pour le sujet k et $\hat{d}_{ijk}$ la distance euclidienne.

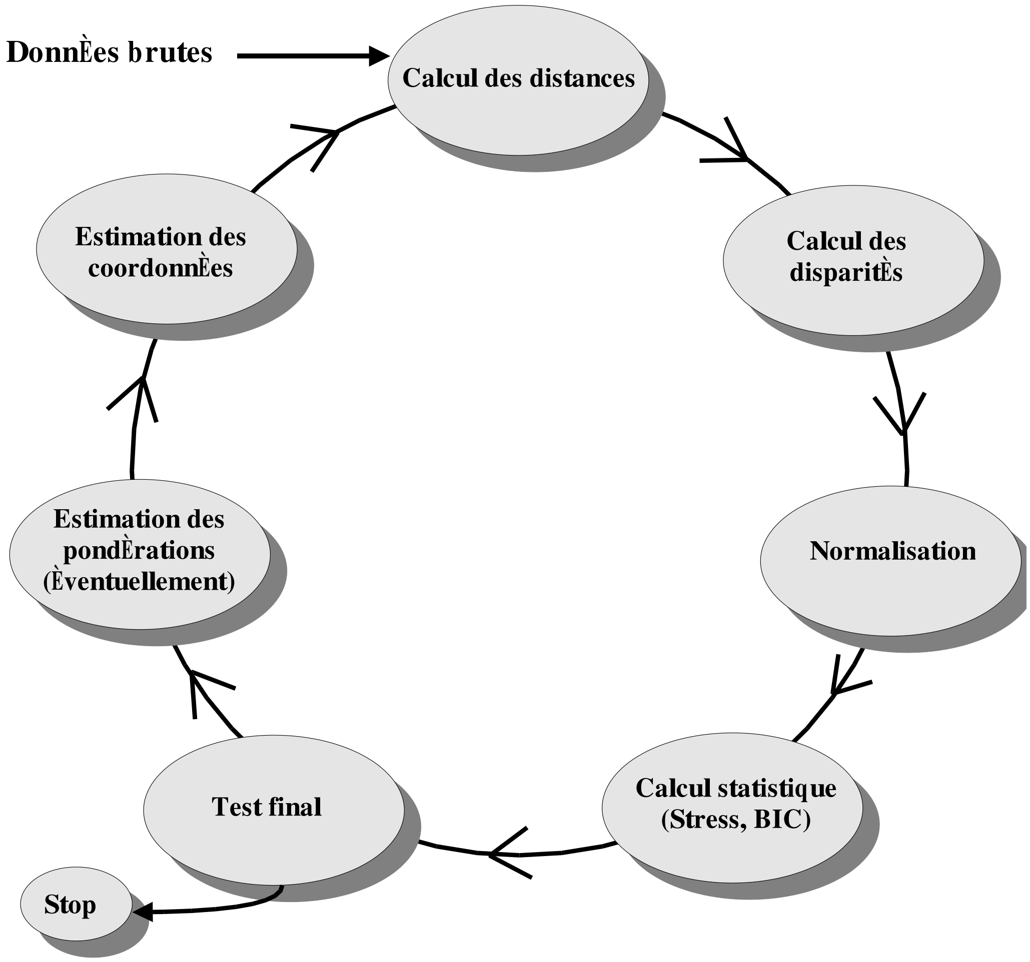

FIG. 4.1 – *Algorithme principal de calcul de l'analyse multidimensionnelle.*

Les dissemblances ainsi ajustées (disparités) sont converties en produits scalaires. L'estimation de la configuration initiale pour r dimensions est ensuite obtenue par calcul des valeurs propres. Cette configuration est ensuite normalisée pour que l'origine soit au centre de gravité de la configuration et que la dispersion soit constante.

Deuxième étape

Nous décrirons ici un modèle de distance euclidienne. Il en existe d'autres (Attneave [2]) inspirés du modèle plus général de Minkowsky dont le modèle euclidien est un cas particulier. Le calcul des distances constitue véritablement la première étape de l'algorithme. Cette étape consiste à calculer les distances à partir des coordonnées calculées à la boucle précédente ou pendant l'étape d'initialisation. Les distances sont obtenues à partir des coordonnées x_{ir} du stimulus i dans la dimension r avec la formule :

$$d_{ij} = \sum_{r=1}^{R}[x_{ir} - x_{jr}]^2 \tag{4.1}$$

Troisième étape

L'étape suivante est le calcul des dissemblances. Il s'agit de réajuster les disparités aux données de dissemblance. Pour cela la fonction de correspondance entre disparités et dissemblances va être optimisée. La méthode la plus courante est la transformation monotone de Kruskal [39] au sens des moindres carrés. Elle assure que, lors de la transformation, les distances soient dans le même ordre que les données, ce qui n'est en général pas le cas.

Quatrième étape

A cette étape le programme calcule un indice statistique, si l'estimation dépasse un seuil critique, le modèle est correct et le programme s'arrête, sinon le programme continue. Les calculs d'ajustement et d'erreurs dépendent essentiellement des fonctions de correspondance entre les données et les disparités. Ainsi les fonctions de Kruskal [39] n'ont aucune contrainte statistique sur l'erreur, elles doivent seulement ajuster les distances aux disparités par la méthode des moindres carrés. Pour ce type de fonction, le calcul d'erreur se fait à partir du stress décrit ci-dessous. Pour des fonctions telles que celles employées par Ramsay dans Multiscale II [60], il est possible d'utiliser des méthodes de calcul d'erreur plus fines tel que le maximum de vraisemblance. Parmi les indices utilisés à ce stade, voici les principaux :

- le stress : le stress est une mesure d'ajustement qui se calcule à partir des différences entre disparités (calculées par une fonction monotone sur les données) et les distances calculées dans le programme. Il existe différentes façons de calculer le stress :

$$S_1 = \left[\frac{\sum_i \sum_j (d_{ij} - \hat{d}ij)^2}{\sum_i \sum_j d_{ij}^2} \right]^{\frac{1}{2}} \quad et \quad S_2 = \left[\frac{\sum_i \sum_j (d_{ij} - \hat{d}ij)^2}{\sum_i \sum_j (d_{ij} - \bar{d})} \right]^{\frac{1}{2}} \tag{4.2}$$

  où $\bar{d}$ représente la moyenne des distances entre tous les points (Kruskal et Carroll, [40]).
- Le pourcentage de variance expliquée : la variance est calculée à partir du carré de la moyenne des corrélations sur l'ensemble des sujets, entre les produits scalaires calculés sur les données initiales et ceux estimés par le modèle. Cette mesure d'ajustement est exprimée en fonction du nombre de dimensions imposé à l'analyse, la courbe présentant un coude au niveau du nombre de dimensions le mieux approprié.
- Le maximum de vraisemblance : il est calculé sous sa forme logarithmique (Log V : Logvraissemblance) à partir des données, des disparités et de la fonction de correspondances. Ce type de calcul est défini par exemple dans le cas de fonctions de correspondance définies comme suit $\delta_{ij} = g(d_{ij}) + \text{rand}$ (Winsberg et Carrol [83]) où g est une fonction monotone et $rand$ une distribution normale d'erreur. Etant donné qu'il est basé sur un calcul de dérivé, il ne peut pas s'employer pour des fonctions comme celle de Kruskal où la dérivée peut être nulle. En complément, pour tenir compte du nombre de paramètres du modèle et de l'hypothèse de petites dimensions de l'espace, nous utiliserons plutôt l'indice BIC (Schwartz [71]) :

$$BIC = -2 \log L + 2_\sigma \log(N \frac{J(J-1)}{2}) \tag{4.3}$$

où N est le nombre de sujets et J le nombre de stimuli. Des études ont montré depuis la fiabilité de cet indice, excepté pour les modèles avec pondération par sujet. Les modèles présentant les plus faibles valeurs de BIC seront les modèles retenus comme présentant la meilleure configuration spatiale des stimuli.

Les différents modèles

Il existe plusieurs programmes d'analyse multidimensionnelle, tous itératifs. Le paragraphe qui suit présente les différents types de modèles. L'affinement progressif de ces modèles a été essentiellement effectué sur le calcul des distances, le calcul des fonctions de correspondances entre disparités et données ainsi que sur les calculs statistiques d'arrêts qui en découlent. Les deux derniers modèles (EXSCAL et CLASCAL) sont des modèles hybrides. Ils combinent à la fois des structures spatiales classiques et des structures non spatiales telles que des analyses en clusters (McQueen [52], Diday [16], Carroll et Arabie [8]) ou des analyses arborées (Barthélemy et Guénoche [3], Sattath et Tversky [68], Cunningham [13]).

– MDSCAL (MultiDimensional SCALing, Torgerson [79], Gower [26]) : dans ce modèle les distances sont euclidiennes, les fonctions de correspondance sont linéaires et monotones. L'espace rendu est invariant en rotation.

– MULTISCALE (Ramsay [59, 60]) : la particularité de ce modèle tient dans le calcul des distances réalisé suivant une fonction de puissance pouvant varier d'un sujet à l'autre. Cette fonction rend possible pour ce modèle l'utilisation de la méthode du maximum de vraisemblance.

– INDSCAL (INdividual Differences SCALing, Carroll et Chang [9]) : dans ce modèle les distances sont calculées par la méthode de pondération des sujets. Elle consiste à calculer les distances suivant la formule :

$$d_{ijk} = \left[\sum_{r=1}^{R} w_{kr}[x_{ir} - x_{jr}]^2 \right]^{\frac{1}{2}} \tag{4.4}$$

où w_{kr} est la pondération du sujet k sur la dimension r (Wish et Carroll [85]). Ce modèle a la particularité de bloquer en rotation l'espace rendu par le programme. L'indice BIC, défini au paragraphe précédent, s'est avéré peu pertinent pour ce type de modèle toutefois cette propriété de bloquer les axes s'avère très utile pour déterminer l'orientation des axes d'un autre modèle. Un tel modèle a permis d'identifier les dimensions privilégiées des auditeurs dans leurs jugements.

- EXSCAL (EXtended euclidian multidimensional SCALing, Winsberg et Caroll [83]) : une autre méthode consiste à calculer un modèle euclidien étendu développé par Winsberg et Carroll [83].

$$d_{ij} = \left[\sum_{r=1}^{R} (x_{ir} - x_{jr})^2 + s_i + s_j \right]^{\frac{1}{2}} \tag{4.5}$$

où x_{jr} est la coordonnée du $j^{ième}$ stimulus sur la $r^{ième}$ dimension ; d_{ij} est la distance entre le $i^{ième}$ et le $j^{ième}$ stimulus et s_i est le carré de la coordonnée du $i^{ième}$ stimulus sur la dimension spécifique de ce stimulus, appelé spécificité du stimulus i. Ce modèle est un modèle hybride qui permet de prendre en compte les spécificités particulières des stimuli telles qu'elles pourraient être exprimées dans les jugements de catégorisation. Ce modèle, postulant l'existence de dimensions propres à un stimulus permet d'expliquer la variabilité des jugements de proximité ne pouvant être représentés dans les dimensions communes.

- CLASCAL (approche de classes latentes, Winsberg et De Soete [84]) : ce modèle permet non plus de pondérer les coordonnées des dimensions pour chaque sujet mais de regrouper les sujets par classe. Il conserve ainsi les avantages des axes bloqués par les pondérations et réduit considérablement le nombre de paramètres à traiter. La distance séparant le stimulus i du stimulus j pour les sujets appartenant à la classe t s'exprime par :

$$d_{ijt} = \left[\sum_{r=1}^{R} w_{tr}[x_{ir} - x_{jr}]^2 \right]^{\frac{1}{2}} \tag{4.6}$$

On parle de classes latentes parce que les classes ne sont pas définies à l'avance. Une extension du modèle permet de prendre en compte les similarités ainsi que leurs pondérations par classes de sujets [51]. La distance entre le stimulus i et le stimulus j pour les sujets appartenant à la classe t s'exprime par :

$$d_{ijt} = \left[\sum_{r=1}^{R} w_{tr}(x_{ir} - x_{jr})^2 + v_t(s_i + s_j) \right]^{\frac{1}{2}} \tag{4.7}$$

où x_{ir} est la coordonnée du stimulus i dans la dimension r, w_{tr} la pondération sur la dimension r par les sujets de la classe t, s_i la spécificité du stimulus i et v_t représente l'importance perceptive globale accordée par les sujets de la classe t à toutes les spécificités. C'est ce modèle que nous utiliserons dans les analyses de nos expériences de dissemblance.

4.2 Bootstrap

4.2.1 Présentation

Le bootstrap est une méthode informatique de traitement statistique qui rend possible, par exemple, l'estimation de l'écart-type d'un paramètre statistique [20]. Lorsque l'on cherche à estimer un paramètre $s(x)$ sur un n-échantillon $x = (x_1,x_2,...,x_n)$ avec la réalisation d'une distribution inconnue F, on définit un ensemble de n-échantillon $x^* = (x_1^*,x_2^*,...,x_n^*)$ qui sont des rééchantillonnages de x où chaque observation x_i à une probabilité $1/n$. A un fichier bootstrap x^* on fait correspondre une application bootstrap de $s(x^*)$. L'estimateur bootstrap de l'écart-type ne nécessite aucun calcul théorique, il est valable quelle que soit la complexité mathématique de l'estimateur.

4.2.2 Application à l'analyse multidimensionnelle

Nous avons entrepris d'utiliser la technique du bootstrap sur la distribution des sujets. Ainsi, avant chaque traitement des données, le panel de sujets est retiré (tirage avec remise) en conservant le nombre total d'individus. Cette procédure pourrait donner par exemple trois fois le sujet 1, zéro fois le sujet 2, zero fois le sujet 3, deux fois le sujet 4, etc. A chaque retirage le choix des sujets est aléatoire et enregistré de sorte que l'on puisse s'assurer d'une distribution normale sur le total des tirages pour chaque sujet. L'application exploitée ici à chaque échantillon est l'analyseur EXSCAL [83]. Cet analyseur impose des contraintes pour chaque tirage nécessitant un réajustement des valeurs trouvées. Le réajustement à effectuer est double : il s'agit à la fois d'une rotation et d'une homothétie. La rotation est effectuée à l'aide de la rotation procustienne. L'homothétie est réalisée par normalisation. Le repère cible, vers lequel chaque rotation est effectuée, est l'espace obtenu par le modèle avec poids. Les axes ainsi obtenus sont les plus pertinents quant à leur interprétation en terme d'attributs perceptifs. Les nuages correspondant à chaque stimulus dans l'espace perceptif peuvent être représentés par des ellipses contenant 95 % des points. L'angle de l'axe principal des ellipses est calculé par régression linéaire sur l'ensemble des points. Les ellipses sont calculées ensuite sur la base de l'écart-type $se(s(x^*))$, et sont déterminées à 95 % dans l'intervalle $[\theta - 1.96.\sigma; \theta + 1.96.\sigma]$ où θ représente la moyenne. Ceci suppose une distribution normale des données, d'où l'intérêt de faire un grand nombre de tirages. A titre d'exemple la figure 4.2 montre les tirages effectués pour une expérience préliminaire avec vingt sujets et seize stimuli.

4.2.3 Interprétation

L'autre intérêt de cette méthode est descriptif pour comprendre certaines inversions dans les jugements par les sujets. En effet on peut constater que les stimuli dont les nuages se coupent correspondent aux stimuli qui ont été inversés dans le plan perceptif par rapport au plan physique. Ainsi il est possible de détecter quels sont les sons qui sont confondus perceptivement dans les jugements de dissemblance. Par exemple, si deux stimuli S_1 et S_2 se retrouvent dans un même nuage à la suite d'une analyse bootstrap, deux observations sont possibles :

- la discrimination entre ces deux sons n'a pas été réalisée par les sujets, ce qui implique que les jugements de chacun d'eux avec les autres

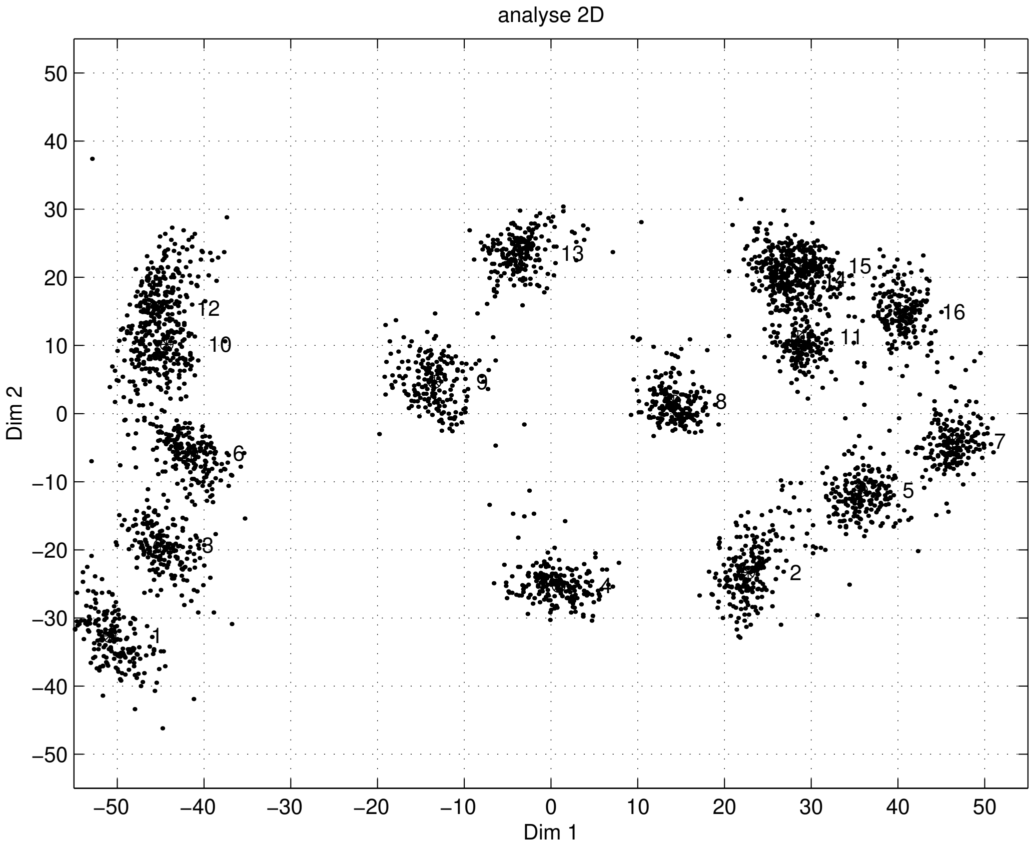

FIG. 4.2 – *Nuages de points pour 250 tirages.*

stimuli de l'espace ont été réalisés de la même façon,

– l'inversion de l'une ou plusieurs des coordonnées des deux stimuli dans l'espace perceptif peut ainsi être interprétée par la présence des sons dans le même nuage en comparaison avec d'autres paires de sons tout aussi proches mais dont les nuages n'ont pas d'intersections.

4.3 Analyse régressive

Pour chaque expérience les liens entre les différents paramètres seront faits à partir de l'analyse des corrélations entre eux. A partir de celles-ci nous représenterons les régressions simples ou multiples entre les différents paramètres à l'aide de *scattergram*. Ces figures représentent les courbes mettant en relation un paramètre dépendant que l'on cherche à décrire avec un (régression simple) ou plusieurs (régression multiple) paramètres indépendants. Le critère choisi pour estimer la qualité du modèle est le R^2 appelé le coefficient de détermination. C'est la proportion de variabilité de la variable dépendante expliquée par le (ou les) variable(s) indépendante(s) avec une valeur maximale de 1. Ainsi un R^2 de 0.80 signifie que 80 % de la variation de la variable dépendante est expliquée par le (ou les) variable(s) indépendante(s). Un R^2 proche de un 1 peut-être obtenu en rajoutant plusieurs variables indépendantes dans le modèle. Toutefois si le nombre de variables indépendantes dans le modèle est proche du nombre d'observations, il faut interpréter le R^2 avec beaucoup de précautions. Nous utilisons le R^2 ajusté qui tient compte du nombre de variables en ajustant une pénalité au R^2 selon le nombre de variables mises en jeu. Cependant nous ne nous servirons que d'une ou deux variables indépendantes dans notre étude, limitant ainsi les risques relatifs à l'interprétation.

4.4 Conclusion

Nous venons d'aborder les techniques d'analyse multidimensionnelle par lesquelles nous obtenons les représentations spatiales du timbre des stimuli. Nous avons vu qu'elles se sont largement développées et permettent de mieux comprendre les stratégies de réponses des sujets. Ces espaces seront ensuite comparés avec les paramètres analytiques décrits au paragraphe 2.2.3 caractérisant acoustiquement les sons. Nous souhaitons dans cette thèse explorer aussi la voie du stimuli naturel en contrôlant les sons depuis la source grâce à la synthèse par modèle physique ; c'est cette méthode que nous allons maintenant décrire.

Chapitre 5

Synthèse par modèle physique

5.1 Introduction

Nous avons utilisé deux programmes de synthèse sonore par modèle physique. Ces deux programmes ont été réalisés a l'ENST par V. Doutaut [18] et C. Lambourg [43] sous la direction d'A. Chaigne. Le premier programme est basé sur un modèle physique de barres et simule des sons de barres frappées avec la possibilité d'avoir des sons harmoniques (barres à section variable) pour simuler les instruments de percussion à clavier. Le deuxième programme, élaboré par la suite, simule des sons de collision plaque-sphère avec un modèle physique plus complexe. Les modèles physiques, la résolution numérique des équations ainsi que la validation de ces programmes ayant été réalisés suivant des méthodes similaires, nous présenterons les deux simultanément avec les spécificités de chacun. Il nous a paru important, étant donné qu'ils ont été les principaux outils du travail, de présenter d'une part les modèles physiques sur lesquels ils ont été construits, et d'autre part la validité des modèles numériques. Les équations données ici ainsi que l'ensemble des travaux sur ce sujet se trouvent dans les articles de Chaigne et Doutaut [10, 19] ainsi que dans les articles de Chaigne et Lambourg [11, 44].

Nous présenterons successivement les modèles physiques, la loi de Hertz qui décrit le contact entre l'excitateur et l'objet physique et enfin le rayonnement acoustique.

5.2 Les modèles physiques

Les notations utilisées pour décrire les paramètres physiques dans les équations sont reportées dans le tableau 5.1 pour le modèle de barres et dans le tableau 5.2 pour les plaques.

Dans cette partie, les équations de la mécanique de chaque systèmes, sont présentées, ainsi que leurs spécifications, notamment dans la prise en compte des phénomènes de pertes énergétiques primordiaux dans notre étude.

Les phénomènes liés aux pertes énergétiques sont présents dans tous les phénomènes physiques. En particulier n'importe quel système physique soumis à une excitation va dissiper au cours du temps cette énergie reçue sous forme d'échanges internes ou externes. Lorsqu'une barre ou une plaque est excitée par un maillet, une partie de cette énergie reçue est dissipée dans l'air sous forme d'une onde de pression, c'est-à-dire un son. Cette onde externe dont le rayonnement est provoqué par la structure vibrante fait partie des échanges externes. Les pertes énergétiques sont contenues aussi dans le son, et, par conséquent, sont perçues par l'auditeur sous forme d'un affaiblissement du niveau sonore et parfois de modifications du timbre du son. La synthèse sonore, quelle que soit sa méthode, a pour but de reproduire les sons le plus fidèlement possible, c'est-à-dire tels qu'ils seront perçus par les auditeurs sans distinction des vrais sons. Il est donc primordial dans la synthèse par modèle physique d'introduire les phénomènes de pertes énergétiques. Toutefois la justification physique ainsi que la formulation mathématique de ces phénomènes peuvent être parfois très complexes ; il faut donc utiliser des modèles généraux plus simples. La validité de ces modèles tient ensuite, d'une part aux mesures physiques expérimentales, d'autre part à la validation perceptive par l'auditeur. Les modèles présentés ici pour l'amortissement sont d'ordre 1 pour les barres et d'un ordre supérieur pour les plaques.

5.2.1 Le modèle de barre

Le système physique de la barre est composé d'une barre à section constante (ou non) et à bords libres, d'un maillet à tête sphérique et d'un résonateur. Le schéma du système complet est représenté figure 5.1.

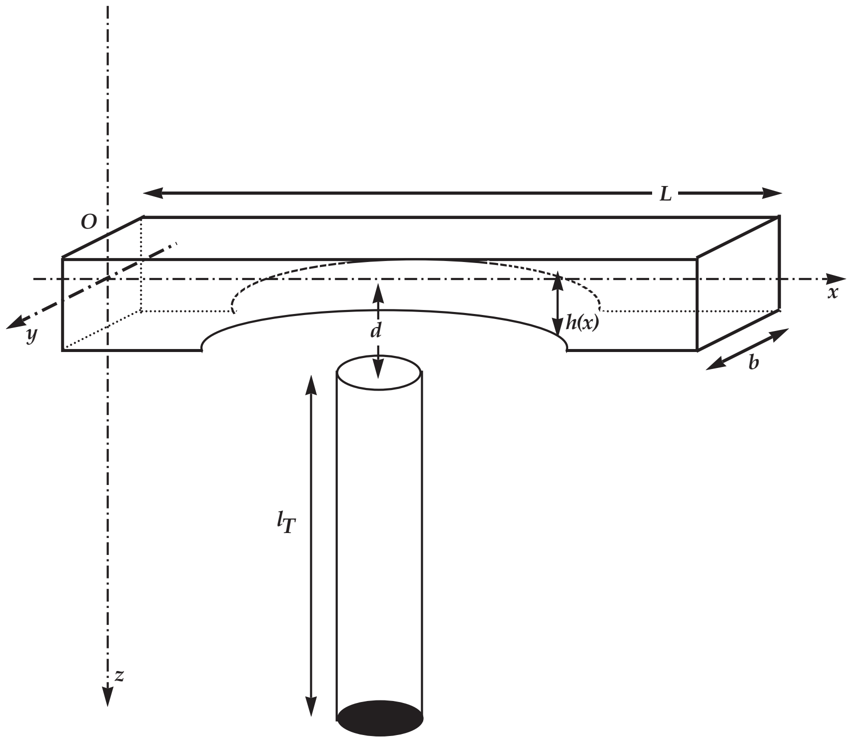

FIG. 5.1 – *Géométrie de la barre.*

Les barres simulées sont des barres sur supports élastiques. Les équations des barres mono-dimensionnelles d'Euler-Bernoulli sont utilisées pour décrire la propagation des ondes dans le matériau que l'on suppose homogène et isotrope. Ce modèle est affiné par la prise en compte de l'amortissement des ondes de flexion à travers un modèle de pertes viscoélastiques. Les déplacements transversaux (selon z) $W(x,t)$ à la surface de la barre sont régis par les équations suivantes :

$$\left\{ \begin{array}{l} M(x,t) = -E\left(1+\eta\dfrac{\partial}{\partial t}\right) I(x)\ \dfrac{\partial^2 W(x,t)}{\partial x^2} \\ \\ \dfrac{\partial^2 W}{\partial t^2}(x,t) = \dfrac{1}{\rho S(x)}\dfrac{\partial^2 M}{\partial x^2}(x,t) - \gamma_B\dfrac{\partial W}{\partial t}(x,t) - \dfrac{\chi}{M_B}W(x,t) + f(x,x_0,t) \end{array} \right. . \tag{5.1}$$

Symboles	Définitions (unités)
M	moment de flexion (N)
$W(x,t)$	Déplacements (m)
E	Module d'Young (MPa)
I	Moment d'inertie (x^4)
ρ	Densité de la barre (Kg/m^3)
c_0	Célérité de l'air (m/s)
S	Surface de la barre (m^2)
L	Longueur de la barre (m)
b	Largeur de la barre (m)
h	Hauteur de la barre (m)
M_B	Masse de la barre (Kg)
$f(x,x_0,t)$	Densité de force (m/s^2)
γ_B , η	Coefficients du modèle d'amortissement
α	Facteur d'amortissement ($/s$)

TAB. 5.1 – *Paramètres utilisés pour le modèle de barre.*

Afin de simuler des barres accordées, c'est-à-dire des barres où les modes excités sont des multiples entiers d'une fondamentale, le programme traite le cas des barres à section variable $S(x)$. Dans ces conditions :

$$S(x) = bh(x) \text{ et } I(x) = bh(x)^3/12. \tag{5.2}$$

Les conditions aux limites sont celles d'une barre à bords libres :

$$\frac{\partial^2 W(x,t)}{\partial x^2} = \frac{\partial^3 W(x,t)}{\partial x^3} = 0 \Big|_{x=0,L} . \tag{5.3}$$

Ceci revient à prendre comme hypothèse que le son d'une barre frappée n'est dû qu'aux seules ondes de flexion dans le plan (xz). On néglige ainsi les contributions des ondes de flexion et longitudinales, comme les ondes de flexion dans le plan (xy). Des mesures dans le passé ont montré que les hypothèses du modèle d'Euler-Bernoulli induisent une erreur maximale de 2 % dans les basses fréquences, c'est-à-dire pour des fondamentaux en dessous de 1200 Hz, ce qui est acceptable.

Le matériau est pris homogène et isotrope ce qui n'est qu'une approximation pour le bois qui est clairement anisotrope. Toutefois les barres de bois utilisées dans la fabrication des xylophones sont coupées selon la direction des fibres, ce qui rend l'équation 5.1 acceptable pour des ondes amorties où E correspond au module d'Young longitudinal.

L'amortissement dans les barres est représenté par deux termes. Le premier représentant les pertes viscoélastiques est proportionnel au coefficient η ; la relation contrainte-déformation devient dans ces conditions :

$$\sigma(x,t) = E\left(\epsilon(x,t) + \eta \frac{\partial \epsilon}{\partial t}(x,t) \right). \tag{5.4}$$

Le second terme d'amortissement de l'équation 5.1 représentant les pertes fluides est proportionnel au cofficient γ_B.

Ces deux coefficients, η et γ_B, sont estimés à partir de mesures du temps de décroissance des partiels. Les solutions de l'équation 5.1, décomposées sur la base des ondes planes, permettent d'écrire les solutions sous la forme d'un produit d'une fonction spatiale et d'une fonction temporelle. On définit pour caractériser l'amortissement le facteur d'amortissement α comme l'inverse du temps pour lequel l'amplitude d'une onde diminue de la quantité $1/e$. L'hypothèse d'ondes sinusoïdales amorties comme solutions de l' équation 5.1 impose alors la relation amortissement-fréquence suivante :

$$\alpha(f) = \frac{\gamma_B}{2} + 2\pi^2 \eta f^2. \tag{5.5}$$

Cette relation permet ensuite de calculer les coefficients du modèle d'amortissement en mesurant la valeur de α pour chaque partiel sur de vraies barres.

Enfin l'hypothèse de barres à bords libres ignore la présence de supports sous forme de fines cordes sur lesquelles les barres reposent. Les fréquences de résonnance induites par la présence de ces cordes sont très basses, autour de 20 Hz et très amorties, et donc ne perturbent pas les fréquences naturelles de la barre.

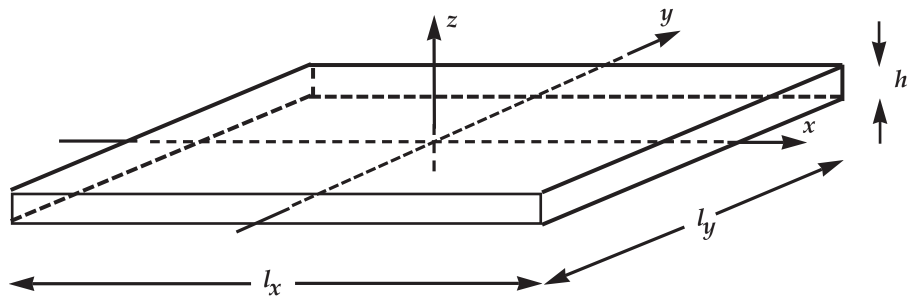

FIG. 5.2 – *Géométrie de la plaque.*

5.2.2 Le modèle de plaque

Les équations du modèle général

Le système physique du modèle de plaque est composé d'une plaque mince à bords libres et d'une sphère qui entre en collision avec la plaque. Le schéma du système complet est représenté figure 5.2.

Symboles	Définitions (unités)
M,M_y,M_z	Moment de flexion (N)
$W(x,t)$	Déplacement transversale de la plaque (m)
D_i	Constantes de rigidité (N/m^2)
ρ	Densité de la plaque (Kg/m^3)
ρ_a	Densité de l'air (Kg/m^3)
c_0	Célérité de l'air (m/s)
S	Surface de la plaque (m^2)
h	Epaisseur de la plaque (m)
f_z	Pression exercée sur la plaque (m/s^2)
R_{vn}, s_n	Paramètres du modèle de pertes viscoélastiques
R_1, c_1	Paramètres du modèle de pertes thermoélastiques
a_n, b_n	Paramètres du modèle de pertes par rayonnement acoustique
w_c	Pulsation de coupure (rad/s)

TAB. 5.2 – *Paramètres utilisés pour le modèle de plaque.*

Les hypothèses sont celles de la théorie des plaques minces. La relation contrainte-déformation pour les mouvements de flexion des plaques a d'abord été établie puis l'écriture de la relation fondamentale de la dynamique pour ce type de déformation a permis d'écrire les équations des déplacements transversaux à la surface d'une plaque. Le système suivant d'équations régit les vibrations de flexion dans le domaine de Laplace et pour une plaque fine isotrope infinie :

$$\left\{\begin{array}{l} \begin{pmatrix} \tilde{M}_x \\ \tilde{M}_y \\ \tilde{M}_{xy} \end{pmatrix} = -h^3 \begin{pmatrix} \tilde{D}_1(s) & \tilde{D}_{2(s)}/2 & 0 \\ \tilde{D}_2(s)/2 & \tilde{D}_1(s) & 0 \\ 0 & 0 & \tilde{D}_4(s) \end{pmatrix} \begin{pmatrix} \dfrac{\partial^2 \tilde{W}}{\partial^2 x} \\ \dfrac{\partial^2 \tilde{W}}{\partial^2 y} \\ \dfrac{\partial^2 \tilde{W}}{\partial x \partial y} \end{pmatrix} \\ \\ \rho h s^2 \tilde{W} = \dfrac{\partial^2 \tilde{M}_x}{\partial x^2} + \dfrac{\partial^2 \tilde{M}_y}{\partial y^2} + 2\dfrac{\partial^2 \tilde{M}_{xy}}{\partial x \partial y} + \tilde{f}_z(x,y,z) \\ \\ \text{avec } \tilde{D}_2 = 2\tilde{D}_1 - \tilde{D}_4 \end{array}\right. . \qquad (5.6)$$

Dans cette expression s est la variable de Laplace. Le terme $f_z(x,y,z)$ représente les forces extérieures au système, c'est-à-dire l'intervention de l'excitateur que nous expliciterons plus loin. Les termes en $\tilde{D}$ vont permettre d'introduire les pertes internes que nous allons maintenant décrire.

