

**ETUDE PSYCHOACOUSTIQUE
DE LA PERCEPTION
DES SONS DE TRAFIC**

Contrat établi entre

S.N.C.F. (Direction de la Recherche – Dp^t RV)
et
I.R.C.A.M. (Equipe Psychoacoustique)

– Rapport Final –

Etude effectuée entre le 1^{er} novembre 96 et le 30 avril 97.

SOMMAIRE

PRESENTATION DU PROJET	3
I- EGALISATION EN SONIE	4
I-a. Présentation	4
I-b. Stimuli	4
I-c. Appareillage	5
I-d. Procédure	6
I-e. Résultats	6
II- ETUDE DES DISSEMBLANCES	7
A- METHODE EXPERIMENTALE	7
II-A-1. Sujets d'expérience	7
II-A-2. Stimuli	7
II-A-3. Appareillage	7
II-A-4. Procédure	7
II-A-5. Résultat de la tâche d'identification	9
II-A-6. Examen de données sémantiques recueillies	10
B- ANALYSE DES DONNEES DE DISSEMBLANCE	10
II-B-1. Première série d'analyses	10
II-B-1-a. Principe de l'analyse multidimensionnelle	10
II-B-1-b. Etapes de l'analyse multidimensionnelle	11
II-B-1-c. Premiers résultats	12
II-B-2. Deuxième série d'analyse	15
II-B-2-a. Principe des analyses de cluster	15
II-B-2-b. Mise en oeuvre des analyses de cluster	17
II-B-2-c. Analyses multidimensionnelles partielles	17
III- ETUDE DE DOMINANCE	20
A- METHODE EXPERIMENTALE	20
III-A-1. Echelle sémantique de l'expérience	20
III-A-2. Informations sur les sujets	20
III-A-3. Stimuli	21
III-A-4. Appareillage	21
III-A-5. Procédure	21
III-A-6. Examen des données sémantiques recueillies	22
B- ANALYSE DES DONNEES DE DOMINANCE	22
III-B-1. Exploitation directe des données de dominance	22
III-B-2. Validation de l'échelle de désagrément	23
IV- ANALYSE ACOUSTIQUE DES SONS - CALCUL DES CORRELATIONS	25
A- CADRE THEORIQUE	25
IV-A-1. Analyse et traitements	25
IV-A-2. Paramètres physiques de référence	27
IV-A-3. Principe du calcul de corrélation	30
B- RECHERCHE DES PARAMETRES ACOUSTIQUES	30
IV-B-1. Détails des premières investigations	30
IV-B-2. Le résultat significatif	33
CONCLUSION	35
BIBLIOGRAPHIE	36
ANNEXES	39

PRESENTATION DU PROJET

L'étude en cours a pour but de mettre à jour expérimentalement les structures perceptives qui sous-tendent la comparaison de divers bruits de trafic. Les échantillons sonores, fournis par la S.N.C.F., sont issus d'enregistrements de trafic ferroviaire, routier et aérien.

La méthode employée s'articule en plusieurs points :

- un test de dissemblance où les sons sont présentés par paire à des auditeurs qui doivent évaluer la distance perceptive entre les deux sons d'une même paire ; les données expérimentales recueillies sont ensuite traitées afin de révéler la structure perceptive qui sous-tend ces jugements de dissemblance.

- un test de dominance qui consiste à faire classer, par des auditeurs, les stimuli sonores en fonction de la dominance d'un attribut sémantique, en l'occurrence le désagrément.

- une phase d'analyse acoustique qui cherche à dégager les paramètres physiques des signaux qui sont le mieux corrélés avec les dimensions perceptives.

Auparavant, les sons font l'objet d'une égalisation en sonie. Cette procédure est nécessaire afin que les jugements de dissemblance et de dominance soient indépendants de ce paramètre acoustique qui s'avère prépondérant dans le processus d'appréciation subjective, et que l'émergence de facteurs plus fins intervenant dans la perception auditive soient facilitée.

Les résultats principaux obtenus à l'issue de l'étude peuvent se résumer en deux grands points : d'une part, le rôle prépondérant de l'identification dans les jugements de dissemblance, et d'autre part, la détermination de facteurs spectraux qui semblent expliquer les jugements de désagrément

I- AGALISATION EN SONIE

I-a. Présentation

Le but de cette phase préliminaire d'égalisation en sonie est d'aboutir à un corpus d'échantillons sonores présentant la même intensité sonore subjective. Par définition même de la sonie, cette tâche ne peut s'effectuer en égalisant les niveaux sonores physiques des sons (L_{eq} , par exemple) puisque la détermination de l'intensité subjective de sons complexes est forcément le résultat d'un processus perceptif.

Dans le cas présent, la tâche est accomplie par les chercheurs de l'équipe "Perception et Cognition Musicales" et des personnes internes à l'I.R.C.A.M., habitués à ce genre d'exercice et ne présentant aucune déficience auditive.

I-b. Stimuli

12 échantillons sonores sont initialement livrés par la S.N.C.F. :

- son 1 : TGV Sud-Est, ligne classique (160 km/h) ;
- son 2 : TGV Sud-Est, ligne à grande vitesse (270 km/h) ;
- son 3 : TGV Réseau, ligne classique (160 km/h) ;
- son 4 : TGV Réseau, ligne grande vitesse (270 km/h) ;
- son 5 : Train Corail (160 km/h) ;
- son 6 : Train marchandise, diesel (80 km/h) ;
- son 7 : Train marchandise, électrique (80 km/h) ;
- son 8 : Autoroute fluide, sans poids lourd ;
- son 9 : Autoroute fluide, avec poids lourds ;
- son 10 : Autoroute chargée, vitesse réduite ;
- son 11 : Voie urbaine, démarrage au feu ;
- son 12m : *Décollage avion.*

La durée des échantillons est normalisée (entre 8 et 10 secondes) au moyen d'une extinction artificielle ('fade out', durée : 1 sec.) et également – pour les sons n° 8, 9, 10 – d'une montée artificielle ('fade in', durée : 1 sec.). Un son d'étalonnage de 1000 Hz à 0 dB (électrique) est livré avec les échantillons de l'étude. En complément, des données techniques sur les sons et les conditions d'enregistrement sont fournies par la S.N.C.F. (cf. annexe B).

Par ailleurs, le son n° 1 présente une anomalie qui semble provenir du traitement ultérieur à la prise de son. C'est pourquoi nous concevons, à partir de l'enregistrement original, un nouvel échantillon n° 1 correspondant à ce stimulus et respectant toutes ses caractéristiques (durée, rampes, etc. ...). L'absence du défaut dans le nouvel échantillon confirme notre hypothèse et permet de remplacer le son n° 1.

Cela étant, le nombre de stimuli s'avère être insuffisant pour la solidité de l'analyse multidimensionnelle qui va suivre (le nombre de stimuli détermine le nombre maximum de dimensions à tester (cf. §II-B-1-a) ; d'autant plus que le son n° 12m est issu d'un enregistrement monophonique, ce qui constitue une spécificité trop importante par rapport aux autres sons pour qu'il fasse partie du test.

La base de données sonores initiale est donc enrichie de nouveaux sons provenant, d'une part, de l'exploitation d'autres enregistrements effectués par les preneurs de son de la S.N.C.F. pour cette même campagne et, d'autre part, de la consultation des archives de la société DIASONIC qui possède notamment des bruits de trafic aérien.

A partir des bandes originales, l'élaboration des nouveaux échantillons – effectuée à l'I.R.C.A.M. sur le logiciel *Audiosculpt* – répond aux mêmes normes que les précédentes en ce qui concerne la durée et les montées et extinctions artificielles.

Une seconde série de 10 échantillons est ainsi constituée ; après écoute et concertation communes, 5 d'entre eux sont finalement retenus, ce qui porte à 16 le nombre définitif de stimuli sonores pour l'étude :

- son 12 : Flux urbain sur boulevard en sens unique ;
- son 13 : Airbus A320, décollage ;
- son 14 : Airbus A320, atterrissage ;
- son 15 : RER A , 60 km/h ;
- son 16 : Autobus, départ.

En outre, il s'avère que les 16 sons stéréophoniques de l'étude ne présentent pas la même spatialisation : pour certains, la source se déplace de gauche à droite alors que pour d'autres, c'est l'inverse. Cette inhomogénéité de localisation nous oblige à effectuer une ultime manipulation sur les échantillons – en fait, une inversion des canaux – pour que la sensation de mouvement de la source sonore soit identique pour l'ensemble des stimuli : un passage de droite à gauche. Les échantillons ainsi 'renversés' sont les suivants : sons n° 1, 2, 3, 4, 5, 6 et 14.

I-c. Appareillage

Les données sonores numériques du support DAT sont importées directement sur le disque dur d'une NeXT Cube équipée d'une carte ISPW et des logiciels MAX et PsiExp. Ces applications permettent de mettre en oeuvre la démarche expérimentale en jouant les sons suivant une procédure donnée et en recueillant les réponses des sujets par l'intermédiaire d'une interface graphique et d'une souris. La reproduction des sons emploie des convertisseurs ProPort et un amplificateur stéréo Canford. Enfin, après une concertation entre l'I.R.C.A.M. et la S.N.C.F., il est décidé d'effectuer l'écoute en cabine audiométrique, sur un casque AKG K1000 (et non pas sur une paire d'enceintes).

En outre, on effectue un étalonnage de cette chaîne de diffusion ce qui permet, d'une part, de placer tous les sujets au même niveau d'écoute et, d'autre part, d'avoir accès au niveau physique réel entendu par l'auditeur.

Pour ce faire, on utilise couramment une sinusoïde pure de 1 kHz à 94 dB SPL (délivrée, par exemple, par un pistonphone). Ce signal est restitué à un niveau sonore physique que l'on règle avec les potentiomètres de l'amplificateur pour un confort d'écoute optimum et que l'on mesure à l'aide d'un sonomètre.

Ainsi, soit le signal de référence 'ref_signal' dont l'onde de pression temporelle est de la forme :

$$p(t) = P_{\max} \cdot \sin(\omega t),$$

le niveau en dB SPL est donné par l'expression :

$$L(\text{dB SPL}) = 20 \cdot \log_{10} \left(\frac{P_{\text{eff}}}{P_0} \right), \quad \text{avec } P_{\text{eff}} = \frac{P_{\text{max}}}{\sqrt{2}} \text{ et } P_0 = 2 \cdot 10^{-5} \text{ Pascal.}$$

On a donc $P_{\text{max}} = P_0 \sqrt{2} \cdot 10^{\frac{94}{20}} \approx 1.4176$ qui permet de déduire le facteur de calibration :

$$\text{calib} = \frac{1.4176}{\max(\text{ref_signal})}.$$

I-d. Procédure

La procédure d'égalisation repose sur la méthode d'ajustement : l'échantillon présentant la sonie la moins élevée est désigné comme référence (en l'occurrence, le son n° 3). A chaque présentation, cet échantillon de référence est confronté à un autre échantillon dont on peut varier le niveau en temps réel. Le sujet ajuste le niveau du son variable de manière à ce qu'il soit perceptivement égal à celui du son de référence, répétant la présentation de la paire autant de fois que nécessaire pour atteindre l'égalité perceptive en niveau.

I-e. Résultats

L'égalisation en sonie a été effectuée par 19 personnes entre le 27 novembre et le 3 décembre 1996. Les résultats généraux donnant le coefficient d'égalisation trouvé par chaque sujet et pour chaque échantillon sont présentés en annexe A.

On effectue ensuite un calcul de moyenne et de variance pour obtenir un coefficient d'égalisation moyen, et son écart-type, pour chaque stimulus. On convertit enfin ces résultats en dB pour connaître l'atténuation équivalente et la tolérance (cf. annexe A).

La valeur importante de certains écart-types – par exemple, 4.2 dB pour le son n° 14 – nous amène à dégager, parmi les 19 sujets du test, une population optimale qui donne globalement les résultats les plus cohérents, c'est-à-dire les écart-types les plus faibles. Cette recherche basée sur l'examen des réponses individuelles aboutit à la suppression de 4 sujets dans le calcul de la moyenne et de la variance et permet de contenir les écart-types dans un intervalle allant de 1 à 3 dB. Les valeurs d'atténuation constituant les résultats définitifs du test d'égalisation en sonie ainsi que les valeurs de L_{eq} qui en résultent se trouvent en annexe A.

II- STUDE DES DISSEMBLANCES

L'objectif de cette étude est de déterminer l'espace perceptif multidimensionnel qui sous-tend les jugements sur l'ensemble des échantillons égalisés en sonie. Cet espace, élaboré à partir des distances perceptives entre les objets, permet de dégager les dimensions perceptives prépondérantes intervenant dans le processus de perception des sons proposés.

A- Méthode expérimentale

II-A-1. Sujets d'expérience

Le test de dissemblance s'est effectué entre le 23 décembre 1996 et le 11 janvier 1997. La population d'auditeurs est constituée de 30 personnes issues d'une liste de volontaires recueillie lors des journées 'Portes Ouvertes' de l'I.R.C.A.M. . Les sujets, 15 hommes et 15 femmes, âgés de 17 à 61 ans, se répartissent de la manière suivante :

- 13 personnes ont moins de 25 ans (6 femmes, 7 hommes),
- 7 personnes ont entre 25 et 35 ans (4 femmes, 3 hommes),
- 5 personnes ont entre 35 et 45 ans (1 femme, 4 hommes),
- 5 personnes ont plus de 45 ans (4 femmes, 1 homme).

En outre, étant donné l'aspect environnemental de l'étude, il nous a semblé opportun de retenir comme critère le lieu d'habitation des sujets ; ainsi, la population d'auditeurs est constituée de 15 personnes habitant Paris 'intra-muros' et 15 personnes habitant en banlieue parisienne (départements 91, 92, 93, 94, 95).

Par ailleurs, les sujets ont tous reçu une indemnité pour leur participation.

II-A-2. Stimuli

Les stimuli utilisés sont les mêmes que ceux qui ont fait l'objet de la procédure d'égalisation en sonie. Une liste récapitulative des sons et de leur nature se trouve en annexe B.

II-A-3 Appareillage

L'appareillage est identique à celui utilisé pour l'égalisation en sonie.

II-A-4. Procédure

Les échantillons sonores sont présentés à l'auditeur par paire (120 paires) et celui-ci doit se prononcer sur la dissemblance perceptive entre les deux sons d'une même paire.

Pour cela, le sujet possède, sur un écran d'ordinateur, un potentiomètre horizontal qui correspond à une échelle continue de jugement allant de "très similaire", à gauche, à "très dissemblable", à droite. Il effectue son jugement en déplaçant un curseur sur le potentiomètre, à l'aide de la souris. Le sujet peut écouter autant de fois qu'il le désire la paire de sons, puis, une fois qu'il a validé son jugement – en activant sur l'écran un bouton prévu à cet effet –, une nouvelle paire est automatiquement présentée.

Les jugements sont codés sur une échelle continue allant de 0 ("très similaire") à 1 ("très dissemblable"). A chaque essai, l'ordinateur stocke en mémoire le nom des échantillons formant la paire et le jugement du sujet.

L'ensemble des paires de sons différents est présenté en ordre aléatoire pour chaque auditeur. Les sons de chaque paire ne sont présentés que dans un seul ordre choisi au hasard lors de chaque essai.

Dans une phase préliminaire, le sujet a écouté tous les échantillons dans un ordre aléatoire, afin de se faire une idée de la gamme de variation possible. Ensuite, il s'est familiarisé avec la tâche sur une dizaine d'essais d'entraînement tirés au hasard de l'ensemble des paires de l'étude.

La tâche principale, portant sur les 120 paires, a duré en moyenne 76 minutes avec un minimum de 46 minutes (sujet n° 27) et un maximum de 109 minutes (sujet n° 20). Cette forte variation de la durée d'expérience est révélatrice d'une première différence de stratégie parmi les sujets qui optent soit pour un jugement direct sans réécoute, soit pour un jugement réfléchi où chaque paire est systématiquement réécoutée au moins une fois ; en moyenne, la plupart des sujets répondent très rapidement sur des paires "faciles" et s'attardent uniquement sur les paires qui leur semblent litigieuses.

A l'issue du test de dissemblance, une tâche d'identification est effectuée. Puis, le sujet donne ses impressions sur la séance en répondant à trois questions : longueur, difficulté et intérêt du test. Il remplit ensuite un questionnaire qui apporte des informations sur son environnement sonore quotidien et qui comporte une ultime question où il lui est demandé d'expliquer la stratégie de jugement employée et éventuellement de définir les critères perceptifs utilisés (un exemplaire de tous les documents concernant cette expérience se trouve en annexe C).

A la lecture des premières réactions, il ressort globalement que le test d'écoute a été perçu comme "intéressant" d'une part, et "divertissant" voire "amusant" d'autre part.

Les avis sont davantage partagés en ce qui concerne la longueur du test : environ la moitié des personnes a estimé que le test était un peu long alors que les autres l'ont considéré de durée convenable. Par ailleurs, toute opinion confondue à propos de la durée, les sujets ont généralement observé une pause au cours du test principal et plus rarement deux pauses et plus.