Modélisation des phénomènes dissipatifs

Dans ce paragraphe nous présentons les différents modèles de pertes pris en compte ainsi que les modèles numériques correspondants pour les phénomènes viscoélastiques, thermoélastiques et les pertes par rayonnement. Les pertes énergétiques sont intégrées en remplaçant les constantes de rigidité D_i des équations de propagations par des termes de rigidité complexe $\tilde{D}_i$ (équation 5.6) dans le domaine de Laplace :

$$\tilde{D}_i = D_i(1 + \tilde{d}_{ive}(s) + \tilde{d}_{ithe}(s) + \tilde{d}_{ray}(s)), \qquad (5.7)$$

Le terme $\tilde{d}_{ive}$ décrit les pertes viscoélastiques et dans quelle mesure le travail fourni au cours d'une déformation est stocké sous forme d'énergie reconvertible ou dissipée intrinsèquement. C'est la principale source de pertes dans le verre en dessous d'une fréquence critique f_c décrite plus loin. L'expression de la formulation différentielle de la relation contrainte-déformation viscoélastique s'écrit [11] :

$$\sum_{n=0}^{N} Q_n \frac{\partial^n \sigma_{ij}(t)}{\partial t^n} = \sum_{m=0}^{N} P_{ijklm} \frac{\partial^m \epsilon_{ij}(t)}{\partial t^m}, \tag{5.8}$$

dans laquelle les coefficients Q_n et P_{ijklm} sont appelés coefficients de viscoélasticité. On utilise ensuite pour l'expression du $\tilde{d}_{ive}$ une classe particulière de la représentation différentielle correspondant au modèle de Wiechert à deux cellules [11] (un ressort en parallèle avec deux modèles de Maxwell constitués chacun d'un ressort et d'un amortisseur en série). Pour ce type de modèle $\tilde{d}_{ive}$ s'écrit :

$$\tilde{d}_{ive}(s) = \sum_{m=1}^{2} \frac{sR_{im}}{s + s_m}, \tag{5.9}$$

où R_im et s_m sont des paramètres réels intrinsèques au matériau.

Le terme $\tilde{d}_{ithe}$ représente les pertes thermoélastiques dues à un couplage entre les ondes de flexion et les ondes thermiques dans l'épaisseur de la plaque. C'est la principale source de pertes pour l'aluminium en basses fréquences ($f < f_c$). La loi de comportement thermoélastique ainsi que l'équation de diffusion de la chaleur ont permis d'établir dans un modèle approché de plaques minces isotropes l'expression de $\tilde{d}_{ithe}$:

$$\begin{aligned} \tilde{d}_{ithe}(s) &= \frac{\varphi^2 \zeta_{th} s}{D_1(1+\tau_{th}s)} (i = 1{,}2) \text{ et } \tilde{d}_{4the}(s) = 0, \\ \text{avec} \quad \zeta_{th} &= \frac{8T_0h^2}{\pi^6 K_z}, \ \tau_{th} = \frac{\rho C_\epsilon h^2}{\pi^2 K_z} \text{ et } \varphi = \gamma(D_1 + \frac{D_2}{2}), \end{aligned} \tag{5.10}$$

et pour lesquels γ est le coefficient de dilatation thermique, T_0 la température absolue, K_z la conductivité thermique dans l'épaisseur et C_ϵ la chaleur spécifique à déformation constante.

On remarquera qu'en posant $R_1 = \dfrac{8\varphi^2 T_0}{\pi^4 D_1 \rho c}$ et $c_1 = \dfrac{\pi^2 K_z}{\rho c}$ l'expression de $\tilde{d}_{ithe}$ s'écrit alors :

$$\tilde{d}_{ithe}(s) = \frac{R_1 s}{s + c_1/h^2}. \tag{5.11}$$

On peut noter dans cette expression que les pertes thermiques diminuent avec h^2.

Le terme $\tilde{d}_{iray}$ représente les pertes dues à un couplage (mécanique ou acoustique) entre la plaque et les milieux extérieurs. L'hypothèse principale consiste à supposer que l'influence de la charge du fluide sur une plaque de dimensions finies peut être modélisée asymptotiquement pour les modes d'ordre élevé en effectuant un développement de Padé de la rigidité, à partir du cas d'une plaque infinie.

Le modèle utilisé pour la représentation de ces pertes est valable surtout au dessus d'une fréquence critique f_c qui s'écrit :

$$f_c = \frac{c^2}{2\pi}\sqrt{\frac{\rho}{h^2 D_1}}. \tag{5.12}$$

Dans la gamme des fréquences inférieures à f_c les autres pertes sont prépondérantes.

Avec le formalisme de la viscoélasticité le modèle de pertes par rayonnement acoustique pour les plaques isotropes s'écrit :

$$\tilde{d}_{iray}(s) = \frac{2\rho_a c}{w_c \rho h} \frac{\sum_{m=1}^{3} b_m \left(\frac{s}{w_c}\right)^m}{\sum_{n=0}^{3} a_n \left(\frac{s}{w_c}\right)^n} \quad \text{avec} \quad w_c = c^2\sqrt{\frac{\rho_a}{h^2 D_1}}. \tag{5.13}$$

On remarquera que l'épaisseur est le seul paramètre géométrique influant dans le calcul des pertes. Les pertes par rayonnement acoustique (par le calcul de la fréquence critique) sont inversement proportionnelles à h et la contribution des pertes thermoélastiques est inversement proportionnelle à h^2. Seule la contribution des pertes viscoélastiques est indépendante de h.

Cas particuliers du verre et de l'aluminium

Selon les matériaux, les différentes contributions des pertes dues aux échanges internes sont plus ou moins importantes. Nous nous sommes intéressés dans nos travaux à deux matériaux isotropes, le verre et l'aluminium, dont les principales pertes n'ont pas les mêmes origines. Nous présentons, d'une part le modèle de verre où les pertes viscoélastiques sont prépondérantes pour $f < f_c$ et, d'autre part le modèle d'aluminium où les principales pertes sont thermoélastiques. Les paramètres présents dans les modèles numériques ci-dessous ont été obtenus à partir de mesures expérimentales et permettent de chiffrer les contributions relatives de D_1 (compression-dilatation) et D_4 (torsion) dans les pertes. Ces paramètres caractéristiques des différentes contributions ont été calculés en utilisant un algorithme d'optimisation qui permet de calculer les coefficients des différents modèles numériques afin que les courbes amortissement-fréquences soient les plus proches possibles des mêmes courbes mesurées expérimentalement sur les plaques.

Dans le cas du verre, le modèle complet exploitable pour la synthèse est décrit par l'expression de $\tilde{D}_i(jw)$ qui comprend trois termes : le premier représente le terme constant de raideur, le deuxième l'amortissement viscoélastique et le troisième les pertes par rayonnement acoustique. L'expression complète de $\tilde{D}_i(jw)$ est donnée par :

$$\tilde{D}_i(jw) = D_i \left(1 + \sum_{n=1}^{2} \left(\frac{jwR_{vn}}{jw + s_n} \right) + \frac{2\rho_a c_0}{w_c \rho h} \frac{\sum\limits_{m=1}^{3} b_m \left(\frac{jw}{w_c} \right)^m}{\sum\limits_{n=1}^{3} a_n \left(\frac{jw}{w_c} \right)^n} \right). \tag{5.14}$$

Dans le cas de l'aluminium, le modèle complet exploitable pour la synthèse contient lui aussi trois termes. Le premier terme représente le terme constant de raideur, le deuxième terme l'amortissement thermoélastique et le troisième terme les pertes par rayonnement acoustique. Les expressions complètes de $\tilde{D}_1(jw)$ et de $\tilde{D}_4(jw)$ sont données par :

$$\tilde{D}_1(jw) = D_1 \left(1 + \frac{jwR_1}{jw + c_1/h^2} + \frac{2\rho_a c}{w_c \rho h} \frac{\sum_{m=1}^{3} b_m \left(\frac{jw}{w_c}\right)^m}{\sum_{n=1}^{3} a_n \left(\frac{jw}{w_c}\right)^n} \right) \tag{5.15}$$

et

$$\tilde{D}_4(jw) = D_4 \left(1 + \frac{2\rho_a c}{w_c \rho h} \frac{\sum_{m=1}^{3} b_m \left(\frac{jw}{w_c}\right)^m}{\sum_{n=1}^{3} a_n \left(\frac{jw}{w_c}\right)^n} \right). \tag{5.16}$$

5.2.3 Calcul de la pression rayonnée

Modèle pour les barres [10]

La discrétisation spatiale de la barre permet d'associer chacun de ses points à une sphère oscillante et ainsi de calculer par sommation le champ de pression total dû à la barre. La contribution $p_i(r_i,\theta,t)$ de chaque sphère oscillante à la position $x_i = i\Delta x$ au point d'observation $M(r,\theta)$ est donnée par la relation :

$$\begin{aligned} p_i(r_i,\theta,t) = &\frac{3p_0}{8\pi} bh\Delta x \cos(\theta_{x_i})[\frac{1}{r_{x_i}^2}\frac{\partial^2 W}{\partial t^2}(x_i,t - r_{x_i}/c_0) \\ &+ \frac{1}{r_{x_i} c_0}\frac{\partial^3 W}{\partial t^3}(x_i,t - r_{x_i}/c_0)]. \end{aligned} \tag{5.17}$$

Le principe de superposition permet ensuite de calculer la pression totale due à la barre p_B au point d'observation $M(r,\theta)$ en sommant les différentes contributions :

$$p_B(r,\theta,t) = \sum_{i=1}^{N} p_i(r_i,\theta,t). \tag{5.18}$$

Pour le tube, on utilise un modèle de tube non bafflé. Une extrémité est close (condition d'impédance infinie), l'autre ouverte (condition d'impédance de Levine et Schwinger [10]). L'extrémité ouverte du résonateur est assimilée à une sphère pulsante. La pression p_T du tube est donnée par l'équation :

$$p_T(r_T,t) = \frac{S_T}{4\pi r_T}\frac{\partial p_r}{\partial z}(d,t - r_T/c_0), \tag{5.19}$$

où S_T et r_t représentent la section et le rayon du tube respectivement.

Enfin, la pression totale au point d'observation $M(r,\theta)$ est donnée par sommation des deux contributions, la pression due à la barre et celle due au résonateur :

$$p(r,\theta,t) = p_B(r,\theta,t) + p_T(r,\theta,t). \tag{5.20}$$

Modèle pour les plaques

La méthode utilisée pour les plaques est basée sur la résolution de l'intégrale de Rayleigh. Cette méthode permet de calculer la pression en un point quelconque de l'espace en considérant uniquement le profil de vitesse normal à la source. La méthode numérique est de type BEM (éléments finis de frontières) basée sur l'intégrale de Rayleigh, formulation simple de l'équation fondamentale de l'acoustique dans le cas des structures planes infinies :

$$p(\vec{r},t) = -\frac{\rho_a}{2\pi}\int\int_S dS_0 \frac{1}{|\vec{r}-\vec{r}_0|}\frac{\partial v_n(\vec{r}_0,t - |\vec{r}-\vec{r}_0|/c)}{\partial t} \tag{5.21}$$

La pression est calculée simplement en intégrant numériquement l'accélération transversale de la plaque retardée et pondérée en fonction de la distance qui sépare le point source du point d'observation.

5.2.4 Loi de Hertz

Doutaut et Chaigne ont réalisé des simulations par différences finies de collisions barre-sphère qui ont conduit à de très bonnes simulations de sons d'instruments de percussions à clavier [18]. En particulier la méthode temporelle était tout à fait adaptée pour simuler ce type d'interaction, fondamentalement non-linéaire, et de durée très brève. On présentera ici la loi de Hertz utilisée pour cette étude ainsi que son introduction dans les deux modèles, barres et plaques.

Interaction barre-maillet

L'impact du maillet sur la barre se traduit dans l'expression du terme de densité de force $f(x,x_0,t)$ de l'équation 5.1. Le phénomène est non propagatif. La densité de force peut donc s'exprimer comme le produit de deux fonctions à variables séparées, en temps et en espace. La fonction $f(x,x_0,t)$ se décompose alors en deux termes : le premier est une densité temporelle de force et le deuxième correspond à une fenêtre spatiale d'excitation :

$$f(x,x_0,t) = \frac{F(t)}{\rho S}\, g(x - x_0), \tag{5.22}$$

où $F(t)$ représente la force d'impact et g la distribution de force sur le lieu de l'excitation $(2\delta x)$ dont l'intégrale vaut 1. $F(t)$ est donnée par la loi de Hertz :

$$F(t) = K|\zeta(t) - W(x_0,t)|^{3/2}, \tag{5.23}$$

où $\zeta(t)$ est le déplacement de la tête du maillet et K son coefficient de raideur.

Interaction plaque-sphère

Les contacts plaque-sphère sont décrits à travers les déplacements relatifs suivants : $W(x,y,t)$ représente le déplacement transversal d'un point de la surface de la plaque non déformée dans son épaisseur et $W_s(t)$ la position du point de la sphère non déformée située le plus près du plan neutre de la plaque. Ces déplacements sont les solutions du système d'équations suivant :

$$\left\{\begin{array}{lcl} \rho h \dfrac{\partial^2 W}{\partial t^2} & = & \dfrac{\partial^2 W}{\partial x^2} + \dfrac{\partial^2 W}{\partial y^2} + 2\dfrac{\partial^2 W}{\partial x \partial y} - g(x - x_0, y - y_0, t) F_h(t) \\ F(t) & = & \left\{\begin{array}{lll} F_h(t) = & [(W(x_0,y_0,t) - W_s(t))/k_h]^{3/2} & \text{si } W_s(t) < W(x_0,y_0,t) \\ & 0 & \text{sinon} \end{array}\right. \\ \dfrac{\partial^2 W_s(t)}{\partial t^2} & = & \dfrac{F_h(t)}{m_s} \end{array}\right. , \tag{5.24}$$

où m_s désigne la masse de la sphère, (x_0,y_0) le centre de l'impact et $g(x-x_0,y-y_0,t)$ une fonction de l'espace et du temps dont l'intégrale suivant x et y vaut 1.

5.2.5 Modèle numérique [19, 44]

La méthode des différences finies a été retenue pour ces deux programmes de synthèse parce qu'elle s'adapte particulièrement bien aux exigences du problème nécessitant une résolution temporelle sur des domaines de géométrie régulière type barres ou plaques. Le choix du modèle numérique a été déterminé en vue d'un compromis entre le temps de calcul et la précision de l'application numérique pour l'application audio.

5.3 Utilisation des programmes

Dans ce chapitre nous avons décrit la synthèse par modèle physique qui nous permet désormais de contrôler les paramètres géométriques et élastiques de la source. Nous disposons maintenant des outils pour explorer la démarche expérimentale qui a guidé nos expériences. Ceci afin d'essayer de répondre aux questions posées par les différentes méthodes d'analyse de notre perception. C'est cette démarche que nous allons maintenant présenter.

Chapitre 6

Approche expérimentale et cadre théorique

6.1 Cadre général

L'environnement sonore dans lequel nous menons cette étude est celui des barres, éventuellement accordées, et des plaques. Les deux attributs perceptifs, immédiatement perçus par les sujets à l'écoute de ces objets mécaniques, sont la hauteur des sons et leur brillance. Les sujets sont en effet sensibles en premier lieu à un certain nombre d'attributs perceptifs des sons qui leur permettent de les comparer entre eux et éventuellement de les reconnaître. Une façon d'appréhender notre sensibilité est de mesurer notre échelle perceptive sur ces différents attributs à travers l'expérience et ceci par le biais de l'analyse multidimensionnelle décrite dans la section 4.1 du chapitre 2. Nous désignerons par attribut perceptif toute dimension rendue par les sujets à travers l'expérience, distinguant ainsi la dimension perceptive du paramètre analytique choisi pour la caractériser objectivement.

Dans notre étude nous avons cherché à décrire quels sont les paramètres physiques (acoustiques ou mécaniques) auxquels nous sommes le plus sensibles lors d'une écoute méthodique de sons percussifs de barres ou de plaques. Les tâches pour les sujets se situent donc au niveau du processus d'extraction des propriétés auditives. Nous nous sommes posés la question suivante : décryptons-nous simplement et méthodiquement chacun des paramètres acoustiques reçus par l'oreille interne ou faisons-nous appel à des processus plus complexes, sensibles à d'autre paramètres, lesquels pourraient être contenus dans la source? Nous avons choisi pour cela un protocole expérimental qui permet de mettre en liaison les différents jeux de paramètres et d'en comparer les corrélations. La figure 6.1 décrit le protocole expérimental depuis la construction des sons jusqu'à l'analyse des attributs perceptifs rendus par l'analyse multidimensionnelle.

Se pose alors un problème plus large : si nous sommes directement sensibles à notre environnement, c'est-à-dire capables d'en extraire des informations propres à sa structure physique, nous devrions pouvoir le lire dans le signal acoustique lui-même, qui n'est que le résultat d'une excitation mécanique de la source. La pratique courante, consistant à chercher dans le signal des indices acoustiques capables de décrire le mieux possible les paramètres perceptifs, est à priori exhaustive puisque toute l'information sonore est forcément contenue dans le signal. Cette hypothèse exclut tous les phénomènes de contexte que l'on cherche en effet à éviter dans l'expérience mais surtout elle suppose que nous savons ce qu'il faut chercher dans le signal acoustique pour en retirer l'information pertinente. Or le plus souvent une même source, même très simple comme les barres ou les plaques, peut donner lieu à une grande quantité de signaux acoustiques très différents. Il peut donc s'avérer difficile de regrouper des informations sur la source de départ. L'auditeur lui-même, s'il est bien évidemment sensible aux paramètres analytiques proprement dits, n'en écoute pas moins en premier lieu la source lors de son interprétation. Il faudrait donc se placer de son point de vue d'abord, et partir de la perception de la source pour trouver ce qui est vraiment pertinent au sens de celle-ci dans le signal acoustique. Nous avons vu dans le chapitre 3 que deux théories apparemment opposées sur ce sujet sont en réalité très complémentaires, l'une rappelant "d'où vient le vent sonore", la source, l'autre rappelant nos moyens, l'analyse acoustique des signaux. C'est l'ensemble des études, du point de vue de la source et du signal, qui donnera une cohérence aux études sur notre perception de l'environnement.

Il faut, pour chercher dans la source elle-même, parfaitement maîtriser et

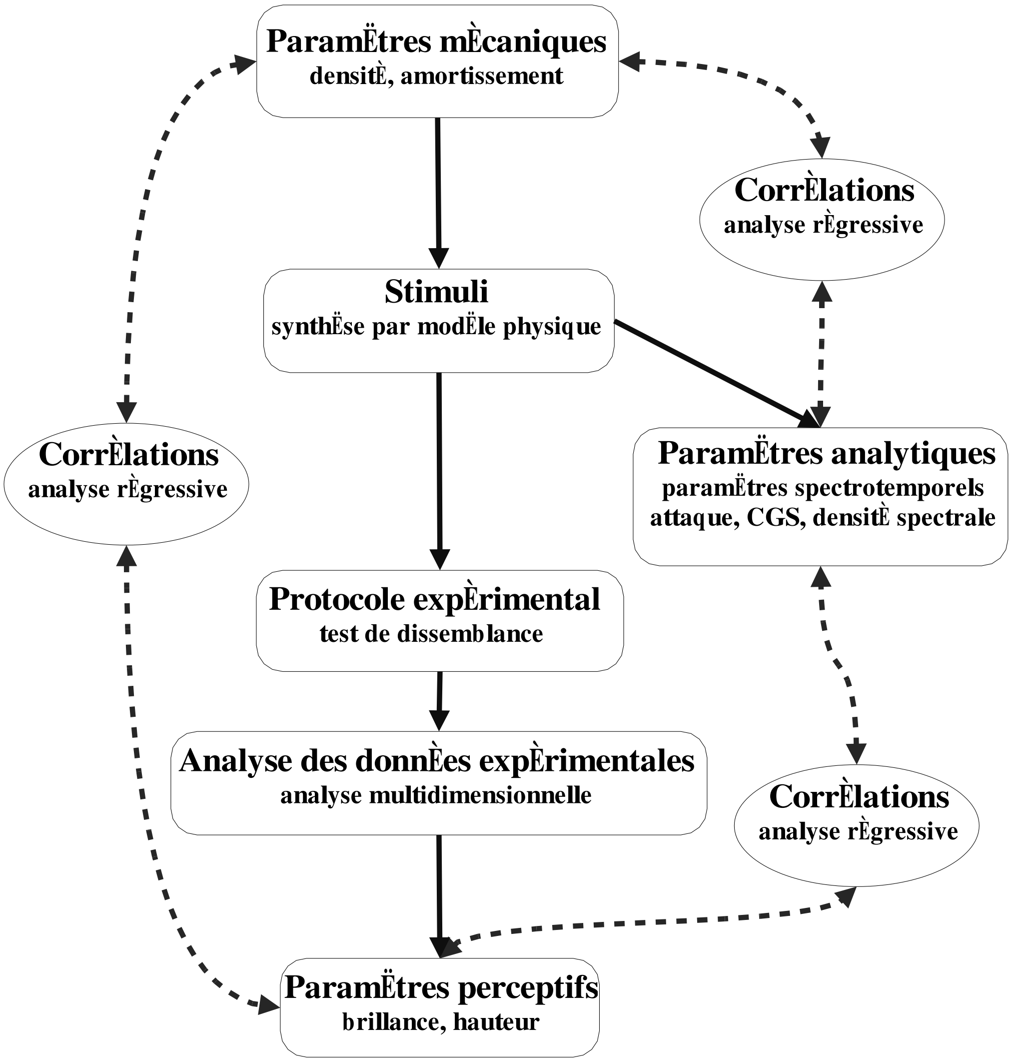

FIG. 6.1 – *Les sons sont construits avec les programmes de synthèse par modèle physique puis proposés aux sujets pour une expérience de dissemblance, les dimensions révélées par l'analyse multidimensionnelle sont ensuite corrélées avec les paramètres mécaniques d'une part et les paramètres analytiques des signaux d'autre part.*

même contrôler les différents paramètres des sources pour en mesurer l'influence. Ce qui est en général impossible : la forme et la structure même des matériaux étant en effet trop complexes pour être décrites, a fortiori simulées, par la physique. C'est le cas d'une table d'harmonie de guitare par exemple dont la forme et la composition fibreuse du matériau rendent l'étude mécanique difficile. Nous disposons, comme nous l'avons montré dans le chapitre 5, de puissants codes de synthèse capables de simuler les sources telles qu'elles produisent les sons. Ce seront les outils de base utilisés pour construire les stimuli des expériences.

6.2 Méthode générale

Nous allons décrire dans cette partie le protocole général de nos expériences et tous les points communs à chacune des trois expériences qui constituent notre étude.

6.2.1 Conception, expériences préliminaires

Pour chaque étude, nous expérimentons à travers une expérience pilote des séries de paramètres mécaniques pour lesquels nous déterminons sur quel domaine nous allons les faire varier afin d'avoir *in fine* des sons perceptivement distincts deux à deux. Tous les sons sont ensuite systématiquement égalisés en sonie après une expérience d'égalisation en niveau réalisée grâce au personnel bienveillant (et bien entendant) de l'IRCAM ou de l'ENST.

Stimuli

Notre étude a bénéficié de deux outils de synthèse puissants développés par Chaigne, Doutaut et Lambourg à l'ENST, permettant de simuler sur un modèle physique validé des sons de barres ou de plaques excités par un maillet.

6.2.2 Paramètres analytiques

Les paramètres analytiques choisis pour décrire les sons sont le spectre, le centre de gravité spectral et la constante d'amortissement décrite dans la section 2.2 du chapitre 2 que nous affinerons pour l'étude des plaques.

Matériel

Les expériences ont été réalisées conjointement à l'IRCAM et à l'ENST. Nous présentons ici le matériel utilisé dans l'un et l'autre cas. A l'IRCAM les sujets sont placés dans une cabine audiométrique à double paroi Soluna S1. Les sons sont générés par une station de travail NeXT équipée d'une carte ISPW et du logiciel Max [47]. Ils sont convertis ensuite par des convertisseurs ProPort avant d'être amplifiés par un amplificateur stéréo Canford et diffusés en écoute diotique (même signal aux deux oreilles), à travers un casque audio ouvert AKG-1000. Chacune des expériences est calibrée par un son pur de 1000 Hz à 70 dB. Le sujet est dans la cabine insonorisée et répond au test par l'intermédiaire du programme PsiExp [77] qui assure l'interface entre l'utilisateur et la station NeXT sur laquelle sont implantés les sons. Ce programme assure également la diffusion des stimuli et l'enregistrement des résultats. Dans le cas de l'ENST, les expériences ont été réalisées dans des conditions similaires : les sons sont diffusés par l'intermédiaire d'une station Sun. La diffusion des sons, l'interface de l'expérience et l'enregistrement des données sont réalisés par un programme conçu par nos soins à cet effet, à partir des modules graphiques de Matlab. Les résultats sont ensuite enregistrés et compilés en une seule matrice ordonnée, dite matrice de dissemblance pour chaque sujet, constituant la partie principale du fichier d'entrée du programme d'analyse multidimensionnel CLASCAL.

Procédure

Dans nos expériences, nous avons proposé deux tâches à nos sujets. Pour les deux expériences sur les barres et pour la première partie de l'expérience sur les plaques, la tâche proposée est un jugement de dissemblance. Quant à la deuxième partie de l'expérience sur les plaques, elle consiste en une tâche d'identification. Nous détaillerons pour chaque expérience, le nombre de sons, la tâche précise proposée au sujet ainsi que la durée de l'expérience.

Technique d'analyse avec l'analyseur CLASCAL

L'analyse des données expérimentales (données de proximité rendues par les sujets lors des tâches de dissemblance) se fait à partir de l'analyseur multidimensionnel CLASCAL présenté succinctement dans la section 4.1 du chapitre. L'analyse régressive est réalisée à partir du logiciel *Statview*. Les écarts-types sur les dimensions sont évalués à partir de l'analyse par bootstrap.

Dans cette partie nous présentons en détail l'analyseur CLASCAL qui a servi pour chacune des expériences de dissemblances. Nous commençerons par rappeler le modèle de distances utilisé dans CLASCAL puis nous énoncerons les différentes étapes de l'analyse pour choisir le meilleur modèle spatial.

Comme nous l'avons évoqué, plusieurs modèles de distance existent dans ce type d'analyse. Les distances associées aux similarités par des fonctions linéaires sont reliées aux coordonnées des stimuli par un modèle de distance. Nous utilisons le modèle euclidien étendu, développé par Winsberg et De Soete [84], avec adjonction de pondérations par classes latentes (McAdams et al. [51]). Dans ces conditions on rappelle que la distance entre le stimulus i et le stimulus j pour les sujets appartenant à la classe t s'exprime par :

$$d_{ijt} = \left[\sum_{r=1}^{R} w_{tr}(x_{ir} - x_{jr})^2 + v_t(s_i + s_j) \right]^{\frac{1}{2}}, \qquad (6.1)$$

.

où x_{ir} est la coordonnée du stimulus i dans la dimension r, w_{tr} la pondération sur la dimension r par les sujets de la classe t, s_i la spécificité du stimulus i et v_t représentant l'importance perceptive globale accordée par les sujets de la classe t à toutes les spécificités.

Le modèle hybride CLASCAL [84] tel que nous l'avons présenté dans la section 4.1.4 permet de regrouper les sujets par classes. Les classes sont dites latentes puisque non définies à l'avance. Le principe est que la fonction de distance entre deux stimuli est pondérée de façon identique pour un groupe de sujets appartenant à la même classe. Ce modèle a donc, comme le modèle de base avec pondération, l'avantage de bloquer les axes en rotation et permet de réduire considérablement l'ensemble des données à traiter comme dans le modèle traditionnel EXSCAL [9] avec pondération par sujet.

Nous ne connaissons pas à priori le nombre de classes latentes. La procédure habituelle pour choisir le nombre de classes consiste à tester un modèle spatial successivement sur T classes et $T+1$ classes et à retenir le meilleur. Dans le cas de CLASCAL, un modèle mixte, le ratio de logvraisemblances utilisé pour tester le modèle à T classes contre le modèle à $T+1$ classes n'est pas distribué asymptotiquement comme un $\chi 2$ avec le nombre de degrés de liberté connus [84]. Aussi les tests de ratio de vraisemblances ainsi que d'autres critères plus complets comme le BIC, qui renvoit à des conditions régulières, ne peuvent être utilisés pour le modèle avec classe. En conséquence nous utilisons une procédure de test Monte Carlo proposée par Hope [31]. La procédure de choix du modèle se déroule comme suit : soit $\hat{\theta}$ le maximum de vraisemblance estimé pour θ dans un modèle à T classes où θ est le vecteur des paramètres. A partir d'une population à T classes avec le paramètre θ, un nombre $(S-1)$ d'échantillons Monte Carlo de taille N sont tirés aléatoirement. Le modèle est estimé avec T et $T+1$ classes pour chacun des échantillons réalisés et nous comparons ensuite les modèles à T classes et $T+1$ classes avec les méthodes statistiques de vraisemblance. La solution à T classes est rejetée au niveau de confiance α en faveur du modèle à $T+1$ classes si la valeur du rapport de vraisemblance pour les données excède $S(1-\alpha)$ des valeurs statistiques obtenues avec les échantillons Monte Carlo.

Pour des modèles fortement non linéaires comme ceux de l'analyse multidimensionnelle, les indices statistiques tels que l'AIC où le BIC ne peuvent pas être utilisés correctement pour déterminer le modèle spatial le mieux adapté, c'est-à-dire les dimensions du modèle spatial ainsi que la présence ou l'absence de spécificités. Winsberg et De Soete (communication privée) recommandent la procédure suivante :

1. Choisir le nombre approprié de classes latentes en utilisant la procédure de Hope sur le modèle nul (modèle avec les dissemblances moyennes rendues par l'expérience comme données de base).
2. Choisir deux modèles spatiaux à comparer en sélectionnant les deux modèles aux plus faibles BIC [71]. Ensuite le programme détermine le

modèle spatial par la procédure de Hope suivante : un tirage de $(S - 1)$ échantillons Monte Carlo de taille N est effectué en ajustant sur le meilleur modèle. Nous comparons alors le rato de vraissemblance pour les deux modèles obtenus avec ceux obtenus sur les échantillons Monte Carlo.

3. Vérifier le nombre approprié de classes latentes en utilisant la procédure de Hope comme dans la première étape avec le modèle spatial sélectionné à la deuxième étape.
4. Si le nombre de classes latentes sélectionné à l'étape 3 est identique à celui de l'étape 1, l'analyse est terminée. Dans le cas contraire il faut répèter l'étape 2 avec le nombre de classes sélectionné à l'étape 3, et répéter ensuite l'étape 3 avec le nouveau modèle sélectionné.

Nous effectuerons ensuite un calcul d'erreur par la méthode du Bootstrap (voir chapitre 4.2) afin de déterminer une approximation de la variabilité des distances entre sons dans le modèle choisi. Pour les retirages nous imposerons un modèle à une classe à l'analyseur et nous estimerons ensuite sur l'ensemble des valeurs la variabilité des jugements des sujets entre eux ainsi que l'évaluation des écarts-types sur chaque dimension de l'espace perceptif.

Notations

Les notations utilisées dans la description des résultats de l'analyse MDS sont présentées dans le tableau 6.1 .

Notations	Correspondances
ns	sans spécificités
sp	avec spécificités
dl	degrés de liberté
N	nombre de dimensions
LLH	logvraisemblance

TAB. 6.1 – *Notations utilisées dans les différentes étapes de l'analyse CLASCAL.*

Un modèle spatial est représenté par $XDyy$ où X désigne le nombre de dimensions du modèle et yy vaut ns si c'est un modèle sans spécificité et sp

si c'est un modèle avec spécificités. Ainsi écrit, le modèle $2Dsp$ désigne un modèle à deux dimensions avec spécificités.

Présentation des expériences

Nous allons décrire dans les prochains chapitres les expériences réalisées successivement sur les barres et les plaques. Dans un premier temps, nous chercherons à caractériser perceptivement l'influence de l'amortissement à travers un modèle physique simple développé pour les barres, puis nous validerons les relations trouvées à travers une deuxième expérience en cherchant de plus à affiner les résultats sur l'influence de la hauteur dans ce type de jugements. Ensuite nous étudierons un modèle plus complexe d'amortissement développé pour la synthèse des plaques. Enfin, toujours pour les plaques, nous analyserons plus précisément la caractérisation des matériaux y compris au niveau de l'excitateur.