Enfin, sur la difficulté du test, les opinions sont variables : certains le trouvent "facile" ou, au contraire, "pas évident", ou bien encore "difficile au départ puis de plus en plus facile" (effet d'accoutumance aux sons), d'autres considèrent que cela demande un réel "effort de concentration", d'autres encore ont hésité sur des paires "délicates" sur lesquelles ils ne sont pas sûrs de leurs réponses, d'autres enfin se sont rendus

compte de l'évolution de leur stratégie de jugement au cours du test et, à ce titre, pensent s'être contredits à plusieurs reprises, etc. ...

Par ailleurs, des observations d'ordre plus général sont intéressantes à prendre en compte : la difficulté à se détacher de l'identification, la sensation que certains sons sont artificiels dans la mesure où ils n'ont pas de réalité écologique (par exemple, un extrait de 10 secondes de trafic routier fluide), la bonne qualité globale de la prise de son et, pour certains, la perception d'une différence d'intensité entre certains sons.

II-A-5. Résultat de la tâche d'identification

La tâche d'identification revêt, dans le cadre de notre étude, un intérêt de tout premier ordre dans la mesure où elle permet de faire la preuve de l'intuition préalable que nous avons à propos de la reconnaissance des stimuli sonores.

En effet, étant donné la nature très marquée des sons de l'expérience et du fait également qu'il s'agisse de sonorités familières à la plupart des gens, il paraissait inévitable que les sujets identifient les sources sonores des échantillons de l'expérience ; les résultats de la tâche d'identification (cf. annexe D) confirme radicalement cette hypothèse puisque, en moyenne, le taux de reconnaissance des stimuli s'élève à 85 %.

Ainsi, malgré la consigne – systématiquement rappelée – d'essayer de passer outre le réflexe d'identification pour s'intéresser uniquement à la sonorité propre de l'échantillon, il semble évident que le processus de reconnaissance des sources sonores a joué un rôle prépondérant dans l'évaluation de la dissemblance chez un grand nombre de sujets. Ces constatations sont d'ailleurs confortées par les remarques qui relatent la "difficulté à se détacher de l'identification" (cf. §II-A-6).

En effet, la reconnaissance des sources sonores associées aux stimuli induit la catégorisation des sons suivant la nature de ces sources : en d'autres termes et de manière simplifiée, l'auditeur classe par automatisme 'les trains avec les trains' et 'les voitures avec les voitures'. Il effectue ensuite son jugement de dissemblance en étant probablement inconsciemment conditionné par cette structuration instantanée et l'on peut concevoir qu'il rencontre alors quelques difficultés à se focaliser sur des critères perceptifs plus généraux à travers tous les échantillons. Bien entendu cette description est un scénario moyen mais au vu des résultats d'analyse que nous présenterons plus loin (cf. §II-B-1), il semble que cela ait été la stratégie utilisée par beaucoup de sujets.

D'un point de vue théorique, la notion de catégorisation est pertinente dans la mesure où elle fait référence à certains modèles de perception : ceux qui postulent que la représentation mentale du stimulus sonore passe très tôt, dans le processus de reconnaissance, d'une nature continue à une nature discrète ; l'information sensorielle étant décomposée en un certain nombre de traits discrets qui sont ensuite intégrés avant d'être comparés à des catégories mémorisées.

Ceci étant, ces modèles sont en contradiction avec l'approche multidimensionnelle que nous avons choisie puisque celle-ci s'appuie précisément sur une représentation continue – un *continuum* – du stimulus au niveau des paramètres psychoacoustiques le décrivant. L'ambiguïté de la situation apparaît nettement dans les résultats de l'analyse multidimensionnelle (cf. §II-B-1) et nous amène à considérer d'autres types d'analyses reposant sur des modèles non-spatiaux qui formalisent la similarité selon un appariement de propriétés discrètes partagées par les stimuli : il s'agit essentiellement des analyses de cluster qui sont présentées et détaillées au §I-B-2.

II-A-6. Examen des données sémantiques recueillies

La description verbale, par les sujets, des critères perceptifs utilisés présente, d'une manière générale, la stratégie de jugement employée, et fournit, le cas échéant, des indices précieux sur la nature des paramètres physiques susceptibles d'expliquer les dimensions perceptives. C'est la raison pour laquelle la dernière question du questionnaire a fait l'objet d'une retranscription intégrale présentée en annexe E.

A la lecture de ces remarques, on note immédiatement que beaucoup de qualificatifs utilisés pour décrire les échantillons sonores appartiennent au domaine de description de l'aspect temporel : la constance, la répétition, l'accélération, la vitesse de défilement, l'évolution temporelle, le mouvement, la stabilité, la progression, la coulée du son, le dynamisme, la continuité/discontinuité, la montée/descente du bruit, etc. ...

Par ailleurs, des impressions se rapprochant davantage d'une écoute musicale des sons se traduisent par : aigu, grave, timbre, rythme, tonalité, hauteur, etc. ...

Enfin, quelques notions originales sont intéressantes à retenir : lourdeur/légèreté, sons unis/non unis, homogénéité, son produit par un seul émetteur/somme de sons émis par plusieurs sources, son isolé/non isolé, agressivité.

B- Analyse des données de dissemblance

II-B-1. Première série d'analyses

II-B-1-a. Principe de l'analyse multidimensionnelle

Les résultats d'expérience sont arrangés sous forme matricielle $D(i,j)$ (demi-matrice triangulaire, sans diagonale) qui donne accès, pour chaque sujet, à la valeur δ_{ij} de la dissemblance entre les objets i et j ; cette matrice constitue la donnée d'entrée principale du logiciel d'analyse.

Puis des algorithmes de calcul permettent d'obtenir un espace géométrique où les objets sont positionnés en fonction de leurs dissemblances mutuelles. Pour construire cette représentation mathématique, la théorie de l'analyse suppose l'existence d'une relation monotone entre les valeurs expérimentales de dissemblance δ et les distances calculées d afin de conserver une cohérence dans la variation des deux ensembles. En d'autres termes, lorsque deux objets sont jugés éloignés, le modèle les écarte d'une distance en rapport avec cette perception et suit cette même logique avec deux objets jugés très proches.

D'un point de vue mathématique, il existe plusieurs modèles de calcul pour exprimer la distance entre deux points de l'espace d'analyse. Les principaux sont :

- le modèle euclidien classique où l'on suppose que les stimuli présentent des attributs communs et où la distance d_{ij} est calculée grâce aux projections x_{in} et x_{jn} des objets i et j sur les N dimensions de l'espace, suivant la formule :

$$d_{ij} = \sqrt{\sum_{n=1}^N (x_{in} - x_{jn})^2} .$$

- le modèle euclidien pondéré qui inclut dans son code de calcul des pondérations propres à un sujet et une dimension. En fait, par l'intermédiaire de coefficients de pondération w_{ns} , cette approche tient compte de l'importance perceptive que le sujet s attribue à la dimension n mais, par la même occasion, fait disparaître l'invariance rotationnelle qui existe dans le modèle précédent. On définit donc, pour le sujet s , la distance sous la forme :

$$d_{ijs} = \sqrt{\sum_{n=1}^N w_{ns} \cdot (x_{in} - x_{jn})^2} .$$

- le modèle euclidien étendu où l'on suppose que les stimuli ont à la fois des dimensions perceptives partagées et une spécificité propre qui traduit la présence plus ou moins saillante d'une caractéristique non-généralisable à l'ensemble des stimuli (par exemple, un parasite haute fréquence). La distance entre deux objets s'écrit alors, en tenant compte, ici encore, de l'effet pondérateur induit par chaque sujet :

$$d_{ijs} = \sqrt{\sum_{n=1}^N w_{ns} \cdot (x_{in} - x_{jn})^2 + v_s \cdot (S_i + S_j)}, \text{ où } S_i \text{ est la spécificité de l'objet } i,$$

v_s est la pondération sur l'ensemble
des spécificités par le sujet s

Le logiciel EXSCAL [Winsberg & Carroll, 1989] permet d'utiliser ce modèle pour l'analyse multidimensionnelle avec, en outre, la possibilité de choisir un calcul avec/sans pondération et avec/sans spécificité.

Le logiciel CLASCAL [Winsberg & De Soete, 1993a], quant à lui, est fondé sur ce même modèle hybride qui permet à la fois de prendre en compte un modèle spatial, avec spécificités, comme ceux décrits plus haut, mais aussi d'évaluer l'existence de classes de sujets en calculant des pondérations de classes et non plus des pondérations de sujets isolés.

II-B-1-b. Etapes de l'analyse multidimensionnelle

Pour la présente étude, l'analyse des données est effectuée avec l'algorithme CLASCAL qui permet de découvrir les classes latentes et donc, le cas échéant, de classifier les sujets. Dans un premier temps, on étudie la structure des classes latentes de sujets. Pour être fiable, cette analyse requiert un minimum de 10 sujets par classe ; ayant 30 sujets, nous testerons donc l'existence de 1 ou 2 classes.

Simultanément, le programme procède à la détermination de la configuration spatiale des données : le nombre de dimensions partagées et l'inclusion éventuelle des spécificités (cf. §II-B-1-a). Le choix entre les différents modèles spatiaux est d'abord basé sur le critère d'information BIC qui prend en compte l'ajustement entre le modèle et les données brutes (estimé par le Logarithme de Vraisemblance), le nombre de paramètres du modèle et le nombre d'observations (nombre de données). Le modèle ayant la plus petite valeur de BIC est considéré comme le plus approprié. En général, pour une analyse qui ne présente pas un problème de minima locaux, ce critère diminue avec l'augmentation du nombre de paramètres du modèle jusqu'à un certain point puis commencent à augmenter. En fait, le BIC constitue un premier également indicatif dans la sélection du modèle approprié et l'on complète cette démarche avec

des simulations Monte-Carlo entre les modèles qui semblent être de bons candidats à la description des données. Ceci étant, pour déterminer définitivement la configuration optimale d'analyse, on se fonde également sur l'interprétabilité des dimensions et spécificités fournies par le modèle.

Si l'analyse des classes latentes indiquent plus d'une classe, chacune des dimensions sera pondérée par les différentes classes et ainsi chacun des axes de la représentation spatiale sera perceptivement orienté.

Si au contraire, l'analyse des classes latentes indiquent une seule classe, les configurations produites par l'analyse multidimensionnelle seront invariantes par rotation. Dans ce cas, on effectue une analyse supplémentaire avec EXSCAL dans la configuration spatiale optimale choisie (estimant les poids pour tous les sujets), afin de trouver une position de l'espace convenable pour interpréter les résultats obtenus : ce modèle multi-classes est peu économique en lui-même, mais il donne l'orientation des axes. Ensuite nous utilisons cette solution comme configuration cible et nous effectuons une rotation et mise à l'échelle (rotation dite 'procrustienne') de la configuration à classe unique vers la configuration cible en minimisant la moyenne des erreurs au carré entre les deux configurations. La procédure procrustienne se limite aux dimensions partagées et n'est pas appliquée aux spécificités. Elle représente en fait une simple rotation des coordonnées et ne modifie nullement la structure des relations entre les objets.

Finalement, l'analyse multidimensionnelle fournit les coordonnées de chaque objet du test dans l'espace perceptif correspondant à la configuration optimale d'analyse.

II-B-1-c. Premiers résultats

En pratique, les premières analyses multidimensionnelles sont menées avec l'ensemble des données (16 stimuli - 30 sujets) suivant une procédure en 4 étapes :

- 1^{ère} étape : on estime le nombre de classes latentes de sujets ;
- 2^{ème} étape : avec le nombre de classes trouvé à la 1^{ère} étape, on teste le modèle spatial en faisant varier le nombre de dimensions et le paramètre de spécificité : la valeur de l'indice BIC permet de dégager 1, 2 ou 3 bons candidats ;
- 3^{ème} étape : on compare les bons candidats trouvés à la 2^{ème} étape pour définir le meilleur d'entre eux, en ce qui concerne l'interprétabilité des dimensions ;
- 4^{ème} étape : avec le modèle spatial optimal dégagé à la 3^{ème} étape, on revérifie le nombre de classes définit à la 1^{ère} étape.

Le résultat obtenu est une configuration à 1 classe et 2 dimensions avec spécificités. Il se visualise sous la forme suivante :

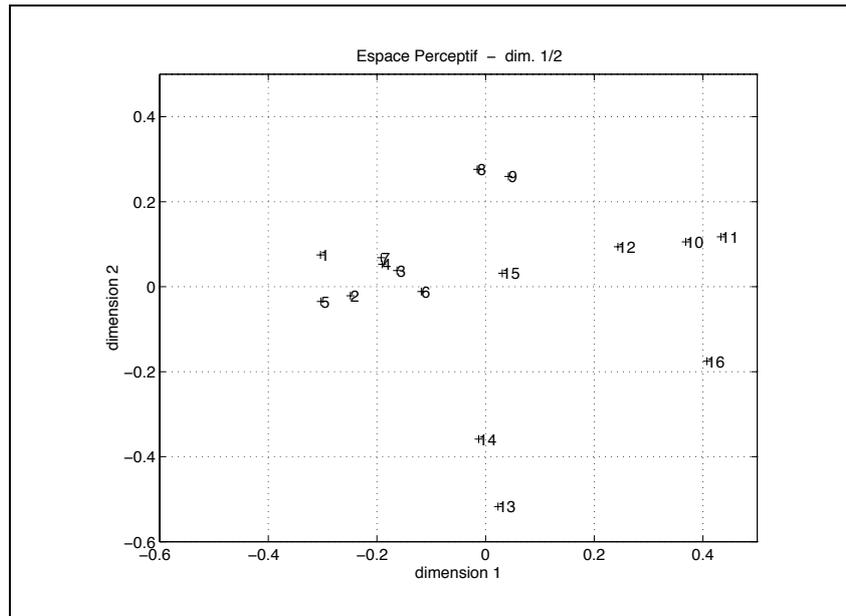


Fig. 1 : Solution optimale de l'analyse multidimensionnelle (16 stim.- 30 suj.)
Projection de l'analyse de cluster - 17 suj. (cf. §II-B-2-b) sur l'analyse multidim.

Sachant que cette représentation ne fait état que des dimensions partagées, il est important d'y adjoindre la valeur des spécificités pour chaque stimulus (en fait, la racine carré des spécificités pour que les valeurs soient homogènes à une distance et donc comparables aux coordonnées ci-dessus) :

son 1	son 2	son 3	son 4	son 5	son 6	son 7	son 8
0.00	0.22	0.31	0.21	0.07	0.32	0.21	0.23

son 9	son 10	son 11	son 12	son 13	son 14	son 15	son 16
0.16	0.11	0.16	0.24	0.00	0.36	0.55	0.35

Par ailleurs, il est intéressant de forcer légèrement ce résultat pour obtenir une représentation tridimensionnelle, ce qui nous amène à un modèle à 1 classe, 3 dimensions sans spécificité de la forme suivante :

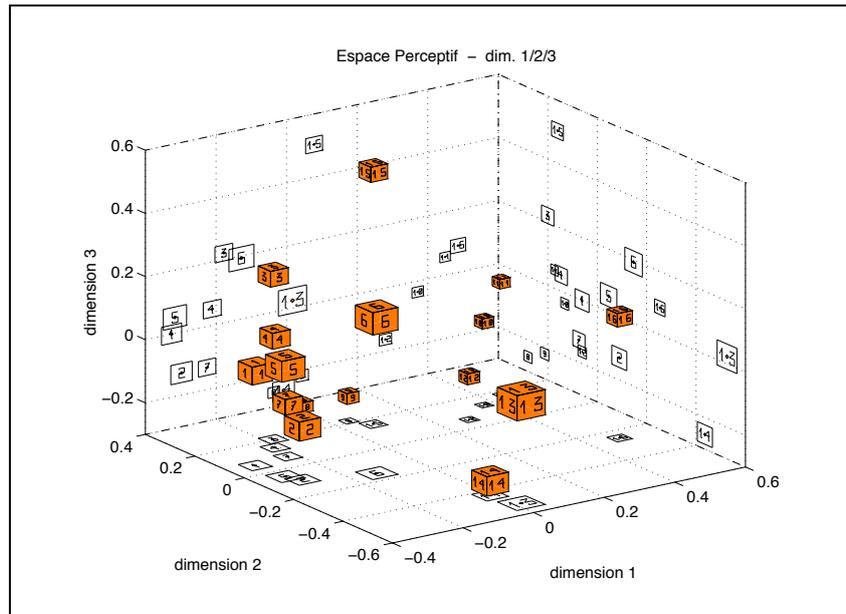


Fig. 2 : Solution forcée de l'analyse multidimensionnelle (16 stim.- 30 suj.)

Cette figure permet de constater la robustesse de la première configuration à 2 dimensions, puisque la projection de l'espace à 3 dimensions sur le plan (1,2) redonne quasiment la même solution ; la troisième dimension peut expliquer, quant à elle, partiellement certaines spécificités (les différentes projections de l'espace sont fournies en annexe F)

La première observation que l'on peut faire sur ce résultat global rejoint les idées amorcées au moment de l'exploitation de la tâche d'identification (cf. §II-A-6) : l'espace perceptif optimal, à 2 dimensions, présente une forte caractéristique de catégorisation et, qui plus est, les classes de stimuli correspondent quasiment aux différents modes de transport : ferroviaire, aérien, routier, urbain. Cette distribution 'par paquets' et la disposition 'en croix' des objets rend l'espace mal défini dans certaines zones et entraîne, par conséquent, une difficulté à généraliser l'interprétation des axes perceptifs en terme de paramètres physiques : dans l'espace multidimensionnel de la fig. 1 (cf. §II-B-1-c), le fait d'interpréter la première dimension, par exemple, permettrait difficilement de conclure puisque la structure 'en croix' empêche de vérifier si cette interprétation est dépendante ou non de la seconde dimension.