Chapitre 7

Expérience 1 : barres à section constante

7.1 Introduction

Dans cette première expérience sur les barres nous avons voulu observer les corrélations entre les différents jeux de paramètres, analytiques, mécaniques et psychologiques. Pour cela nous faisons varier deux paramètres physiques dans le programme de synthèse puis nous proposons les sons ainsi construits aux sujets dans le cadre d'une expérience de dissemblance. L'analyse des résultats de ces expériences nous permet ensuite de reconstruire un espace de timbre. Puis nous comparons les trois jeux de paramètres, les coordonnées issues de l'espace perceptif, les paramètres analytiques issus des analyses spectrotemporelles des sons et les paramètres physiques avec lesquels nous avons construit les sons. Le programme de synthèse de barres permet de faire varier indépendamment un grand nombre de paramètres mécaniques et par conséquent nous avons dû faire un choix. Une expérience pilote nous a montré que les paramètres élastiques et géométriques influent sur la hauteur des sons. Nous avons donc retenu un paramètre élastique. Les paramètres d'amortissement influent sur le timbre du son et la décroissance de l'enveloppe temporelle et nous avons également choisi l'un d'eux. Le détail de ces choix est présenté plus loin.

7.2 Méthode

7.2.1 Stimuli

Les stimuli sont décrits, d'une part à travers les paramètres mécaniques à partir desquels ils ont été calculés, et d'autre part à travers les paramètres analytiques (spectrotemporels) qui en sont issus.

Choix des paramètres mécaniques

Les paramètres géométriques et élastiques de la barre (E, I, ρ, h, S) influent sur les composantes spectrales du son, donc sur la hauteur de celui-ci. Nous faisons varier l'un d'entre eux (ρ) et maintenons les autres constants (tableau 7.1). La contribution des paramètres du maillet n'est pas étudiée ici et ceux-ci sont aussi maintenus constants (tableau 7.1).

Paramètres physiques inchangés			
barres			
module d'Young	E	9.29×10^9	N/m^2
moment d'inertie	I	2.6×10^{12}	m^4
amortissement fluide	γ_B	8.0	$/s$
hauteur	h	0.02	m
longueur	L	0.20	m
section	S	8×10^{-3}	m^2
maillet			
masse	m_m	0.012	kg
rayon	R_m	0.012	m
raideur	K	4.2×10^8	$Nm^{-3/2}$

TAB. 7.1 – *Tableau des paramètres constants dans le modèle avec symboles, valeurs et unités*

Les paramètres d'amortissement (γ_B, η) influent sur la distribution spectrale et l'enveloppe temporelle du son. Comme il est rappelé dans l'équation 7.1 le paramètre γ_B est un terme correctif sur les pertes dans les barres. Nous avons donc choisi de faire varier le paramètre η qui influe également directement sur la décroissance temporelle en fonction de la fréquence :

$$\alpha(f) = \frac{\gamma_B}{2} + 2\pi^2\eta f^2. \tag{7.1}$$

Les deux paramètres choisis pour construire l'ensemble des sons pour les tests perceptifs sont donc ρ et η. La densité est variée linéairement dans l'intervalle $[300,\ 900\ kg/m^3]$. Nous faisons ensuite varier l'amortissement dans l'intervalle $[1.0 \times 10^{-10}\ ,\ 2.0 \times 10^{-7} rad/s]$ ainsi que le montre la figure 7.1 afin d'obtenir un espacement relativement homogène des amortissements du point de vue perceptif. En effet, cette étude a fait l'objet d'une expérience pilote préliminaire qui a permis d'évaluer le seuil de discrimination entre deux sons afin qu'ils ne soient pas jugés identiques et de déterminer la gamme de variation adéquate pour construire les sons de l'expérience. La gamme de variation du paramètre η a été déterminée pour avoir le plus de sons possibles dissemblables dans le plan physique constitué par la variation des deux paramètres ρ et η. Ainsi pour les valeurs au-delà des valeurs limites de l'amortissement η, les sons sont inchangés du point de vue perceptif : les sons très amortis sont perçus comme des clics et les sons peu amortis sont perçus comme étant constants en niveau sur la durée que nous nous sommes fixée (1 seconde). Les valeurs de ρ ont été choisies pour qu'il y ait moins d'un quart de ton comme écart sur la hauteur et afin d'être au-dessus du seuil de discrimination même pour des sujets peu sensibles à la hauteur. Les valeurs exactes des seizes couples (ρ,η) sont reportées dans le tableau 7.4. De plus pour chercher ensuite les fonctions capables de lier le mieux possible les paramètres physiques variés avec les résultats de l'expérience, nous avons besoin d'un grand nombre de données d'entrée distinctes. La seule contrainte demandée pour le choix de ces valeurs étant d'obtenir une famille de sons dont toute paire est facilement discriminée. Il est alors possible de choisir une distribution aléatoire de ceux-ci dans l'espace physique repéré par les intervalles de valeurs de ρ et η (fig 7.1). Pour cela nous avons choisi une distribution aléatoire sur $16 * 16 = 256$ coordonnées possibles des sons dans l'espace physique densité-amortissement. La distribution (fig 7.1) permet d'avoir seize sons avec des valeurs différentes de ρ et η pour chacun d'eux et pour lesquels chaque paire est discriminable.

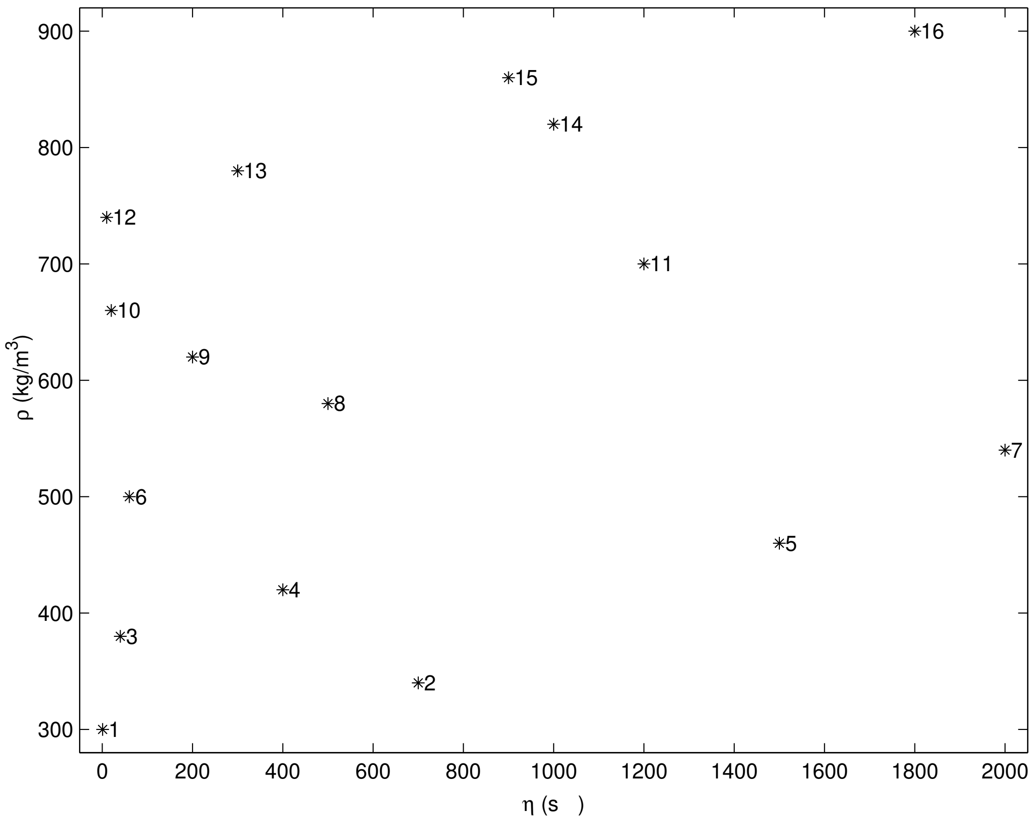

FIG. 7.1 – *Représentation des sons dans l'espace physique (η,ρ)*

Calcul des paramètres analytiques

Les paramètres analytiques qui décrivent les stimuli sont la fréquence fondamentale, le centre de gravité spectral et la décroissance temporelle. Nous avons répertorié l'ensemble de ces paramètres pour chaque son dans le tableau 7.4. Le spectre des sons est composé de quatre fréquences prédominantes qui correspondent aux modes propres de flexions classiques d'une barre (modèle à une dimension). Une fréquence supplémentaire apparaît sur le spectre à 440 Hz et est consécutive de la présence d'un résonateur dans le modèle pour les utilisations du logiciel dans le cas des barres accordées. Nous ne la développerons pas ici. Cette fréquence n'est pas prise en compte dans les analyses de résultats pour trois raisons. Tout d'abord le niveau de cette fréquence est de 65 dB en dessous de la fréquence la plus élevée : elle n'est donc probablement que très faiblement perçue. Et puis, la décroissance temporelle de cette fréquence calculée par la méthode Esprit [45] est de 24 s^{-1} ce qui correspond à une perte en gain de -3 dB en 27 ms, soit à peine plus que le temps d'attaque estimé à 20 ms. Enfin cette fréquence est commune à tous les sons et ne peut donc être prise en compte pour évaluer les dissemblances entre les sons qui est la tâche de l'expérience. La figure 7.2 donne à titre d'exemple le spectre du son 10.

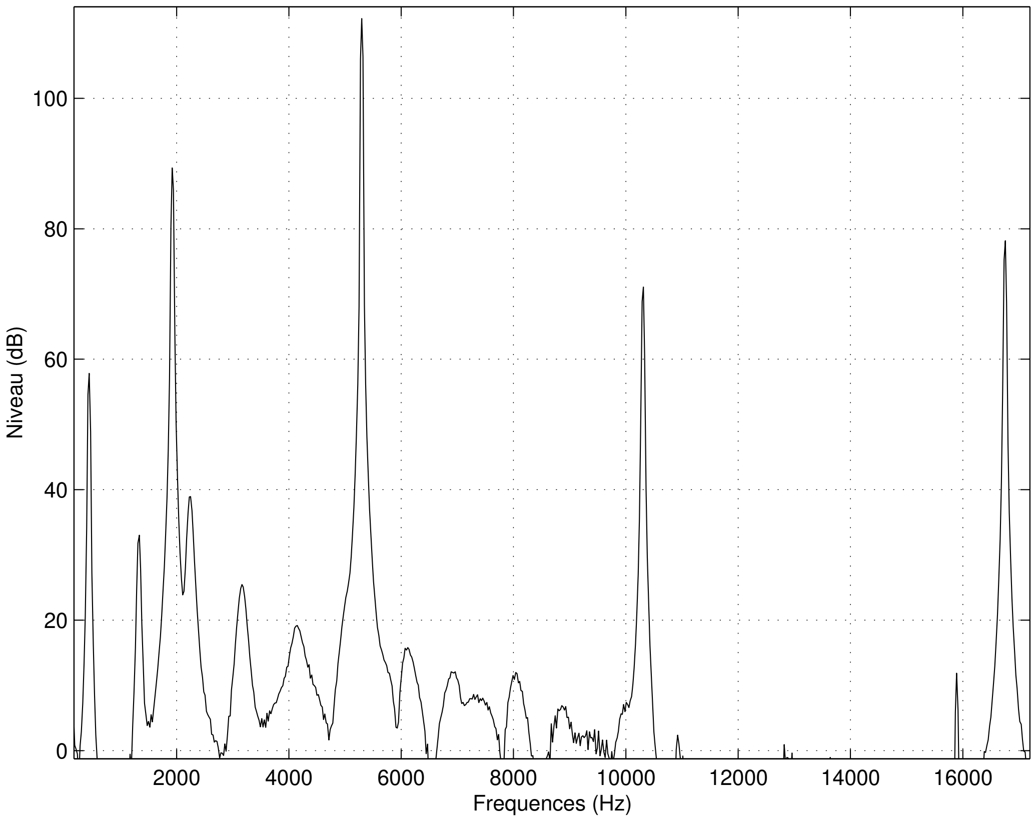

FIG. 7.2 – *Spectre du son* 10 *calculé sur* 85 *ms (2048 points) à partir de la fin de l'attaque (*23 *ms)*

La figure 7.3 représente l'enveloppe temporelle des deux stimuli situés aux extrémités sur l'axe des amortissements, c'est-à-dire le plus amorti et le moins amorti. Les sons ont été égalisés en sonie par une expérience préliminaire, de sorte que les niveaux *rms* des sons ne puissent pas influer sur les jugements des sujets. La durée des sons est fixée à une seconde, ce qui correspond à une perte de -28 dB pour le son ayant l'amortissement le plus faible ; ceci est suffisant pour évaluer l'amortissement perceptivement et répondre à l'expérience. Les sons sont échantillonnés à 48 kHz et sont enregistrés en mono 16 bits linéaire.

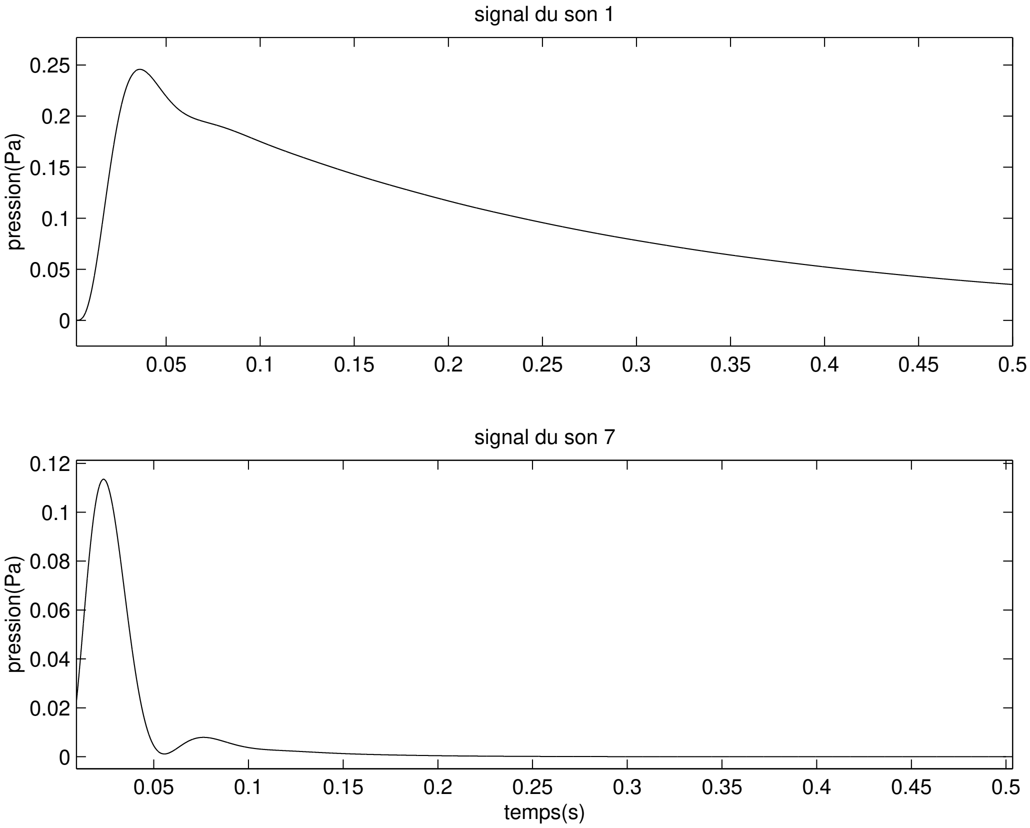

FIG. 7.3 – *Evolution temporelle pour les sons* 1 *et* 7 *représentant respectivement le plus faible et le plus fort amortissement.*

7.2.2 Sujets

Le groupe expérimental est constitué de dix-huit sujets, âgés de 21 à 40 ans (Moyenne = 26 ans) et provenant tous de la région parisienne. Le groupe ne comporte aucun musicien professionnel et est composé de huit femmes et dix hommes. Les sujets sont payés pour leur participation à l'expérience. Aucun n'a indiqué avoir un problème d'audition.

7.2.3 Procédure

La tâche est d'évaluer une distance perceptive (jugement de dissemblance) entre deux stimuli et ceci pour les 120 couples distincts possibles pris dans l'ensemble des 16 sons. A chaque essai, le sujet doit juger une paire d'échantillons choisie aléatoirement et sans répétition. Un silence d'une seconde sépare les deux sons. Le sujet doit alors déplacer un potentiomètre horizontal allant de "très similaire" à gauche à "très dissemblable" à droite. Il effectue le déplacement à l'aide d'une souris avec la possibilité de réécouter la paire de sons autant de fois qu'il le désire. Une fois son choix effectué, il valide sa réponse enregistrée sur un fichier sous la forme d'un nombre compris entre 0 et 1 (0 pour très similaires et 1 pour très dissemblables). L'expérience en moyenne dure un peu moins d'une heure par sujet.

7.3 Résultats

Nous détaillons maintenant les étapes du choix du modèle avec l'analyseur CLASCAL afin de déterminer le nombre de classes, la dimensionalité de l'espace perceptif et la présence de spécificités.

Choix du nombre de classes

Pour la première étape, le programme effectue le test Monte Carlo pour déterminer le nombre de classes sur le modèle nul. Il choisit deux classes.

Choix des modèles spatiaux

Nous testons maintenant les modèles de une à cinq dimensions, avec et sans spécificités, avec deux classes latentes.

Les deux meilleurs modèles spatiaux choisis selon l'indice BIC (donnés dans le tableau 7.2) nous donne 2Dsp puis 4Dns.

Test Monte Carlo sur le modèle spatial

Un test Monte Carlo opposant les modèles 2Dsp et 4Dns ne rejette pas le modèle 2Dsp. Il est donc retenu.

	N	LV	BIC	dl
Modèle sans spécificités	1	-9771	19682	18
	2	-9420	19103	34
	3	-9283	18952	50
	4	-9183	**18878**	66
	5	-9170	18974	82
	N	LV	BIC	dl
Modèle avec spécificités	1	-9469	19208	35
	2	-9232	**18859**	51
	3	-9204	18926	67
	4	-9204	19050	83
	5	-9179	19124	99
modèle nul	BIC = 19016			

TAB. 7.2 – *Test BIC pour le modèle à deux classes. N désigne le nombre de dimensions, LV le logvraissemblance et dl le nombre de degrés de liberté*

Vérification du nombre de classes

On recommence l'étape 1 avec cette fois le modèle spatial retenu à la place du modèle nul. Désormais le programme ne rejette pas le modèle à une classe. Il faut donc recommencer la deuxième étape avec le nouveau nombre de classes.

Calcul des nouveaux modèles spatiaux

Le programme recommence le choix du modèle spatial avec cette fois une classe de sujets.

Les modèles choisis sur les BIC donnés dans le tableau 7.3 sont maintenant les modèles 2Dns puis 3Dns, le test Monte Carlo ne rejette pas le modèle 2Dns. La nouvelle vérification du nombre de classes ne rejette pas le modèle à une classe. Le modèle définitif est donc **le modèle à deux dimensions sans spécificités et avec une seule classe latente**.

	N	LV	BIC	dl
Modèle sans spécificités	1	-9790	19704	16
	2	-9264	**18760**	31
	3	-9266	**18865**	46
	4	-9255	18935	61
	N	LV	BIC	dl
Modèle avec spécificités	1	-9305	18868	32
	2	-9292	18940	47
	3	-9266	18988	62
	4	-9257	19062	77

TAB. 7.3 – *Test BIC pour le modèle à une classe*

7.4 Discussion

7.4.1 Discussion sur les résultats bruts

Le meilleur modèle est donné pour un espace à deux dimensions sans spécificités et une classe. L'absence d'efficacité du modèle avec spécificités s'explique par la bonne homogénéité du groupe de sons. Ce modèle est invariant en rotation (voir paragraphe 4.1.4). Le modèle avec pondérations individuelles pour chaque sujet est donc calculé pour orienter et fixer les axes psychologiquement à travers l'ensemble des sujets. Ce modèle sert de cible pour réorienter les axes et ceci après chaque tirage de bootstrap, de sorte que chaque résultat issu d'un tirage des données est réorienté de la même façon. La rotation procustienne utilisée pour réorienter les axes est composée d'une rotation pour une orientation homogène des axes et d'une homothétie pour que les échelles soient homogènes entre les espaces. Les données rendues sont ensuite représentées dans un plan que nous appellerons l'espace perceptif, représentée fig 7.4. La détermination des écarts-types est réalisée par l'analyse bootstrap 4.2. Dans ce plan les axes représentent les dimensions perceptives sur lesquelles les sujets ont estimé les dissimilarités entre les sons. Deux points espacés sur la figure 7.4 représentent deux stimuli jugés dissimilaires par les sujets. A ce stade de l'analyse nous pouvons procéder à quatre observations :

- le nombre de dimensions de l'espace perceptif est le même que le nombre de paramètres variés pour construire les stimuli.

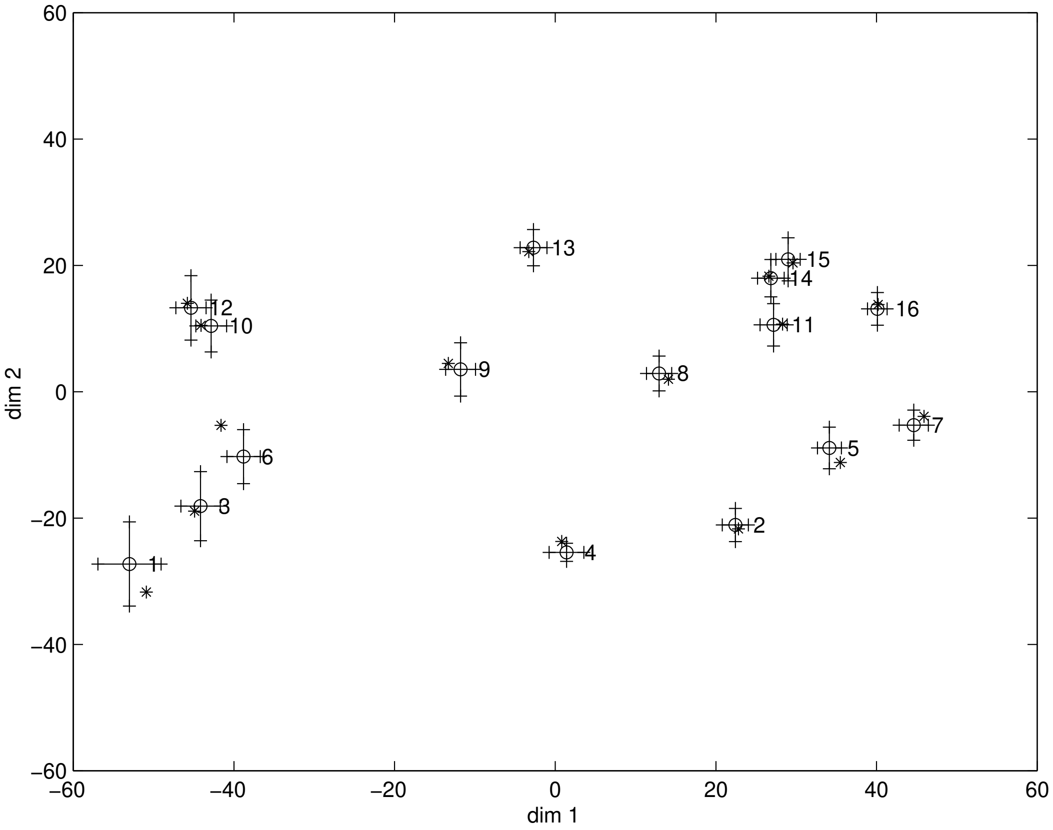

FIG. 7.4 – *Représentation des sons dans l'espace perceptif. La dimension 1 représente le timbre et la dimension 2 la hauteur des sons. Les écarts-types ont été calculés à partir d'une analyse bootstrap sur 250 tirages. Les * symbolisent les résultats de CLASCAL, les o représentent les valeurs moyennes après calcul par bootstrap*

– l'écoute des sons suivant chaque axe de l'espace perceptif 7.4 révèle deux attributs de notre perception sonore : la hauteur et un attribut du timbre sensible à l'amortissement. Celui-ci est décrit par l'axe des abscisses que nous appellerons $dim1$ et la hauteur par l'axe des ordonnées que nous désignerons par $dim2$.
– puisque les écarts perceptifs perçus entre deux stimuli construits, soit sur deux valeurs de ρ consécutives, soit sur deux valeurs de η consécutives, sont sensiblement équivalents (au dessus du seuil), nous pouvons comparer la gamme de variation des deux jeux de coordonnées correspondant aux deux dimensions. En effet la gamme correspondant aux

coordonnées des stimuli sur la premiére dimension est presque deux fois plus grande que celle correspondant à la deuxième dimension (la largeur suivant $dim2$ représente 57% de la largeur suivant $dim1$). C'est aussi valable pour les écarts-types estimés sur chacune des deux dimensions représentant la variance dans les jugements de dissemblance entre sujets et qui sont significativement plus élevés sur la dimension 2. De tout celà nous pouvons en déduire que l'attribut du timbre représente le premier critère de jugement pour les sujets devant la hauteur. Ce résultat est d'autant plus surprenant qu'il est usuel de considérer la hauteur comme un attribut facile et donc prépondérant dans les tâches de jugements en absence de variations de sonie.

– certains stimuli jugés trop proches par les sujets ont donné lieu à des confusions provoquant des inversions dans la représentation spatiale sur les deux dimensions. Il faut néanmoins distinguer sur l'espace perceptif rendu par l'analyse les sons proches et confondus des sons proches et non confondus par les sujets. Pour cela l'analyse par bootstrap permet d'observer les nuages de points et montre qu'en effet, certains nuages se retrouvent partiellement confondus tandis que d'autres sont nettement séparés alors que les distances entre les points représentant les stimuli sont équivalentes. C'est la représentation de la figure 7.5 qui ne montre toutefois que les confusions sur les deux coordonnées. Par exemple les sons 11, 14 et 15 ne sont pas ordonnés entre eux comme ils le sont physiquement et la distribution des points pour ces trois stimuli ne forme qu'un seul nuage, comme si les jugements par paires entre chacun de ces sons et les autres stimuli présents dans l'espace étaient réalisés de la même façon.

7.4.2 Analyse des corrélations

Nous allons maintenant discuter des résultats bruts de l'analyse et de l'analyse des corrélations entre les différents paramètres, acoustiques analytiques et perceptifs.

L'information acoustique relative aux paramètres physiques est véhiculée par les paramètres analytiques : le centre de gravité spectral et la constante

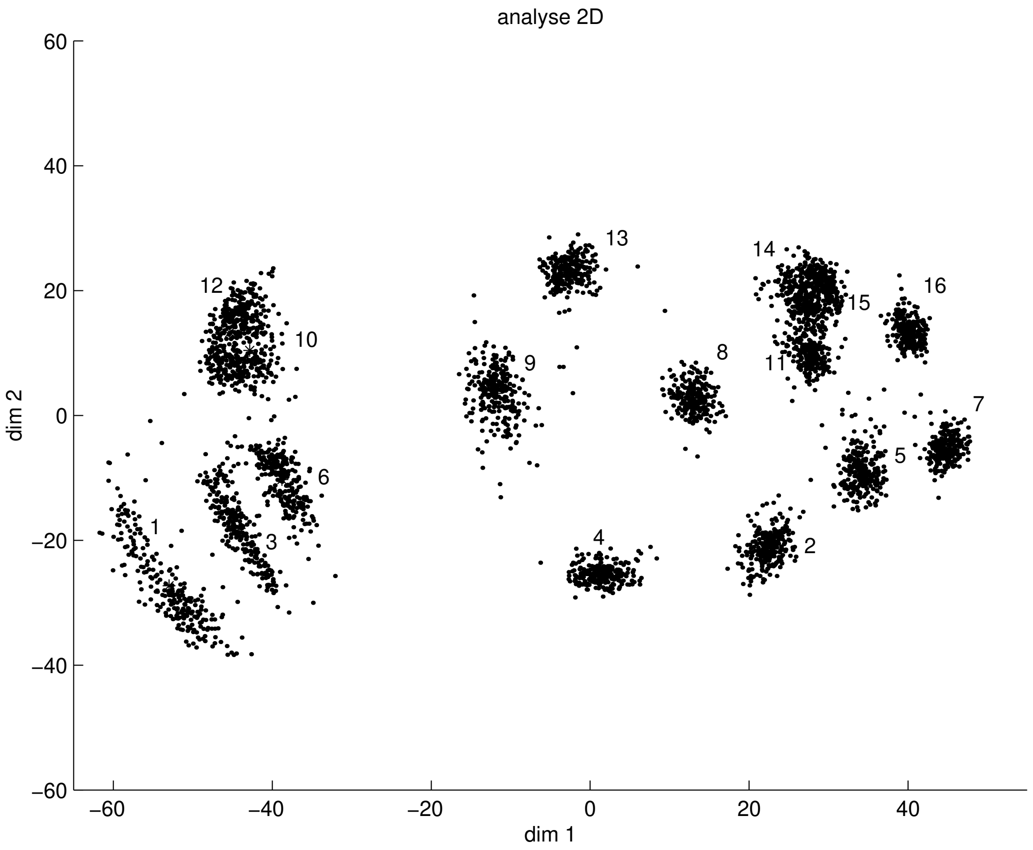

FIG. 7.5 – *Représentation des nuages de sons dans l'espace perceptif, le calcul est fait par la méthode bootstrap sur 250 tirages*

d'amortissement décrivent les variations du paramètre du modèle d'amortissement tandis que la fréquence du premier partiel décrit la variation de la densité du matériau. L'ensemble des paramètres acoustiques, analytiques et perceptifs sont répertoriés dans le tableau 7.5.

Afin de déterminer les relations entre les différents paramètres nous avons calculé les corrélations entre les trois ensembles de paramètres : les paramètres analytiques (f_1, CGS, α), les paramètres physiques (ρ, η) ainsi que les données perceptives rendues par l'expérience et l'analyse MDS (dim1, dim2). Le tableau 7.5 donne les corrélations entre les différents jeux de paramètres.

Pour ajuster les courbes de corrélations, et augmenter ainsi les coefficients de corrélation, nous avons calculé les valeurs sur $\eta^{0.3}$ et $\log \alpha$. L'exposant 0.3 a été obtenu après calcul d'ajustement sur plusieurs fonctions (log et

	Mécanique		Analytique			Perceptif	
Sons	$\rho_{(kg/m^3)}$	η $_{(s*10^{10})}$	$CGS_{(\mathrm{Hz})}$	$f_{1(\mathrm{Hz})}$	$\alpha_{(\mathrm{rad/s})}$	$dim1$	$dim2$
1	300	1	5473	2857	4.0	-50.9	32.2
2	340	700	3713	2685	26.0	22.8	22.9
3	380	40	5362	2543	6.5	-44.9	19.4
4	420	400	4893	2419	26.0	0.8	25.1
5	460	1500	3291	2313	42.0	35.5	11.3
6	500	60	5254	2220	7.0	-41.6	5.8
7	540	2000	4142	2138	80.0	45.9	4.1
8	580	500	4639	2062	28.0	14.1	-2.0
9	620	200	4934	1992	14.5	-13.3	-4.5
10	660	20	5225	1922	5.0	-44.1	-10.5
11	700	1200	4122	1874	46.0	28.3	-11.2
12	740	10	4991	1816	45.0	-45.8	-14.6
13	780	300	4440	1781	18.0	-3.3	-23.3
14	820	1000	3892	1734	39.0	26.6	-19.1
15	860	900	3888	1693	36.0	29.6	-21.5
16	900	1800	3543	1658	61.0	40.2	-14.6

TAB. 7.4 – *Récapitulatif des résultats et données. Le tableau contient successivement les numéros des sons, leurs caractéristiques physiques (ρ et η), les paramètres analytiques (CGS, f_1 et α), et les coordonnées des sons dans l'espace perceptif ($dim1$ et $dim2$).*

	$\eta^{0.3}$	$\log\alpha$	CGS	$dim1$	$\sqrt{\rho}$	f_1	$dim2$
$\eta^{0.3}$	1.00						
$\log\alpha$	0.99	1.00					
CGS	-0.95	-0.93	1.00				
$dim1$	0.99	0.98	-0.96	1.00			
$\sqrt{\rho}$	0.34	0.33	-0.37	0.41	1.00		
f_1	-0.35	-0.33	0.33	-0.40	-0.98	1.00	
$dim2$	0.08	0.10	-0.07	-0.01	-0.98	0.96	1.00

TAB. 7.5 – *Coefficients de corrélation entre les paramètres mécaniques (ρ,η), analytiques (CGS, α et f_1) et perceptifs ($dim1$,$dim2$).*

puissance). Compte tenu de la forme de la courbe, le coefficient retenu est celui pour lequel l'ajustement est le meilleur. Ces ajustements ont-ils une signification physique, analytique? Nous reviendrons sur cette interprétation lors de l'expérience suivante de validation.

Afin de visualiser ces corrélations nous avons calculé les régressions simples et multiples entre les différents paramètres. Nous donnerons pour chaque régression la représentation graphique (scattergram) de la relation, la relation numérique correspondante entre les paramètres et le coefficient de détermination multiple R^2 correspondant à la variance expliquée (voir section 4.3).

Remarques sur les corrélations croisées

Les corrélations croisées entre les paramètres CGS, η, α et la dimension 2 sont très faibles (autour de 0.3), de même pour les corrélations entre ρ, f_1 et la dimension 1. Les paramètres analytiques sont donc bien spécifiques à chaque dimension trouvée par l'analyse MDS. La corrélation entre $dim1$ et $dim2$ est quant à elle extrêmement faible, $r = 0.01$, ce qui rend ces deux dimensions parfaitement orthogonales et nous permet de les traiter séparément.

Dimension 1

- Corrélation entre les paramètres analytiques (CGS,α) et la dimension 1
  Les régressions de CGS, α sur la $dim1$ sont reportées dans le tableau 7.6 et montrent que le timbre est décrit à travers ces deux paramètres et non par le seul CGS introduisant ainsi le paramètre temporel α. Ainsi pour le groupe de sons présentés dans cette expérience il apparaît que la dimension 1 sera décrite le mieux possible en terme de corrélation en prenant en compte les deux paramètres. Cette corrélation entre les trois paramètres peut être calculée par régression multiple sur les deux paramètres CGS et α (fig 7.7). Le calcul de cette corrélation $R^2 = 0.99$ révèle la plus forte valeur trouvée pour cette expérience. La dimension 1 regroupe donc des aspects spectraux (CGS) et temporels (α) de la perception.
- Corrélation entre le terme d'amortissement η et les paramètres analytiques (CGS,α)
  Le lien entre les paramètres physiques et les paramètres analytiques est indépendant de l'expérience et n'est régi que par les équations de la physique et leurs formulations numériques. Ainsi le CGS et la décroissance temporelle sont consécutifs de la valeur imposé de l'amor-

$dim1 = -103.05 + 81.0 * \log(\alpha)$	$R^2 = 0.98$
$dim1 = 205 - 0.046 * CGS$	$R^2 = 0.91$
$dim1 = 19.5 + 62.4 * \log(\alpha) - 0.013 * CGS$	$R^2 = 0.99$

TAB. 7.6 – *Régression entre les paramètres analytiques et la dimension dim*1*, valeurs des* R^2 *ajustés correspondant*

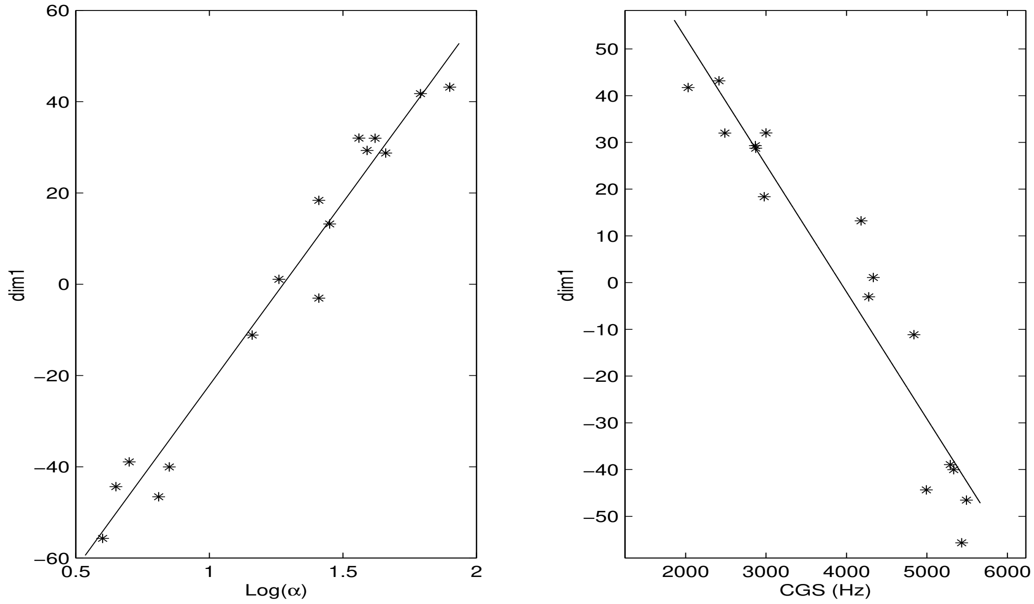

FIG. 7.6 – *Représentation des corrélations entre la dimension perceptive dim*1 *et les paramètres analytiques* (α *,CGS) séparément*

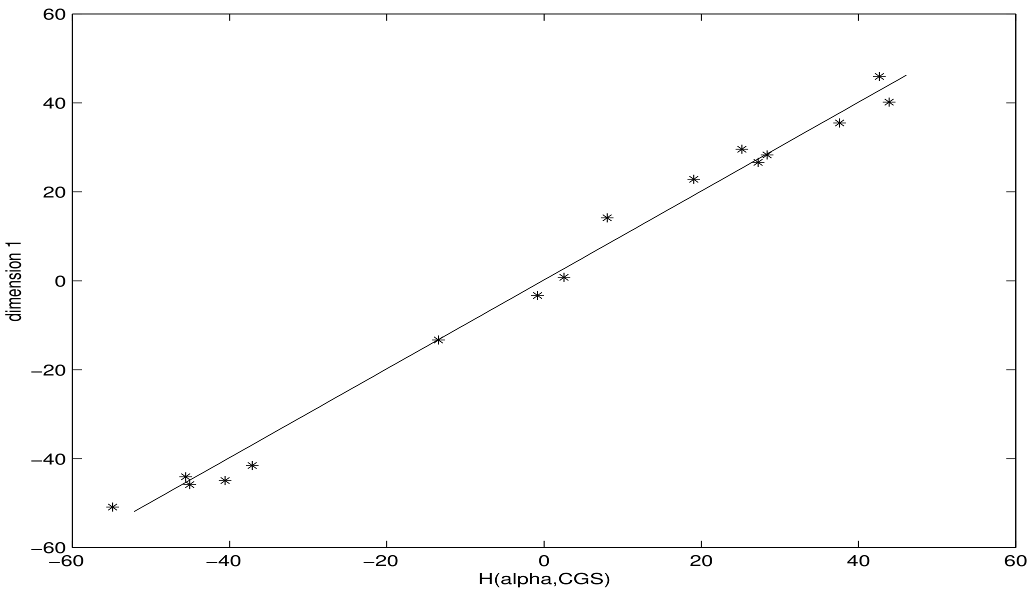

FIG. 7.7 – *Régression multiple de α et CGS sur la dimension perceptive $dim1$*

tissement d'où de fortes corrélations séparées : $R^2 = 0.87$ entre η et le CGS et $R^2 = 0.91$ entre η et α. Nous avons calculé ensuite la double corrélation par régression multiple sur les deux paramètres analytiques CGS et α. Le résultat $R^2 = 0.95$ montre encore une fois une très forte corrélation cette fois-ci prévisible. L'ensemble est répertorié dans le tableau

$\eta = 0.26 + 0.3 * \log \alpha$	$R^2 = 0.91$
$\eta = 0.61 - 0.001 * CGS$	$R^2 = 0.87$
$\eta = 2.5 + 0.14 * \log \alpha - 0.0005 * CGS$	$R^2 = 0.95$

TAB. 7.7 – *Régressions entre les paramètres analytiques et le terme d'amortissement η, valeurs des R^2 correspondants.*

Dans ce cas, comme dans le précédent nous pouvons remarquer que l'apport supplémentaire en variance expliquée par adjonction du CGS est relativement faible, 0.004 pour ce cas et 0.01 pour le cas précédent. Il est donc difficile de savoir quelle est la contribution réelle des deux paramètres.

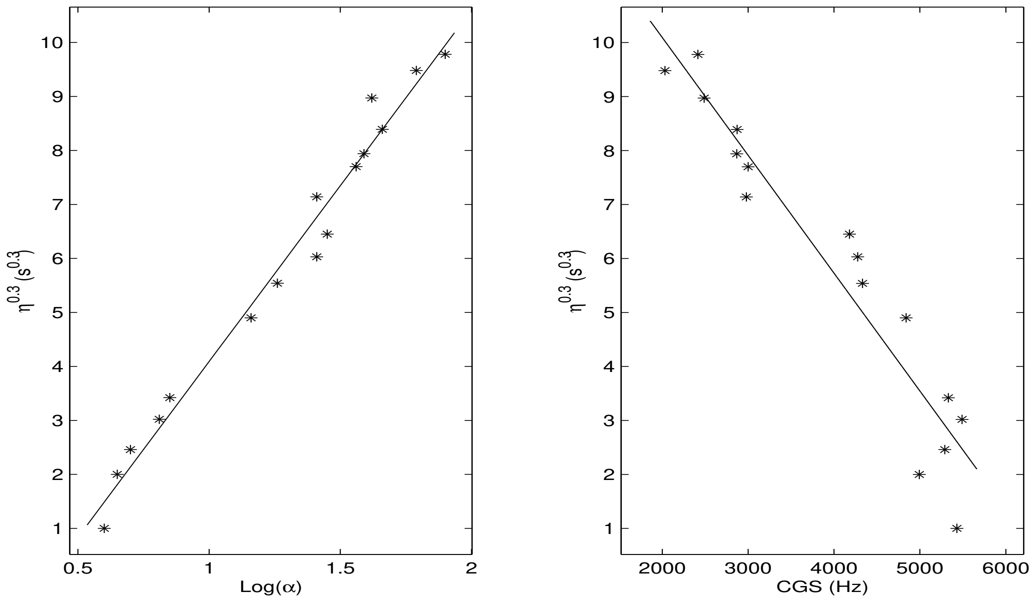

FIG. 7.8 – *Représentation des corrélations entre le terme d'amortissement* η *et les paramètres analytiques* (α *,CGS).*

– Corrélation entre le paramètre mécanique η et la dimension 1 de l'espace perceptif.

  Le contrôle du paramètre d'amortissement η lors de la construction des sons nous permet ici d'établir un lien entre les variations de celui-ci et la dimension 1 de l'espace perceptif. L'analyse régressive donne la variance expliquée $R^2 = 0.97$. Cette très forte corrélation confirme la relation triangulaire des liens entre les trois jeux de paramètres. La figure 7.9 montre la relation entre η et $dim1$, les valeurs sont répertoriées dans le tableau 7.8.

$dim1 = -72.3 + 97.7 * \eta^{0.3}$	$R^2 = 0.97$

TAB. 7.8 – *Tableau de régression entre le paramètre mécanique* η *et la dimension perceptive* $dim1$*, valeurs du* R^2 *correspondant*

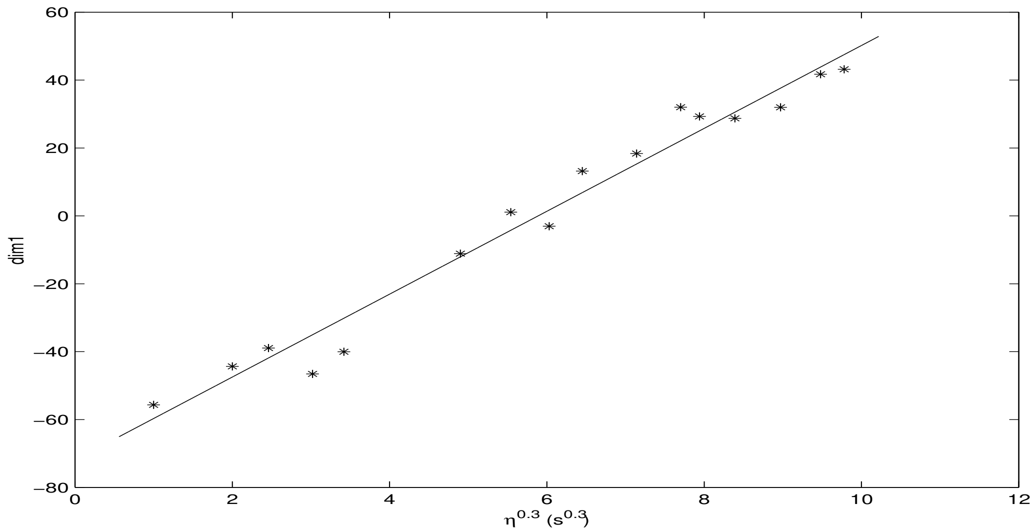

FIG. 7.9 – *Représentation de la corrélation entre la dimension perceptive* $dim1$ *et le paramètre d'amortissement* η

Dimension 2

La dimension 2 de l'espace perceptif est très bien corrélée avec la fréquence f_1 et s'interprète comme la dimension représentant la hauteur du son. Il existe une relation physique directe entre f_1 et $\sqrt{\rho}$; de fait la corrélation entre les deux est de un. Nous ne considérons donc que les corrélations entre ces deux paramètres et la dimension 2 rendue par l'analyse multidimensionnelle (tab 7.9):

$dim2 = 52 - 0.09 * \sqrt{\rho}$	$R^2 = 0.92$
$dim2 = -95.5 + 0.045 * f_1$	$R^2 = 0.94$

TAB. 7.9 – *Régressions entre* ρ, f_1 *et la dimension perceptive* $dim2$ *et valeurs du* R^2 *correspondant.*

Les corrélations très fortes avec f_1, d'une part, et ρ, d'autre part confirment l'interprétation de la dimension 2 comme axe de la hauteur. L'ordre des sons suivant $dim2$ correspond à l'ordre des sons suivant la fréquence f_1 à l'exception de trois d'entre eux (fig 7.4). Deux sons ont été inversés, les sons numérotés 13 et 16 sur lesquels nous pouvons faire deux remarques : l'inversion se situe tout d'abord dans la limite de leurs écarts-types, enfin le CGS

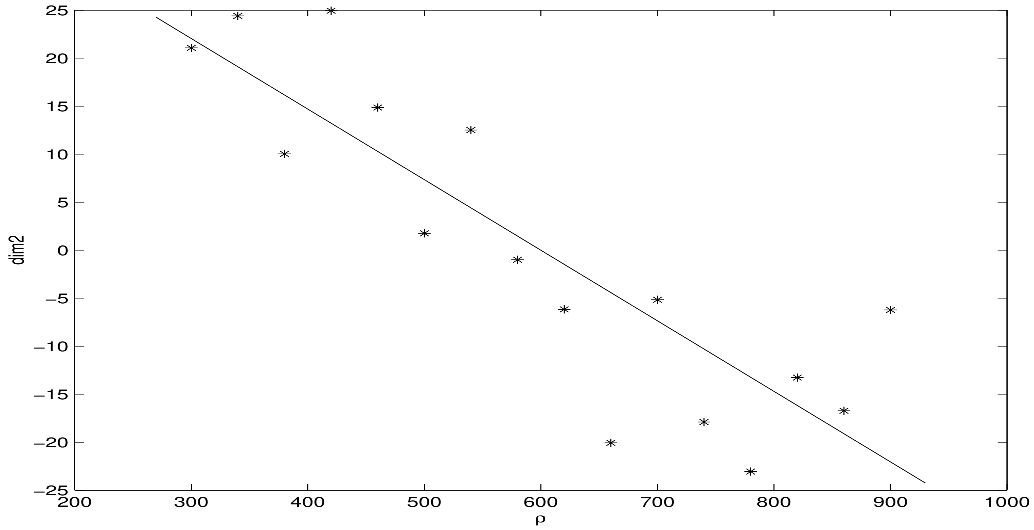

FIG. 7.10 – *Représentation de la corrélation entre la dimension perceptive dim2 et la densité du matériau* ρ

du son 16 est plus élevé que celui du 13 pouvant induire une confusion dans le jugement de hauteur.
Comme nous l'avons précédemment remarqué, la hauteur a été jugée sur un intervalle plus étroit comparativement à l'amortissement, faisant de celle-ci un critère de jugement secondaire. Nous pouvons l'expliquer par le faible écart tonal entre les valeurs extrêmes des fréquences fondamentales qui se situent entre 1658 et 2857 Hz soit un peu moins d'un quart de ton entre deux stimuli successifs.

7.5 Conclusion

A la fin de cette première expérience, l'hypothèse d'une relation triangulaire des liens entre les différents jeux de paramètres est confirmée et appuyée par de solides corrélations. Premièrement, les paramètres analytiques décrivent bien les attributs perceptifs auxquels les sujets sont sensibles pendant l'expérience. Deuxièmement, ces dimensions perceptives sont aussi fortement corrélées avec les paramètres mécaniques qui varient lors de la synthèse sonore. Ces résultats montrent une relation forte et précise entre un terme d'amortissement introduit dans le modèle numérique et une dimension directement perçue par l'auditeur. La voie de l'analyse de la source est donc bien ouverte. On peut aussi remarquer qu'il faut plusieurs paramètres analytiques et un seul paramètre mécanique pour décrire la dimension 1. Ceci va absolument dans le sens de nos hypothèses pour lesquelles l'analyse de la source va permettre de cibler dans quelle mesure les paramètres analytiques sont pertinents et exhaustifs pour décrire tel ou tel attribut. En ce qui concerne les deux dimensions spécifiquement, il nous faut maintenant valider la relation entre l'amortissement et l'attribut du timbre puis tester la sensibilité des auditeurs sur un jeu de sons de barres accordées afin d'observer l'effet de la hauteur tonale dans les jugements de dissemblance.

Chapitre 8

Expérience 2 : barres à sections variables

8.1 Introduction

Lors de la première expérience sur les barres nous avons observé de fortes corrélations entre les paramètres physiques, analytiques et mécaniques. Ces corrélations sont soutenues par des fonctions de régression entre les différents jeux de paramètres. Du point de vue de notre étude nous nous sommes spécialement intéressés pour cette deuxième expérience à la fonction de régression entre le terme d'amortissement η qui caractérise le modèle d'amortissement de la source et la dimension 1 de l'espace perceptif représentant l'attribut auditif du timbre, ceci afin d'affiner les liens entre le terme d'amortissement et un attribut privilégié du timbre sensible à l'amortissement. De plus la saillance de cet attribut apparaît plus forte dans les jugements des auditeurs que pour la hauteur. Afin d'estimer là aussi l'importance relative de ces deux attributs et l'influence des corrections de hauteurs réalisées par le programme, nous avons choisi de faire varier à nouveau la hauteur des sons.

8.2 Méthode

8.2.1 Stimuli

Choix des paramètres d'entrée du code de synthèse

Pour cette expérience nous avons choisi de faire varier l'amortissement logarithmiquement afin de retrouver une fonction de correspondance avec le timbre, similaire à celle de la première expérience, et de s'assurer que les espaces entre les sons correspondent perceptivement à un écart supérieur au seuil de discrimination. Pour obtenir une meilleur corrélation entre η et $dim1$ nous avons ajusté la courbe correspondante à l'aide d'un exposant 0.3 que nous espérons retrouver ici, si ce coefficient a un sens pour l'échelle de perception des sujets. Les courbes de valeurs du paramètre d'amortissement η pour les deux expériences sont reportées sur la figure 8.1.

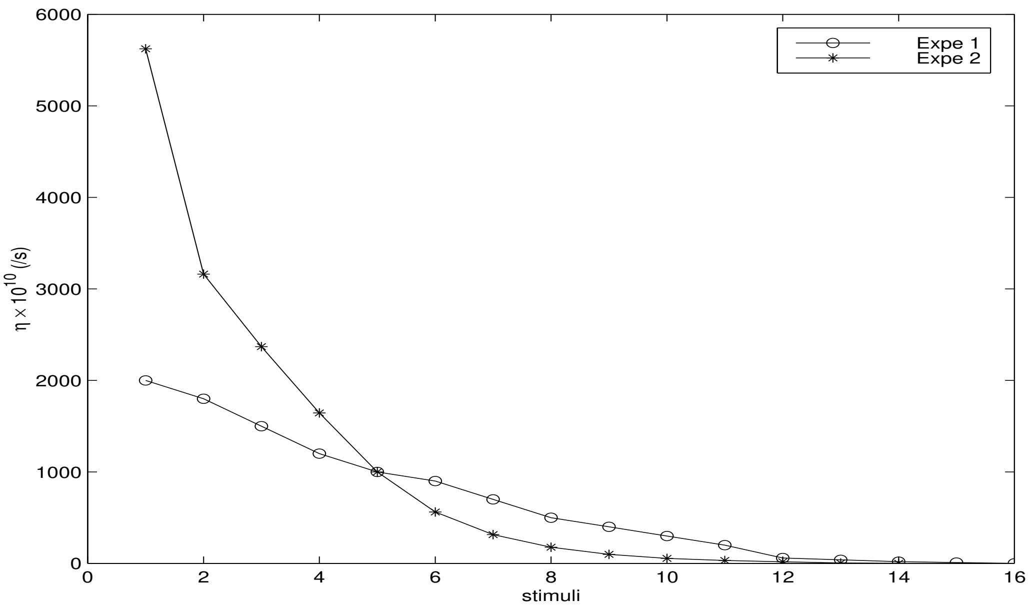

FIG. 8.1 – *Choix des paramètres mécaniques. Les stimuli sont renumérotés dans l'ordre de l'amortissement décroissant.*

Le programme de synthèse de barres par modèle physique prévoit la synthèse de barres à section variable permettant d'ajuster les partiels du spectre de fréquences sur les harmoniques de la fondamentale. La longueur de la barre permet d'ajuster la valeur de la fondamentale tandis que la partie creusée de la barre permet d'aligner les partiels sur les harmoniques de la fondamentale. C'est la base des techniques utilisées par les facteurs d'instrument lors de la fabrication. Les stimuli de cette expérience sont donc des sons de barres accordées type marimba. Les fréquences imposées et les notes correspondantes sont reportées dans le tableau 8.1 alors que les sons sont représentés dans l'espace physique décrit par la figure 8.2

Sons	η $(s * 10^{10})$	$f_{1(\mathrm{Hz})}$	Note
8	17.8	369.9	fa4d
1	0.3	415.3	sol4d
5	3.2	440.0	la4
12	177.8	466.2	la4d
2	0.6	493.9	si4
6	5.6	523.2	do5
11	100	554.4	do5d
4	1.8	587.4	re5
14	562.2	621.2	re5d
10	56.2	659.2	mi5
3	237	698.0	fa5
13	316.2	740.0	fa5d
9	31.6	784.0	sol5
7	10	880.0	la5

TAB. 8.1 – *Distribution des valeurs du paramètre d'amortissement η et des fréquences imposées au programme.*

Calcul des paramètres analytiques

Les paramètres analytiques calculés pour établir les corrélations sont les mêmes que pour les barres à section constante, à savoir le facteur d'amortissement (coefficient de décroissance α de l'enveloppe temporel) et le CGS. La fréquence de la fondamentale étant imposée au programme, nous nous sommes contentés de nous assurer qu'elle était respectée pendant la synthèse. Les temps d'attaque des sons sont compris entre 16 et 19 ms ce qui est insuffisant pour être perçus différemment ; l'attaque n'est donc pas prise en compte dans les paramètres analytiques pouvant décrire les jugements des sujets.

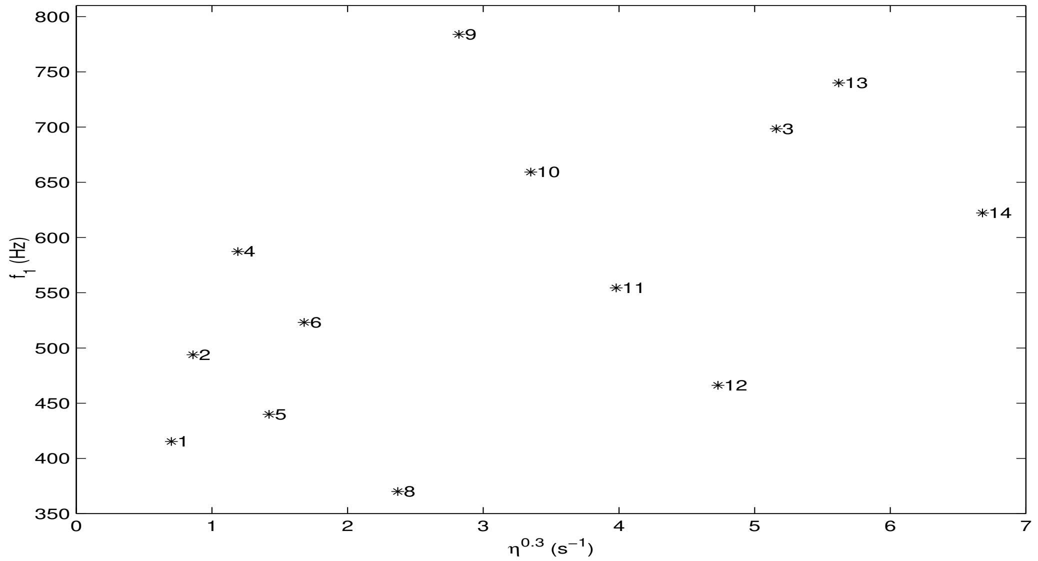

FIG. 8.2 – *Espace physique des stimuli. Les paramètres variés sont la fondamentale et le paramètre d'amortissement* η.

8.2.2 Matériel, sujets, procédure

L'expérience a eu lieu à l'ENST et a été précédée d'une expérience pilote sur 9 sons et 5 sujets ainsi que d'une expérience d'égalisation en sonie réalisée également avec 5 sujets. L'expérience durait environ une heure et a été réalisée par 22 sujets (14 hommes, 8 femmes) âgés de 22 à 40 ans ne comportant aucun musicien professionnel. La tâche était la même que pour l'Expérience 1, à savoir un test de dissemblance réalisée sur 14 sons, ce qui fait 105 paires de sons à juger.

8.3 Résultats

Nous détaillons maintenant les étapes du choix du modèle avec l'analyseur CLASCAL.

Choix du nombre de classes

Pour la première étape le programme effectue le test Monte Carlo pour déterminer le nombre de classes sur le modèle nul. Il choisit deux classes.

Choix des modèles spatiaux

Nous testons maintenant le modèle de 1 à 4 dimensions, avec ou sans spécificités, avec deux classes latentes.

	N	LV	BIC	dl
modèle	1	-8925	17972	16
sans	2	-8336	16899	30
spécificités	3	-8257	16846	44
	4	-8244	16928	58
	N	LV	BIC	dl
modèle	1	-8394	17022	31
avec	2	-8269	16878	45
spécificités	3	-8247	16941	59
	4	-8242	17035	73
Modèle nul		-8250	17891	

TAB. 8.2 – *Test BIC pour le modèle à deux classes*

Les deux meilleurs modèles spatiaux choisis selon l'indice BIC (donnés dans le tableau 8.2) nous donne 3Dns puis 2Dsp .

Test Monte Carlo sur le modèle spatial

Un test Monte Carlo opposant les modèles 2Dsp et 3Dns rejette le modèle 3Dns, mais un test opposant 2Dns et 2Dsp ne rejette pas 2Dns.

Les classes

On recommence l'étape 1 avec cette fois le modèle spatial 2Dns retenu à la place du modèle nul. Le programme confirme le nombre de classes. Le modèle définitif est donc **le modèle à deux dimensions avec spécificités et deux classes**.

Nous donnons dans le tableau 8.3 le poids relatif de chaque classe pour chaque dimension. Il s'agit des coefficients $w_t r$ introduits dans le modèle de calcul des distances décrit par l'équation 6.1. Cette répartition nous permet de constater que les classes sont faites autour du choix du paramètre jugé le plus saillant par les sujets. Ainsi la classe 1 constituée par deux tiers des sujets a jugé les sons principalement sur la dimension 2, et la classe 2 constituée d'un tiers des sujets, sur la dimension 1. Le tableau 8.3 récapitule ces résultats. La probabilité de chaque sujet d'appartenir effectivement à une classe est sans équivoque comme le montre le tableau 8.4

classe	Ns	$dim1$	$dim2$
1	14	0.55	1.70
2	7	1.45	0.29

TAB. 8.3 – *Répartition du poids des classes sur chaque dimension, chaque classe contenant Ns sujets.*

Nous pouvons remarquer que les pondérations sont proportionnellement plus élevées sur la classe 2 indiquant que les auditeurs concentrés sur la dimension 2 le sont davantage que pour la classe 1 sur la dimension 1.

1	0.0000	1.0000
2	0.0000	1.0000
3	0.0000	1.0000
4	1.0000	0.0000
5	0.0000	1.0000
6	1.0000	0.0000
7	0.0021	0.9979
8	0.0000	1.0000
9	1.0000	0.0000
10	1.0000	0.0000
11	0.9997	0.0003
12	1.0000	0.0000
13	1.0000	0.0000
14	1.0000	0.0000
15	0.9901	0.0099
16	1.0000	0.0000
17	1.0000	0.0000
18	1.0000	0.0000
19	1.0000	0.0000
20	1.0000	0.0000
21	0.0000	1.0000

TAB. 8.4 – *Probabilité à posteriori des sujets d'appartenir à chacune des deux classes.*

Récapitulatifs des résultats et données

Les stimuli sont représentés sur l'espace à deux dimensions dans la figure 8.3 que l'on peut comparer à l'espace physique des stimuli représentés sur la figure 8.2.

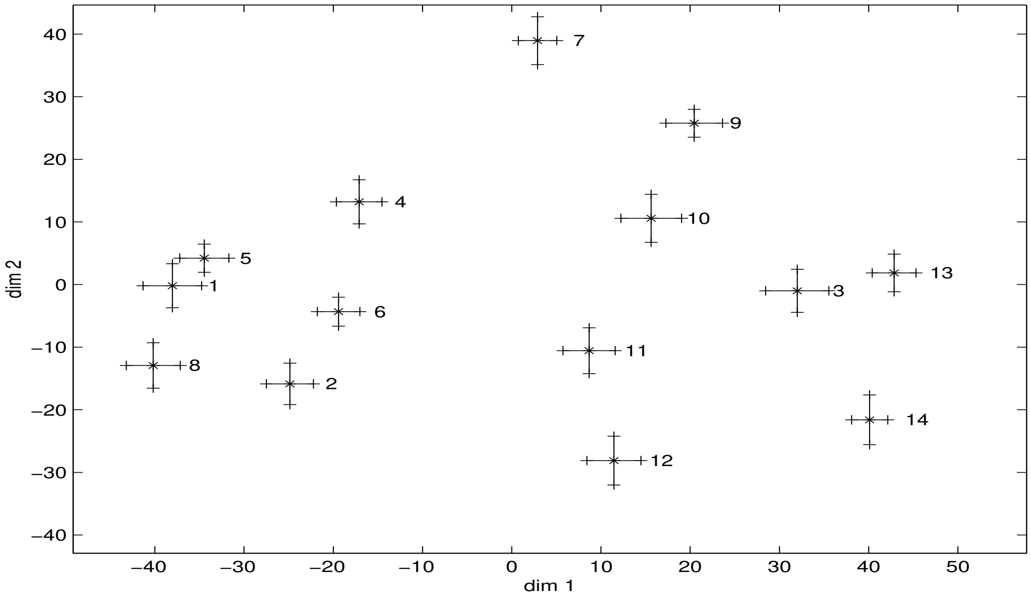

FIG. 8.3 – *Espace perceptif rendu par l'analyse MDS, les écarts-types sont estimés par l'analyse bootstrap sur 250 tirages.*

Le modèle spatial étant complet et déterminé, nous allons maintenant comparer les résultats avec ceux obtenus dans la première expérience sur les barres. Les comparaisons sont faites par rapprochement direct entre les espaces perceptifs ainsi que par l'analyse des corrélations. Le tableau 8.5 donne l'ensemble des données analytiques, mécaniques et perceptives.

8.4 Discussion

8.4.1 Discussion sur les résultats bruts

L'espace rendu par l'analyse est un espace à deux dimensions. Ainsi, comme pour la première expérience, le nombre de paramètres physiques variés pour construire les stimuli correspond à la dimensionnalité de l'espace

	Mécanique		Analytique		Perceptif	
Sons	$f_{1(Hz)}$	η $(s*10^{10})$	CGS(Hz)	α(rad/s)	$dim1$	$dim2$
14	622	562.3	2223	5.6	48.8	-1.0
13	740	316.2	2868	7.1	43.6	-19.1
3	698	237.0	3500	7.4	33.1	-12.2
12	466	177.8	3933	7.7	29.1	19.9
11	554	100.0	4096	7.2	20.0	9.1
10	659	56.2	4231	6.1	8.8	-12.9
9	784	31.6	4332	5.2	4.9	-29.0
8	370	17.8	5859	5.6	-25.0	31.0
7	880	10.0	5022	4.5	-18.5	-35.5
6	523	5.6	5412	4.3	-18.9	9.1
5	440	3.2	6039	4.2	-36.9	11.8
4	587	1.8	5470	4.0	-27.2	-7.6
2	493	0.6	5192	3.9	-28.2	21.3
1	415	0.3	6382	3.9	-33.6	14.9

TAB. 8.5 – *Tableau des paramètres analytiques, mécaniques et perceptifs*

perceptif. Le modèle est un modèle pondéré à deux dimensions. Les axes sont donc bloqués en rotation. Nous pouvons donc comparer directement les espaces physiques (fig 8.2) et perceptifs (fig 8.3) et remarquer visuellement une bonne correspondance qui laisse supposer des corrélations convenables entre les paramètres de chaque espace.

La gamme de valeurs occupée par la dimension 2 est encore sensiblement inférieure à celle occupée par la dimension 1 comme dans la première expérience avec une gamme de 83 pour la dimension 1 et de 66 pour la dimension 2. Nous avons observé en effet dans la première expérience que les sujets prenaient l'attribut du timbre correspondant à la sensation d'amortissement comme premier critère. Dans cette expérience nous avons forcé les sons, par le modèle de barres à sections variables, à être accordés. Chaque son étant séparé d'un demi-ton nous pouvions espérer cette fois-ci une plus grande sensibilité sur la hauteur. De plus la sensation d'amortissement est moins forte à l'oreille du fait que la fondamentale conserve cette fois un amortissement assez faible. Ceci est dû au modèle de barres à section variable qui réajuste la longueur de la barre pour forcer la valeur de la fondamentale.

8.4.2 Analyse des corrélations

Le tableau 8.6 rend compte des corrélations entre les jeux de paramètres.

	$\eta^{0.3}$	CGS	α	$dim1$	f_1	$dim2$
$\eta^{0.3}$	1.00					
CGS	-0.94	1.00				
α	0.82	-0.71	1.00			
$dim1$	0.97	-0.97	0.84	1.00		
f_1	0.34	-0.51	0.21	0.41	1.00	
$dim2$	-0.23	0.39	-0.11	-0.30	-0.97	1.00

TAB. 8.6 – *Tableau des corrélations entre le paramètre mécanique η, les paramètres analytiques (CGS, α et f_1) et perceptifs ($dim1$, $dim2$)*

L'exposant 0.3 choisi dans la première expérience afin d'obtenir de meilleurs corrélations sur le paramètre mécanique η a été calculé par itération. Le tableau 8.7 donne les coefficients de corrélation en fonction de l'exposant choisi et montre que dans les deux expériences l'exposant 0.3 est retenu pour obtenir le meilleur coefficient de corrélation.