Par ailleurs, une première écoute des sons suivant chacune des dimensions ne suggère pas clairement l'existence d'un indice psychoacoustique variant continûment sur l'un ou l'autre des axes.

Néanmoins la dimension 1 pourrait s'exprimer en terme de 'lisse/perturbé' et être relié à un paramètre relatif à l'enveloppe temporelle du signal. Cela étant, certains stimuli viennent infirmer cette logique d'écoute : le son 15 (RER) s'intègre difficilement dans ce *continuum*, peut-être à cause de sa forte spécificité (0.55), et l'attribut commun aux sons 8, 9 (trafic autoroutier), 15 (RER), 14 et 13 (trafic aérien) qui sont presque confondues sur cette dimension, est très difficile à extraire.

De la même manière, la compréhension de la seconde dimension passe par la définition d'un attribut commun aux sons 1 à 7 (trafic ferroviaire), 10 à 12 (trafic

urbain) et 15 (RER) et le petit nombre d'éléments disposés par ailleurs sur cette dimension ne permet pas d'avoir une écoute comparative fiable laissant le problème, pour l'instant, sans solution.

A ce stade de l'étude, il semble donc opportun d'affiner notre analyse sur les données de dissemblance. Plusieurs directions sont prises :

- opérer d'autres types d'analyses sur les stimuli pour confirmer la nature catégorielle des données de dissemblance (cf. §II-B-2-b),
- déterminer des groupes de sujets présentant des caractéristiques de jugement proches afin d'examiner la stabilité de la solution générale à deux dimensions (cf. §II-B-2-c),
- s'intéresser à des groupes de stimuli correspondant aux sous-ensembles de l'espace perceptif et étudier leur structure interne (cf. §II-B-2-c).

Pour ce faire, nous nous intéressons, dans un premier temps, à une nouvelle famille d'analyse qui, à l'inverse de l'approche multidimensionnelle, repose sur un modèle non spatial : l'analyse de cluster.

II-B-2. Deuxième série d'analyses

II-B-2-a. Principe des analyses de cluster

L'analyse théorique des similarités peut être menée en termes de traits communs plutôt qu'en termes de dimensions continues. Une telle conception aboutit à représenter les relations de proximité entre des objets non plus par des distances les séparant dans un espace géométrique mais en déterminant des groupes de stimuli (clusters) qui partagent les mêmes traits. Cela revient soit à diviser l'ensemble initial d'objets en plusieurs sous-ensembles relativement homogènes éventuellement hiérarchiquement organisés à l'intérieur desquels les objets sont perçus comme similaires, soit à déterminer la meilleure structure en arbre dans laquelle la proximité correspond à la hauteur ou à la longueur du lien joignant deux objets.

L'interprétation de telles structures consiste à identifier les traits que partagent les objets appartenant à un même sous-groupe. Autrement dit, ce type d'analyse nous permet de déterminer les traits ou propriétés d'un domaine de stimuli donné.

On utilisera, pour notre étude, deux approches différentes :

- Les clusters additifs correspondent à une organisation non hiérarchique des clusters mais dont les sous-groupes d'objets peuvent présenter des recouvrements partiels [Arabie & Shepard, 1973 ; Carroll, 1976 ; Shepard & Arabie, 1979 ; Carroll & Arabie, 1979 ; Shepard, 1980 ; Carroll & Arabie, 1979 ; Arabie & Carroll, 1980 ; De Sarbo, 1982].

L'hypothèse de base de cette méthode est que tous les objets ont des propriétés auxquelles sont associées des poids (ou pondérations). La similarité perceptive S_{ij} entre deux stimuli i et j correspond à la somme des poids associés à ces propriétés discrètes, W_k , que les deux stimuli partagent. Dans ce cas, un sous-groupe – ou cluster – d'objets est la réunion des objets possédant le même trait ou la même propriété. Un tel modèle peut s'écrire sous la forme :

$$S_{ij} = \sum_k W_k P_{ik} P_{jk}$$

où $P_{ik} = 1$ si l'objet i possède la propriété k et W_k est une valeur de pondération non-négative représentant la saillance de la propriété k . Si les deux objets sont dans un sous-ensemble donné (c'est-à-dire qu'ils partagent le même trait k), alors $P_{ik} P_{jk} = 1$, mais si l'un des deux objets appartient à un autre sous-ensemble le produit est nul.

- L'analyse de cluster hiérarchique, appelée également arbre ultramétrique selon le type de représentation, où dans ce cas, le recouvrement des sous-groupes est possible et les objets peuvent appartenir à plusieurs clusters de différents niveaux hiérarchiques [Johnson, 1967 ; Hartigan, 1967, 1975 ; Carroll, 1976 ; Shepard, 1980 ; Carroll & Arabie, 1979].

Un arbre ultramétrique représente la similarité entre deux objets, correspondant aux noeuds extérieurs uniquement, par la hauteur du lien les joignant. Les objets les plus similaires sont reliés au plus bas niveaux de l'arbre et les objets les plus dissemblables au niveau le plus haut. Par exemple, sur la fig. 3 ci-dessous, l'objet C possède beaucoup de caractéristiques communes avec l'objet D mais peu avec l'objet B qui, quant à lui, est plus semblable à l'objet E.

Un tel arbre est appelé 'ultramétrique' car il doit satisfaire l'inégalité ultramétrique selon laquelle pour tout x, y et z :

$$d(x,y) \leq \max\{d(x,z), d(y,z)\}, \quad \text{où } d(x,y) \text{ est la distance entre } x \text{ et } y.$$

Autrement dit, deux objets reliés à un noeud donné doivent être tous deux également distants d'un troisième objet auquel ils sont reliés à un noeud supérieur. Ceci implique que tous les objets sont à égal distance de la racine, correspondant au noeud le plus haut de l'arbre (cf. figure ci-dessous).

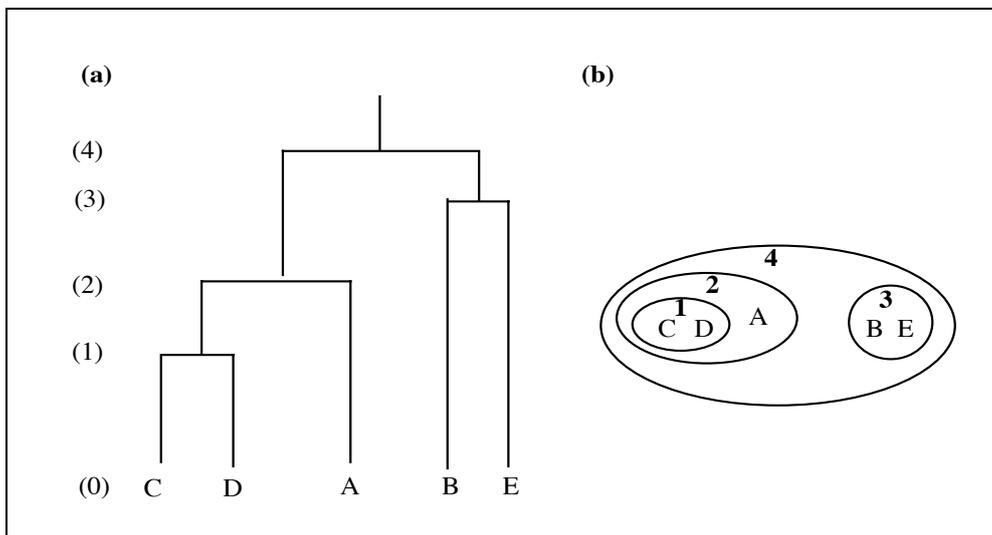


Fig. 3 : Illustration de la représentation en arbre ultramétrique

II-B-2-b. Mise en oeuvre des analyses de cluster

Les données sont d'abord analysées par un algorithme de clusters additifs (INDCLUS). Les premiers résultats obtenus semblent difficilement exploitables quant à la composition d'éventuels groupes de stimuli, ou d'éventuels groupes de sujets. C'est la raison pour laquelle nous nous tournons vers le procédé de représentation par arbre ultramétrique qui paraît mieux adapté au cadre de notre étude.

Dans un premier temps, on construit un arbre ultramétrique de la population des 30 sujets, en considérant, comme données d'entrée de l'analyse, la matrice de corrélation entre les individus (équivalente, dans ce cas, à une matrice de similarité). Cela permet de dégager un premier sous-ensemble de 17 sujets (1^{er} sous-groupe) : 1 à 3, 5, 8, 14 à 16, 18, 19, 21 à 25, 27, 30 parmi lesquels 11 sujets peuvent encore être extraits (2^{ème} sous-groupe): 3, 5, 8, 15, 16, 18, 19, 22, 24, 25, 30 (voir annexe G). Ces sous-populations de sujets plus cohérents entre eux vont permettre de réanalyser les données initiales (cf. II-B-2-c) et d'étudier la stabilité de la première solution obtenue.

Notons, au passage, que l'examen des questionnaires de chacun de ces sujets ne fait apparaître aucune caractéristique commune tant au niveau des critères de base (age, sexe, lieu d'habitation) que de l'environnement sonore quotidien.

Par ailleurs, la représentation en arbre ultramétrique des 16 stimuli, avec les données des 30 sujets, nous donne cette fois la structure catégorielle des échantillons sonores au sens de l'analyse de cluster (voir document en annexe H). On observe alors une très forte similitude entre la composition des clusters issus de cette analyse et celle des clusters qui se dessinent visuellement dans l'espace perceptif : seul le son n° 3 (TGV Réseau, ligne classique) se distingue légèrement puisque dans l'arbre ultramétrique, il est plus proche du son n° 15 (RER) que du groupe des sons n° 1 à 7 (trafic ferroviaire) alors que c'est plutôt le contraire dans l'espace perceptif.

Ceci étant, cette similitude se renforce si l'on choisit d'effectuer l'analyse de cluster toujours sur les 16 stimuli mais avec les données des 17 sujets sélectionnés précédemment (cf. fig. 1) ; et l'ambiguïté constatée à propos du son n° 3 disparaît si l'on considère uniquement les 11 sujets du 2^{ème} sous-groupe (cf. annexe H).

Ces observations constituent une preuve supplémentaire dans l'affirmation de la structure catégorielle discontinue représentée au sein de l'espace perceptif multidimensionnel.

II-B-2-c. Analyses multidimensionnelles partielles

Il s'agit ici d'effectuer des analyses multidimensionnelles partielles par rapport aux données de départ (16 stimuli, 30 sujets), pour, d'une part, étudier la stabilité de l'espace perceptif à 16 stimuli (sous-groupes de sujets) et, d'autre part, déterminer éventuellement la représentation continue d'un sous-ensemble de stimuli.

L'analyse multidimensionnelle effectuée sur la population réduite à 17 puis 11 sujets (cf. §II-B-2-b) donne la même configuration optimale de résultats que celle qui a été obtenue avec la population totale : 1 classe, 2 dimensions avec spécificités. En outre, les trois espaces perceptifs (cf. §II-B-2-c et annexe I) ne présentent que des variations mineures qui peuvent être apparentées à du bruit expérimental.

Ces deux derniers résultats finissent donc de confirmer la stabilité de la solution élaborée au §II-B-1 et confirment l'allure catégorielle de l'espace perceptif.

Parallèlement à cela, des analyses multidimensionnelles sont menées sur différents sous-groupes de stimuli constitués à partir des regroupements qui se déduisent visuellement de l'espace perceptif (cf §II-B-1-c - fig. 1) :

- un ensemble à 12 stimuli : sons n° 1 à 12 ;
- un ensemble à 11 stimuli : sons n° 1 à 7, sons n° 10 à 12, son n° 15 ;
- un ensemble à 10 stimuli : sons n° 1 à 9, son n° 15 ;
- un ensemble à 7 stimuli : sons n° 1 à 7 ;

afin de tenter de répondre à la question suivante : y-a-t-il une structure fine continue à l'intérieur de certaines catégories ?

A ce propos, il est important de noter que ces analyses s'effectuent sous l'hypothèse d'une équivalence entre les résultats de dissemblance issus d'un test à 11 stimuli, par exemple, et ceux extraits du test global et concernant ces mêmes 11 stimuli. En s'affranchissant de cette hypothèse, on peut dès lors se pencher sur les résultats obtenus avec les différentes compositions décrites précédemment.

La configuration optimale d'analyse pour les ensembles à 12, 11 et 10 stimuli est un espace à 2 dimensions, avec spécificités sachant qu'avec ce nombre réduit d'objets, il est de toute façon impossible d'aller au-delà de 2 dimensions, pour des raisons de fiabilité de l'algorithme d'analyse.

En outre, les espaces perceptifs respectivement obtenus (cf. annexe J) n'apportent pas d'informations vraiment nouvelles par rapport à la solution globale du §II-B-1 dans la mesure où la structure des clusters 'visuels' de celle-ci est quasiment conservée ; ceci, en considérant que seule compte la variation des données perceptives sur un axe de l'espace et que l'on peut donc opérer des renversements symétriques sur chacun d'entre eux. Par ailleurs, les petites dilatations observées peuvent s'interpréter en termes de bruit expérimental pas suffisamment significatif pour être exploitable.

L'analyse du sous-ensemble à 7 stimuli (trafic ferroviaire) paraît, quant à elle, plus riche d'enseignements. En effet, l'espace à 2 dimensions obtenu révèle une distribution sensiblement différente de celle de l'analyse globale examinée sur les mêmes stimuli :

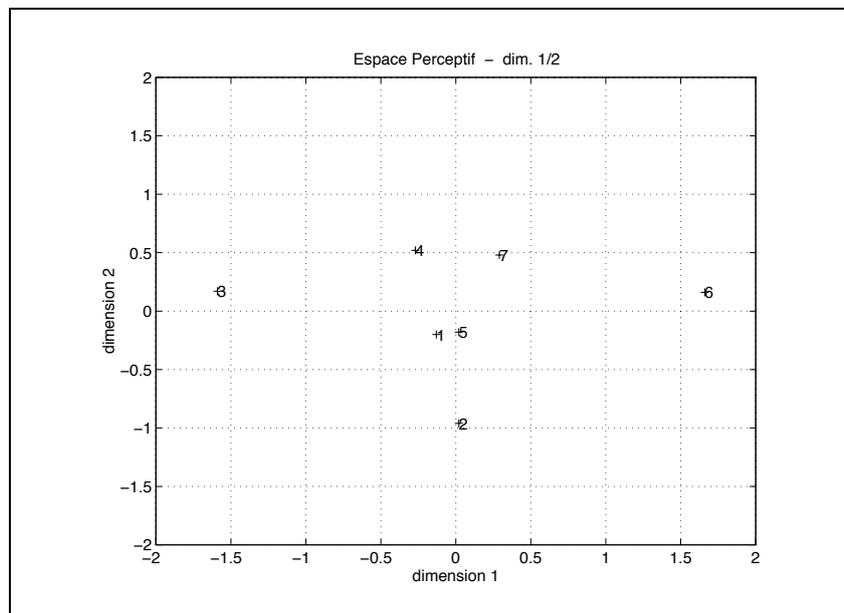


Fig. 4 : solution de l'analyse multidimensionnelle sur les sons de trafic ferroviaire

La différence importante entre l'analyse à 16 stimuli et celle à 7 stimuli (trafic ferroviaire) montre que le fait de focaliser l'analyse uniquement sur les éléments d'une catégorie a provoqué une réorganisation de celle-ci. De ce fait, on peut tout à fait concevoir que cette structure bidimensionnelle dévoilée par l'analyse partielle a été écrasée dans l'analyse globale par un attribut perceptif beaucoup plus saillant que l'on peut définir, par exemple, comme l'identification des sources sonores.

Enfin, l'ensemble de l'étude de dissemblance présentée dans ce chapitre met à jour, pour les sons de l'étude, une structure perceptive très fortement catégorielle due essentiellement à des facteurs cognitifs telle que la reconnaissance des sources sonores. Les analyses de clusters, précisément fondées sur le concept d'une organisation catégorielle de la perception auditive, confirment cette conclusion. Par ailleurs, l'étude d'un sous-ensemble du corpus d'échantillons sonores (sons de trafic ferroviaire), plus homogène et donc moins assujéti à ce facteur de reconnaissance semble faire apparaître des dimensions perceptives davantage continues ; néanmoins, cette analyse partielle souffre du nombre restreint de stimuli (7 sons) ce qui empêche l'établissement de conclusions fiables ainsi que l'approfondissement de l'étude en termes d'interprétation de l'espace perceptif par des paramètres physiques.

III. ETUDE DE DOMINANCE

L'objectif de cette seconde étude était initialement de déterminer le lien entre un facteur de dominance ressenti chez les sujets et les paramètres physiques qui sous-tendent l'espace perceptif multidimensionnel ; la dominance se traduit par un attribut sémantique et peut être exprimée, par exemple, en terme de préférence, de confort sonore ou de gêne.