	η	$\eta^{0.1}$	$\eta^{0.2}$	$\eta^{\mathbf{0.3}}$	$\eta^{0.4}$	$\eta^{0.5}$
SV	0.856	0.953	0.971	**0.975**	0.968	0.957
SC	0.894	0.961	0.981	**0.985**	0.981	0.971

TAB. 8.7 – *Tableau des coefficients de corrélations entre la dimension 1 déterminée pour chaque expérience et le paramètre mécanique varié η. SC désigne l'expérience sur les barres à section constante et SV désigne l'expérience sur les barres à section variable*

Cet exposant n'est donc pas dû au hasard mais reflète l'échelle de perception des barres par les sujets. Un grand nombre de paramètres physiques de la nature sont perçus logarithmiquement (luminosité, sonie, ..). Il semble donc que ce soit le cas aussi du modèle d'amortissement pris en compte dans le programme. Ce résultat est d'autant plus intéressant qu'il valide la relation forte trouvée avec les barres à section variable avec le même coefficient 0.3 dans un contexte sonore très différent (valeurs de l'amortissement, barres accordées). De plus la gamme du paramètre temporel α est cette fois-ci beaucoup plus étroite du fait d'un amortissement faible de la fondamentale pour chaque son. Toutefois la saillance et la fonction de correspondance du paramètre η avec la dimension perceptive sont maintenues. Nous sommes donc directement sensibles à ce paramètre et ceci d'autant plus que la corrélation avec le paramètre analytique α est moins forte. Nous pouvons même remarquer que la meilleur corrélation analytique est obtenue après un calcul par l'analyse ESPRIT [45] des coefficients α fréquence par fréquence et reste inférieure à la valeur trouvée pour le coefficient η. Le paramètre analytique le plus pertinent serait dans ce cas la décroissance temporelle de la deuxième harmonique (voir le tableau 8.8), ce qui a nécessité une étude détaillée tandis que la corrélation avec η est immédiate et plus forte.

	enveloppe	f_1	f_2	f_3	f_4	f_5
R($dim1$, α)	0.840	0.605	**0.878**	0.871	0.867	0.812

TAB. 8.8 – *Coefficient de corrélation R pour l'enveloppe temporelle et chacun des harmoniques.*

Dimension 1

- Corrélation entre les paramètres analytiques (CGS,α) et la dimension 1
  La corrélation avec le CGS est la plus forte, montrant ainsi que le timbre décrit par la dimension 1 est proche de la brillance. La corrélation avec le coéfficient α est moins forte, comme nous l'avons déjà évoqué au paragraphe précédent, mais elle est néanmoins significative confirmant ainsi que la dimension perceptive $dim1$ est décrite à travers les deux paramètres analytiques. Ceci est montré par la régression multiple du paramètre dépendant $dim1$ sur les deux paramètres indépendants CGS et α dont le R^2 est le plus fort coefficient de régression : $R^2 = 0.98$ (voir la figure 8.4).
- Corrélation entre le paramètre mécanique η et la dimension 1 de l'espace perceptif
  La corrélation de 0.97 confirme le lien direct entre le paramètre η et la dimension 1 de l'espace perceptif. La validation de l'exposant 0.3 montre à nouveau la sensibilité de l'auditeur à la variation du paramètre $\eta^{0.3}$.

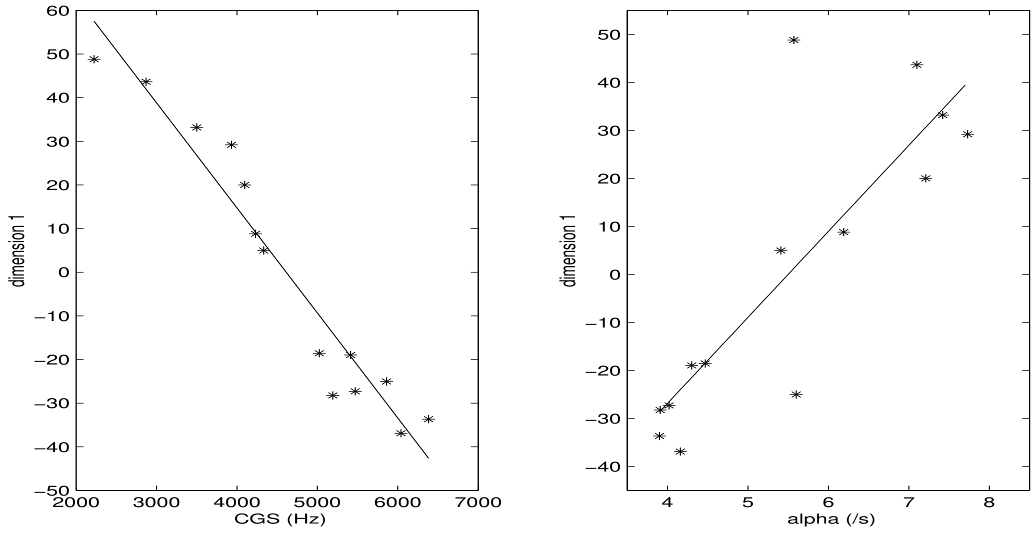

FIG. 8.4 – *Représentation de la corrélation entre la dimension perceptive dim*1 *et la régression multiple sur les deux paramètres analytiques CGS et* α.

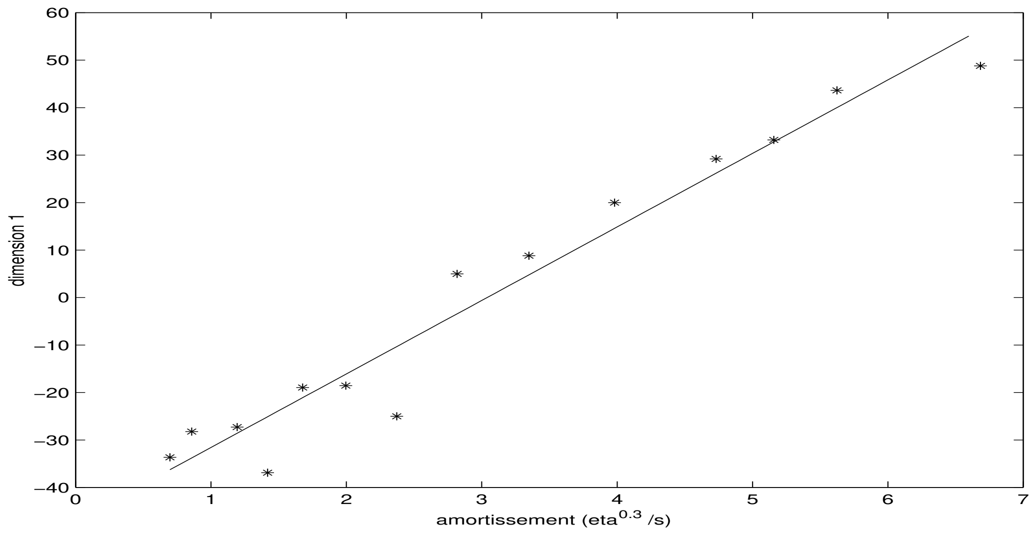

FIG. 8.5 – *Représentation de la corrélation entre la dimension perceptive dim*1 *et le paramètre d'amortissement* η.

Dimension 2

La corrélation très forte entre la dimension 2 et la fréquence imposée au programme lors de la synthèse des stimuli confirme la sensibilité des auditeurs aux variations de la hauteur. Cependant, les barres sont cette fois-ci accordées et donc la hauteur des stimuli est à priori encore plus saillante que pour l'expérience 1. On peut donc s'attendre à ce que les sujets focalisent sur cet attribut dans leurs jugements ce que semble confirmer la répartition des sujets dans les classes. En effet nous pouvons remarquer que deux tiers des auditeurs ont fait leurs jugements principalement sur la hauteur, comme le montre les pondérations du tableau 8.3. Néanmoins de nombreuses inversions de points sont présentes sur la dimension perceptive, comme le montre la figure 8.6.

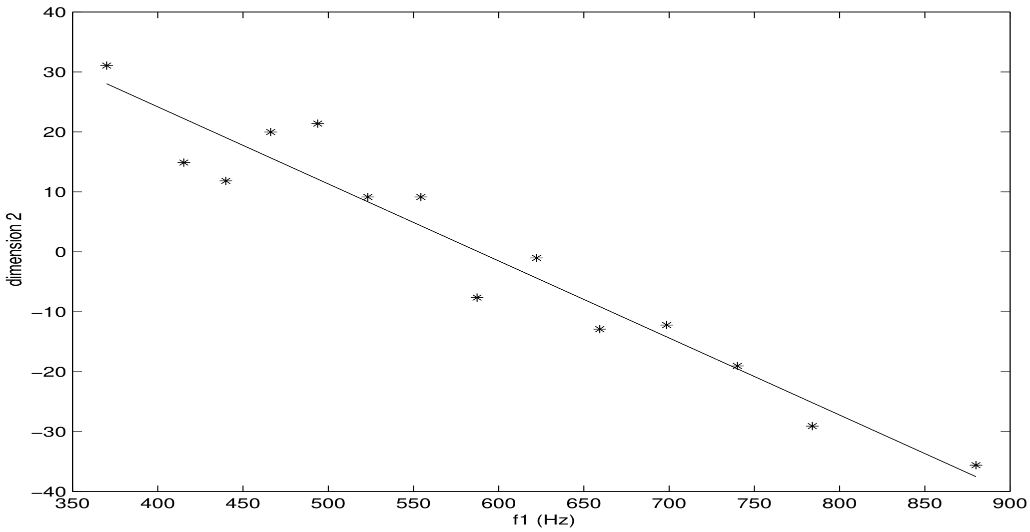

FIG. 8.6 – *Représentation de la corrélation entre la dimension perceptive dim2 et la fréquence imposée au programme* f_1.

8.5 Conclusion

Cette expérience sur les barres à sections variables nous a permis de valider les résultats de l'expérience 1. Ainsi nous pouvons maintenant confirmer que les auditeurs ont une sensibilité directe à l'amortissement mécanique tel qu'il est décrit dans le modèle de synthèse. De plus cette corrélation forte et directe semble au moins autant appropriée que les paramètres analytiques, nécessitant d'être composés entre eux et étudiés dans le choix précis du calcul. Il paraît donc naturel de se référer a l'amortissement mécanique pour décrire le timbre contenu dans la dimension 1. Nous avons également montré que l'exposant 0.3, trouvé lors de la première expérience pour ajuster au mieux les corrélations entre le paramètre analytique d'amortissement η et la dimension 1, est confirmé par cette nouvelle expérience. Il a donc une signification psychologique pour l'auditeur. Cette expérience a permis aussi de confirmer que la hauteur n'était pas un attribut systématiquement prédominant dans le jugement des sujets, même quand les stimuli présentent des dissemblances aussi fortes qu'un écart d'un demi-ton entre deux stimuli consécutifs.

Chapitre 9

Etude des plaques

9.1 Introduction

La prise en compte des pertes décrite au chapitre 5 s'est révélée déterminante pour une synthèse perceptivement réussie du matériau. Nous avons vu dans notre étude sur les barres qu'avec un modèle simple de l'amortissement nous pouvons décrire une dimension perceptive saillante pour les auditeurs. Il est donc naturel de s'intéresser maintenant aux plaques qui ont nécessité un modèle plus complexe pour l'amortissement, afin de s'assurer que ce phénomène physique est pertinent pour décrire le timbre des stimuli. Nous décrirons l'influence des différents phénomènes dissipatifs afin de mieux comprendre leurs rôles respectifs dans la caractérisation perceptive des matériaux. Pour celà, après avoir rappelé les modèles d'amortissement et leurs significations physiques pour chaque matériau, nous décrirons le modèle hybride utilisé pour la synthèse des stimuli de nos expériences. Dans cette étude nous observons tout d'abord la fiabilité des hypothèses sur la sensibilité des auditeurs aux phénomènes mécaniques inhérents au système physique. Selon ces hypothèses nous serions directement sensibles à différents phénomènes mécaniques en particulier à la dissipation de l'énergie. Ensuite nous observerons à travers une deuxième expérience le rôle de ces phénomènes physiques dans la reconnaissance des matériaux par les auditeurs.

Symboles	Définitions
α	Facteur d'amortissement $(/s)$
P_α	Pente de la courbe $\alpha(f)$ $(rad/s/Hz)$
R_i, s_i	Paramètres des modèles d'amortissement (rad/s)
l_x, ,l_y, h	Paramètres géométriques des plaques (m)
c_i	$(E_i/\rho)^{1/2}$ célérité des ondes (m/s)
m	Masse du maillet (kg)
V	Vitesse du maillet (m/s)
K	Raideur du maillet $(m/N^{-2/3})$

TAB. 9.1 – *Symboles utilisés dans la description des modèles physiques.*

9.2 Présentation des stimuli

9.2.1 Modèle mécanique : rappel du modèle de plaque

Dans l'hypothèse des faibles amortissements, le terme dissipatif $\tilde{D}_i$ introduit dans l'équation 9.1 est donné par la somme des trois termes correspondant aux différents phénomènes physiques [11] du modèle d'amortisement. Dans ce cas le plus général des plaques isotropes, les rigidités complexes sont données par :

$$\tilde{D}_i = D_i(1 + \tilde{d}_{ive}(s) + \tilde{d}_{ithe}(s) + \tilde{d}_{ray}(s)), \tag{9.1}$$

où les différentes contibutions sont modélisées par :

$$\left\{\begin{array}{lcll}
\tilde{d}_{ive}(s) & = & \displaystyle\sum_{m=1}^{2} \frac{sR_{im}}{s+s_m} & \text{pertes viscoélastiques} \\
\tilde{d}_{ithe}(s) & = & \displaystyle\frac{R_1 s}{s+c_1/h^2} & \text{pertes thermoélastiques} \\
\tilde{d}_{iray}(s) & = & \displaystyle\frac{2\rho_a c}{w_c \rho h} \frac{\displaystyle\sum_{m=1}^{3} b_m \left(\frac{s}{w_c}\right)^m}{\displaystyle\sum_{n=0}^{3} a_n \left(\frac{s}{w_c}\right)^n} & \text{pertes par rayonnement}
\end{array}\right. \tag{9.2}$$

Les paramètres des modèles sont définis dans le tableau 5.2 du chapitre 5.

9.2.2 Modèle des pertes dissipatives pour le verre

Le phénomène de pertes dissipatives dans le verre est dit de relaxation. Il représente dans quelle mesure le travail fourni au cours d'une déformation est stocké sous forme d'énergie mécanique reconvertible ou dissipée intrinsèquement. Les coefficients R_i et s_i du modèle dans l'équation 9.2 représentent la facilité du matériau à transmettre avec ou sans retard une énergie mécanique. Ce modèle de pertes pour le verre est dit proportionnel. Dans ces conditions le facteur d'amortissement prédit par le modèle sur les plaques en verre peut se calculer simplement en fonction des modes propres et des constantes R_i et s_i :

$$\alpha_{iv}(\omega_i) = \frac{\omega_i}{2}\mathcal{IM}\left(\sum_{i=1}^{2}\frac{j\omega R_i}{j\omega + s_i}\right) + \frac{R_f}{2} \tag{9.3}$$

La fonction $\alpha_{iv}(\omega_i)$ du modèle viscoélastique est quasi linéaire dans la bande de fréquences qui nous intéresse comme le montre la figure 9.1. Nous représentons donc la contribution des pertes viscoélastiques par la pente de la courbe simulée P_V.

9.2.3 Modèle des pertes dissipatives pour l'aluminium

Dans le cas de l'aluminium les termes de pertes dans l'expression des rigidités complexes sont dus à la thermoélasticité et au rayonnement :

$$\left\{\begin{array}{lcl} \tilde{d}_1(s) & = & \tilde{d}_{ithe}(s) + \tilde{d}_{ray}(s) \\ \tilde{d}_4(s) & = & \tilde{d}_{ray}(s) \end{array}\right. \tag{9.4}$$

Les pertes thermoélastiques résultent de couplages entre les ondes de flexion et les ondes de chaleur dans l'épaisseur de la plaque. Le coefficient R_1 de l'équation 9.2 de ce modèle thermoélastique correspond à la contribution de deux paramètres, un coefficient thermique correspondant à la sensibilité du matériau et un coefficient de conductivité thermique. Nous représenterons donc la contribution des pertes thermoélastiques par le paramètre R_1 dont la correspondance exacte avec les paramètres physiques est donnée dans l'équation 5.10 du chapitre 5.

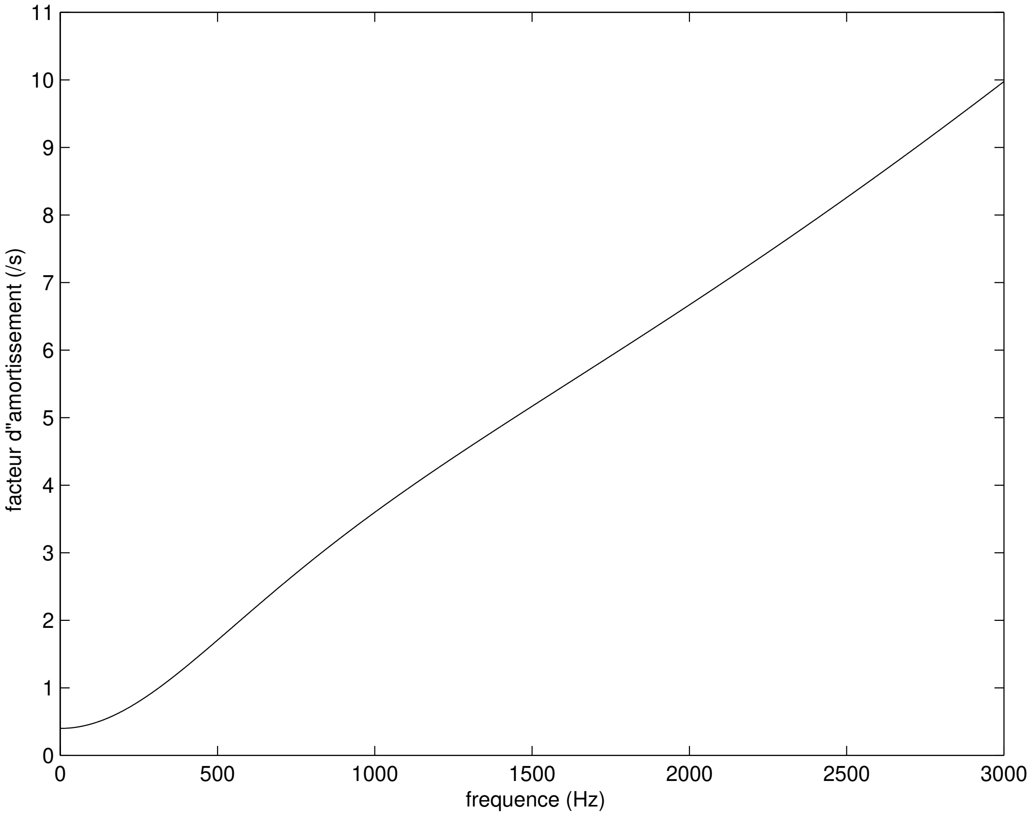

FIG. 9.1 – *Courbe représentant le facteur d'amortissement théorique calculé sur le modèle de pertes viscoélastiques en fonction de la fréquence.*

9.2.4 Paramètres du modèle de synthèse

Le modèle de plaque nous permet de contrôler séparément les contributions des différentes pertes énergétiques dont il est établi qu'elles sont indépendantes (voir chapitre 5). Pour en évaluer l'influence globale et séparée, nous avons simulé des matériaux hybrides en interpolant les deux modèles viscoélastiques et thermoélastiques caractéristiques des deux matériaux verre et aluminium. L'indépendance des deux modèles d'amortissement d'une part, et la caractérisation des matériaux verre et aluminium par ces deux modèles d'autre part, nous permettent d'étudier avec un même jeu de stimuli deux aspects essentiels de ce type d'objets vibrants : la perception de l'amortissement et la caractérisation des matériaux par celui-ci.

Contrôle des paramètres d'amortissement

Nous avons rajouté dans le programme un modèle de rigidité complexe prenant en compte simultanément les termes $\tilde{d}_{ithe}$ et $\tilde{d}_{ive}$ par interpolation. Ceci nous permet également d'explorer au-delà des contributions normales pour chacun des matériaux, en simulant ainsi un "ultra-verre" et un "ultra-aluminium". Les valeurs des constantes R_i et s_i des modèles de pertes interpolés sont celles issues du programme de synthèse et estimées sur deux plaques en verre et aluminium. Leurs valeurs sont données dans le tableau 9.2. Le terme $\tilde{d}_{iray}$ n'est pas pris en compte dans l'interpolation puisqu'il est commun aux deux matériaux.

Modèle de pertes	Valeurs des paramètres	
Thermoélastique	$R_1 = 0.0545$	$s_1 = 31000\ rad/s$
Viscoélastique	$R_1 = 1.625 \times 10^{-3}$	$s_1 = 5180.0\ rad/s$
	$R_2 = 1.962 \times 10^{-3}$	$s_2 = 55100.0\ rad/s$
Rayonnement	$a_0 = 1.1669$	$b_1 = 0.0620$
	$a_1 = 1.6574$	$b_2 = 0.5950$
	$a_2 = 1.5528$	$b_3 = 1.0272$
	$a_3 = 1.0$	$\rho_a = 1.2\ kg/m^3$
	$c_a = 344\ m/s$	

TAB. 9.2 – *Tableau des paramètres utilisés dans les modèles de pertes, les paramètres sont ceux des équations 9.2.*

Le tableau 9.3 donne la fonction complète du modèle hybride d'amortissement dans lequel le paramètre H est le facteur d'interpolation. Les sons simulés vont au-delà des valeurs limites caractéristiques de l'aluminium ou du verre. Ceci est exprimé par les valeurs de $H > 1$ ou < 0 .

Pour réduire le nombre de paramètres mécaniques caractéristiques des contributions respectives des pertes viscoélastiques et thermoélastiques nous calculons deux nouveaux paramètres issus de la fonction d'interpolation :

- nous avons vu que la fonction $\alpha_{iv}(\omega_i)$ peut être approximée dans la gamme des basses fréquences par la pente P_V de la courbe simulée. La contribution des pertes viscoélastiques dans la fonction d'interpolation sera approximée par le paramètre P_V, produit de H et de la pente P_V.

Valeurs de H	modèle de rigidité complexe	matériau simulé
$H < 0$	$\tilde{d} = (1 - H) * \tilde{d}_{ive}$	ultraverre
$H = 0$	$\tilde{d} = \tilde{d}_{ive}$	verre
$0 < H < 1$	$\tilde{d} = H * \tilde{d}_{ithe} + (1 - H) * \tilde{d}_{ve}$	hybride
$H = 1$	$\tilde{d} = \tilde{d}_{ihe}$	aluminium
$H > 1$	$\tilde{d} = H * \tilde{d}_{the}$	ultraaluminium

TAB. 9.3 – *Détail de la fonction d'interpolation ainsi que des fonctions d'extrapolation utilisées pour construire les matériaux hybrides. Il est à noter que* $H = 0$ *correspond au verre et* $H = 1$ *correspond à l'aluminium.*

– de même le modèle de pertes thermoélastiques peut être caractérisé par le paramètre R_1 défini dans l'équation 9.2. Nous calculerons comme paramètre mécanique caractéristique des pertes thermoélastiques dans nos stimuli le paramètre HR_1, pondération du paramètre R_1 par le facteur d'interpolation H.

Contrôle des paramètres élastiques

Les constantes de rigidités D_1 et D_4 correspondent respectivement aux modes de flexion et de torsion, comme nous l'avons rappellé dans les équations 5.6 d'une plaque isotrope infinie. Les valeurs correspondantes pour les plaques en aluminium ou en verre ont été estimées par les approximations classiques les reliant aux fréquences des premiers modes qui ont été mesurées. Inversement, si on modifie leurs valeurs de façon continue, on obtiendra une variation continue des spectres. Nous avons choisi de faire varier les constantes de rigidités D_i autour des valeurs correspondant aux deux matériaux verre et aluminium afin d'obtenir comme pour l'amortissement des matériaux hybrides mais du point de vue élastique. Nous reportons dans les tableaux récapitulatifs B de l'annexe B les valeurs des quantités $c_i = \sqrt{E_i/\rho}$ variées linéairement. Les c_i sont variés de la même façon, on pourra donc les représenter dans les analyses indifféremment par c_1 ou c_2. L'analyse spectrale des stimuli ainsi calculés confirme du reste une relation parfaite (corrélation de 1) entre les fréquences des premiers modes et les valeurs $c_i = \sqrt{E_i/\rho}$. Les sons de plaques, à l'inverse des sons de barres, sont beaucoup plus riches spectralement et fortement inharmoniques. Le choix de la variation linéaire n'est donc pas basé sur la hauteur perçue des sons mais uniquement sur une répartition linéaire des premiers modes. Les valeurs, comme pour l'amortissement, sont réparties autour des valeurs correspondant aux deux matériaux verre et aluminium reportées dans le tableau 9.4. Les 16 stimuli sont programmés sur 16 valeurs de H et 16 valeurs des paramètres élastiques. L'ensemble des stimuli est représenté sur la figure 9.2.

Verre	Aluminium
$D_1 = 6700 \times 10^6\ MPa$ $D_4 = 10270 \times 10^6\ MPa$ $\rho = 2550\ kg/m^3$	$D_1 = 6500 \times 10^6\ MPa$ $D_4 = 8600 \times 10^6\ MPa$ $\rho = 2660\ kg/m^3$
amortissement fluide $R_f = 0.4\ /s$	

TAB. 9.4 – *Paramètres élastiques introduits dans les modèles pour les plaques en verre et en aluminium, le terme constant d'amortissement fluide R_f est pris identique pour les deux matériaux.*

Choix du maillet

Pour l'étude sur les plaques nous avons voulu explorer une autre possibilité du programme de synthèse qui est le choix du matériau du maillet. Les sons hybrides présentés précédemment ont donc été synthétisés en deux groupes : un jeu de plaques frappées par un maillet en caoutchouc et un jeu de plaques frappées par un maillet en bois. Le choix de ces matériaux pour l'excitateur parmi ceux proposés dans le programme vient de l'écart perceptif entre les deux excitations, qui est maximal pour ces deux matériaux. Les constantes introduites dans les deux modèles d'excitation sont répertoriées dans le tableau 9.5 ainsi que les paramètres géométriques des plaques simulées. On peut remarquer que les deux modèles d'excitateurs sont séparés par la valeur de la constante de raideur des maillets K. Les tâches proposées aux sujets seront donc systématiquement doublées pour juger chaque groupe de sons selon l'excitateur.

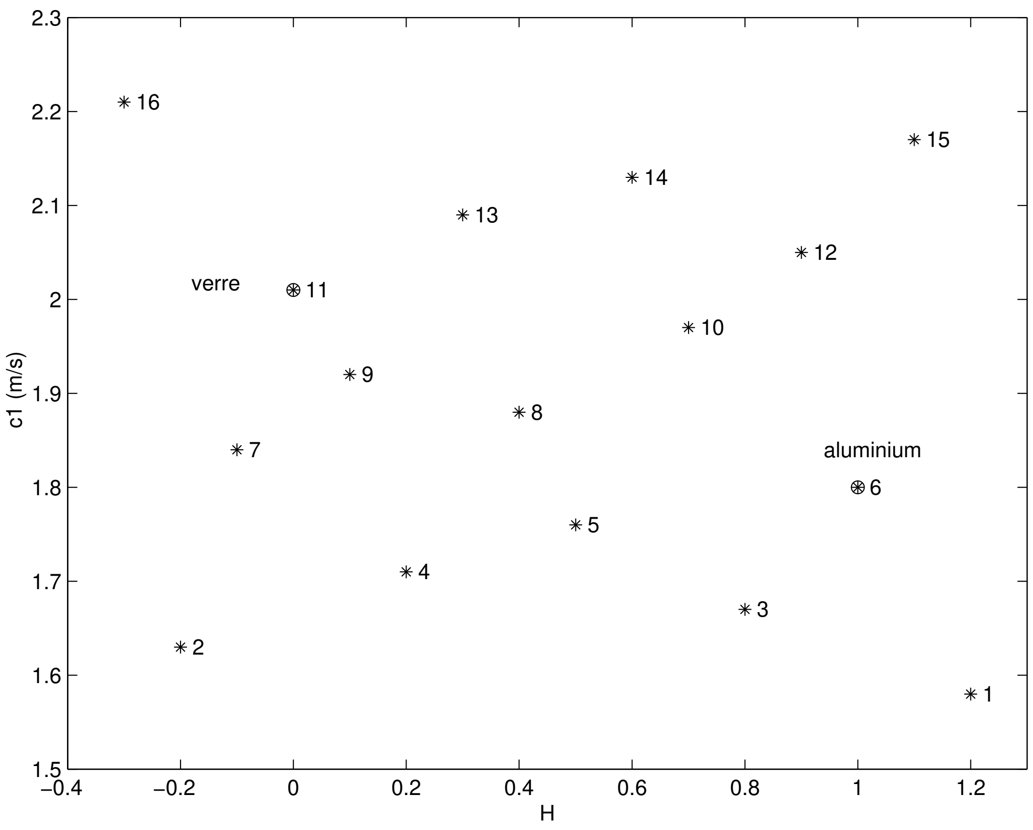

FIG. 9.2 – *Représentation spatiale des stimuli dans le plan amortissement-élasticité. L'amortissement est représenté par la valeur H du coefficient d'interpolation et l'elasticité est représenté par c_1. Les stimuli 6 et 11 correspondent aux valeurs exactes de la synthèse respectivement pour l'aluminium et le verre.*

Paramètres du maillet			Géométrie de la plaque
	bois	caoutchouc	
$m\ (kg)$	3×10^{-3}	3×10^{-3}	largeur $= 22\ cm$
$K\ (m/N^{-2/3})$	42.0×10^{7}	1.2×10^{7}	longueur $= 23\ cm$
$V\ (m/s)$	1.0	1.0	épaisseur $= 2\ cm$

TAB. 9.5 – *Valeurs des constantes introduites dans le modèle de synthèse au chapitre 5 du maillet en bois d'une part, et du maillet en caoutchouc, d'autre part.*

9.2.5 Analyse spectrotemporelle

Pour cette partie nous devons tout d'abord identifier quels sont les paramètres analytiques pertinents pour caractériser nos signaux. Les plaques, contrairement aux barres, ont un spectre très riche et un modèle d'amortissement complexe. Nous allons donc redéfinir les paramètres analytiques pour cette étude. De plus nous observerons les différences des résultats pour les deux séries bois et caoutchouc. L'allure spectrale des stimuli de référence verre et aluminium est représentée figure 9.3 pour les deux maillets bois et caoutchouc.

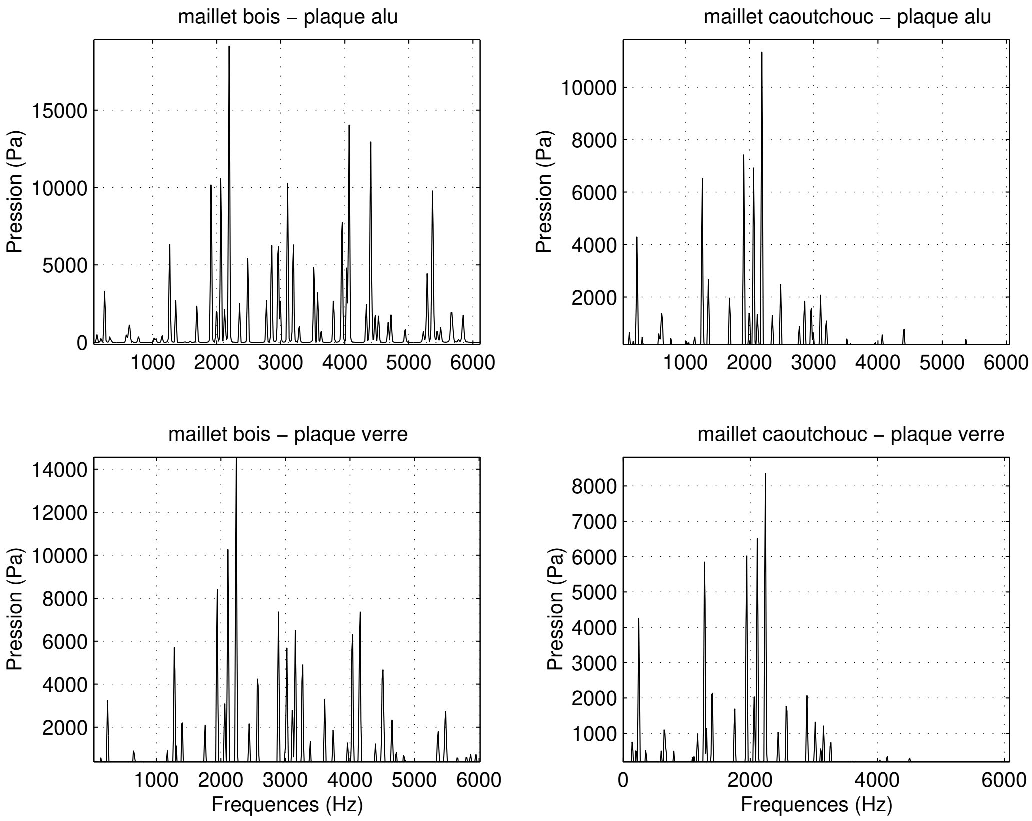

FIG. 9.3 – *Représentation spectrale des plaques* 6 *et* 11 *caractéristiques de l'aluminium et du verre, frappées par des maillets en bois et en caoutchouc. Les niveaux sont alignés sur les maximums tels qu'ils sont présentés aux sujets. Les spectres ont été calculés sur une fenêtre temporelle de* 22 *ms à partir de* 1 *ms du début du son* ($N = 1024\,points,\ f_e = 48kHz$).