Cependant, la structure catégorielle de l'espace perceptif révélée au chapitre précédent et la difficulté qui en découle pour relier les axes perceptifs de l'espace global à des paramètres physiques, nous oblige à bouleverser la méthodologie théorique de l'étude et à exploiter différemment les valeurs de dominance : la théorie de l'analyse de dominance [De Soete & Winsberg, 1993b] nécessite l'existence de "facteurs objectifs" (paramètres physiques et psychoacoustiques) pour caractériser les échantillons sonores, or dans le cas présent, les données de dominance vont être exploitées indépendamment de la relation à un paramètre physique grâce à un algorithme qui permet de construire une hiérarchie entre les sons lisible sur la seule échelle sémantique choisie comme base de l'expérience (cf. §III-B-1). L'axe perceptif ainsi construit sera alors plus pertinent et donc plus facilement interprétable en termes d'indices psychoacoustiques.

A- Méthode expérimentale

III-A-1. Echelle sémantique de l'expérience

Le facteur de dominance a été établi, comme convenu, en collaboration avec la S.N.C.F. : il définit une échelle de désagrément pour les échantillons sonores de l'étude. Ainsi, la question sur laquelle repose le test de dominance est la suivante :
" Parmi ces deux sons, lequel trouvez-vous le plus désagréable ? "

III-A-2. Informations sur les sujets

Le test de dominance s'est effectué entre le 27 janvier et le 12 mars 1997. Il porte, cette fois, sur 60 sujets car l'échelle de jugement n'est plus continue – comme pour le test de dissemblance – mais binaire – "son 1" ou "son 2" – ce qui demande davantage de données pour obtenir une certaine stabilité des résultats. Les auditeurs sont recrutés dans la même base de données que pour l'étude de dissemblance (cf. §II-A-1). Les sujets, 30 hommes et 30 femmes, âgés de 17 à 61 ans, se répartissent de la manière suivante :

- 24 personnes ont moins de 25 ans (12 femmes, 12 hommes),
- 19 personnes ont entre 25 et 35 ans (9 femmes, 10 hommes),
- 9 personnes ont entre 35 et 45 ans (3 femmes, 6 hommes),
- 8 personnes ont plus de 45 ans (5 femmes, 3 hommes).

En outre, comme pour le test de dissemblance, il nous a semblé opportun de retenir comme critère le lieu d'habitation des sujets ; ainsi, la population d'auditeurs

est constituée de 30 personnes habitant Paris 'intra-muros' et 30 personnes habitant en banlieue parisienne et province (départements 91, 92, 93, 94, 95, 35). Par ailleurs, parmi les 60 personnes qui ont effectué ce test, 24 d'entre elles avaient déjà participé à l'expérience précédente concernant la dissemblance. Enfin, les sujets ont été indemnisés d'une somme forfaitaire pour leur participation.

III-A-3. Stimuli

Les stimuli utilisés sont les mêmes que ceux qui ont fait l'objet de la procédure de test de dissemblance. Une liste récapitulative des sons et de leur nature se trouve en annexe B.

III-A-4. Appareillage

L'appareillage est identique à celui utilisé pour l'étude de dissemblance.

III-A-5. Procédure

Au début de la séance le sujet entend tous les échantillons sonores dans un ordre quelconque. Ensuite toutes les paires d'échantillons sont présentées suivant une séquence aléatoire, dans les deux ordres.

A chaque essai le sujet entend la paire d'échantillons une seule fois et doit choisir celui qui domine l'autre sur l'échelle choisie. Son choix est enregistré en cliquant sur le bouton correspondant de l'interface graphique de PsiExp (SI-c). Pour l'ensemble des sujets, une seule demi-matrice triangulaire sans diagonale est construite, elle représente la proportion de sujets qui ont préféré un échantillon à un autre.

Une phase préliminaire comportait l'écoute des 16 sons de l'expérience et une séance d'entraînement sur 10 paires.

La tâche principale, portant sur les 120 paires, a duré en moyenne 48 minutes avec un minimum de 43 minutes (sujets n° 24, 40, 43) et un maximum de 57 minutes (sujets n° 19, 54).

Puis, le sujet donne ses impressions sur la séance en répondant à deux questions – longueur et difficulté du test – et peut ensuite faire des remarques sur l'expérience qu'il vient de passer – pertinence de la question, nature des réponses (instinctive, affective, analytique), etc. ... Il remplit enfin le même questionnaire que celui utilisé pour le test de dissemblance mais cette fois-ci sans la dernière question – concernant la stratégie de jugement – car celle-ci est abordée oralement et présente un intérêt mineur pour l'analyse des données de dominance (un exemplaire de tous les documents nouveaux concernant cette expérience se trouve en annexe K).

III-A-7. Examen des données sémantiques recueillies

L'examen des commentaires nous conforte, en premier lieu, dans le choix de l'échelle sémantique : le mot 'désagréable' est globalement jugé "adapté" et "pertinent" même si quelques personnes estiment qu'ils ont répondu en inversant la question et si quelques autres ont proposé les mots "irritant", "nuisible" ou "dérangeant".

Par ailleurs, en moyenne l'expérience n'est perçue ni trop longue, ni trop ennuyeuse et, ce qui vient conforter ces impressions, les sujets font très souvent une réponse de type affective ("j'aime plus les trains que les avions", "je n'aimerais pas habiter près d'un aéroport", etc. ...) et beaucoup moins souvent une réponse instinctive ou analytique.

Enfin, on retrouve ici des phénomènes observés lors du test de dissemblance comme l'association du son et de la source sonore, l'influence de l'ordre des échantillons dans la paire (effet de mémoire ou d'impression instantanée) ou bien, ce qui est plus spécifique à cette expérience, la mise en contexte du son avant d'effectuer un quelconque choix. Une retranscription des remarques se trouve en annexe L.

B- Analyse des données de dominance

III-B-1. Exploitation directe des données de dominance

Les données de dominance ont été traitées indépendamment de l'existence de paramètres physiques, grâce à une procédure qui permet de construire une hiérarchie entre les stimuli, à partir de leur proportion de dominance mutuelle.

Cette procédure est décrite par l'algorithme de Bradley-Terry-Luce [David, 1988] qui calcule le pourcentage de dominance pour chaque son – équivalent au nombre de fois où le son k a été choisi 'dominant' dans sa confrontation avec chacun des autres sons – et qui effectue une normalisation sur une échelle de 1 (le plus dominant) à -1 (le moins dominant). D'une manière plus précise, la méthode de Bradley-Terry-Luce fait l'hypothèse que la distribution des variables statistiques autour de la 'vraie' valeur de désagrément – pour chaque stimulus – peut s'approcher par une loi hyperbolique de variance constante. Ainsi, la probabilité π_{ij} que le stimulus i domine le stimulus j est donnée par l'équation suivante :

$$\pi_{ij} = \frac{1}{2 * \left(1 + \tanh\left(\frac{1}{2} * (\theta_i - \theta_j) \right) \right)}$$

où θ_i est la 'vraie' valeur de désagrément du stimulus i , mentionnée plus haut.

Les probabilités π_{ij} sont estimées à partir de la proportion de jugements observés p_{ij} ; en inversant l'équation précédente et en imposant la condition de somme nulle pour le résultat global, on obtient alors une valeur estimée de θ_i sous la forme :

$$\hat{\theta}_i = \log \left(\prod_{j \neq i} \left(\frac{p_{ij}}{p_{ji}} \right) \right)$$

Les valeurs de désagrément obtenues par ce calcul sont les suivantes :

Son	Etiquette	Valeur de désagrément	
			- désagréable
15	RER	- 0.28	
14	avion - atterrissage	- 0.27	
13	avion - décollage	- 0.22	
12	flux urbain	- 0.16	
9	autoroute fluide - avec PL	- 0.16	
16	autobus - départ	- 0.13	
11	voie urbaine - démarrage	- 0.10	
8	autoroute fluide - sans PL	- 0.09	
10	autoroute chargée	- 0.01	
7	train marchandise - électrique	0.01	
6	train marchandise - diesel	0.09	
5	train Corail	0.09	
4	TGV Réseau - ligne grande vitesse	0.16	
3	TGV Réseau - ligne classique	0.28	
2	TGV SudEst - ligne grande vitesse	0.30	
1	TGV SudEst - ligne classique	0.49	
			+ désagréable

L'examen et l'écoute de cette 'échelle de désagrément' appellent plusieurs remarques :

- l'effet de catégorisation est encore visible sur ce nouvel axe de perception : en effet, mis à part le son n° 15, tous les sons de trafic ferroviaire (n° 1 à 7) se situent en bas de l'échelle {'plus désagréable'}, les sons de trafic routier (n° 8 à 12 et n° 16) sont dans la partie médiane et les sons de trafic aérien (n° 13 et 14), dans la partie haute {'moins désagréable'}.
- cette organisation qui peut sembler surprenante à la première lecture, se justifie assez facilement à l'écoute où l'on perçoit beaucoup plus nettement que pour les axes de l'espace perceptif un *continuum* dans la variation des sons le long de l'échelle de désagrément.
- les premiers indices sur la nature des paramètres physiques liées à cette variation semblent s'orienter vers le domaine spectral ou spectro-temporel.

III-B-2. Validation de l'échelle de désagrément

Afin d'étudier la stabilité et la variabilité du résultat obtenu, nous procédons à deux analyses supplémentaires : d'une part, une analyse en composantes principales qui examine l'homogénéité des sujets vis-à-vis de leurs réponses, et d'autre part, une technique de 'bootstrap' qui crée une population fictive à partir de la population réelle et permet ainsi d'étudier la distribution des données de dominance.

L'analyse en composantes principales [Lebart, Morineau & Fenelon, 1979] nous donne la position de chaque sujet suivant les deux premières composantes principales et le pourcentage de variance expliquée par chacune d'elles (cf. annexe M).

On peut alors interpréter le résultat par le fait que les 11 sujets qui possèdent une première composante négative (demi-plan gauche de l'espace) présentent une inhomogénéité de jugement par rapport à la majorité des sujets qui se situe dans un faisceau opposé. Pour cette raison, il est opportun de concevoir une sous-population homogène incluant tous les sujets appartenant au demi-plan droit (cf. annexe M).

Une nouvelle analyse de type Bradley-Terry-Luce avec ce panel restreint donne le résultat suivant :

Son	Etiquette	Valeur de désagrément	
			- désagréable
15	RER	- 0.43	
13	avion - décollage	- 0.35	
14	avion - atterrissage	- 0.35	
9	autoroute fluide - avec PL	- 0.19	
12	flux urbain	- 0.16	
11	voie urbaine - démarrage	- 0.16	
16	autobus - départ	- 0.15	
10	autoroute chargée	- 0.09	
8	autoroute fluide - sans PL	- 0.08	
7	train marchandise - électrique	0.00	
6	train marchandise - diesel	0.14	
5	train Corail	0.15	
4	TGV Réseau - ligne grande vitesse	0.19	
3	TGV Réseau - ligne classique	0.41	
2	TGV SudEst - ligne grande vitesse	0.41	
1	TGV SudEst - ligne classique	0.66	
			+ désagréable

La comparaison des deux échelles (cf. §III-B-2) montre que le profil général de l'axe de perception n'est pas modifié par la nouvelle analyse mais, par contre, que les positions des stimuli sur l'échelle sont davantage marquées, ce qui renforce l'idée que les 11 sujets supprimés suivant le critère décrit précédemment présentait des types de jugements contraires à la majorité et contraignait la solution : leur suppression a augmenté la résolution et la dynamique de l'échelle de désagrément.

La technique de 'bootstrap' permet, quant à elle, de simuler des résultats d'expérience à partir de ceux obtenus au cours du test de dominance : par un tirage au sort, avec remise, dans l'ensemble des 60 sujets 'réels', on constitue une nouvelle population de 60 sujets qui constitue un deuxième test de dominance fictif. L'opération répliquée un grand nombre de fois – en l'occurrence 100 fois – permet, après analyse, de créer une distribution autour de chaque valeur de désagrément et d'étudier la variance de ces données.

Le résultat présenté en annexe N montre que les intervalles de confiance sont suffisamment faibles pour admettre une certaine sélectivité sur l'échelle de désagrément et donc que la relation d'ordre telle qu'elle a été obtenue avec l'algorithme de Bradley-Terry-Luce sur les 49 sujets (cf. précédemment) constitue une solution fiable pour la suite de l'étude.

IV- ANALYSE ACOUSTIQUE DES SONS ET CALCUL DES CORRELATIONS

A- Cadre théorique

Un processus empirique permettant d'écouter les sons par rapport à leur position relative sur l'échelle de désagrément donne des repères perceptifs sur le type de variation qui caractérise cette dimension subjective.

Puis, à partir d'hypothèses sur les indices physiques ou psychoacoustiques possibles, on effectue des analyses pour dégager des valeurs quantitatives et l'on calcule la corrélation entre chaque indice et les coordonnées sur la dimension subjective. Une bonne corrélation indique si le paramètre physique concerné est un bon candidat pour prédire quantitativement cette dimension.

Ces analyses, basées sur la technique de transformation de Fourier à court terme moyennée, permet d'obtenir une représentation spectrale des signaux. De plus, on tient compte, dans cette phase de calcul, de la réponse de l'oreille en fonction du niveau et des fréquences en considérant les différentes pondérations classiques utilisées (dBA, dBB et dBC).

IV-A-1. Analyse et traitements

- Méthode de Welch

Une approche usuelle permettant de réduire la variance sur un ensemble d'estimations spectrales d'un signal – et donc d'obtenir un meilleur résultat – est de moyenner sur différentes estimations indépendantes. C'est pourquoi, étant donné le caractère quasi-stationnaire des sons, chacun d'eux subit une analyse spectrale par la méthode de Welch [Oppenheim & Schaffer, 1975].

Dans cette approche, chaque séquence d'échantillons numériques $x(n)$, $0 \leq n \leq N - 1$, est divisée en K segments de M échantillons, tel que $N = K \cdot M$:

$$x_i(n) = x(n + iM - M) \quad , 0 \leq n \leq M - 1 \quad , 1 \leq i \leq K .$$

De plus, à chaque segment est appliquée une fenêtre $w(n)$ qui permet un calcul direct du spectre par transformée de Fourier discrète (TFD). Ainsi, la densité spectrale pour chaque segment i s'écrit :

$$P_i(\omega) = \frac{1}{MU} \cdot \left| \sum_{n=0}^{M-1} x_i(n) \cdot w(n) \cdot e^{-j\omega n} \right|^2 ,$$

et l'estimation de la densité spectrale totale est :

$$P(\omega) = \frac{1}{K} \cdot \sum_{i=1}^K P_i(\omega) , \text{ où } K \text{ et } MU \text{ sont les facteurs de normalisation :}$$

- K = facteur relatif au nombre de fenêtres,
- M = facteur de normalisation dû à la TFD,
- U = facteur de normalisation dû à la fenêtre :

$$U = \frac{1}{M} \cdot \sum_{n=0}^{M-1} w^2(n).$$

En augmentant le nombre de fenêtres K, avec N fixe, la variance diminue mais M – et donc la résolution spectrale – diminue également. Par conséquent, il y a un compromis à trouver entre la variance de l'estimation selon Welch et la résolution fréquentielle. Ce choix dépend d'une connaissance préalable du signal à analyser. Car, par exemple, si l'on sait que le spectre du signal présente une série de raies que l'on veut mettre en évidence, il faut choisir une valeur de M assez grande pour avoir la résolution fréquentielle désirée. Cette résolution est d'ailleurs quantifiable en fonction de la valeur de M :

- 2¹⁷ points ---> résolution de 0.33 Hz,
- 2¹⁴ points ---> résolution de 2.7 Hz,
- 2⁹ points ---> résolution de 85 Hz.

Dans le premier cas, la variance est trop grande alors que dans le troisième cas, c'est la résolution fréquentielle qui est trop faible. Cela étant, ce dernier cas de figure permet tout de même d'avoir une bonne représentation de la forme globale du spectre et de faire apparaître les différents formants.

Pour des représentations spectrales fines, on opte généralement pour un fenêtrage de 2¹⁴ points qui offre un bon compromis entre résolution et variance.

- Courbes isosoniques et pondération A, B et C

Le niveau d'un bruit en décibels linéaires représente bien une grandeur physique mais ne rend pas compte de l'impression sonore ressentie – la sonie –, dont le calcul reste difficile, surtout dans le cas de sons complexes et présentant des fluctuations temporelles. A partir de l'étude de la sensibilité de l'oreille aux sons purs continus, une famille de courbes isosoniques a été établie. Chaque courbe représente les niveaux nécessaires pour que deux sons de fréquences différentes donnent la même impression de sonie. Dans ces courbes la référence est un son pur de 1 kHz pour différents niveaux. La difficulté commence si l'on compare ce son pur avec un son complexe de spectre soit continu, soit formé de nombreuses raies, puisque ces courbes ne prennent pas en compte les effets de masquage mutuel entre composantes spectrales voisines.

Cependant, une bonne approximation est donnée, à partir de ces courbes, en appliquant aux mesures une pondération en fonction de la fréquence, au moyen d'un filtre dont la forme correspond à celle des courbes isosoniques la plus proche du niveau considéré.

Trois courbes de filtrage, qui reproduisent, avec des profils simplifiés, les atténuations présentées par les courbes isosoniques passant par 40, 70 et 100 dB à 1 kHz ont été établies comme normes. Ces courbes conduisent aux niveaux exprimés en dBA, dBB et dBC.

A partir des tableaux donnant les niveaux de pondération en dB pour chacune de ces trois courbes, on obtient, par régression polynomiale de quatrième ordre, une formulation analytique pour chacune des pondérations avec une erreur maximum inférieure à 0.5 dB. En posant $x = \log_{10}(f(\text{Hz}))$, on obtient :

- pour la pondération en dBA:

$$y = -184.78 + 160.66x - 55.73x^2 + 10.07x^3 - 0.83x^4 \text{ (dBA)},$$

– pour la pondération en dBB :

$$y = -147.34 + 171.11x - 77.82x^2 + 16.64x^3 - 1.41x^4 \quad (\text{dBB}),$$

– pour la pondération en dBC:

$$y = -99.08 + 150.59x - 85.13x^2 + 21.2x^3 - 1.96x^4 \quad (\text{dBC}).$$

En pratique, le niveau en dBA, qui ne devrait être appliqué que pour les niveaux faibles, donne un assez bon classement des bruits même forts et, pour cette raison, il est le plus souvent employé.

IV-A-2. Paramètres physiques de référence

Les paramètres pris en compte, dans un premier temps, sont inspirés d'une étude semblable effectuée précédemment sur des bruits automobiles [Susini, McAdams et Winsberg, 1997a&b]. Ils se réfèrent à plusieurs grandeurs physiques caractéristiques :

- le niveau RMS ou valeur efficace du niveau sonore :

Les sons ayant été égalisés en sonie, logiquement aucune dimension ne devrait être corrélée avec ce paramètre. Dans le but de vérifier son indépendance avec les jugements des sujets, on mesure l'énergie efficace de chacun des signaux en évaluant la valeur RMS (*root-mean-square*) avec différentes pondérations (dBA, dBB et dBC) et en tenant compte de la calibration de l'appareillage effectuée initialement (cf. §I-c) :

$$\text{RMS}(p(t)) = \sqrt{\frac{1}{T} \int_0^T [p(t)]^2 dt}.$$

Le niveau d'un signal sonore peut également être appréhender par un paramètre noté L_{eq} – niveau sonore équivalent – qui revient à mesurer le niveau RMS en dB sur la durée totale du son.

- le rapport de l'énergie périodique sur l'énergie non-périodique :

Afin d'analyser d'un point de vue perceptif les différentes contributions de la partie harmonique et de la partie non harmonique de chaque son, on dissocie ces deux parties par analyse/synthèse additive.

Une fois la partie harmonique et non harmonique obtenue pour chaque échantillon et pour chaque voie (gauche, droite), on évalue leur rapport énergétique, en considérant les différentes pondérations (en dBA, dBB et dBC) :

$$\text{RAPHn} = \frac{\text{RMS}_h}{\text{RMS}_n}.$$

- le centre de gravité spectral :

Dans les études de classification et d'analyse des timbres musicaux, une des dimensions les plus saillantes, décrite verbalement par la "brillance", est fortement liée au centre de gravité spectral [Grey et Gordon, 1978]. En effet, la répartition de

l'énergie en fonction du contenu fréquentiel d'un son joue un rôle important dans la perception.

Afin de rendre compte des effets de l'oreille externe et moyenne, et du traitement cochléaire, on peut décrire quantitativement la brillance d'un son comme le centre de gravité de l'ensemble des échantillons spectraux (ici $N = N_e$) en considérant les pondérations A, B et C d'une part, et en adoptant une démarche plus physiologique quant au fonctionnement de l'oreille, d'autre part :

- Mesure physique

L'expression du centre de gravité spectral (CG en Hz) utilisée est :

$$CG = \frac{\sum_{i=1}^{N_e} f(i) \cdot L(i)}{\sum_{i=1}^{N_e} L(i)} \quad \text{avec} \quad L(i) = 20 \log_{10} \left(\frac{p_i}{p_0} \right), \text{ calculé en dBA, dBB et dBC}$$

$$\text{et } f(i) = i \cdot \frac{f_e}{2N_e}, \text{ pour } 0 \leq i \leq N_e \quad (f_e = 44.1 \text{ kHz})$$

- Approche physiologique

Dans cette approche, nous considérons le système auditif périphérique à deux niveaux : le filtrage dû à l'oreille externe et moyenne (atténuation des fréquences basses et élevées) et une analyse spectrale au niveau de la membrane basilaire, modélisée par une série de filtres auditifs dont la largeur de bande a été estimée par différentes méthodes psychoacoustiques [Zwicker & Feldtkeller, 1981 ; Moore & Glasberg, 1983, 1990 ; Zwicker & Fastl, 1990].

L'oreille externe (conque, canal auditif, tympan) introduit des résonances produisant une augmentation du niveau sonore au niveau du tympan tout particulièrement dans la bande fréquentielle de 2 à 7 kHz [Pickles, 1988]. L'oreille moyenne (chaîne des osselets) joue le rôle d'un adaptateur d'impédance entre air et liquide. Aussi, elle peut être vue comme un autre filtrage fréquentiel passe-bande (essentiellement passe-bas). Les effets cumulés de l'oreille externe et de l'oreille moyenne sont modélisés par un filtre passe-bande de deuxième ordre de type Butterworth avec 450 Hz et 8500 Hz pour fréquences de coupure à -3 dB.

En ce qui concerne la cochlée, sa non-linéarité, sa réponse en fonction du niveau, les divers effets de masquage fréquentiel et temporel rendent sa modélisation difficile. Cependant, des résultats ont permis de montrer que certaines modélisations se rapprochent de la réalité en considérant des stimuli relativement simples (sons purs, bruits à bande étroite).

Dans notre approche, nous suivons la formulation de Moore et Glasberg (1983) selon laquelle la largeur du filtre auditif est représentée par un filtre équivalent de forme rectangulaire, ERB, (*equivalent rectangular bandwidth*) dont la largeur (ERB, en Hz) s'exprime en fonction de la fréquence centrale du filtre (f, en Hz) par la relation :

$$ERB = 24.7 * (0.00437 * f + 1).$$

Dans un premier temps on évalue l'énergie RMS_i (valeur efficace) dans chaque filtre ERB en fonction de la fréquence centrale f_i du filtre i correspondant, puis on calcule le centre de gravité spectral "physiologique" CGERB, en pondérant chacune des fréquences centrales par le niveau d'énergie obtenu dans le filtre correspondant :

$$\text{CGERB} = \frac{\sum_{i=1}^{N_{\text{filt}}} (\text{RMS}_i \cdot f_i)}{\sum_{i=1}^{N_{\text{filt}}} \text{RMS}_i} \quad \text{avec } N_{\text{filt}} = \text{nombre de filtres.}$$

- la déviation spectrale :

Un des indices qui s'est révélé important dans la perception des instruments de musique est la structure fine de l'enveloppe spectrale [Krimphoff, 1993 ; Krimphoff, McAdams et Winsberg, 1994]. On calcule ici la déviation des amplitudes des harmoniques par rapport à l'enveloppe spectrale globale. Cette enveloppe est calculée en prenant la moyenne des amplitudes en dBA de trois partiels adjacents :

$$\text{DEV} = \sum_{k=1}^Q \left| L(k) - \frac{L(k-1) + L(k) + L(k+1)}{3} \right|, \quad \text{avec } Q = \text{nombre d'harmoniques.}$$

Cependant, à la différence des sons musicaux, les sons de l'étude ne possèdent pas – ou quasiment pas – de caractère harmonique ; la définition théorique des points de l'enveloppe moyenne comme position moyenne entre 3 partiels consécutifs ne peut donc être appliquée.

Sur la base d'une technique utilisée pour caractériser les signaux sonores générés par différents types d'applaudissements [Repp, 1986], le calcul du profil spectral moyen s'effectue par analyse en composantes principales des données spectrales des sons : le spectre est calculé sur un nombre de points donné et la matrice représentant la valeur de chacun de ces échantillons pour tous les stimuli constitue la donnée d'entrée de l'analyse. Il en ressort la valeur des données spectrales sur la première composante principale. Si cette composante possède un taux significatif de variance expliquée (supérieur à 50%), on dispose alors d'un profil-type de spectre utilisable pour le calcul du paramètre de déviation spectrale.

- la décroissance spectrale par rapport à la fondamentale :

La démarche consiste à observer l'allure des enveloppes spectrales de chacun des sons suivant les dimensions considérées. On détermine les pics spectraux correspondants aux harmoniques du signal, puis l'on effectue une interpolation de ces pics par une fonction 'spline', généralement de degré 3.

Une observation des enveloppes spectrales ainsi obtenues permet d'étudier les différentes décroissances qui s'opèrent par paliers plus ou moins rapides et dans des bandes de fréquences différentes selon les sons. C'est un comportement qui peut être assimilé à un déplacement formantique.

Pour rendre compte de cette décroissance, on détermine le facteur 'DEC' qui évalue la somme des pentes entre les différents harmoniques et la fondamentale, notée \tilde{N} :

$$\text{DEC} = \sum_{k=k_{\text{inf}}}^{k_{\text{sup}}} \frac{\left(L(k \cdot \tilde{N}) - \frac{L(k \cdot \tilde{N})}{c} \right)}{\tilde{N} \cdot (k - 1)}, \quad \text{avec } c = 4 \text{ (facteur d'ajustement).}$$

Le calcul est effectué de $k=2$ à $k=17$, et afin d'évaluer la décroissance dans les aigus par rapport à la fondamentale on calcule également ce facteur pour k allant de 17 à 100.

IV-A-3 Principe du calcul de corrélation

Le coefficient de corrélation (noté R) permet d'évaluer le degré de correspondance linéaire entre deux ensembles de valeurs appariés. La valeur absolue de R varie entre 0, correspondant à une absence de relation linéaire entre les deux variables (physique et perceptive), et 1, pour une relation linéaire parfaite. Le signe du coefficient de corrélation correspond à une orientation de la dimension physique par rapport à l'axe perceptif. La significativité de ce coefficient dépend du nombre de degrés de liberté de la corrélation : $dl = \text{nombre d'objets du test} - 2$ (la valeur '2' venant du fait que la pente et l'abscisse à l'origine de la droite de régression sont fixes).

Le carré du coefficient de corrélation R^2 exprime, quant à lui, la proportion de variance dans les coordonnées perceptives expliquée par l'approximation linéaire du paramètre physique. Afin d'évaluer la stabilité de la corrélation entre deux jeux de valeurs qui décrivent les sons, on pourra avoir recours à une technique dite 'jack-knife' qui consiste à enlever chaque échantillon à tour de rôle et de recalculer R . On exprime ainsi la variabilité de R . Si un son est le principal responsable d'une forte corrélation, le coefficient R doit baisser de façon importante quand ce son est enlevé. Enfin, il est important de noter que la corrélation n'est pas une relation causale.

B- Recherche des paramètres acoustiques

IV-B-1. Détails des premières investigations

Un ensemble de facteurs physiques, ainsi que leur corrélation avec l'échelle de désagrément, sont calculés sur la base théorique développée aux paragraphes précédents (cf. §IV-A) : les différents paramètres – notés en **gras** – sont décrits dans la suite du paragraphe (Un récapitulatif se trouve également en annexe O). Ils sont calculés systématiquement pour les **voix droite et gauche** (sauf mention contraire) et les résultats numériques sont présentés dans les annexes O1 à O5.

On s'intéresse au préalable à quelques paramètres concernant l'intensité physique des sons (annexe O1) : **rmsdB**, **rmsdBcut2TO4**, **rmsdBcut4TO6**, **rmsdBcut6TO8** sont calculés en considérant, d'une part, le niveau RMS globale sur toute la durée du signal (équivalent au L_{eq}) et le niveau du signal compris respectivement entre 2-4 secondes, 4-6 secondes et 6-8 secondes. On calcule également un niveau RMS interaural moyen à partir de la moyenne du signal des deux voix : **mixrms**.

Les faibles taux de corrélations obtenus avec l'échelle de désagrément sont cohérents avec l'égalisation en sonie des stimuli qui a eu lieu au début de l'étude (cf. §I-e) ; et ceci malgré une légère dissymétrie des résultats entre voix droite et gauche – le niveau de la voix droite corrèle davantage – qui peut s'expliquer peut-être par un effet d'antériorité puisque les sources sonores passaient effectivement de droite à gauche durant les tests d'écoute.

Par ailleurs, du fait de la relative difficulté du problème posé, différentes personnes – équipe psychoacoustique de l'IRCAM et personnes travaillant autour du projet à la SNCF – ont écouté les sons afin de recouper les remarques à propos de la nature la plus probable des paramètres traduisant l'échelle de désagrément. Il ressort de cette étape que, comme nous l'avions initialement pressenti (cf. §II-B-1-c), les caractéristiques spectrales ou spectro-temporelles des sons semblent les plus propices à fournir les informations attendues.

C'est pourquoi, nous avons concentré nos efforts sur le moyen de décrire les échantillons sonores vis-à-vis de ces deux approches (fréquence et temps-fréquence), en déclinant plusieurs paramètres de référence :

- le Centre de Gravité Spectral (annexe O2) :

- **cgs_r15, cgs_r11, cgs_r8** :

- la résolution spectrale étant liée au nombre de points 2^r que compte la fenêtre d'analyse FFT, le CGS est issu d'un spectre calculé en dB avec différentes résolutions : 170 Hz ($r = 8$), 20 Hz ($r = 11$), et 1 Hz ($r = 15$).

- L'examen à posteriori de l'influence de la taille de la fenêtre d'analyse sur le calcul du CGS (cf. annexe P) nous amène à considérer dans les calculs suivants une fenêtre à 2^{15} points, c'est-à-dire 1 Hz de résolution fréquentielle pour un taux d'échantillonnage de 44,1 kHz.

- **cgsA, cgsB, cgsC** :

- la pondération rendant compte de l'influence conjuguée de l'oreille externe et de l'oreille moyenne (filtre passe-bande centré environ autour de 4000 Hz) est prise en compte et le CGS est calculé cette fois sur un spectre en dBA, dBB et dBC.

- **cgsERB** :

- la modélisation du système auditif interne nous amène à filtrer le signal initial par un banc de 38 filtres ERB dont les caractéristiques sont détaillées en annexe Q ; on calcule ensuite le CGSERB telle qu'il est décrit au §IV-A-2.

- **cgs_SR16** :

- la bande de fréquence est restreint à [0 – 8000 Hz], par sous-échantillonnage des signaux à 16 kHz, et le CGS est calculé à partir des informations spectrales comprises dans cette zone.

La baisse significative du taux de corrélation calculé avec ce dernier paramètre (env. 0.5 pour le 'cgsdB_r15' ou le 'cgsdBA_r15' et 0.2 pour le 'cgsdB_r15_SR16') incite à penser que globalement le CGS n'est pas un bon corrélat acoustique avec l'échelle de désagrément. En effet, le fait de restreindre la bande de fréquence à 8 kHz dans le calcul du CGS équivaut à s'intéresser à une zone plus pertinente du point de vue de l'audition que la zone qui s'étend jusqu'à 20 kHz : si le CGS était un indice physique réellement lié à la perception du désagrément, son influence serait renforcée dans cette bande de fréquence où l'oreille est la plus sensible.

Par ailleurs, en examinant les valeurs individuelles du 'cgsdB_r15_SR16', on s'aperçoit que leur variation est faible et qu'elles sont comprises dans un intervalle d'environ 400 Hz, avec une grande majorité entre 3400 Hz et 3600 Hz. Cette dernière observation amène à la conclusion que le CGS est certainement plus adapté à des signaux présentant des composantes harmoniques marquées et que, pour des signaux très bruités – comme c'est le cas dans notre étude –, le CGS n'est pas suffisamment fin pour discriminer les différences de caractéristiques spectrales entre les sons.

- **cgscut2TO4, cgscut4TO6, cgscut6TO8** :

- afin d'examiner les fluctuations temporelles du CGS, on s'intéresse au signal entre 2-4 secondes, 4-6 secondes et 6-8 secondes et l'on calcule le CGS sur le spectre du signal compris dans chacune de ces fenêtres.

La similitude des résultats entre les trois fenêtrages et avec le résultat global ('cgsdB_r15') montre que les fluctuations spectrales au cours du temps sont à prendre en compte avec une résolution temporelle beaucoup plus petite que 2 secondes (cf. plus bas).

- la déviation spectrale (annexe O3) :

– **spectypdBdev, spectypdBAdev** :

en appliquant la méthode détaillée au §IV-A-2, on calcule les spectres des 16 sons – en dB et en dBA – sur 129 points. On obtient, après analyse, la valeur des 129 données spectrales sur la première composante principale qui explique plus de 90 % de la variance des données et qui permet de définir un profil-type de spectre (en dB, et en dBA). On peut alors calculer une déviation de chacun des spectres par rapport à ces données.