Caractéristique de l'amortissement

- Facteur d'amortissement

  Pour la caractérisation de l'amortissement nous avons calculé les valeurs du facteur d'amortissement α représentant la décroissance temporelle de l'enveloppe. Toutefois les corrélations avec les paramètres physiques d'une part, et les résultats expérimentaux d'autre part, ont montré que même si les corrélations étaient satisfaisantes ($r^2 = 0.8$), un autre paramètre est beaucoup plus représentatif en terme de corrélation. Il s'agit de la pente des courbes amortissement-fréquence calculées pour chaque son par la méthode ESPRIT [45]. Ainsi nous remplaçons le paramètre α calculé sur l'enveloppe par un paramètre plus fin calculé sur les facteurs d'amortissement pour chaque fréquence α_i. Ainsi les sons hybrides construits peuvent être caractérisés par la pente de leurs courbes d'amortissement assimilables à des fonctions affinées dans la bande de fréquences qui nous intéresse ($\leq 5000Hz$).

- Facteur de perturbations *std*

  Le modèle d'amortissement thermoélastique induit une courbe $\alpha_i(f)$ beaucoup plus perturbée dans les très basses fréquences ($\leq\ 500Hz$. La figure 9.4 montre à titre d'exemple cinq sons représentatifs de la gamme d'amortissement variée et témoigne bien des différences de régularités de pentes entre les stimuli les plus amortis ($H \leq 0$) et les moins amortis ($H \geq 1$) dans la région spectrale jusqu'à 5 kHz. Afin de s'assurer de cette spécificité du modèle thermoélastique nous avons calculé la régression linéaire sur chaque courbe $\alpha_i(f)$ pour la bande [0, 500 Hz] où les perturbations sont significatives (fig 9.4). Cette autre caractéristique du modèle thermoélastique caractérise l'aluminium tandis que l'amortissement viscoélastique prépondérant dans le verre induit des courbes beaucoup plus régulières. Cette affinement du modèle a permis un bien meilleur rendu acoustique des matériaux. Le paramètre *std* désigne l'erreur calculée sur cette régression et qui rend compte des perturbations autour de la droite de régression indépendament de la pente. L'erreur est calculée par la fonction Matlab *polyval* à partir des paramètres issus du calcul de regression calculés par la fonction *polyfit*.

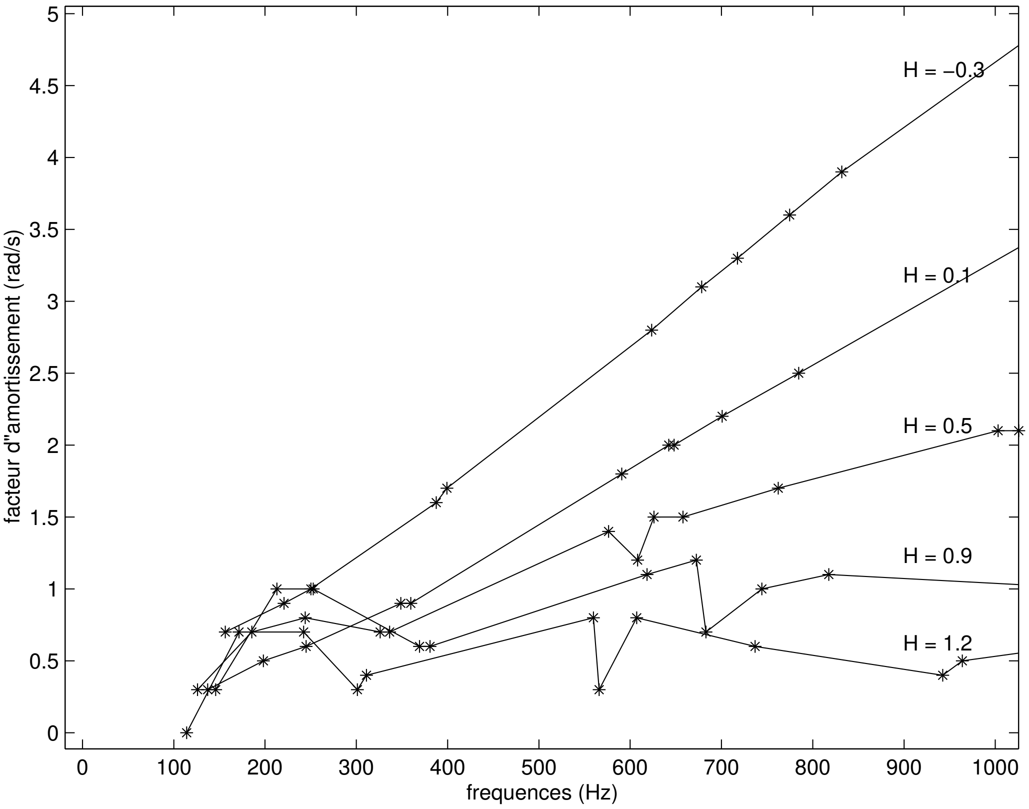

FIG. 9.4 – *Représentation des facteurs d'amortissement calculés pour chaque mode et pour 5 stimuli du moins amorti (le bas de la figure) jusqu'au plus amorti (le haut de la figure). Les courbes sont annotées du facteur d'interpolation H introduit dans le modèle hybride d'amortissement.*

La pente des fonctions affines ajustées sur les courbes, que nous désignerons par P_α est néanmoins suffisante pour décrire l'ensemble des stimuli et sera la caractéristique principale des signaux pour l'amortissement.

Calcul de paramètres liés à la hauteur

Les sons de plaques sont inharmoniques et ils présentent un spectre riche en fréquences. L'estimation de la hauteur est donc plus complexe que pour les barres. Nous calculons ici trois fréquences pouvant éventuellement décrire une hauteur spectrale perçue par les sujets pendant l'expérience. Il s'agit de :

- f_1, la fréquence la plus basse du spectre,
- f_{per} la fréquence calculée sur la fonction d'autocorrélation. La fonction d'auto-corrélation correspond à la transformée de Fourier inverse de la fonction puissance ; la périodicité de cette fonction correspond au lieu le plus élevé de l'enveloppe spectrale,
- et f_{max}, la fréquence dont le niveau est le plus élevé correspondant au pic spectral.

Remarques sur l'influence des maillets

Les courbes amortissement-fréquences sont rigoureusement identiques pour les plaques frappées avec un maillet en bois et celles frappées avec un maillet en caoutchouc dans la gamme $[0 - 4000\ Hz]$. Au dessus nous observons la présence de fréquences à un niveau significativement élevé pour les plaques frappées avec un maillet en bois. Les fréquences élevées des plaques excitées avec un maillet en caoutchouc ont un niveau plus faible et sont très amorties et donc moins saillantes dans l'attaque. Les stimuli étant égalisés en sonie pour les expériences, nous ne tiendrons pas compte des différences de niveaux. En résumé, le maillet en caoutchouc revient à un filtre passe-bas avec une fréquence de coupure aux environs de 4000 Hz (voir la figure 9.5).

Protocole expérimental

Le groupe expérimental est constitué de vingt deux sujets, âgés de 20 à 32 ans (Moyenne = 24) et provenant tous de la région parisienne. Le groupe ne comporte aucun musicien professionnel et est composé de douze femmes et dix hommes. Les sujets sont indemnisés pour leur participation à l'expérience. Aucun n'a indiqué avoir un problème d'audition. La totalité de l'expérience dure $1h45$ en moyenne et se déroule dans une cabine audiométrique. L'ensemble du matériel est décrit au chapitre 6.

Deux expériences sont proposées aux sujets. La première comporte une tâche de jugement de dissemblance et la seconde une tâche d'identification.

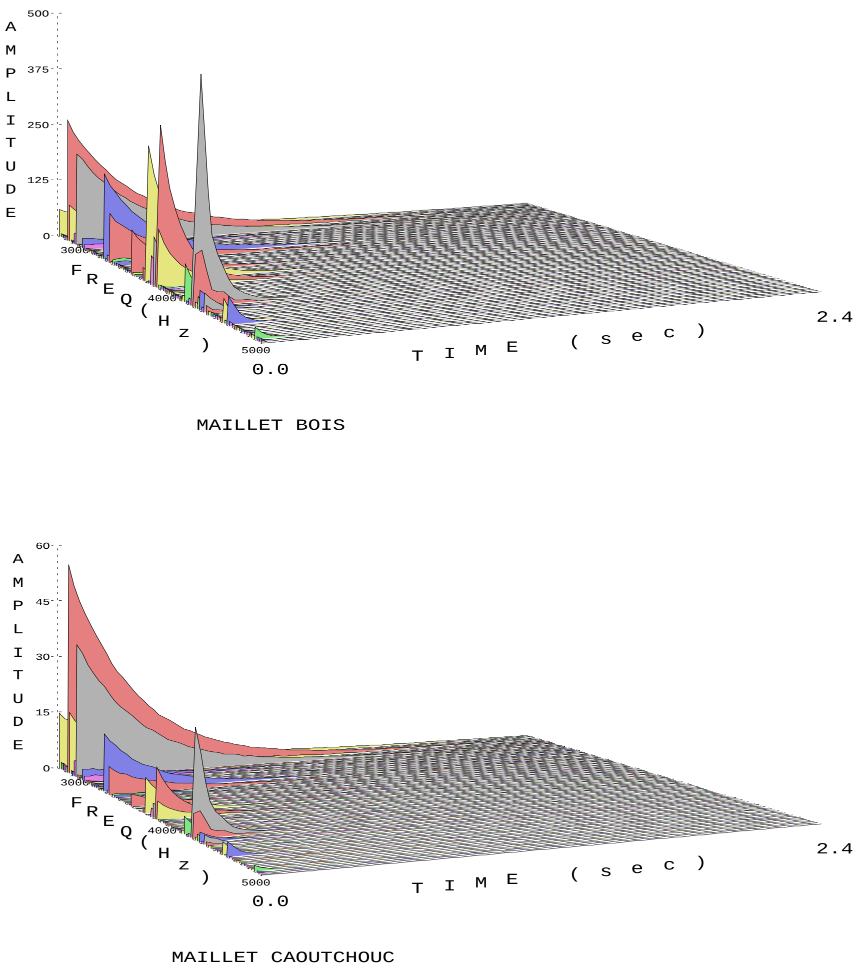

FIG. 9.5 – *Visualisation 3D de la gamme de fréquences [*$3000 - 5000\ Hz$*]. Les modes excités par les plaques frappées avec le maillet dur (bois) sont moins amortis que ceux excités par le maillet caoutchouc. Les niveaux sont ajustés sur le maximum d'amplitude de la gamme de fréquence, les stimuli étant égalisés eux-mêmes en sonie avant l'expérience.*

Les sujets ont toujours passé l'expérience de dissemblance avant celle d'identification afin de ne pas créer un éventuel biais cognitif dans l'exécution de la première.

9.3 Expérience 3 : jugements de dissemblance

9.3.1 Présentation

Les sujets doivent évaluer la dissemblance relative entre les 136 paires possibles des 16 stimuli proposés pour chaque maillet y compris les paires de stimuli identiques, soit 272 jugements au total. Les sons correspondant aux deux maillets étaient présentés dans des blocs séparés. La moitié des auditeurs ont commencé par le maillet bois et l'autre moitié par le maillet caoutchouc. Nous donnons maintenant les résultats de l'analyse multidimensionnelle calculés avec l'analyseur CLASCAL [84] décrit aux chapitres 6 et 4.1.

9.3.2 Résultats

Choix du nombre de classes

Le nombre de classes est déterminé par une série de tests Monte Carlo sur le modèle nul. Il choisit trois classes pour les deux séries de sons (maillets en caoutchouc et en bois) [1].

Choix des modèles spatiaux

Nous testons maintenant le modèle de une à quatre dimensions, avec et sans spécificités, avec trois classes latentes pour chaque série de stimuli.

Les deux meilleurs modèles spatiaux choisis selon l'indice BIC (voir tableau 9.6) nous donnent 2Dsp puis 3Dns pour la série bois et 3Dns puis 2Dns pour la série caoutchouc.

Test Monte Carlo sur le modèle spatial

Les tests Monte Carlo opposant les modèles rejettent le modèle 3Dns contre le modèle 2Dsp pour la série maillet bois et rejettent le modèle 2Dns contre le modèle 3Dns pour la série maillet caoutchouc. Ces deux modèles sont donc retenus.

1. La détermination du nombre de classes pour le caoutchouc a nécéssité de faire deux fois la procédure d'analyse, puisque le nombre de classes estimé sur le modèle nul était 2 alors que celui sur le modèle spatial choisi par ensuite était 3.

		Maillet bois		Maillet caoutchouc		
	N	LV	BIC	LV	BIC	dl
Modèle sans spécificités	1	-11671	23501	-12063	24283	20
	2	-11071	22434	-10991	**22273**	37
	3	-10757	**21939**	-10899	**22223**	54
	4	-10789	22138	-11004	22568	71
	N	LV	BIC	LV	BIC	dl
Modèle avec spécificités	1	-11064	22429	-11485	23269	38
	2	-10720	**21874**	-11129	22691	55
	3	-10696	21961	-10990	22548	72
	4	-10752	22205	-10954	22609	89
Modèle nul		-9920	22700	-10527	23914	89

TAB. 9.6 – *Test BIC pour les modèles à trois classes.*

Les classes

Le programme confirme le nombre de classes pour chaque série sur le modèle spatial choisi. Les modèles définitifs sont : **un modèle à deux dimensions avec spécificités et trois classes pour la série maillet bois et un modèle à trois dimensions sans spécificités et trois classes pour la série maillet caoutchouc**.
Nous donnons dans le tableau 9.7 le poids relatif de chaque classe pour chacune des dimensions y compris pour les spécificités du modèle 2Dsp.

	Maillet bois			Maillet caoutchouc		
Classe	dim_1	dim_2	spécificités	dim_1	dim_2	dim_3
1	0.60	1.57	1.99	1.57	0.20	0.03
2	0.93	0.86	0.17	0.65	2.00	2.25
3	1.47	0.57	0.84	0.78	0.80	0.72

TAB. 9.7 – *Répartition du poids des classes sur chaque dimension.*

La répartition complète des sujets pour chaque classe est donnée dans l'annexe B. Le tableau 9.8 montre que la répartition des sujets se fait différemment pour les deux séries de sons. En effet, au mieux, il y a seulement cinq sujets en commun entre les classes.

Classes	B_1	B_2	B_3
C_1	1	2	3
C_2	2	5	2
C_3	3	0	2

TAB. 9.8 – *Nombre de sujets en commun entre les classes. L'intersection B_i et C_j désigne le nombre de sujets en commun entre la classe i de la série bois et la classe j de la série caoutchouc.*

La figure 9.6 représente le poids de chaque classe pour chacune des dimensions 1 et 2 dans les espaces correspondant aux deux maillets. Cette figure montre une bonne correspondance en terme de poids à travers les espaces. En revanche il n'y a pas de correpondance en terme de sujets puisque les classes correspondantes $C_1 - B_3$, $C_2 - B_1$ et $C_3 - B_2$ ont respectivement trois, deux et zéro sujets en commun. Il semblerait donc que les mêmes sujets n'utilisent pas les mêmes stratégies de réponse pour les deux séries de sons.

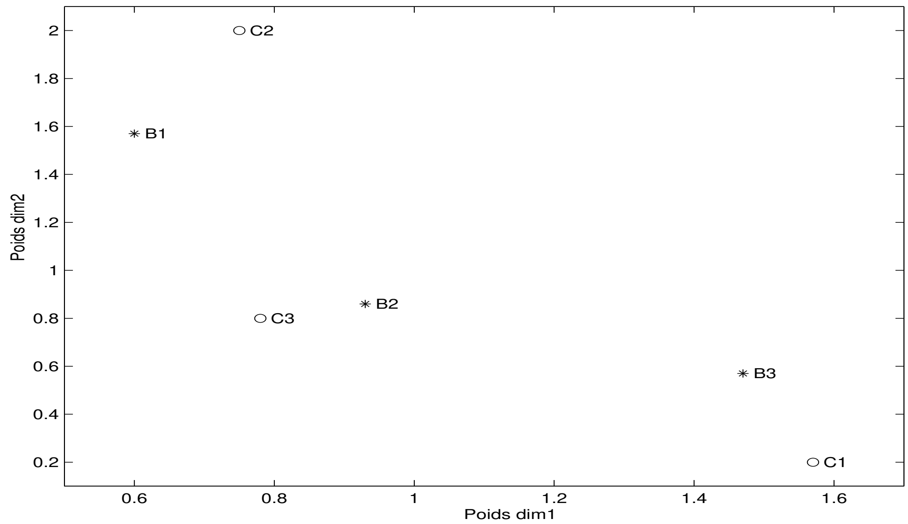

FIG. 9.6 – *Espace des poids pour chaque classe sur les dimensions 1 et 2 (C=caoutchouc, B=bois).*

Les espaces perceptifs

La figure 9.7 décrit l'espace perceptif des deux premières dimensions rendues par l'analyse MDS pour les deux maillets. A première vue, en comparaison avec la figure 2, on peut remarquer la préservation approximative de l'espace physique avec toutefois plus de perturbations dans l'ordre des sons sur la dimension liée à l'élasticité. L'ensemble des paramètres mécaniques et analytiques des sons est répertorié dans l'annexe B. Nous procéderons ci-dessous à l'analyse des corrélations entre ces paramètres.

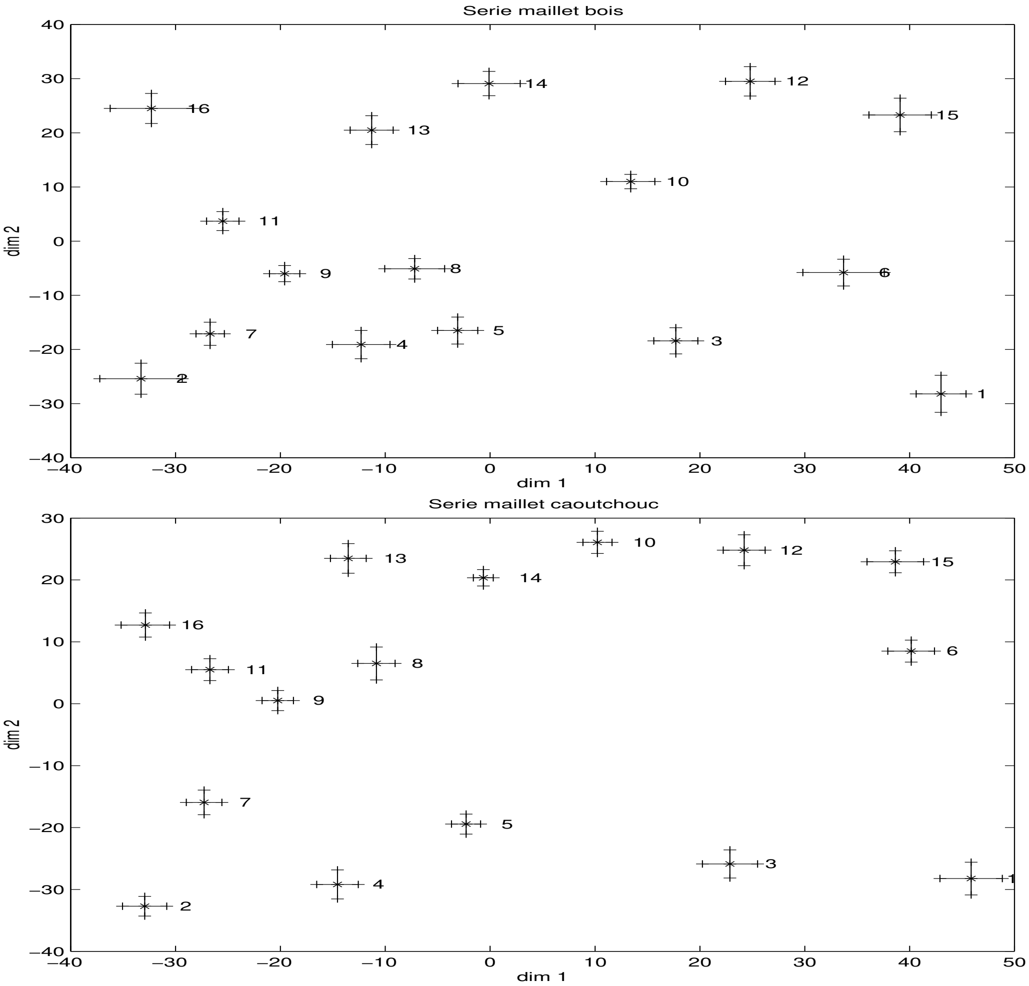

FIG. 9.7 – *Espaces perceptifs correspondant aux deux premières dimensions rendues par l'analyse MDS pour les deux séries de sons (maillet bois en haut, maillet caoutchouc en bas). Les écarts-types ont été calculés par la procédure bootstrap sur 250 tirages (la répartition des points est dans l'annexe A, figures A.1 et A.2).*

9.3.3 Corrélations

Dimension 1

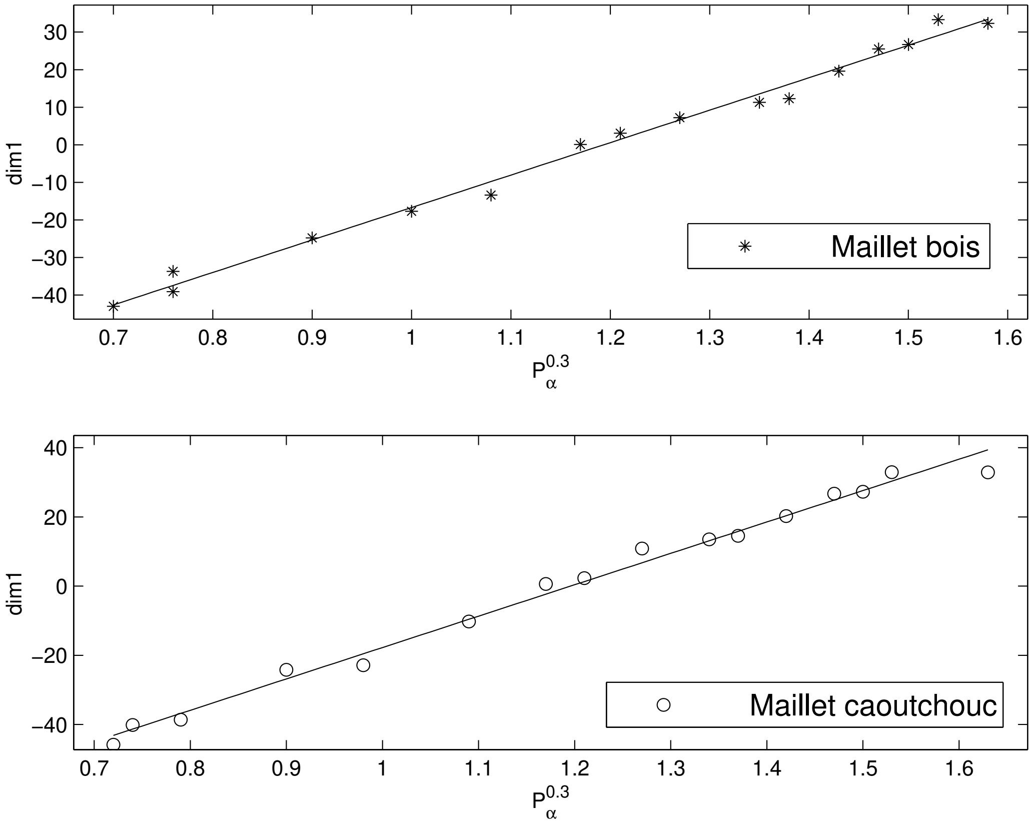

FIG. 9.8 – *Corrélation entre $P_\alpha^{0.3}$, paramètre analytique représentant la pente de la courbe $\alpha_i(f)$ et la dimension perceptive dim1.*

La dimension 1 représente la sensibilité de l'auditeur à l'amortissement, ou tout du moins aux effets sensibles de celui-ci sur le timbre des stimuli. Le tableau 9.9 montre de très fortes inter-corrélations entre les paramètres.

Nous constatons en premier lieu que toutes les corrélations sont très élevées ($\geq$ 0.9). L'ensemble des paramètres est ainsi covarié. Il n'est pas possible de faire une régression multiple en l'absence de paramètres indépendants, au moins partiellement ($r < 0.80$). Nous ne pouvons pas non plus choisir un paramètre plus qu'un autre pour décrire la sensibilité des auditeurs du fait

dim1	maillet bois					
	H	HR_1	$\sqrt{P_V}$	$P_\alpha^{0.3}$	std	$CGS_{(\mathrm{Hz})}$
H	1.00					
HR_1	0.98	1.00				
$\sqrt{P_V}$	-0.97	-0.98	1.00			
$P_\alpha^{0.3}$	-0.99	-0.99	0.97	1.00		
std	0.95	0.97	-0.96	-0.96	1.00	
CGS	0.94	0.93	-0.94	-0.94	0.91	1.00
$dim1$	-0.99	-0.99	0.98	1.00	-0.95	-0.94
	maillet caoutchouc					
H	1.00					
HR_1	0.98	1.00				
$\sqrt{P_V}$	-0.97	-0.98	1.00			
$P_\alpha^{0.3}$	-0.99	-0.98	0.98	1.00		
std	0.95	0.96	-0.95	-0.95	1.00	
CGS	0.94	0.92	-0.93	-0.91	0.88	1.00
$dim1$	-0.99	-0.98	0.97	0.99	-0.94	-0.90

TAB. 9.9 – *Tableau des corrélations entre les paramètres mécaniques* (HR_1, $\sqrt{Pv}$)*, analytiques* (P_α, *std*, CGS) *et la dimension perceptive* ($dim1$).

de leur covariation presque parfaite pour les deux ensembles de stimuli.

En ce qui concerne la variation des paramètres mécaniques, nous pouvons donc dire que seul le facteur d'interpolation H (voir figure 9.9) rend compte de la variation perceptive des sujets, les paramètres dérivés spécifiques aux deux modèles de pertes HR_1 et P_V n'apportant aucune information supplémentaire. On notera que la meilleur corrélation pour le paramètre P_V est obtenue avec un exposant 1/2. Les intercorrélations nous empèchent de pouvoir mesurer séparément l'influence des deux modèles de pertes. Il faudrait pour cela des stimuli pour lesquels les paramètres analytiques correspondants P_α et std sont décorrélés.

La meilleur corrélation avec la dimension perceptive $dim1$ est $r = 1.00$ pour le bois et $r = 0.99$ pour le caoutchouc avec le paramètre P_α, ce qui confirme que ce paramètre est pertinent pour décrire la dimension correpondant au timbre de nos stimuli (voir la figure 9.8 qui décrit la corrélation

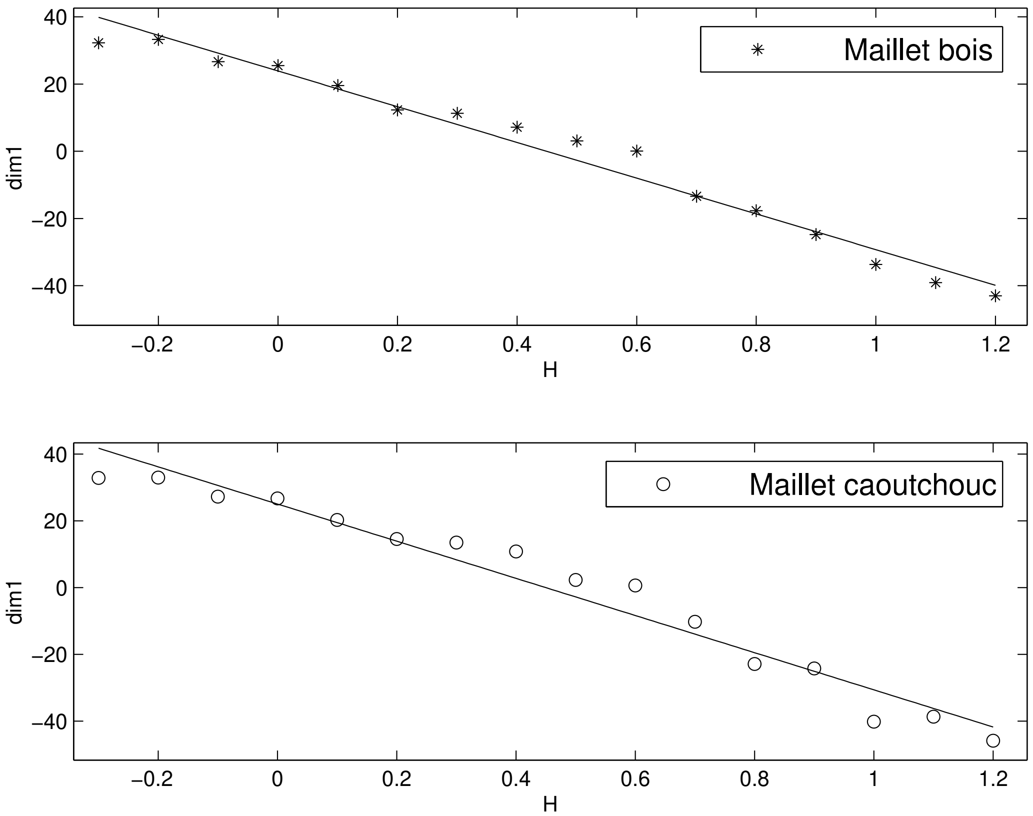

FIG. 9.9 – *Corrélation entre H, coefficient d'interpolation du modèle de perte, et la dimension perceptive dim1.*

$dim1 - P_\alpha$) par rapport, par exemple, au paramètre α calculé directement sur l'enveloppe temporelle. Sur ce point nous pouvons constater : premièrement, la caractérisation des sons par P_α vient de la démarche calculatoire du modèle de synthèse ; en effet c'est en mesurant les courbes d'amortissement des plaques que les paramètres constants du modèle d'amortissement ont été calculés ; deuxièmement, la corrélation de 0.99 a été obtenue avec $P_\alpha^{\mathbf{0.3}}$. Le coefficient de puissance retenu, 0.3, est le même que pour le paramètre d'amortissement du modèle de barre η (voir chapitres 7 et 8), confirmant cette valeur dans l'étude de la perception de l'amortissement sur deux types d'objets physiques, barres et plaques.

dim2	maillet bois				
	$\sqrt{\frac{E_1}{\rho}}$	$\sqrt{\frac{E_2}{\rho}}$	f_1	f_{per}	f_{max}
$\sqrt{\frac{E_1}{\rho}}$	1.00				
$\sqrt{\frac{E_2}{\rho}}$	1.00	1.00			
f_1	1.00	1.00	1.00		
f_{per}	0.97	0.96	0.96	1.00	
f_{max}	0.98	0.98	0.98	0.96	1.00
$dim2$	0.95	0.95	0.95	0.92	0.97
	maillet caoutchouc				
$\sqrt{\frac{E_1}{\rho}}$	1.00				
$\sqrt{\frac{E_2}{\rho}}$	1.00	1.00			
f_1	1.00	1.00	1.00		
f_{per}	0.95	0.96	0.96	1.00	
f_{max}	0.98	0.98	0.98	0.99	1.00
$dim2$	0.87	0.88	0.88	0.92	0.93

TAB. 9.10 – *Tableau des corrélations entre les paramètres mécaniques* ($\sqrt{(E_i/\rho)}$)*, analytiques* (f_1, f_{per}, f_{max}) *et la dimension perceptive* $dim2$*.*

Dimension 2

Le tableau 9.10 récapitule les corrélations pour la dimension 2. En premier lieu, on voit que la dimension 2 de l'analyse MDS est indépendante des paramètres des modèles d'amortissement : les corrélations entre la dimension 2 et les paramètres HR_1, p_V, $p_\alpha^{0.3}$, sdt et CGS sont toutes inférieures à 0.2.

La dimension 2 est en revanche corrélée avec les paramètres élastiques variés c_i qui influent directement sur le spectre du signal. Nous pouvons donc nous attendre, là encore, à ce que la dimension 2 corresponde à la hauteur perçue par les sujets. Toutefois le contenu spectral des plaques est cette fois-ci très riche et la simple écoute des stimuli montre qu'il est difficile d'évaluer immédiatement une hauteur perçue. Aussi a-t-on calculé trois paramètres analytiques sur les signaux pouvant correspondre à la hauteur perçue et qui sont, rappelons-le, la fréquence la plus basse du spectre f_1, la fréquence calculée sur la fonction d'auto-corrélation f_{per} et le pic spectral f_{max}.

A ce stade nous observons les corrélations entre les différents paramètres et nous constatons tout d'abord une parfaite corrélation entre les paramètres mécaniques et la fréquence f_1, conséquence directe du modèle mécanique, et qui est donc prévisible. Lorsque nous examinons maintenant les corrélations entre le paramètre perceptif $dim2$ et les trois fréquences calculées sur le signal, il apparaît que toutes sont corrélées fortement avec des valeurs du coefficient de corrélation supérieures à 0.9. Bien que, pour le bois comme pour le caoutchouc, la fréquence f_{max} ait la plus forte corrélation avec $dim2$, les facteurs de corrélation sont trop proches pour pouvoir conclure sur la prédominance d'un paramètre sur les autres.