- la variation spectrale (annexe O4) :

– **variatspect** :

le spectre du signal est calculé à travers une fenêtre temporelle de 200 ms qui se déplace par pas de 100 ms. On obtient ainsi, pour un son de 8 à 10 secondes, environ une cinquantaine de spectres instantanés correspondant à chaque pas temporel. Le calcul de la corrélation entre deux spectres instantanés consécutifs et finalement de la moyenne de ces coefficients de corrélations nous donne un indicateur de la variabilité du spectre au cours du temps. Le calcul peut se résumer par l'expression suivante :

$$r_{p,p-1} = \frac{\sum_{k=1}^N A_k(t_{p-1}) \cdot A_k(t_p)}{\sqrt{\sum_{k=1}^N A_k^2(t_{p-1})} \cdot \sqrt{\sum_{k=1}^N A_k^2(t_p)}} \Rightarrow \text{variatspect} = \frac{1}{M} \cdot \sum_{p=1}^M |r_{p,p-1}|$$

où N est le nombre de points de calcul du spectre.

M est le nombre de fenêtres temporelles utilisées pour le calcul.

$r_{p,p-1}$ est la corrélation entre les spectres $(A_k)_{k=1..N}$ calculés aux temps t_p et t_{p-1}

- le flux spectral (annexe O4) :

– **specfluct, specfluct_SR16** :

les fluctuations temporelles fines du CGS sont prises en compte par le calcul d'un CGS instantané à partir des spectres instantanés décrit précédemment. Le paramètre se déduit par mesure de l'écart de ces CGS instantanés avec un CGS global selon la formule :

$$\text{specfluct} = \frac{1}{M} \sqrt{\sum_{p=1}^M (\text{CGS}(p) - \text{CGS})^2}$$

où $\text{CGS}(p)$ est le CGS issu du spectre calculé dans la fenêtre p.

- la répartition de l'énergie dans le spectre (annexe O5) :
 - **energmean2500_5000Hz, energmean1000_5000Hz, energmeanERB5bands, energmeanERB14bands, energmeanBF0_4kHz, energmeanHF4_8kHz, energmeanRAP_HFBBF :**

les faibles taux de corrélation obtenus jusqu'à présent et l'examen des profils spectraux des signaux sonores (cf. annexe R) nous amène à considérer des paramètres relatifs à l'énergie moyenne (niveau RMS) comprise dans une bande de fréquence donnée : entre 2500 et 5000 Hz, entre 1000 et 5000 Hz, dans une zone Basse Fréquence (0 – 4 kHz) ou dans une zone Haute Fréquence (4 – 8 kHz).

Dans la même optique, l'exploitation du signal filtré par le banc de filtres ERB permet d'étudier l'énergie moyenne comprise dans un certain nombre de filtres (5 ou 14 filtres) recouvrant tout ou partie de la zone sensible de l'oreille, respectivement [2632 Hz - 4739 Hz] et [949 Hz - 5322 Hz] (cf. annexe Q).

L'élévation notable du taux de corrélation relatif aux paramètres de cette nature nous incite à poursuivre les recherches dans cette voie et à appréhender de manière plus fine le rapport des énergies contenues dans chacun des 38 filtres ERB du modèle relatif au système auditif interne.

IV-B-2. Le résultat significatif

La détermination du paramètre le plus significatif de cette étude repose sur un calcul de régression multiple, aboutissement d'une procédure en plusieurs étapes :

- les données de départ sont les signaux sonores filtrés par un banc de 38 filtres ERB simulant la réponse en fréquence de la membrane basilaire (cf. §IV-A-2). A la sortie de chacun des 38 canaux, on calcule le niveau RMS global sur la durée totale du signal, ce qui nous donne, pour les 16 sons de l'étude, une valeur moyenne de l'énergie contenue dans chaque filtre auditif. Par souci de pertinence, on sélectionne les canaux 2 à 32 qui permettent de couvrir la plage de fréquence 100 – 11000 Hz.

- pour chacun de ces filtres, un calcul de corrélation entre les valeurs de désagrément et les valeurs des niveaux opère une première sélection et permet d'écartier les filtres dont le niveau de sortie présente une faible corrélation (inférieure à 0.5) avec l'échelle de désagrément.

- on effectue avec les filtres ainsi retenus une analyse de cluster par arbre ultramétrique (cf. §II-B-2-a) sur la base de leurs intercorrélations à travers les 16 sons, c'est-à-dire en considérant la corrélation entre les 16 valeurs de niveau pour deux filtres distinctes comme une relation de proximité. Cette étape permet de regrouper les données en groupes de filtres cohérents et de dégager un nombre réduit de paramètres relativement indépendants pour l'algorithme de régression multiple.

Pour les voix gauche et droite, on obtient respectivement 4 et 6 groupes de filtres situés dans des régions fréquentielles différentes : les résultats détaillés de l'analyse sont présentés en annexe S. Au sein de chaque groupe de filtres, on calcule ensuite, pour chaque son, une moyenne des niveaux sur tous les filtres constituant le sous-ensemble (cf. annexe S).

- le calcul de régression multiple est un processus itératif qui examine successivement le rôle apporté par chacun des groupes de filtres dans le calcul de la corrélation avec les valeurs de désagrément et effectue une combinaison linéaire des paramètres les plus influents. Le résultat final se présente sous la forme suivante :

$$\text{Désagrément(gauche)} = 1.245 - 0.026 * B_{2-5} \quad (R = 0.83)$$

$$\text{Désagrément(droite)} = -0.308 + 0.026 * B_{17-24} - 0.019 * B_{6-7} \quad (R = 0.84)$$

où B_{k-m} définit le groupe de filtres regroupant les bandes k à m ;

$$\text{avec } \begin{cases} B_{2-5} & = [118 - 311] \text{ Hz} \\ B_{6-7} & = [313 - 445] \text{ Hz} \\ B_{17-24} & = [1607 - 4217] \text{ Hz} \end{cases}$$

Ces données suggèrent que, pour la voix gauche, plus la bande grave B_{2-5} est forte et moins le son est désagréable, pour la voix droite le désagrément augmente avec le niveau dans la bande medium-aiguë B_{17-24} et diminue avec celui de la bande grave B_{6-7} , et globalement les bandes aiguës sont désagréables et les bandes graves agréables.

Il est intéressant de noter, par ailleurs, que la bande médium-aiguë [1607 Hz – 4217 Hz] correspond quasiment à la zone sensible de l'oreille et que la contribution de la bande grave à gauche ($0.026 * B_{2-5}$) est légèrement plus forte que la contribution de la bande grave à droite ($0.019 * B_{6-7}$).

En résumé et en faisant une moyenne entre voix droite et gauche, il semble que le désagrément soit fortement lié à la présence d'énergie dans une zone de fréquence medium-aiguë et que, à l'inverse, cette sensation de désagrément peut être atténuée grâce à la présence d'énergie dans une zone basse fréquence.

Néanmoins, au vu de la technique employée pour aboutir à cette interprétation et du petit nombre de stimuli, le résultat doit être considéré avec précaution étant donné qu'avec peu de données, les coefficients de corrélation peuvent présenter une grande variabilité.

Ainsi, afin de rendre plus robuste la conclusion majeure de cette étude, nous avons menée une dernière analyse dite de 'jack-knife' sur les données de la régression multiple. Cette méthode consiste à réitérer une procédure d'analyse en supprimant à tour de rôle un élément du corpus de stimuli ; à partir d'un corpus de 16 stimuli, la méthode de 'jack-knife' nous donne donc une série de 16 résultats similaires à celui présenté plus haut et grâce auxquelles on va pouvoir estimer la variabilité du résultat global et montrer l'influence de chacun des stimuli sur ce même résultat global.

L'examen des données issues de ce traitement (cf. annexe T) permet d'envisager une bonne stabilité du résultat global obtenu avec la totalité du corpus de stimuli (cf. précédemment), en s'appuyant notamment sur la variabilité du coefficient de corrélation R qui varie dans [0.82 – 0.88], pour la voix gauche et dans [0.85 – 0.89], pour la voix droite.

CONCLUSIONS

La méthodologie prévue initialement et calquée sur les précédentes études – notamment celles sur les sons d'habitacles de voitures [Susini, McAdams et Winsberg, 1997a&b] – a donc dû être légèrement modifiée au gré des résultats obtenus.

Nous avons été rapidement confrontés à des problèmes nouveaux essentiellement dus à la nature des stimuli sonores : leur caractère très marqué et la culture commune très forte existant autour d'eux a permis une reconnaissance des sources sonores qui a joué le rôle d'un facteur cognitif prééminent et a empêché, dans une large mesure, de concevoir l'espace perceptif qui sous-tend la perception des stimuli en termes d'un ensemble de facteurs perceptifs primaires liés clairement aux paramètres acoustiques. Différentes analyses et traitements de données ont confirmé l'aspect catégoriel du résultat général obtenu.

Malgré cela, un axe subjectif pertinent – l'échelle de désagrément – a pu être dégagé grâce au test de dominance dont on a exploité les résultats de manière plus directe que ce qui avait été prévu au départ.

En outre, une analyse acoustique approfondie sur les sons a permis de trouver une relation de corrélation tout à fait significative (70% de la variance expliquée) entre les valeurs de désagrément et les valeurs d'un paramètre physique relatif à l'énergie contenue dans un certain nombre de filtres auditifs modélisant le système auditif interne.

Finalement, cette étude qui se situe dans l'optique de la perception des sons d'environnement, a mis à jour les problématiques inhérentes à ce thème de recherche notamment en ce qui concerne le choix et le réalisme des stimuli, et la pertinence des méthodes expérimentales employées. Ainsi, les perspectives de ce projet peuvent clairement se concevoir autour de plusieurs axes comme le soin apporté à la phase d'enregistrement des sons, l'homogénéisation du corpus d'échantillons sonores, la mise au point d'un protocole de test spécifique aux sons d'environnement avec éventuellement une mise en situation plus affirmée du sujet, ou bien encore des procédures d'extraction de paramètres acoustiques davantage appropriées aux signaux large bande et relativement inharmoniques que l'on rencontre généralement dans ce genre d'étude.

BIBLIOGRAPHIE

A- Référence citées dans le rapport

1. Analyses multidimensionnelles

- Winsberg S. & Carroll J.D. (1989) : "A quasi-nonmetric method for multidimensional scaling via an extended Euclidean model". *Psychometrika*, **54**, 217-229.
- Winsberg S. & De Soete G. (1993a) : "A latent class approach to fitting the weighted Euclidean model, CLASCAL". *Psychometrika*, **58**, 315-330.

2. Analyses de cluster

- Arabie P. et Carroll J.D. (1980) : "MAPCLUS : a mathematical programming approach to fitting the ADCLUS model". *Psychometrika*, **45**, 211-235.
- Arabie P. et Shepard R.N. (1973) : "Representation of similarities as additive combinations of discrete, overlapping properties". Presented at Math. Psychol. Meet. Montréal.
- Carroll J.D. (1976) : "Spatial, non-spatial and hybrid models for scaling". *Psychometrika*, **41**, 439-463.
- Carroll J.D. et Arabie P. (1979) : "INDCLUS : A three-way approach to clustering". Presented at Meet. Psychom. Soc., Monterey, Calif.
- Carroll J.D. et Arabie P. (1980) : "Multidimensional scaling". In M.R. Rosenzweig and L.W.Porter (Eds.). *Annual Review of Psychology*, Palo Alto : Annual Reviews.
- De Sarbo W.S. (1982) : "GENNCLUS : New model for general nonhierarchical clustering analysis". *Psychometrika*, **47**(4), 449-475.
- Hartigan J.A. (1967) : "Representation of similarity matrices by tree". *Journal of the American Statistical Association*, **62**, 1140-1158.
- Hartigan J.A. (1975) "Clustering algorithms". New York, Wiley.
- Johnson S.C. (1967) : "Hierarchical clustering schemes". *Psychometrika*, **32**, 241-254.
- Shepard R.N. (1980) : "Multidimensional scaling, tree-fitting, and clustering". *Science*, **210**, 390-398.
- Shepard R.N. et Arabie P. (1979) : "Additive clustering : Representation of similarity as combinations of discrete overlapping properties". *Psychological Review*, **86**, 87-123.

3. Analyses de dominance

- Winsberg S. & De Soete G. (1993b) : "A Thurstonian pairwise choice model with univariate and multivariate spline transformation". *Psychometrika*, **58**, 233-256.

4. Méthode de Bradley-Terry-Luce

- David H. A. (1986) : "The method of paired comparisons". Oxford University Press, 2nd Edition.

5. Analyses en composantes principales

- Lebart L., Morineau A. Fenelon J.P. (1979) : "Traitement de données statistiques - Méthodes et programmes". Ed. Dunod, 510 pp.

6. Traitements et analyses acoustiques

- Oppenheim A.V. & Schafer R.W. (1975) : "Digital Signal Processing". Prentice-Hall, NJ.
- Grey J.M. & Gordon J.W. (1978) : "Perceptual effects of spectral modifications on musical timbres". *J. Acoust. Soc. Am.* 63, 1493-1500.
- Zwicker E. & Feldtkeller R. (1981) : "L'oreille : récepteur d'information". [trad. C. Sorin], Masson, Paris.
- Moore B.C.J. & Glasberg B.R. (1983) : "Suggested formule for calculating auditory-bandwidths and excitation patterns". *J. Acoust. Soc. Am.*, 74, 750-753.
- Moore B.C.J. & Glasberg B.R. (1990) : "Derivation of auditory filter shapes from notched-noise data". *Hearing Research*, 47 103-138.
- Zwicker E. & Fastl H. (1990) : "Psychoacoustics: Facts and models". Springer-Verlag, Berlin.
- Krimphoff J. (1993) : "Analyse acoustique et perception du timbre". Mémoire de DEA Acoustique Appliqué, Université du Maine, Le Mans.
- Krimphoff J., McAdams S. & Winsberg S. (1994) : "Caractérisation du timbre des sons complexes. II : Analyses acoustiques et quantification psychophysique". *Journal de Physique*, 4(C5), 625-628.
- Repp B. H. (1987) : "The sound of two hands clapping : an exploratory study". *J. Acoust. Soc. Am.*, 81(4), 1100-1109.
- Susini P., McAdams S. & Winsberg S. (1997a) : "Perceptual charaterisation of vehicule noise". EAA Symposium : Psychoacoustics in industry and university.
- Susini P., McAdams S. & Winsberg S. (1997b) : "Caractérisation perceptive des bruits de véhicules". 4ème Congrès Français d'Acoustique, *Proceedings*, vol. I, pp. 543-546.

B- Références générales

1. Analyses multidimensionnelles

- Bismarck G. von (1974a) : "Timbre of steady sounds: a factorial investigation of its verbal attributes". *Acustica*, 30, 146-159.
- Bismarck G. von (1974b) : "Sharpness as an attribute of imbre of steady sounds". *Acustica*, 30, 159-172.
- Carroll J. D. & Wish M. (1974) : "Models and methods for three-way multidimensional scaling." In D. H. Krantz, R. C. Atkinson, R. D. Luce, P. Suppes (eds.) *Contemporary Dewvelopments in Mathematical Psychology. vol. II Measurements, Psychophysics, and Neural Information Processing*.
- Grey J.M. (1977) : "Multidimensional perceptual scaling of musical timbres". *J. Acoust. Soc. Am.* 61, 1270-1277.

- Iverson P. & Krumhansl C.L. (1991) : "Measuring similarity of musical timbres". *J. Acoust. Soc.Am.* **89**, 1988.
- Kendall R.A., & Carterette E.C. (1991) : "Perceptual scaling of simultaneous zind instrument timbres". *Music Perception* **8**, 369-404.
- McAdams S., Winsberg S., Donnadiou S., De Soete G. & Krimphoff J. (1995) : "Perceptual scaling of synthesized musical timbres: Common dimensions, specificities, and latent subject classes". *Psychological Research/Psychologische Forschung* (sous presse).
- Miller J.R. & Carterette E.C. (1975) : "Perceptual space for musical structures". *J.Acoust. Soc.Am.* **58**, 711-720.
- Plomp R. (1969) : "Timbre as a multidimensional attribute of complex tones". In *Frequency Analysis and Periodicity Detection in Hearing*, R. Plomp & G.F. Smoorenburg (eds.), pp. 397-414. Sijthoff, Leiden.
- Shepard R. N. (1962) : "The analysis of proximities: multidimensional scaling with an unknown distance function". I *Psychometrika*, **27**, 125-140.
- Shepard R. N. (1962) : "The analysis of proximities: multidimensional scaling with an unknown distance function". *Psychometrika*, **27**, 219-246.
- Wedin L & Goude G. (1972) : "Dimension analysis of perception of instrument timbres". *Scandinavian Journal of Psychology*, **13**, 228-240.

2. Analyses de cluster

- Barthélemy J.P., Guénoche A. (1988) : "Les arbres et les représentation des proximités". Masson éd., Paris.
- Cunningham J.P. (1978) : "Free trees and bidirectional trees as representations of psychological distance". *Journal of Mathematical Psychology*, **17**, 165-188.
- Diday E. (1974) : "Optimization in non-hierarchical clustering". *Pattern Recognition*, **6**, 17-33.
- MacQueen J. (1967) : "Some methods for classification and analysis of multivariate observations". In L.M. LeCam et J. Neyman (Eds.), *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, Berkeley : University of California Press, **1**, 281-297.

3. Psychoacoustique

- Botte M.C., Canévet G., Demany L. & Sorin C. (1990) : "Psychoacoustique et perception auditive". INSERM, Paris.
- Daniel P. & Weber R. (1993) : "Calculating psychoacoustical roughness". *Contributions to psychological acoustics VI*, A. Schick (ed.), pp. 251-265, Universität Oldenburg, Oldenburg.
- Daniel P. & Weber R. (1994) : "Evaluating and measurement of roughness of artificial and technical sounds". *Inter-noise 94*, 893-896.
- Grey J.M. & Moorer J.A. (1977) : "Perceptual evaluation of synthesized musical instrument tones". *J. Acoust. Soc. Am.* **62**, 454-462.
- Moore B.C.J. (1989) : "An introduction to the psychology of hearing". 3e édition, Academic Press, London.
- Terhardt E. (1974) : "Perception of sound fluctuations". *Acustica*, **30**, 200-213.

ANNEXES

- **Annexe A** : **Résultats de l'égalisation en sonie** p. 41
 - coefficient d'égalisation trouvé par chaque sujet, pour chaque échantillon
 - coefficient d'égalisation moyen et atténuation en dB correspondante (19 sujets)
 - coefficient d'égalisation moyen, atténuation et L_{eq} correspondants (15 sujets)

- **Annexe B** : **Echantillons sonores de l'étude** p. 43
 - liste récapitulative
 - données techniques annexes

- **Annexe C** : **Documents concernant le test de dissemblance** p. 44
 - feuille de présentation
 - premières réactions et compte-rendu d'identification
 - questionnaire

- **Annexe D** : **Résultats de la tâche d'identification** p. 50
 - pour chaque son, pourcentage de sujets ayant identifié correctement

- **Annexe E** : **Test de dissemblance** p. 51
 - Remarques concernant les critères utilisés dans la stratégie de jugement

- **Annexe F** : **Solution multidimensionnelle 'forcée' de l'analyse globale** p. 53
 - projection dans le plan (1,2)
 - projection dans le plan (1,3)
 - projection dans le plan (2,3)

- **Annexe G** : **Analyse de clusters hiérarchiques sur les sujets** p. 54
 - représentation en arbre ultramétrique

- **Annexe H** : **Analyse de clusters hiérarchiques sur les stimuli** p. 55
 - représentation en arbre ultramétrique (résultat avec 30 sujets)
 - représentation en arbre ultramétrique (résultat avec 17 sujets)
 - représentation en arbre ultramétrique (résultat avec 11 sujets)

- **Annexe I** : **Analyse multidimensionnelle 'partielle'** p. 58
 - résultat obtenu avec le 1^{er} sous-groupe de 17 sujets
 - résultat obtenu avec le 2^{ème} sous-groupe de 11 sujets

- **Annexe J** : **Analyse multidimensionnelle 'partielle'** p. 60
 - résultat obtenu avec l'ensemble à 12 stimuli
 - résultat obtenu avec l'ensemble à 10 stimuli
 - résultat obtenu avec l'ensemble à 9 stimuli

- **Annexe K** : **Documents concernant le test de dominance** p. 62
 - feuille de présentation
 - premières réactions

- **Annexe L** : **Test de dominance** p. 64
 - remarques concernant les critères utilisés dans la stratégie de jugement

- **Annexe M** : **Analyse en composantes principales** p. 67
 - position des sujets sur les 2 premières composantes
 - pourcentage de variance expliquée par les 2 premières composantes

- **Annexe N** : **Technique de bootstrap** p. 68
 - résultat du bootstrap sur la population de 60 sujets, avec 100 répliques
 - résultat du bootstrap sur la population des 49 sujets, avec 100 répliques

- **Annexe O** : **Corrélation entre paramètres physiques et échelle de désagrément** p. 70
 - récapitulatif des paramètres avec leur descriptif

- **Annexe O1** : **Corrélation entre paramètres physiques et échelle de désagrément** p. 71
 - valeurs des paramètres relatifs au niveau RMS
 - corrélations avec les valeurs de désagrément

- **Annexe O2** : **Corrélation entre paramètres physiques et échelle de désagrément** p. 72
 - valeurs des paramètres relatifs au CGS
 - corrélations avec les valeurs de désagrément

- **Annexe O3** : **Corrélation entre paramètres physiques et échelle de désagrément** p. 74
 - profil spectral 'type' (dB et dBA) obtenu par analyse en composantes principales
 - valeurs des paramètres relatifs à la Déviati on Spectrale
 - corrélations avec les valeurs de désagrément

- **Annexe O4** : **Corrélation entre paramètres physiques et échelle de désagrément** p. 76
 - valeurs des paramètres relatifs aux variati ons spectro-temporelles
 - corrélations avec les valeurs de désagrément

- **Annexe O5** : **Corrélation entre paramètres physiques et échelle de désagrément** p. 77
 - valeurs des paramètres relatifs à la répartiti on de l'énergie en fréquence
 - corrélations avec les valeurs de désagrément
 - représentation des droites de régression linéaire

- **Annexe P** : **Influence de la résolution spectrale sur la valeur du CGS** p. 79
 - comparaison des CGS sur un spectre calculé avec $r = 15, 11$ et 8

- **Annexe Q** : **Caractéristiques du banc de filtres auditifs** p. 80
 - fréquence centrale, largeur de bandes et gain des filtres ERB

- **Annexe R** : **Profils spectraux des échantillons sonores** p. 81
 - spectres (dB, voix G) dans l'ordre décroissant du désagrément
 - spectres (dBA, voix G) dans l'ordre décroissant du désagrément

- **Annexe S1** : **Analyse de clusters hiérarchiques des filtres auditifs (ERB) sur la base de leur intercorrélati on à travers les 16 sons - voix GAUCHE** p. 85
 - clustering des filtres dont la corrélation avec l'échelle de désagrément > 0.5
 - niveau moyen des bandes regroupées pour l'analyse de régression multiple

- **Annexe S2** : **Analyse de clusters hiérarchiques des filtres auditifs (ERB) sur la base de leur intercorrélati on à travers les 16 sons - voix DROITE** p. 86
 - clustering des filtres dont la corrélation avec l'échelle de désagrément > 0.5
 - niveau moyen des bandes regroupées pour l'analyse de régression multiple

- **Annexe T** : **Etude de la stabilité des résultats de la régression multiple** p. 87
 - données issues de la méthode de 'jack-knife'

ANNEXE A

Résultats de l'égalisation en sonie :

- coefficient d'égalisation trouvé par chaque sujet, pour chaque échantillon

	sujet 1	sujet 2	sujet 3	sujet 4	sujet 5	sujet 6	sujet 7	sujet 8	sujet 9
son 1	0.542	0.565	0.693	0.534	0.883	0.354	0.643	0.677	0.458
son 2	0.612	0.622	0.714	0.607	0.607	0.503	0.625	0.753	0.586
son 3	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
son 4	0.664	0.654	0.565	0.862	0.755	0.680	0.729	0.716	0.682
son 5	0.534	0.513	0.870	0.625	0.500	0.656	0.503	0.688	0.724
son 6	0.445	0.383	0.513	0.372	0.406	0.477	0.318	0.557	0.469
son 7	0.674	0.534	0.771	0.823	0.693	0.878	0.693	0.758	0.612
son 8	0.990	0.818	1.062	1.232	0.802	0.924	1.221	0.846	1.003
son 9	0.622	0.714	0.578	0.622	0.724	0.810	0.771	0.714	0.794
son 10	0.414	0.401	0.398	0.365	0.445	0.497	0.372	0.286	0.529
son 11	0.357	0.375	0.250	0.513	0.367	0.453	0.365	0.328	0.372
son 12	0.432	0.430	0.521	0.573	0.536	0.581	0.594	0.469	0.547
son 13	0.672	0.516	0.831	0.740	0.401	0.901	0.992	1.128	0.669
son 14	0.365	0.365	0.549	0.393	0.221	0.375	0.305	0.862	0.260
son 15	0.445	0.346	0.336	0.458	0.310	0.255	0.250	0.518	0.385
son 16	0.799	0.513	0.677	0.849	0.380	0.534	0.638	0.602	0.557

sujet 10	sujet 11	sujet 12	sujet 13	sujet 14	sujet 15	sujet 16	sujet 17	sujet 18	sujet 19
0.542	0.755	0.482	0.599	0.661	0.953	0.555	0.492	0.599	0.523
0.615	0.706	0.667	0.479	0.531	0.826	0.661	0.539	0.521	0.521
1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.518	0.578	0.594	0.612	0.576	0.779	0.896	0.612	0.643	0.701
0.497	0.479	0.464	0.680	0.534	0.688	0.794	0.607	0.604	0.516
0.352	0.448	0.289	0.596	0.435	0.615	0.714	0.531	0.401	0.362
0.659	0.560	0.651	0.854	0.544	0.745	1.034	0.690	0.719	0.594
0.708	0.661	0.818	0.943	0.794	0.953	1.091	0.753	0.646	0.651
0.688	0.771	0.547	0.750	0.648	0.906	0.607	0.615	0.596	0.534
0.534	0.352	0.339	0.521	0.443	0.771	0.323	0.482	0.357	0.370
0.508	0.326	0.289	0.586	0.393	0.576	0.378	0.263	0.260	0.312
0.695	0.469	0.409	0.409	0.503	0.516	0.445	0.370	0.365	0.266
0.833	0.680	0.578	0.453	0.602	0.833	0.573	0.453	0.677	0.536
0.591	0.513	0.227	0.482	0.315	0.573	0.253	0.109	0.258	0.195
0.276	0.383	0.336	0.523	0.320	0.576	0.302	0.216	0.320	0.195
0.680	0.607	0.602	0.896	0.664	0.591	0.805	0.328	0.456	0.193

ANNEXE A

Résultat de l'égalisation en sonie :

- coefficient d'égalisation moyen et atténuation en dB correspondante (19 sujets)

	coefficient d'égalisation		atténuation en dB	
	moyenne	écart-type	moyenne	écart-type
son 1	0.606	0.144	-4.353	2.079
son 2	0.615	0.091	-4.217	1.285
son 3	1.000	0.000	0.000	0.000
son 4	0.674	0.100	-3.421	1.284
son 5	0.604	0.115	-4.381	1.658
son 6	0.457	0.110	-6.802	2.093
son 7	0.710	0.126	-2.978	1.548
son 8	0.890	0.179	-1.009	1.753
son 9	0.685	0.100	-3.289	1.271
son 10	0.431	0.110	-7.301	2.226
son 11	0.383	0.101	-8.343	2.311
son 12	0.480	0.099	-6.368	1.799
son 13	0.688	0.194	-3.251	2.463
son 14	0.380	0.180	-8.415	4.199
son 15	0.355	0.107	-8.986	2.623
son 16	0.598	0.178	-4.460	2.606

- coefficient d'égalisation moyen, atténuation et L_{eq} correspondants (15 sujets)

	coefficient d'égalisation		atténuation en dB		niveau L_{eq}
	moyenne	écart-type	moyenne	écart-type	
son 1	0.589	0.128	-4.599	1.895	65,1
son 2	0.594	0.071	-4.520	1.043	67,1
son 3	1.000	0.000	0.000	0.000	63,6
son 4	0.654	0.087	-3.686	1.161	62,8
son 5	0.580	0.113	-4.734	1.699	65,2
son 6	0.418	0.079	-7.583	1.642	64,8
son 7	0.684	0.109	-3.301	1.391	65,1
son 8	0.885	0.191	-1.062	1.884	62,1
son 9	0.678	0.092	-3.376	1.180	64,7
son 10	0.422	0.069	-7.486	1.415	60,6
son 11	0.382	0.096	-8.364	2.204	61,0
son 12	0.489	0.107	-6.222	1.915	58,7
son 13	0.672	0.167	-3.452	2.165	64,6
son 14	0.361	0.125	-8.851	3.037	62,5
son 15	0.343	0.086	-9.302	2.202	63,3
son 16	0.603	0.180	-4.394	2.619	64,6

ANNEXE B

Echantillons sonores de l'étude : • liste récapitulative

- **son 1** : TGV Sud-Est, ligne classique (160 km/h).
- **son 2** : TGV Sud-Est, ligne à grande vitesse (270 km/h).
- **son 3** : TGV Réseau, ligne classique (160 km/h).
- **son 4** : TGV Réseau, ligne grande vitesse (270 km/h).
- **son 5** : Train Corail (160 km/h).
- **son 6** : Train marchandise, diesel (80 km/h).
- **son 7** : Train marchandise, électrique (80 km/h).
- **son 8** : Autoroute fluide, sans poids lourd.
- **son 9** : Autoroute fluide, avec poids lourds.
- **son 10** : Autoroute chargée, vitesse réduite.
- **son 11** : Voie urbaine, démarrage au feu.
- **son 12** : Flux urbain sur boulevard en sens unique.
- **son 13** : Airbus A320, décollage.
- **son 14** : Airbus A320, atterrissage.
- **son 15** : RER A (60 km/h).
- **son 16** : Autobus, départ.

• données techniques annexes

- les enregistrements de trafic ferroviaire [**sons n° 1 à 7**] ont été effectués en pleine campagne (non vallonnée) à une distance de 100 mètres du bord de la voie.
- les enregistrements de trafic autoroutier [**sons n° 8 et 9**] ont été effectués en banlieue parisienne, à une distance de 50 mètres du bord de la chaussée.
- l'enregistrement de voie urbaine [**son n° 11**] a été effectué à Paris, boulevard Saint-Germain, sur la chaussée.
- l'enregistrement de flux urbain [**son n° 12**] a été effectué à Paris à une distance de 5 mètres du bord de la chaussée.
- l'enregistrement des trafics aériens [**sons n° 13 et 14**] a été effectué à Orly à une distance de 365 m./verticale (son n° 13) et 90 m./verticale (son n° 14).
- l'enregistrement du RER A (type 'rouge&bleu') [**son n° 15**] a été effectué à Vincennes, à une distance de 5 mètres du bord de la voie.
- l'enregistrement de l'autobus [**son n° 16**] a été effectué à Paris, à une distance de 1 mètre du bord de la chaussée.
- toutes les prises de son des échantillons S.N.C.F. [**sons n° 1 à 11**] utilisent la méthode d'enregistrement stéréophonique ORTF à 3 micros, placés à une hauteur de 1m70 du sol.
- les prises de son des échantillons DIASONIC. [**sons n° 12 à 16**] utilisent un micro Shoeps (sons n° 12 et 15) ou un micro B&K (sons n° 13, 14 et 16).
- le niveau réel a été mesuré pour le :
 - **son n° 2** : L_{eq} (tps passage) = 93 dBA, L_{max} = 95 dBA.

(à la position du microphone)

- **son n° 13** : L_{max} = 81 dBA
- **son n° 14** : L_{max} = 92 dBA
- **son n° 15** : L_{max} = 92 dBA
- **son n° 16** : L_{max} = 92.8 dBA

ANNEXE C

Documents concernant le test de dissemblance : • feuille de présentation

I.R.C.A.M. - EQUIPE PSYCHOACOUSTIQUE

EXPERIENCE SUR LA PERCEPTION DES SONS DE TRAFIC

Nous nous intéressons à la façon dont les personnes perçoivent les sons de trafic des moyens de transport actuels. Dans cette expérience, les sons ont fait l'objet d'un enregistrement sur site ; ils sont au nombre de 16.

- Pour vous donner une idée d'ensemble, nous vous présentons, au début de l'expérience, tous les sons dans un ordre quelconque.

- L'expérience est un test de dissemblance dans lequel les échantillons sonores sont présentés par paire : le but du test n'est pas d'essayer d'identifier les sons mais plutôt d'utiliser votre impression globale, pour évaluer le degré de différence entre les deux sons de chaque paire et positionner en conséquence le curseur du potentiomètre sur l'échelle ["*très similaire - très différent*"].

Il vous est possible de réécouter la paire de sons autant de fois que vous le désirez avec le bouton de gauche ["*je voudrais écouter encore une fois*"] ; néanmoins, nous vous conseillons de ne pas trop répéter l'écoute d'une même paire (une à trois fois suffisent normalement), pour cela, nous vous proposons la stratégie suivante :

- 1^{ère} écoute, faire un jugement global
- éventuellement, 2^{ème} et 3^{ème} écoute, affiner ce jugement.

Essayer d'utiliser toute l'étendue de l'échelle de jugement au cours de l'expérience. Le curseur se replace aléatoirement au début de chaque nouvelle paire afin de ne pas y privilégier une zone particulière.

Un fois que votre décision est prise sur la valeur de dissemblance entre les deux sons de la paire, vous validez votre résultat en cliquant sur le bouton de droite ["*j'ai fait mon choix*"] et une nouvelle paire de sons se présente automatiquement. Il vous est alors impossible de revenir sur la paire précédente.

- Au préalable, une phase d'entraînement est proposée ; elle s'effectue sur 10 paires choisies aléatoirement dans l'ensemble des échantillons de l'expérience. Profitez de ce moment pour, d'une part, vous familiariser avec la tâche et, d'autre part, commencer à dégager les critères perceptifs que vous utiliser naturellement pour effectuer votre jugement de dissemblance.