Comment savoir alors quel est le paramètre suivi par les auditeurs pour décrire $dim2$? Il faut observer à nouveau les trois fréquences calculées sur le signal ; en particulier, calculer la gamme de fréquence sur laquelle elles sont présentes. La gamme pour f_1 et f_{per} s'étend sur une quinte (rapport de fréquence 1.5) tandis que $\triangle f_{max}$ contient à peine plus d'un demi-ton, (rapport de fréquence 1.13).

Un auditeur supplémentaire, musicien amateur ayant l'oreille absolue, a évalué la hauteur des stimuli directement par sinus glissant en ajustant un son de sinusoïde sur la hauteur perçue pour chaque stimulus. L'ajustement était réalisable au centième de ton, donc supérieur à la sensibilité de l'auditeur, afin qu'il puisse juger de façon continue. Nous constatons alors deux choses. Premièrement, l'auditeur est sensible à une hauteur spécifique et située dans la gamme de fréquence de la première fréquence basse de niveau élevé. Deuxièmement, l'auditeur perçoit un écart maximal de la hauteur entre les stimuli les plus éloignés de 13.2 % de la fréquence soit à peine plus d'un demi-ton. La hauteur perçue par cet auditeur est biaisée par la variation de l'amortissement, induisant chez lui une sensation de hauteur plus élevée pour les sons proches de l'aluminium dont le niveau des harmoniques aigus est plus élevé. Nous avons calculé la pente entre la hauteur spectrale rendue par le sujet et les différentes fréquences estimées sur les signaux rapportées à la gamme de fréquence contenant la hauteur spectrale. De cette façon nous observons que les pentes les plus proches de 1 sont observées entre les pics spectrals calculés pour les deux séries de sons et la hauteur spectrale. Les valeurs des pentes sont reportées dans le tableau 9.11.

	Maillet bois		Maillet caoutchouc	
	pente	r	pente	r
HS et f_1	0.75	0.93	0.70	0.93
HS et f_{per}	0.31	0.91	0.65	0.94
HS et f_{max}	0.88	0.93	0.82	0.97

TAB. 9.11 – *Pente des droites de régression de la hauteur spectrale perçue (HS) par le sujet sur chacune des fréquences calculées et coefficient r de régression correpondant pour les deux séries, bois et caoutchouc.*

A priori nous pouvons donc supposer que les sujets sont plutôt sensibles au pic spectral sans toutefois qu'il y ait une relation parfaite. La richesse spectrale, l'amortissement variable entre les stimuli et la perception de plusieurs hauteurs possibles sont autant d'éléments pouvant biaiser la perception des sujets et donc l'interprétation de la hauteur perçue.

Comparaison des résultats par séries

L'analyse comparée des résultats obtenus avec les séries bois et caoutchouc met en évidence les différences liées à l'excitateur résumés à une perte d'information spectrale à peine sensible dans l'erreur sur les résultats des sujets. La capacité des auditeurs à se concentrer sur la tâche en faisant abstraction de l'excitateur est donc soulignée ici. De plus l'analyse spectrale a montré que si les différences à l'écoute sont très marquées pour les deux excitateurs, les différences spectrales sont beaucoup plus fines. Il s'agit en effet d'un changement de l'amortissement sur une gamme étroite de fréquences et accompagné d'un apauvrissement du spectre en hautes fréquences pour le caoutchouc. L'analyse MDS a montré des résultats équivalents : moins de 2% d'écart sur les écarts-types et des corrélations supérieures à 0.9 avec les paramètres analytiques et physiques pour les deux séries.
Si l'on observe maintenant les corrélations entre les dimensions 1 et 2 estimées avec chacune des deux séries de sons (maillets bois et caoutchouc) on s'apercoit d'une très forte inter-corrélation des deux espaces. Ceci pour les

deux espaces estimés pour le caoutchouc, 2Dns et 3Dns. Ces résultats reportés dans le tableau 9.12 confirment la très forte dépendance des deux espaces et la très bonne corrélation pour les deux espaces possibles du caoutchouc.

facteur de corrélation r	bois 2Dsp et caou 3Dns	bois 2Dsp et caou 2Dns
$dim1$	0.99	0.99
$dim1$	0.92	0.98

TAB. 9.12 – *Corrélation entre les espaces croisés estimés sur les deux séries (maillet bois et caoutchouc).*

Remarques

Nous avons volontairement écarté la dimension 3 de l'espace perceptif de la série caoutchouc qui ne correspond à aucun critère perceptif et n'est corrélée avec aucun paramètre analytique où mécanique. De plus l'espace à deux dimensions trouvé comme deuxième score de l'indice BIC donne des résultats plus proches de l'espace bois pour la dimension 2.
En ce qui concerne les spécificités obtenues dans l'analyse de la série bois, nous n'avons pas d'explications. Les fortes spécificités (stimuli 4,6 et 15) ne nous apparaissent pas fondamentalement différentes des autres stimuli. Ces aspects mériteraient une étude supplémentaire approfondie.

9.3.4 Conclusion

Les fréquences f_1, f_{ter} et f_{max} sont très bien corrélées avec la dimension 2. La meilleure corrélation est obtenue pour f_{max} avec $r = 0.98$. Néanmoins la hauteur spectrale est estimée sur une bande plus étroite, la pente obtenue entre HS et f_{max} étant de $O.88$. On peut donc difficilement trouver sur quelle fréquence les sujets ont fait leurs jugements. Les corrélations très faibles avec les autres paramètres spectraux type CGS n'ont pas permis d'expliquer un éventuel biais spectral pouvant induire une confusion dans les réponses des sujets. Les sujets ont donc bien montré leur sensibilité à la hauteur et sur une gamme de fréquence très étroite, mais parfaitement lisible sur la dimension 2.

Le paramètre analytique P_α s'est révélé un paramètre très fort pour caractériser l'attribut du timbre correspondant à la perception de l'amortissement. Nous avons là un précieux paramètre de mesure pour prédire une dimension perceptive issue de l'écoute de plaques frappées avec deux maillets différents.

Les paramètres mécaniques n'ont pas apporté de corrélations spécifiques pouvant montrer une sensibilité directe aux pertes thermoélastiques ou viscoélastiques. Toutefois la corrélation avec le paramètre d'interpolation H confirme le lien direct entre le modèle d'amortissement général, le paramètre analytique P_α et la dimension perceptive $dim1$. Le modèle de perte thermoélastique étant fonction de l'épaisseur ($s_1 = c_1/h^2$), lorsque H diminue, cela revient à augmenter l'épaisseur ; des courbes amortissement-fréquences proches du verre pourraient ainsi être obtenues en modifiant l'épaisseur de la plaque en aluminium ($\alpha_{ith} \propto 1/h^2$).

Nous allons observer maintenant la caractérisation des deux matériaux verre et aluminium, avec une tâche d'identification.

9.4 Expérience 4 : identification du matériau

9.4.1 Présentation

Les sujets et les stimuli sont les mêmes que pour l'expérience de dissemblance, seul le protocole expérimental change. Les sujets assistent tout d'abord à une démonstration au cours de laquelle leur sont présentées une plaque en verre et une plaque en aluminium. Ces plaques ont les mêmes caractéristiques géométriques et mécaniques que celles modélisées dans le programme.
Afin de s'assurer de la bonne écoute des auditeurs, nous avons vérifié les courbes amortissement-fréquences des plaques présentées en enregistrant celles-ci en chambre sourde puis en analysant les signaux mesurés par la méthode ESPRIT [45]. Nous avons vérifié ainsi les courbes amortissement-fréquence primordiales dans notre étude, lesquelles constituent la référence pour la perception de l'amortissement par les sujets lors de la tâche d'identification.

L'expérimentateur frappe chacune des plaques avec un maillet puis les sujets commencent l'expérience. Cette procédure est répétée pour chaque série : plaques frappées avec un maillet en bois et plaques frappées avec un maillet en caoutchouc. La tâche d'identification est proposée sous forme d'une classification à choix forcé. Nous sommes donc dans le cas de classements dichotomiques [17]. Chaque stimulus est proposé douze fois à l'auditeur , soit $12 \times 16 = 192$ essais au total en ordre aléatoire et ceci pour chacune des séries. Au total l'auditeur doit faire 384 identifications du matériau. Pour chaque stimulus le sujet exprime si, selon lui, la plaque qu'il vient d'entendre est soit du verre soit de l'aluminium. Pour cela il doit cliquer sur l'un des deux boutons étiqueté "verre" ou "aluminium". La seule consigne donnée aux auditeurs est de faire abstraction de l'excitateur (maillet bois, maillet caoutchouc) qui de toute façon ne varie pas pendant chaque série.

9.4.2 Remarques sur le protocole expérimental

Aucune information n'est donnée aux sujets sur les stimuli qu'ils vont entendre ; ils ignorent ainsi qu'il s'agit de sons de synthèse, en particulier de sons hybrides entre deux matériaux.

Comme ils ont assisté préalablement à une démonstration avec de véritables plaques et de vrais maillets, tous les sujets ont affirmé avoir jugé les stimuli issus de la synthèse comme des sons de plaques enregistrées. Tous affirment être convaincu de réussir la tâche après l'écoute des plaques frappées par l'expérimentateur et perdre ensuite progressivement leur confiance pendant l'expérience.

Une étude sémantique complète aurait été intéressante parallélement à cette tâche d'identification mais impossible à mettre en oeuvre compte-tenu de la trop grande diversité des réponses des sujets. En effet, nous avons interrogé quatre sujets test sur la nature de l'objet vibrant qui leur était proposé, en l'occurence les plaques hybrides de synthèse, et ceci indépendament du maillet. Les matériaux cités sont : l'acier, le fer, la fonte, l'aluminium et le verre. Les objets eux-mêmes frappés qui sont cités, par les sujets, comprennent entre autre : une poêle, une cloche, une plaque, une timbale, une casserole, un verre, un plat, un bidon, une bouteille, un couvercle, un capot, une fenêtre, une grille. La grande diversité des réponses peut venir aussi du fait que les objets synthétisés ne correspondent pas à des objets courants, les plaques se présentant généralement dans la vie courante sous leurs formes encastrés (fenêtre, boîte en aluminium). Les sujets reconnaissent ne pas penser à des matériaux hybrides et lorsque la nature de la matière est difficile à identifier, ils cherchent davantage à identifier un objet et sa forme particulière pouvant donner un son similaire.

9.4.3 Résultats

Les sujets n'ayant que deux choix possibles, verre ou aluminium, nous avons choisi de représenter les réponses sous forme de pourcentage de réponses “aluminium” effectuées par le sujet. Ainsi un score de 15% signifie que le sujet a choisi l'aluminium dans 15% de ses réponses et le verre dans les autres 85% de ses réponses.

Nous avons représenté sur la figure 9.11 les pourcentages de réponses “aluminium” pour la totalité des sujets. Il apparaît immédiatement que leurs jugements sont indépendants du paramètre c_i varié qui influe sur la hauteur, alors que celui-ci était sensible lors de l'expérience de dissemblance.

Les auditeurs ont donc fait leurs jugements sur le timbre des stimuli correspondant à la sensation d'amortissement. Lors de la partie démonstration, les auditeurs se sont concentrés sur la différence d'amortissement pour caractériser chaque plaque. Il est très difficile en effet pour l'expérimentateur de reproduire exactement les mêmes sensations de hauteur sur deux impacts successifs des plaques : les modes sont excités différemment selon les paramètres de vitesse du maillet et du lieu d'excitation qui varient sensiblement pour chaque démonstration. La différence d'amortissement en revanche est beaucoup plus sensible et constitue vraisemblablement le seul critère de jugement des sujets.

Les jugements des auditeurs étant fonction des contributions respectives des pertes, nous représentons les résultats des sujets (% de réponses aluminium) en fonction de la valeur du paramètre d'interpolation H dans la partie supérieure de la figure 9.11.

Le facteur d'interpolation varie linéairement et est perçu comme tel. On peut donc s'attendre à un pourcentage de réponses linéairement croissant entre les extrêmes, soit une perception continue du changement de matière. Les figures C.1 et C.2 dans l'annexe C représentent les scores individuels des sujets en fonction du facteur H caractéristique de la fonction d'interpolation et montrent que les sujets ont réagi de façon non linéaire. La forme générale des courbes est une sigmoïde. Nous pouvons donc caractériser chaque courbe par deux facteurs, le point d'inflexion et la pente de la courbe en ce point. Le point d'inflexion représente la frontière de catégorisation des matériaux et la pente la précision de cette frontière. Une pente plus raide indique un basculement rapide entre catégories pour un petit changement du paramètre H près de la frontière.

9.4.4 Catégorisation

La forme générale des réponses est celle d'une sigmoïde dont une équation est :

$$y = a - b * \arctan\left(\frac{x - x_0}{c}\right)$$

pour laquelle x_0 est le point d'inflexion, $p = -\frac{b}{c}$ représente la pente de la fonction au point d'inflexion. Les réponses vont de 0 à 100 % et définissent les bornes de la fonction.

Pour calculer la réponse globale des sujets nous avons calculé la pente moyenne p_m des courbes au point x_0 moyen, ce qui revient à ajuster les courbes au point moyen de changement de matériau. Nous avons utilisé le logiciel *kaleidagraph* pour l'ajustement des courbes individuelles et nous avons de cette façon calculé les paramètres d'ajustement a, b ,c et x_0 pour chaque série, bois et caoutchouc. Il apparaît que les paramètres de chaque série

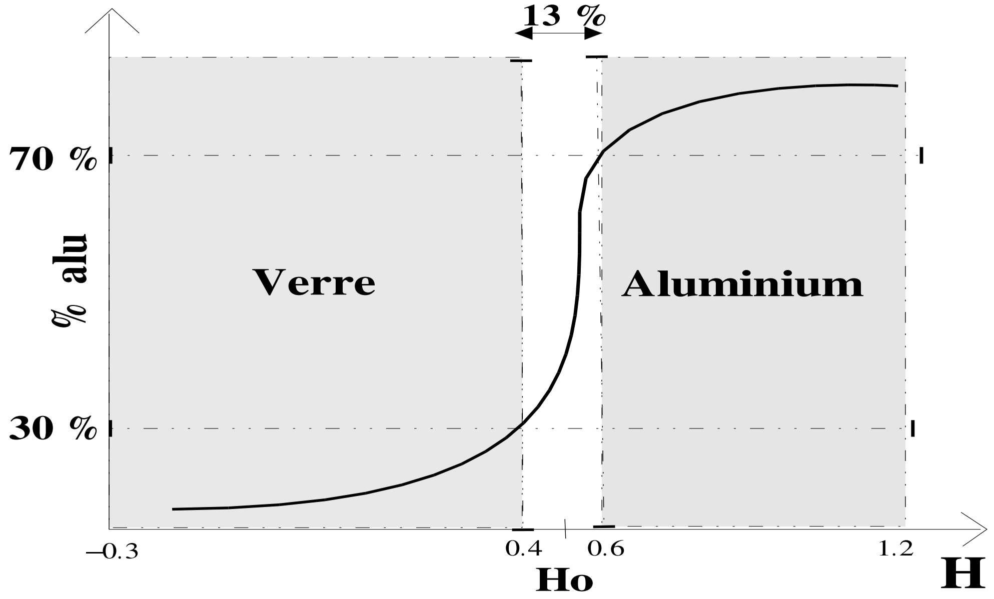

FIG. 9.10 – *Représentation de la sigmoïde moyenne ajustée sur les courbes de réponses. Le changement rapide de perception du matériau montre une catégorisation des stimuli par les sujets.*

sont très proches ; les sujets ont donc des performances équivalentes quel que soit l'excitateur. Enfin, nous avons calculé la gamme sur laquelle les sujets changent de perception du matériau, laquelle correspond à la gamme 30% – 70% de réponses aluminium, $\triangle H$, ainsi que le point d'inflexion, H_0 (voir tableau 9.13).

maillet	$\triangle H$	H_0
bois	0.22	0.39
caoutchouc	0.24	0.42

TAB. 9.13 – *Frontière catégorielle H_0 et gamme de changement $\triangle H$ de perception du matériau.*

Nous pouvons dès lors en conclure deux choses : les sujets changent de perception du matériau sur une bande très étroite qui représente $0.2/1.5 = 13\%$ de la bande totale. Nous pouvons donc parler de catégorisation des matériaux par les sujets. De plus, cette bande n'est pas centrée sur la gamme totale mais décalée vers l'aluminium. La gamme de changement de matière s'effectue en moyenne pour les deux séries de stimuli sur la gamme $H = [0.3,0.5]$ et est centrée sur $H_0 = 0.4$ qui correspond donc pour le modèle de dispersion à $\tilde{d} = 0.4 * \tilde{d}_{ithe} + 0.6 * \tilde{d}_{ve}$. Ces résultats sont reportés sur la figure 9.10 qui représente la sigmoïde moyenne ajustée.

Les réponses des sujets passent très rapidement d'une catégorie à l'autre, cependant nous n'avons pas expérimenté la discrimination intra-catégorielle qui nous aurait permis de tester une perception catégorielle des deux matériaux. Nous ne parlerons donc que de catégorisation des stimuli par les sujets à partir de l'identification d'un continuum de stimuli[2]. Néanmoins nous pouvons remarquer que la dimension perceptive attribuée à la perception de l'amortissement par les sujets lors de l'expérience de dissemblance était corrélée linéairement par les sujets ($r^2 = 0.99$) avec le paramètre d'interpolation H. Les mêmes stimuli soumis aux sujets lors de l'expérience d'identication montrent cette fois une catégorisation des stimuli par les sujets. On peut donc supposer que la structure mentale des sujets pour la perception de ces deux matériaux est catégorielle.

9.4.5 Conclusion

Cette expérience d'identification nous a permis d'appréhender la perception de deux matériaux par les sujets. Les stimuli perçus de façon continue dans l'expérience de dissemblance en fonction du paramètre d'extrapolation ont été cette fois catégorisés par les sujets, les matériaux hybrides n'induisant les sujets en erreur que sur une courte marge (13%). De plus les auditeurs ont répondu de façon identique indépendemment de l'excitateur. Dans ces deux expériences nous avons constaté que le changement de maillet n'a perturbé

2. La perception catégorielle est définie de la façon suivante par Liberman en 1957 : il est possible d'identifier et de discriminer deux objets appartenant à deux catégories distinctes, mais il n'est pas possible de discriminer deux éléments identifiés comme appartenant à la même catégorie. Il y a donc un lien étroit entre la fonction d'identification et celle de discrimination sur un continuum sensoriel donné dans le cas de la perception catégorielle. Ce que l'on peut appeler plus généralement la "catégorisation perceptive" aura une fonction d'identification avec une transition nette entre catégories, mais avec une fonction de discrimination plate.

les sujets que dans l'expérience de dissemblance, et ceci sur une courte marge d'erreur, non pas dans la dispersion inter-sujets qui est équivalente mais dans la répartition spatiale des stimuli. C'est sans doute dû à la perte d'information spectrale dans les stimuli simulés avec un maillet en caoutchouc. Ceci confirme la capacité du système auditif à se concentrer sur la tâche indépendemment de certaines propriétés acoustiques perçus. Dans un cas le sujet doit se concentrer sur le son lui-même en tant qu'objet acoustique (dissemblance), dans l'autre cas il doit se concentrer sur la source qui a donné lieu aux stimuli présentés (identification). Il faut remarquer içi que l'auditeur a donc des performances sur des tâches différentes qui ne sont pas davantage prévisibles avec les paramètres de la source simulée qu'avec les paramètres analytiques. En ce sens la modélisation physique a plutôt confirmé la pertinence des paramètres analytiques.

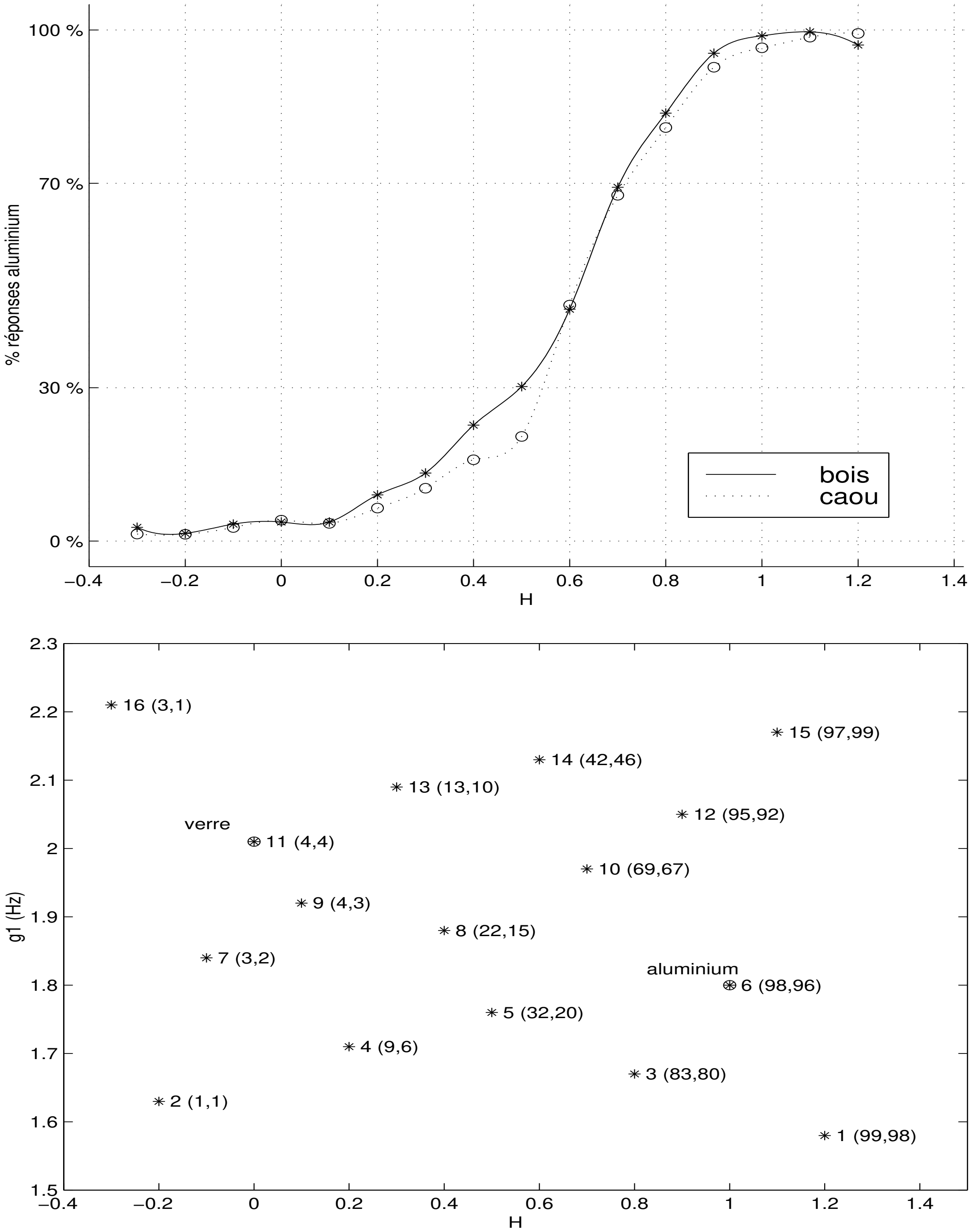

FIG. 9.11 – *En bas, espace physique des stimuli indiquant le pourcentage de réponses aluminium (maillet bois,maillet caoutchouc) pour chaque stimulus. En haut pourcentage de réponses en fonction du paramètre d'interpolation H. Seule la dimension 1 correspondant à l'amortissement est utilisée dans les jugements.*

Chapitre 10

Conclusion

Les expériences présentées au cours de cette thèse nous ont permis d'explorer plusieurs niveaux de correspondance entre des phénomènes physiques précis et tel ou tel attribut de la perception. Et ceci pour des sons émis par des structures vibrantes simples de type barres ou plaques frappées. Nous avons aussi évoqué la perception de deux matériaux à travers les modèles physiques proposés dans la synthèse sonore. Nous conclurons en résumant les principaux résultats obtenus et en soulignant les perspectives pouvant orienter la suite qui, selon toute vraisemblance, devraient conduire à une meilleur compréhension de la perception de notre environnement.

10.1 Synthèse des résultats obtenus

De nombreuses études de psychoacoustique telles que celles menées par Grey [28], Plomp [56], McAdams et al. [51], Krimphoff et al. [36] ont montré qu'un certain nombres de propriétés de la structure acoustique d'un événement sonore complexe pouvaient décrire de façon satisfaisante la sensibilité d'un auditeur lors d'une tâche de dissemblance ou d'identification [51] et aider ainsi à relier la perception d'un son à sa structure spectro-temporelle propre. L'application du modèle CLASCAL [84] a permis de vérifier l'hypothèse de structure spatiale et de bien cerner les stratégies de réponse des sujets en fonction de ces propriétés.

Des études plus récentes comme celle de Lutfi et Oh [48] ont tenté d'établir un lien entre ces paramètres et les matériaux des sources. Mais, ne disposant que de constructions analytiques du signal (synthèse additive), ils n'ont pu établir de lien direct avec les propriétés mécaniques des sources.

Dans notre première expérience sur les barres, nous avons trouvé une très forte corrélation entre un paramètre du modèle d'amortissement et une dimension perceptive. La combinaison d'un paramètre spectral et d'un paramètre temporel avec la dimension perceptive d'une part, et le paramètre d'amortissement d'autre part, a permis d'établir les liens triangulaires entre les trois jeux de paramètres (analytiques, mécaniques et perceptifs) montrant que nous sommes bien dans l'approche écologique décrite au chapitre 6. Nous avons montré aussi la nécessité de prendre en compte l'enveloppe temporelle et donc la dynamique des stimuli, ce qui va dans le sens des travaux précédents d'Iverson et Krumhansl [34].

L'expérience 2 a permis de confirmer nos hypothèses de triangularité des liens et de valider l'exposant 0.3 trouvé comme meilleur exposant en terme de corrélations pour le paramètre mécanique η d'amortissement. Cet exposant, que l'on retrouve dans un autre contexte, n'est donc pas le fruit du hasard mais a une signification perceptive pour l'auditeur. L'origine de cet exposant, probablement psychophysique, n'a pas été établie. Enfin, bien que la hauteur soit plus saillante pour l'auditeur, l'amortissement est resté prédominant dans les jugements des sujets.

Nous avons montré dans l'expérience 3, avec cette fois un modèle physique plus complexe, de très fortes corrélations dans notre relation triangulaire confirmant les résultats obtenus jusqu'alors. Les intercorrélations très fortes entre l'ensemble des paramètres n'ont pas permis de trouver les paramètres mécaniques pouvant décrire spécifiquement le rôle de chacune des contributions. Toutefois, c'est là encore le paramètre utilisé pour contrôler les modèles d'amortissement qui est sous-jacent à la dimension perceptive rendue par les sujets. L'apport supplémentaire de cette partie vient aussi du nouveau paramètre analytique P_α, représentant la pente des courbes amortissement-fréquences, et qui décrit parfaitement la dimension perceptive. Ceci peut avoir des conséquences pour d'autres expériences sur certains objets frappés où l'on chercherait par exemple à identifier une dimension. Nous avons établi ce paramètre analytique grâce à la mécanique qui doit une partie de sa modélisation à l'étude de telles courbes.

A ce stade nous pouvons donc dire que si la théorie écologique ne dispense pas de l'étude analytique des signaux, elle apporte en revanche des voies pour déterminer les paramètres analytiques. Le choix des ces paramètres peut être délicat, une même source pouvant être à l'origine d'un grand nombre de stimuli très différents. Les propriétés physiques ne changent pas et les paramètres analytiques liés permettront de suivre les auditeurs dans leurs jugements.

L'amélioration des modèles physiques permettra nécessairement de mieux comprendre notre perception des objets physiques. Si notre perception est directement sensible et sans cesse reliée à son environnement, c'est par un processus complexe dont l'analyse acoustique n'est qu'un décodage. Et les modèles physiques ne sont que des modèles dans lesquels sont perdues aussi un certain nombre d'informations contenues dans les stimuli naturels. Toutefois l'étude systématique de modifications de modèles physiques, au sens énergétique, et de la correspondance psychoacoustique montrera comment l'individu arrive à extraire, d'une énergie donnée qu'il reçoit de son environnement, les informations dont il a besoin [53]. Et ce par l'analyse de nouveaux paramètres analytiques issus de la modélisation physique.

L'expérience d'identification a permis de montrer la capacité des auditeurs à concentrer leur attention sur la nature de la structure excitée, la plaque, en faisant abstraction de l'excitateur, le maillet. Nous avons montré ainsi qu'un même jeu de stimuli proposé sur une tâche d'identification donne lieu à une catégorisation des stimuli. Les auditeurs ont donc changé leur stratégie montrant qu'ils sont capables d'extraire d'un même signal plusieurs informations selon la tâche. Nos travaux vont donc dans le sens de ceux de Lakatos, McAdams et Caussé [42] et de la recherche d'invariants, le paramètre H d'extrapolation serait donc un invariant pour la perception de l'amortissement et P_α le paramètre analytique correspondant.

10.2 Perspectives

L'exposant trouvé dans l'étude sur les barres mériterait, maintenant que l'on sait qu'il a une signification perceptive, une étude approfondie pour trouver sa signification, physique ou physiologique. Les paramètres analytiques capables de décrire les différentes contributions énergétiques doivent être validés dans des conditions de décorrélation et mériteraient pour cela une étude supplémentaire, notamment dans le rôle respectif de ces contributions dans la reconnaissance des matériaux. L'isolement progressif des différents phénomènes physiques inhérents aux structures permettra progressivement d'isoler les invariants perceptifs s'ils existent et de les caractériser analytiquement.

Plus généralement, l'affinement des modèles physiques et l'augmentation de la puissance des machines informatiques vont permettre la modélisation de structures de plus en plus complexes. Ainsi, la généralisation à des matériaux anisotropes et aux formes curvilignes quelconques, permettra de précieuses simulations en matière de lutherie ou pour la synthèse musicale appliquée à la composition par exemple. La mécanique d'éléments simples tels que les barres ou les plaques servant de bases à des structures comme celle d'une table d'harmonie d'instruments de musique ou une structure de véhicule. Il sera ainsi possible d'analyser et de prévoir le rayonnement d'une structure. En particulier, l'étude de matériaux anisotropes particulièrement utilisés dans l'industrie pour leurs propriétés physiques, permettra de prévoir le rayonnement des structures. On peut alors imaginer modifier la composition de certains matériaux afin d'obtenir un son dont on saura à l'avance sa représentation mentale dans une situation d'excitation précise.

Annexe A

Expérience 3 : distribution des résultats MDS sur 250 tirages bootstrap

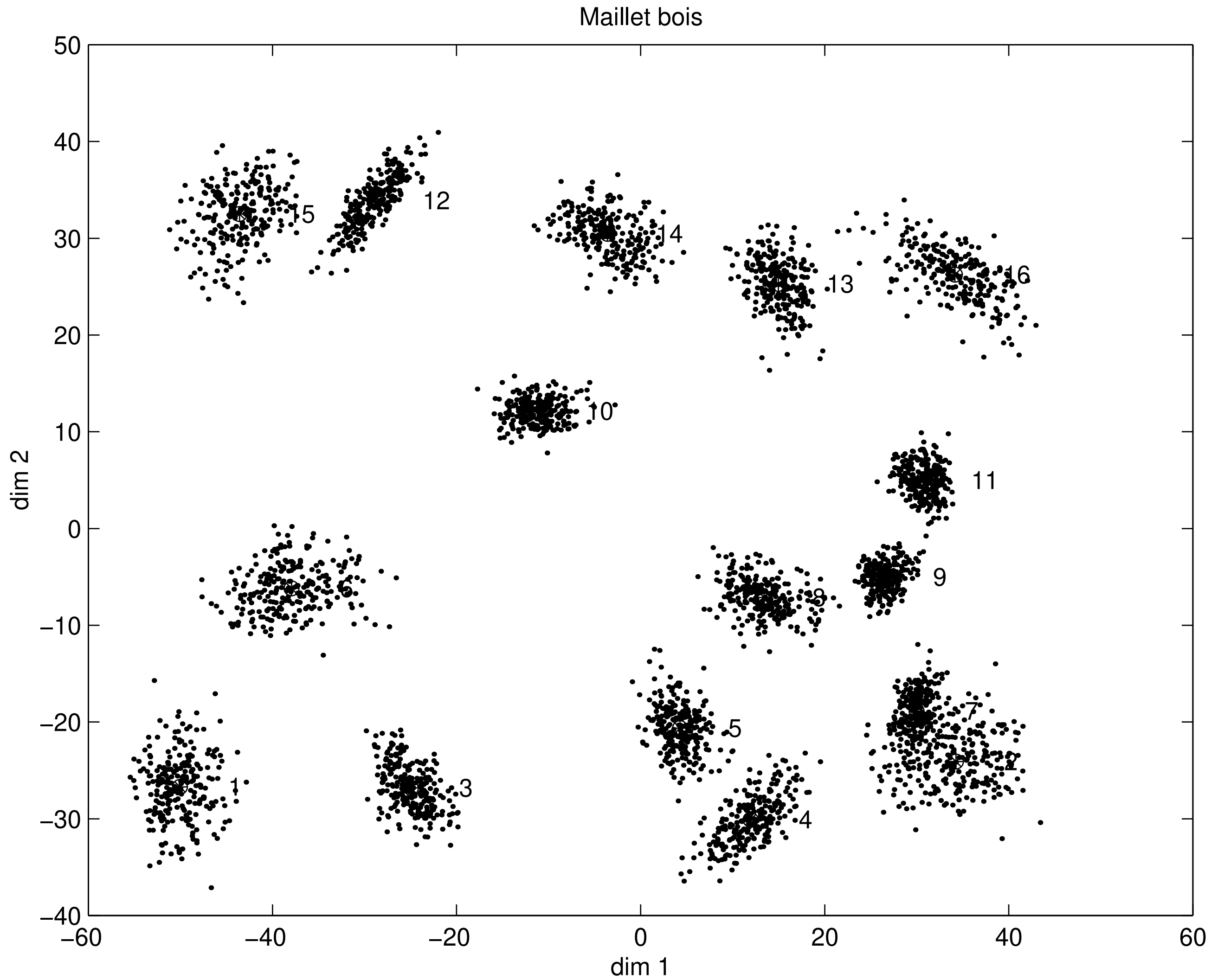

FIG. A.1 – *Bootstrap : distribution des nuages sur 250 tirages, maillet bois.*

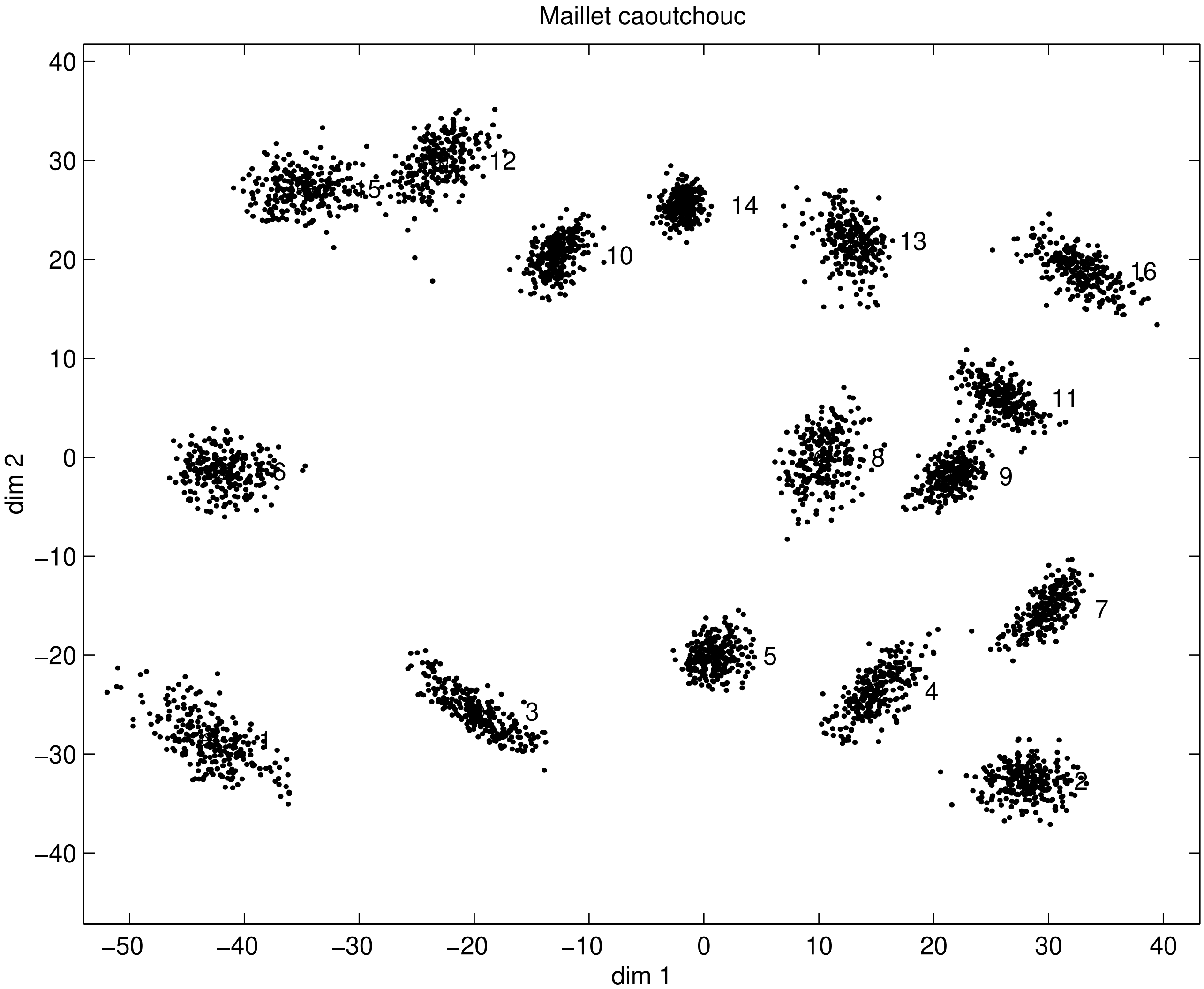

FIG. A.2 – *Bootstrap : distribution des nuages sur 250 tirages, maillet caoutchouc.*

Annexe B

Expérience 3 : répartition par classes des sujets et récapitulatif des données expérimentales

	Maillet bois			Maillet caoutchouc		
Sujets	c1	c2	c3	c1	c2	c3
1	0.00	0.00	1.00	1.00	0.00	0.00
2	0.00	1.00	0.00	0.00	1.00	0.00
3	1.00	0.00	0.00	0.00	0.00	1.00
4	0.00	0.00	1.00	0.00	0.00	1.00
5	0.00	0.00	1.00	0.00	0.00	1.00
6	1.00	0.00	0.00	0.00	0.00	1.00
7	0.00	1.00	0.00	0.00	1.00	0.00
8	1.00	0.00	0.00	1.00	0.00	0.00
9	0.00	1.00	0.00	0.00	1.00	0.00
10	0.99	0.00	0.01	0.00	0.00	1.00
11	1.00	0.00	0.00	0.00	0.00	1.00
12	0.00	1.00	0.00	1.00	0.00	0.00
13	0.00	1.00	0.00	0.00	1.00	0.00
14	1.00	0.00	0.00	0.00	0.00	1.00
15	0.00	0.00	1.00	1.00	0.00	0.00
16	0.00	0.00	1.00	1.00	0.00	0.00
17	0.00	1.00	0.00	1.00	0.00	0.00
18	1.00	0.00	0.00	0.00	1.00	0.00
19	1.00	0.00	0.00	0.00	0.00	1.00
20	0.00	0.00	1.00	0.00	1.00	0.00
21	0.00	1.00	0.00	0.00	1.00	0.00
22	0.01	0.00	0.99	0.00	1.00	0.00

TAB. B.1 – *Répartition des sujets à l'intérieur des classes pour chaque série (maillet bois, maillet caoutchouc)*

dim1	Paramètres						
	Mécaniques			Analytiques			Perceptifs
	H	$HR_1 10^3$	$\sqrt{P_V 10^3}$	$P_\alpha^{0.3} 10^3$	std	$CGS_{(\mathrm{Hz})}$	$dim1$
No sons	**Maillet bois**						
1	1.2	6.00	0.00	0.70	0.39	7079	43.0
2	-0.20	0.00	4.38	1.53	0.05	6447	-33.3
3	0.80	4.36	0.73	1.00	0.33	6943	17.7
4	0.20	1.09	2.92	1.38	0.10	6655	-12.3
5	0.50	2.73	1.82	1.21	0.20	6809	-3.1
6	1.00	5.45	0.00	0.76	0.45	7083	33.7
7	-0.10	0.00	4.01	1.50	0.04	6632	-26.7
8	0.40	2.18	2.19	1.27	0.18	6895	-7.2
9	0.10	0.55	3.28	1.43	0.02	6685	-19.6
10	0.70	3.81	1.09	1.08	0.31	6893	13.4
11	0.00	0.00	3.65	1.47	0.02	6599	-25.5
12	0.90	4.91	0.36	0.90	0.41	6892	24.8
13	0.30	1.64	2.55	1.35	0.13	6654	-11.3
14	0.60	3.27	1.46	1.17	0.27	6818	-0.1
15	1.10	6.54	0.00	0.76	0.36	6968	39.1
16	-0.30	0.00	4.74	1.58	0.05	6573	-32.3
	Maillet caoutchouc						
1	1.20	6.00	0.00	0.72	0.35	2611	45.8
2	-0.20	0.00	4.38	1.53	0.05	2340	-32.9
3	0.80	4.36	0.73	0.98	0.37	2540	22.8
4	0.20	1.09	2.92	1.37	0.07	2447	-14.5
5	0.50	2.73	1.82	1.21	0.20	2532	-2.3
6	1.00	5.45	0.00	0.74	0.45	2660	40.1
7	-0.10	0.00	4.01	1.50	0.04	2422	-27.2
8	0.40	2.18	2.19	1.27	0.18	2526	-10.8
9	0.10	0.55	3.28	1.42	0.02	2458	-20.2
10	0.70	3.81	1.09	1.09	0.28	2549	10.2
11	0.00	0.00	3.65	1.47	0.02	2465	-26.7
12	0.90	4.91	0.36	0.90	0.41	2576	24.2
13	0.30	1.64	2.55	1.34	0.13	2509	-13.5
14	0.60	3.27	1.46	1.17	0.27	2579	-0.6
15	1.10	6.54	0.00	0.79	0.36	2662	38.6
16	-0.30	0.00	4.74	1.63	0.04	2445	-32.8

TAB. B.2 – *Tableau récapitulatifs des paramètres mécaniques, analytiques pour la dimension 1.*

dim2	Paramètres					
	Mécaniques		Analytiques			Perceptifs
No sons	$\sqrt{\frac{E_1}{\rho}}$ (m/s)	$\sqrt{\frac{E_2}{\rho}}$ (m/s)	f_1 (Hz)	f_{per} (Hz)	f_{max} (Hz)	$dim2$
	Maillet bois					
1	1.51	1.58	114	1824	2093	-28.2
2	1.52	1.63	116	1643	2073	-25.4
3	1.53	1.67	120	1803	2131	-18.4
4	1.54	1.71	122	1786	2120	-19.1
5	1.55	1.76	125	1827	2155	-16.5
6	1.56	1.80	128	1917	2202	-5.8
7	1.57	1.84	131	1848	2166	-17.1
8	1.59	1.88	134	2024	2213	-5.1
9	1.60	1.92	137	2108	2213	-6.0
10	1.61	1.97	140	2026	2260	11.0
11	1.62	2.01	142	2183	2248	3.7
12	1.63	2.05	145	2185	2319	29.5
13	1.64	2.09	148	2317	2295	20.5
14	1.66	2.13	151	2377	2330	29.1
15	1.67	2.17	153	2467	2366	23.3
16	1.68	2.21	156	2404	2342	24.5
	Maillet caoutchouc					
1	1.51	1.58	114	1971	2096	-28.2
2	1.52	1.63	116	1925	2073	-32.7
3	1.53	1.67	120	2005	2131	-25.8
4	1.54	1.71	122	2003	2120	-29.1
5	1.55	1.76	125	2044	2155	-19.4
6	1.56	1.80	128	2080	2202	8.5
7	1.57	1.84	131	2059	2166	-15.9
8	1.59	1.88	134	2077	2213	6.5
9	1.60	1.92	137	2084	2213	0.5
10	1.61	1.97	140	2127	2260	26.0
11	1.62	2.01	142	2094	2248	5.5
12	1.63	2.05	145	2175	2319	24.8
13	1.64	2.09	148	2177	2295	23.4
14	1.66	2.13	151	2202	2330	20.3
15	1.67	2.17	153	2219	2366	22.9
16	1.68	2.21	156	2180	2342	12.7

TAB. B.3 – *Tableau récapitulatif des paramètres mécaniques et analytiques pour la dimension 2.*

Annexe C

Expérience 4 : figures des résultats individuels

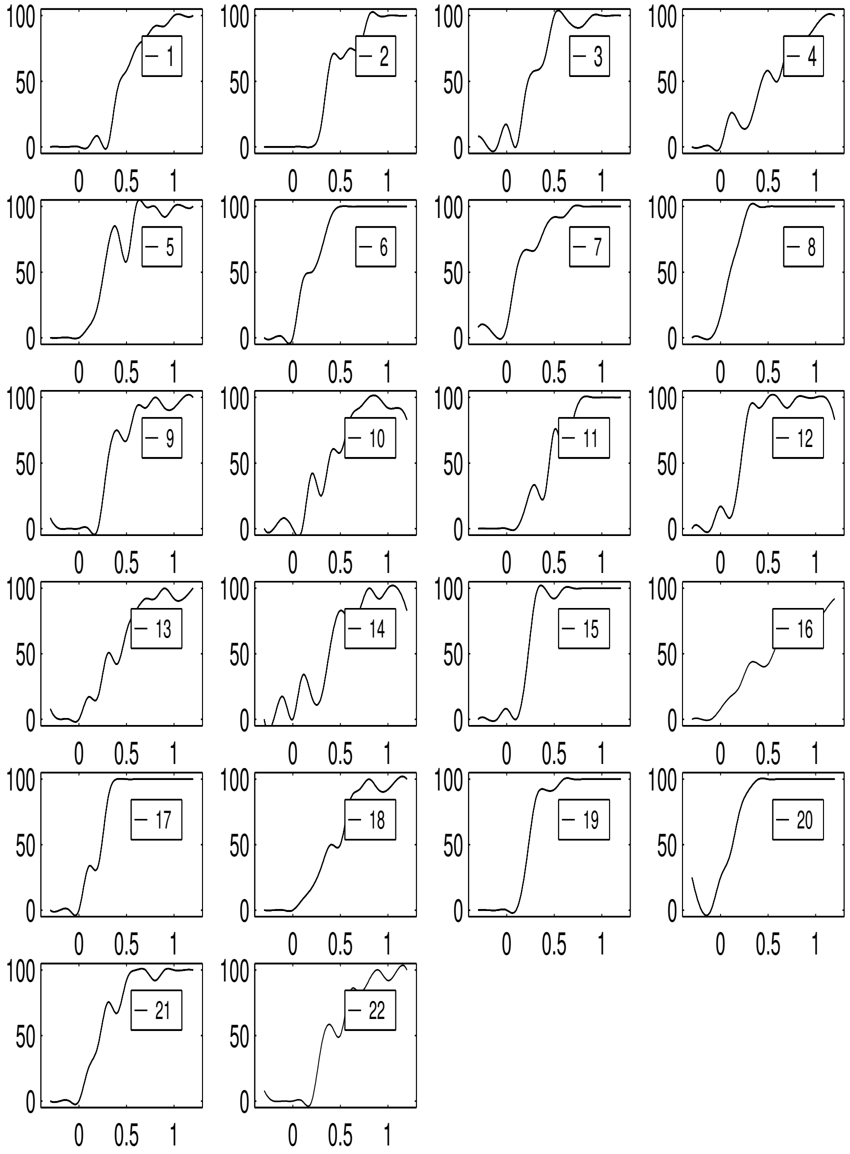

FIG. C.1 – *Résultats individuels des sujets numérotés dans la légende pour la série bois, chaque courbe représente le pourcentage de réponses "aluminium" en fonction du paramètre d'amortissement H.*

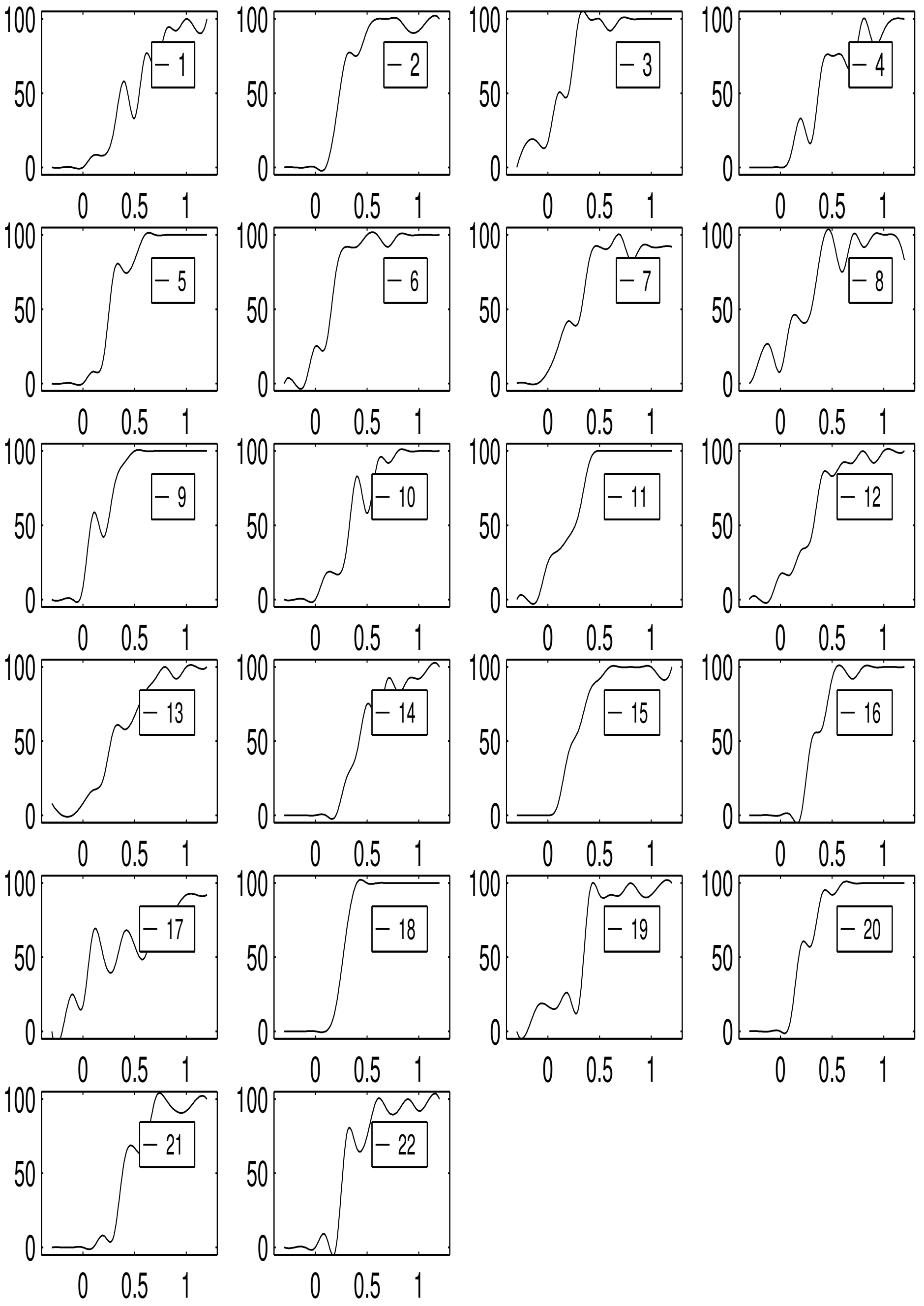

FIG. C.2 – *Résultats individuels des sujets numérotés dans la légende pour la série caoutchouc, chaque courbe représente le pourcentage de réponses "aluminium" en fonction du paramètre d'amortissement H.*

Bibliographie

[1] ANSI. « USA Standard Acoustical Terminology (Including Mechanical Shock and Vibration) ». Rapport Technique S1.1-1960 (R1976), American National Standards Institute, New York, 1960.

[2] F.A. Attneave. « Dimensions and similarity ». *American Journal of Psychology*, 63:516–556, 1950.

[3] J.P. Barthélemy et A. Guénoche. *Les arbres et la représentation des proximités*. Masson, Paris, 1988.

[4] G. von Békésy. *Experiments in Hearing*. McGraw-Hill, New York, 1960.

[5] A.S. Bregman, C. Liao, et R. Levitan. « Auditory grouping based on fyndamental frequency and formant group frequency ». *Canadian Journal of Psychology*, 44(3):400–413, 1990.

[6] M.L. Burton et S.B. Nerlove. « Balanced designs for triades tests : two examples from English ». *Social Science Research*, 5:247–267, 1976.

[7] J.D. Carroll. « Spatial, non spatial and hybrid models for scaling ». *Psychometrika*, 41:439–463, 1976.

[8] J.D. Carroll et P. Arabie. Multidimensional scaling. Dans M.R. Rosenzweig et L.W. Porter, éditeurs, *Annual Review of Psychology*. Annual Reviews, Paolo Alto, 1980.

[9] J.D. Carroll et J.J. Chang. « Analysis of individual differences in multidimensional scaling via an n-way generalization of "Eckart Young" decomposition ». *Psychometrika*, 35:283–319, 1970.

[10] A. Chaigne et V. Doutaut. « Numerical simulations of xylophones. I. Time-domain modeling of the vibrating bars. ». *J. Acoust. Soc. Am.*, 101(1):539–557, 1996.

[11] A. Chaigne et C. Lambourg. « Time-domain simulation of damped impacted plates. Part I. Theory and experiments ». *soumis à J. Acoust. Soc. Am.*, 1999.

[12] M. Clarke, D. Luce, R. Abrams, H. Schlossberg, et J. Rome. « Preliminary experiments on the aural significance of parts of tones of orchestral instruments and choral tones ». *J. Audio. Eng. Soc.*, 11(1):45–54, 1963.

[13] J.P. Cunningham. « Free trees and bidirectionnal trees as representatoins of psychologocal distance ». *Journal of Mathematical Psychology*, 17:165–188, 1978.

[14] H.A. David. *The Method of Pair Comparisons*. Griffin, London, 1963.

[15] L. Demany. « *La perception de la hauteur tonale* ». PhD thesis, Université Pierre et Marie Curie, Paris VI, 1986.

[16] E. Diday. « Optimization in non-hierarchical clustering ». *Pattern Recognition*, 6:17–33, 1974.

[17] S. Donnadieu. « *Représentation mentale du timbre des sons complexes et effets de contexte* ». PhD thesis, Université René Descartes, Paris V, 1997.

[18] V. Doutaut. « *Etude expérimentale et simulations numériques d'instruments de percussion à clavier* ». PhD thesis, Ecole Nationale Supérieure des Télécommunications, Paris, 1996.

[19] V. Doutaut, D. Matignon, et A. Chaigne. « Numerical simulations of xylophones. II. Time-domain modeling of the resonator and of the radiated sound pressure ». *J. Acoust. Soc. Am.*, 104(3):1633–1647, 1998.

[20] B. Efron et R.J. Tibshirami. *An Introduction to the Bootstrap*. Chapman and Hall, 1993.

[21] A. Faure, S. McAdams, et V. Nosulenko. « Verbal correlates of perceptual dimensions of timbre ». Dans *4Th International Conference on Music Perception and Cognition*, Montréal, 1996.

[22] H. Fletcher. « Loudness, pitch and the timbre of musical tones and their relation to the intensity, the frequency and the overtone structure ». *J. Acoust. Soc. Am.*, 6:56–59, 1934.

[23] D.J. Freed. « Auditory correlates of perceived mallet hardness for a set of recorded percussive sound events ». *J. Acoust. Soc. Am.*, 87(1):311–322, 1989.

[24] J.J. Gibson. *The Senses Considered as Perceptual Systems*. Houghton-Mifflin, Boston, 1966.

[25] J.J. Gibson. *The Ecological Approach to Visual Perception*. Houghton-Mifflin, Boston, 1979.

[26] J.C. Gower. « Some distance properties of latent root and vector methods using multivariate analysis ». *Biometrika*, 53:325–338, 1966.

[27] P.E. Green et V.R. Rao. « Rating scales and information recovery - how many scales and response categories to use? ». *Journal of Marketing*, 34:33–39, 1970.

[28] J.M. Grey. « Multidimensional perceptul scaling of musical timbres ». *J. Acoust. Soc. Am.*, 61(5):1270–1277, 1977.

[29] J.M. Grey et J.W. Gordon. « Perceptual effects of spectral modifications on musical timbres ». *J. Acoust. Soc. Am.*, 63(5):1493–1500, 1978.

[30] H.L.F. von Helmholtz. *La théorie physiologique de la musique*. J. Gabay, Paris, 1990.

[31] A.C. Hope. « A simplified Monte Carlo significance test procedure ». *Journal of the Royal Statistical Society*, Series B, 30:582–598, 1968.

[32] A.J.M. Houtsma. Section V: Timbre. Dans S. Nielzen et O. Olsson, éditeurs, *Structure and Perception of Electroacoustic Sounds and Music*, pages 157–159. Excerpta Medica, Amsterdam, 1989.

[33] W. Hutchinson et L. Knopoff. « The acoustic component of Western consonance ». *Interface*, 7:1–29, 1978.

[34] P. Iverson et C.L. Krumansl. « Isolating the dynamic attributes of musical timbre ». *J. Acoust. Soc. Am.*, 94(5):2595–2603, 1993.

[35] S.C. Johnson. « Hierarchical clustering schemes ». *Psychometrika*, 32:241–254, 1967.

[36] J. Krimphoff, S. McAdams, et S. Winsberg. « Caractérisation du timbre des sons complexes. II. Analyses acoustiques et quantification

psychophysique ». *Supplément au Journal de physique III*, 4:625–628, 1994.

[37] C.L. KRUMHANSL. « Why is musical timbre so hard to understand? ». Dans S. NIELZÉN et O. OLSON, éditeurs, *Structure and Perception of Electroacoustic Sound and Music*, pages 43–53, Elsevier (Excerpta Medica 846), Amsterdam, 1989.

[38] J.B. KRUSKAL. « Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis ». *Psychometrika*, 29:1–28, 1964a.

[39] J.B. KRUSKAL. « Non-metric multidimensional scaling: a numerical method ». *Psychometrika*, 29:115–129, 1964b.

[40] J.B. KRUSKAL et J.D. CARROLL. Geometrical models and badness-of-fit functions. Dans P.R. KRISHNAIAH, éditeur, *Multivariate Analysis*, volume II. Academic Press, New York, 1969.

[41] J.B. KRUSKAL et M. WISH. *Multidimensional Scaling*. Sage University Paper, London, 1978.

[42] S. LAKATOS, S. MCADAMS, et R. CAUSSÉ. « The representation of auditory source characteristics: Simple geometric form ». *Perception and Psychophysics*, 59:1180–1190, 1997.

[43] C. LAMBOURG. « *Modèle temporel pour la simulation numérique de plaques vibrantes - Application à la synthèse sonore* ». PhD thesis, Université du Maine, 1997.

[44] C. LAMBOURG, A. CHAIGNE, et D. MATIGNON. « Time-domain simulation of damped impacted plates. Part II. Numerical model and results ». *soumis*, 1999.

[45] J. LAROCHE. « The use of the matrix-pencil method for the spectrum analysis of musical signals ». *J. Acoust. Soc. Am.*, 4:1958–1965, 1993.

[46] J.C.R. LICKLIDER. « “Periodicity” pitch and “place” pitch ». *J. Acoust. Soc. Am.*, 26:945(A), 1954.

[47] E. LINDEMANN, M. PUCKETTE, E. VIARA, M. DE CECCO, et F. DECHELLE. « The architecture of the IRCAM Musical Workstation ». *Computer Music Journal*, 15(3):41–49, 1991.

[48] R.A. LUFTI et E.L. OH. « Auditory discrimination of material changes in a struck-clamped bar ». *J. Acoust. Soc. Am.*, 102(6):3647–3656, 1997.

[49] S. McAdams. La reconnaissance de sources et d'événements sonores. Dans *Penser les sons (S. McAdams et E. Bigand). Psychologie cognitive de l'audition*. PUF, Paris, 1994.

[50] S. McAdams. « Communication personnelle », 1999.

[51] S. McAdams, S. Winsberg, S. Donnadieu, G. De Soete, et J. Krimphoff. « Perceptual scaling of synthesized musical timbres : Common dimensions, specificities, and latent subject classes ». *Psychological Research*, 58:177–192, 1995.

[52] J. McQueen. « Some methods for classification and analysis of multivariate observations ». Dans L.M. le Cam et J. Neyman, éditeurs, *Proceedings of the Fifth Berkeley Symposium on Mathematical Statisitcs and Probability*, volume 1, pages 281–297, Berkeley, 1967. University of Cqlifornia Press.

[53] C.F. Michaels et C. Carello. *Direct Perception*. Century Psychology Series. Prentice-Hall, New York, 1981.

[54] J.R. Miller et E.C. Carterette. « Perceptual space for musical structures ». *J. Acoust. Soc. Am.*, 58:711–720, 1975.

[55] U. Neisser. *Cognitive Psychology*. Prentice-Hall, New York, 1967.

[56] R. Plomp. « Timbre as a multidimensional attribute of complex tones ». Dans R. Plomp et G.F. Smoorenburg, éditeurs, *Frequency Analysis and Periodicity Detection in Hearing*, pages 397–414, Sijthoff, Leiden, 1970.

[57] R. Plomp et H.J.M. Steeneken. « Pitch versus timbre ». Dans *Proc. 7th Internat. Congr. Acoustics*, volume 3, pages 377–380, Budapest, 1971.

[58] D. Pressnitzer, S. McAdams, S. Winsberg, et J. Fineberg. « Tension et rugosité dans une musique non tonale ». Dans *Actes des Journées d'Informatique Musicale*, pages 147–156, Caen, France, 1996.

[59] J.O. Ramsay. « Confidence regions for multidimensional scaling analysis ». *Psychometrika*, 43:145–160, 1978.

[60] J.O. Ramsay. « Some statistical approaches to multidimensional data ». *Journal of the Royal Statistical Society*, 145:145–160, 1982.

[61] R.A. Rasch et R. Plomp. The perception of musical tones. Dans D. Deutsch, éditeur, *The Psychology of Music*, pages 1–24. Academic Press, New York, 1982.

[62] B.H. Repp. « The sound of two hands clapping: An exploratory study ». *J. Acoust. Soc. Am.*, 81(4):1100–1109, 1987.

[63] J.C. Risset. « Paradoxes de la hauteur : le concept de hauteur sonore n'est pas le même pour tout le monde. ». Dans *7eme Congrès d'Acoustique Musicale*, volume 20, page 10, Budapest, 1971.

[64] J.C. Risset. Paradoxes de la hauteur. IRCAM, Paris, 1978.

[65] J.C. Risset et D. Wessel. « Exploration du timbre par analyse et synthèse ». Dans Christian Bourgeois/IRCAM, éditeur, *Le timbre métaphore pour la composition*, Paris, 1991.

[66] J.G. Roederer. *Introduction to the physics and psychophysics of music*. Springer, New York, 2ème édition, 1975.

[67] E.E. Roskam. « The method of triads for nonmetric multidimensional scaling ». *Psychology*, 25:404–417, 1970.

[68] S. Sattah et A. Tversky. « Additive similarity trees ». *Psychometrika*, 42:319–345, 1977.

[69] J.F. Schouten. « The perception of pitch ». *Philips Techn. Rev.*, 5:286–294, 1940.

[70] J.F. Schouten. « The perception of timbre ». Dans *Rep. 6th Internat. Congr. Acoustics*, Article GP-6-2, Tokyo, 1968.

[71] G. Schwarz. « Estimating the dimensions of a model ». *Annals of Statistics*, 6:461–464, 1978.

[72] A. Seebeck. « Beobachtungen uber einige Bedingungen der Entstehung von Tonen ». *Ann. Phys. Chem.*, 53:417, 1841.

[73] R.N. Shepard. « The analysis of proximities : Multidimensional scaling with an unknown distance function. Part I ». *Psychometrika*, 27:125–140, 1962a.

[74] R.N. Shepard. « The analysis of proximities: Multidimensional scaling with an unknown distance function. Part II ». *Psychometrika*, 27:219–246, 1962b.

[75] R.N. SHEPARD. « Representation of structure in similarity data: Problems and prospects ». *Psychometrika*, 39:373–421, 1974.

[76] R.N. SHEPARD et P. ARABIE. « Additive clustering: Representation of similarity as combinations of discrete overlapping properties ». *Psychological Review*, 86:87–123, 1979.

[77] B. SMITH. « PsiExp : an environment for Psychoacoustic Experimentation Using the IRCAM Musical Workstation ». Dans *Musical Perception and Cognition Conference 95*, Berkeley, 1995.

[78] S.S. STEVENS. « A metric for social consensus ». *Science*, 166:530–531, 1966.

[79] W.S. TORGERSON. *Theory and Methods of Scaling.* Wiley, New York, 1958.

[80] W.S. TORGERSON. « Multidimensional scaling of similarity ». *Psychometrika*, 30(4):379–393, 1965.

[81] L. WEDIN et G. GOUDE. « Dimension Analysis of the Perception of Instrumental Timbre ». *Scandinavia Journal of Psychology*, 13:228–240, 1972.

[82] S. WINSBERG et J.D. CARROLL. « A quasi-nonmetric method for multidimensional scaling via an extended Euclidian model ». *Psychometrika*, 53:217–229, 1988.

[83] S. WINSBERG et J.D. CARROLL. A quasi-nonmetric method for multidimensional scaling of multiway data via a restricted case of an extended INDSCAL model. Dans R. COPPI et S. BOLASCO, éditeurs, *Multiway Data Analysis*, pages 405–414. North-Holland, Amsterdam, 1989.

[84] S. WINSBERG et G. DE SOETE. « A latent class approach to fitting the weighted euclidean model, CLASCAL ». *Psychometrika*, 58:315–330, 1993.

[85] M. WISH et J.D. CARROLL. Applications of "INDSCAL" to studies of human perception and judgement. Dans E.C. CARTERETTE et M.P. FRIEDMAN, éditeurs, *Handbook of perception*, volume 2. Academic Press, New York, 1974.

[86] F.W. YOUNG. « TORSCA-9, A Fortran IV program for Shepard-Kruskal multidimensional scaling analysis ». *Behavioral Science*, 13:343–

344, 1968.

[87] F.W. YOUNG. « Nonmetric multidimensional scaling : recovery of metric information ». *Psychometrika*, 35:455–473, 1970.

[88] F.W. YOUNG. Scaling. Dans M.R. ROSENZWEIG et L.W. PORTER, éditeurs, *Annual Review of Psychology*, volume 35, pages 55–81. Annual Reviews, Palo Alto, 1984.

[89] F.W. YOUNG, J.B. KRUSKAL, et J.B. SEERY. « How to use KYST, a very flexible program to do multidimensional scaling and unfolding ». *Bell Laboratories*, 1973.