- Le test de dissemblance globale comprend 120 paires. Il s'effectue, à votre convenance, en une seule fois ou bien avec des pauses (1 ou 2 maximum, pour ne pas rallonger démesurément la durée du test).

Remarques : si vous préférez utiliser le clavier de l'ordinateur plutôt que la souris, il est possible de :

- déplacer le curseur du potentiomètre avec les touches '<' et '>',
- rejouer les paires de son avec la touche 'espace',
- valider votre choix et passer à la paire suivante avec la touche 'retour chariot'.

ANNEXE C

Documents concernant le test de dissemblance :
• premières réactions et compte-rendu d'identification

I.R.C.A.M. - EQUIPE PSYCHOACOUSTIQUE

EXPERIENCE SUR LA PERCEPTION DES SONS DE TRAFIC
- TEST DE DISSEMBLANCE : COMMENTAIRES -

Date : _____

Nom du sujet : _____

N° de l'expérience : _____

Impressions 'à chaud'

• longueur du test : _____

• difficulté du test : _____

• intérêt du test (ennuyeux, divertissant, etc. ...) :

Identification

Son N°	IDENTIFICATION / OBSERVATIONS
• 1	
• 2	
• 3	
• 4	
• 5	
• 6	
• 7	
• 8	
• 9	
• 10	
• 11	
• 12	
• 13	
• 14	
• 15	
• 16	

ANNEXE C

Documents concernant le test de dissemblance : • questionnaire

I.R.C.A.M. - EQUIPE PSYCHOACOUSTIQUE

EXPERIENCE SUR LA PERCEPTION DES SONS DE TRAFIC - RENSEIGNEMENTS CONCERNANT LE SUJET -

• Nom : _____ • Prénom : _____

• Adresse : _____

• Téléphone : _____

• Age : _____ ans • Sexe : _____

• Catégorie socio-professionnelle : _____

• Vous êtes : () droitier () gaucher

• Avez-vous déjà eu des problèmes d'audition ? Si oui :

a) lesquels ? _____

b) quand ? _____

c) pensez-vous avoir encore des problèmes ? Si oui, lesquels ? _____

• Est-ce que votre logement est exposé au bruit :

- () d'une route/autoroute
- () d'une voie ferrée
- () d'un aéroport
- () autre : _____
- () aucun

• Est-ce que vous êtes gêné par ces bruits quand vous êtes chez vous ? Si oui, essayez de quantifier la gêne sur une échelle de 0 (= pas du tout gêné) à 5 (= très gêné) :

	0	1	2	3	4	5
le bruit de la route/autoroute	()	()	()	()	()	()
le bruit de la voie ferroviaire	()	()	()	()	()	()
le bruit de l'aéroport	()	()	()	()	()	()
autre : _____	()	()	()	()	()	()

• **Avez-vous une isolation individuelle :**

- | | |
|---|-------------------------------------|
| <input type="checkbox"/> vitrage épais | <input type="checkbox"/> tentures |
| <input type="checkbox"/> survitrage | <input type="checkbox"/> végétation |
| <input type="checkbox"/> double vitrage | <input type="checkbox"/> aucune |

• **A quelle période de la journée êtes-vous chez vous ?**

	presque jamais	pas souvent	assez souvent	très souvent
le matin	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
dans la journée	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
le soir	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

• **Etes-vous soumis à un environnement bruyant ailleurs (bruit de machines, circulation intensive, concert de rock, walkman écouté fort, etc. ...) ? Si oui :**

	combien d'heures par jour / semaine / mois ?
bruit de machines	_____
circulation intensive	_____
concert de rock	_____
walkman écouté fort	_____
autre : _____	_____
autre : _____	_____

• **Le bruit auquel vous êtes soumis dans la vie courante vous gêne-t-il ? Essayez de quantifier la gêne sur une échelle de 0 (= pas du tout gêné) à 5 (= très gêné) :**

	0	1	2	3	4	5
bruit de machines	<input type="checkbox"/>					
circulation intensive	<input type="checkbox"/>					
concert de rock	<input type="checkbox"/>					
walkman écouté fort	<input type="checkbox"/>					
autre : _____	<input type="checkbox"/>					
autre : _____	<input type="checkbox"/>					

• **Pour aller travailler ou pour d'autres trajets quotidiens, utilisez-vous plutôt :**

	temps de trajet par jour ?
<input type="checkbox"/> le train	_____
<input type="checkbox"/> le métro/ RER	_____
<input type="checkbox"/> le bus	_____
<input type="checkbox"/> la voiture	_____
<input type="checkbox"/> le vélo	_____
<input type="checkbox"/> autre : _____	_____

Remarques : _____

- Pour vos autres déplacements (week-end, vacances, etc. ...), utilisez-vous plutôt :

fréquence ?

- | | |
|--|-------|
| <input type="checkbox"/> le train | _____ |
| <input type="checkbox"/> la voiture | _____ |
| <input type="checkbox"/> l'avion | _____ |
| <input type="checkbox"/> autre : _____ | _____ |

Remarques : _____

- Y-a-t-il quelqu'un dans votre proche entourage qui travaille pour :

- les transports aériens
- les transports ferroviaires
- les transports routiers

- Appartenez-vous à une association :

- de lutte contre le bruit
- pour la protection de l'environnement
- autre : _____

- Voulez-vous ajouter quelque chose vous concernant ?

- Quelles sont les expériences de psychoacoustiques auxquelles vous avez déjà participé ?

- A propos du test de dissemblance que vous venez de passer, pouvez-vous, en quelques mots, expliquer la stratégie de jugement que vous avez employée et définir éventuellement les critères perceptifs que vous avez utilisés ?

Merci de votre collaboration.

ANNEXE D

Résultats de la tâche d'identification :

- pour chaque son, pourcentage de sujets ayant identifié correctement

son 1	son 2	son 3	son 4	son 5	son 6	son 7	son 8
93 %	86 %	83 %	76 %	80 %	60 %	60 %	83 %

son 9	son 10	son 11	son 12	son 13	son 14	son 15	son 16
93 %	96 %	93 %	83 %	96 %	96 %	80 %	100 %

ANNEXE E

Test de dissemblance :

• Remarques concernant les critères utilisés dans la stratégie de jugement

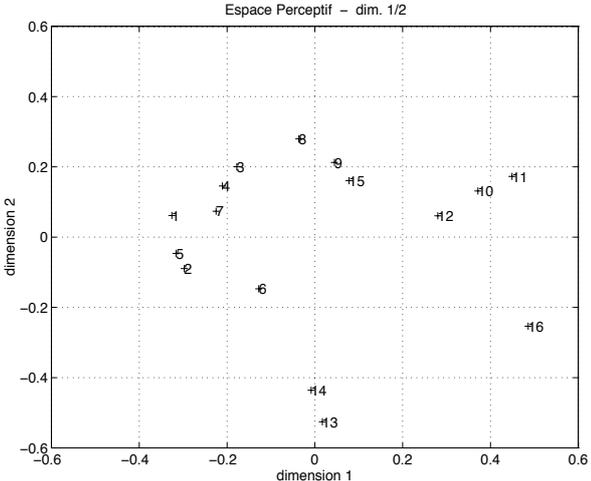
- **Sujet 1 :** " sons constants - sons non constants, répétés - non répétés
sons graduellement augmentés - non graduellement augmentés
sons lourds, pesants - sons relativement plus légers
sons unis - sons non unis (plusieurs bruits) "
- **Sujet 2 :** " intensités des sons, accélération et longueur des sons "
- **Sujet 3 :** " reconnaître les différents véhicules, la vitesse de défilement du son, les fréquences au cours du défilement du son "
- **Sujet 4 :** " nature du bruit (train, avion, voiture) pour mémoriser structure du bruit (aigu, grave, bruits annexes,) "
- **Sujet 5 :** " évolution du son dans la phrase - recherche de rythmes - superposition de sons - sons homogènes, liés - différence entre le son produit par un émetteur et la somme des sons émis par plusieurs sources - recherches des timbres (hauteurs différentes) "
- **Sujet 6 :** " j'ai essayé de me baser sur le timbre plutôt que sur les décibels ou la hauteur "
- **Sujet 7 :** " critère d'ambiance : mouvement, douceur, violence "
- **Sujet 8 :** " rythmes (s'ils existent) - répartition des fréquences - stabilité/variation dans un même son "
- **Sujet 9 :** " si discordances -> différent - autrement, j'ai tenu compte de l'amplitude, la progression du son, les dominantes plutôt aiguës ou graves - j'ai jugé à priori "
- **Sujet 10 :** *sans commentaires*
- **Sujet 11 :** " j'ai essayé de voir si les sons entendus lors des 2 séquences se recoupaient dans leur durée, mais aussi dans leur qualité (métalliques, diffus, tremblant) - agencement des sons (petits sons/sons proéminents) "
- **Sujet 12 :** " fort/pas fort, continu/discontinu, agréable/désagréable, matière (métallique ou vaporeux), net ou noyé dans un écho - réécoute si très semblables ou si écoute inattentive - choix en écoutant, pas après car oubli immédiat "
- **Sujet 13 :** " son grave ou aigu, son isolé ou non isolé, variations temporelles et sonores "
- **Sujet 14 :** " homogénéité du son (présence de bruits de percussions, accélérations multiples) "

- **Sujet 15** : " perception du son dans sa globalité -> situer dans un 'espace' comparaison des 2 sons (positions relatives dans cet espace) affinement (ex. : tonalité, hauteur,) "
- **Sujet 16** : " même sur des sons de provenances différentes, essayer de s'attacher au rythme et à l'intensité du bruit "
- **Sujet 17** : " interprétation des sons en images (taches, points, lignes, courbes de sons = niveaux) + aigus, graves "
- **Sujet 18** : " faire la reconnaissance des aigus et graves - le temps - enregistrer le premier son (dans la tête) par rapport à l'oscillation puis se remémorer celui-ci sur le second "
- **Sujet 19** : " les sons plus ou moins agressifs, plus ou moins forts, la coulée du son et ses tabulations de fréquences (roulement de tambour) "
- **Sujet 20** : intensité, puissance - rythme (planant, agité, chaotique) - hauteur (pas toujours) - dynamisme - impact affectif (gémissant, lumineux, triomphant, calme,) "
- **Sujet 21** : " surtout l'évolution : en fréquence, en amplitude "
- **Sujet 22** : *commentaires sans intérêt*
- **Sujet 23** : " plus ou moins fort - sonorité - évolution temporelle du son "
- **Sujet 24** : " je me suis basé sur les fréquences plus ou moins graves et aiguës, puis sur le mouvement du son "
- **Sujet 25** : " selon les paires, je me suis axée sur le timbre ou sur le rythme, pour les ressemblances et souvent sur les mélodies pour les différences "
- **Sujet 26** : " je n'ai pas cherché à comparer la provenance des sons mais plutôt l'intensité "
- **Sujet 27** : " son continu/ discontinu, grave/aigu, répétitif "
- **Sujet 28** : " étant donnée mes quelques connaissances musicales (5 années de conservatoire), j'ai choisi comme critères le rythme, les crescendo, ... "
- **Sujet 29** : " rythmes, saccades - aigu ou grave - positionnement dans l'espace "
- **Sujet 30** : " reconnaissance de la source sonore possible ou pas - le nombre de sources, la distance, le sens, la vitesse, la montée/descente du bruit ou la stabilité "

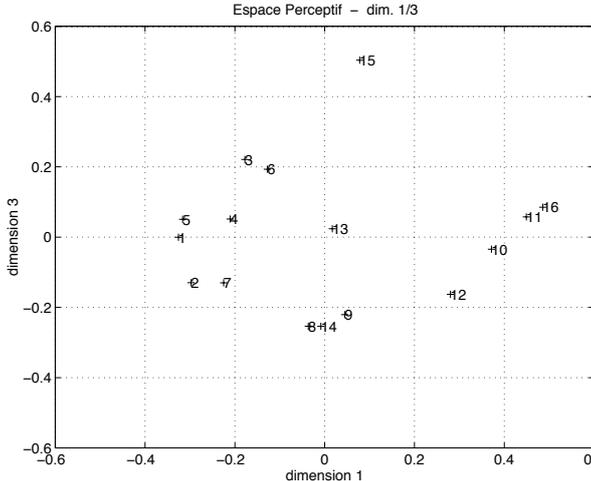
ANNEXE F

Solution multidimensionnelle 'forcée' de l'analyse globale (16 stim. - 30 suj.) :

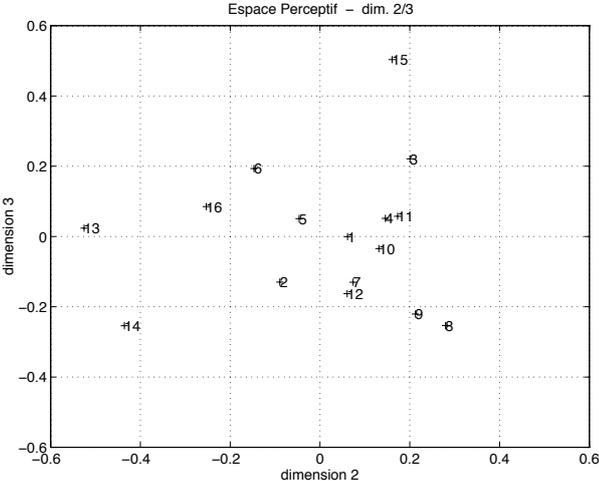
- projection dans le plan (1,2)



- projection dans le plan (1,3)



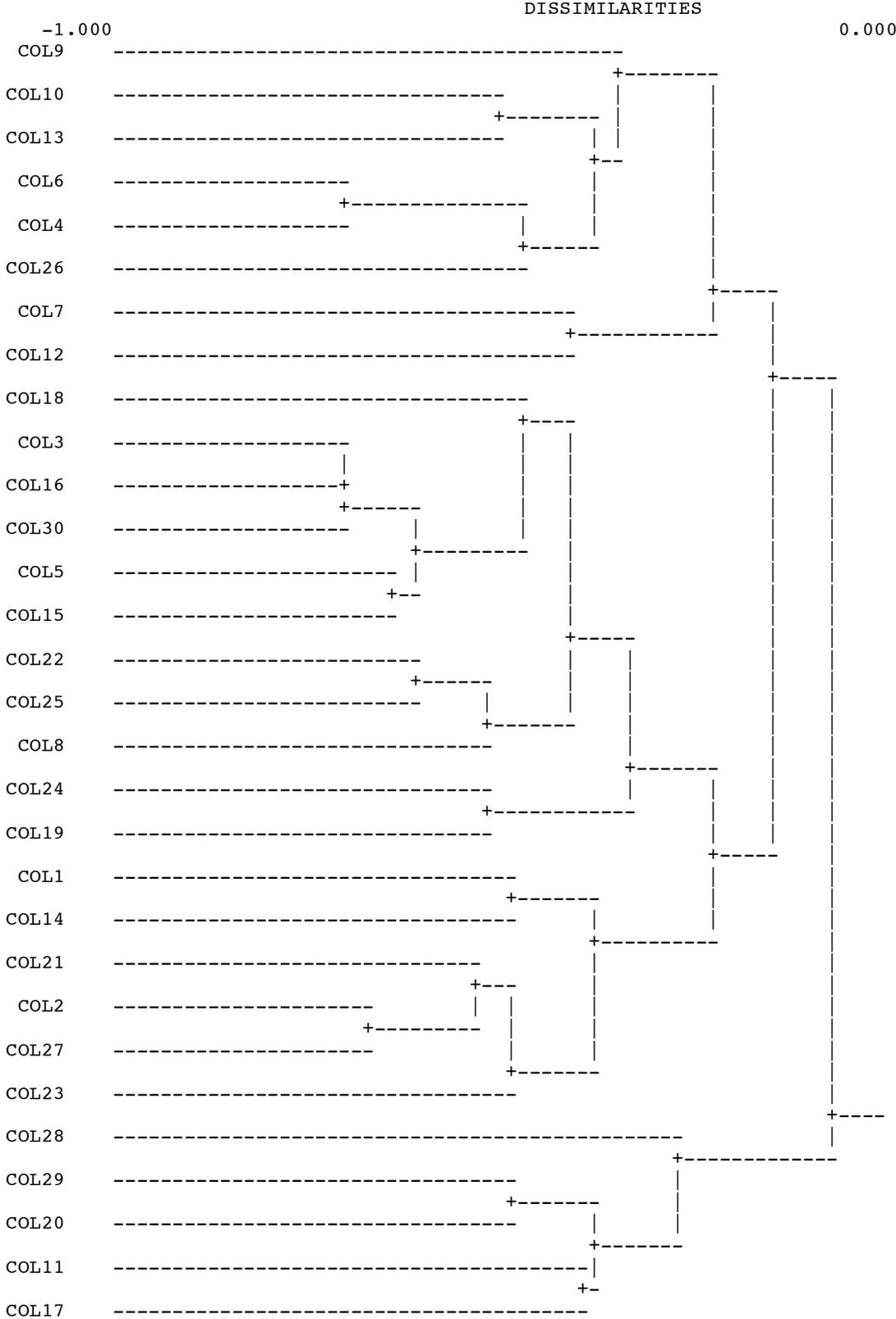
- projection dans le plan (2,3)



ANNEXE G

Analyse de clusters hiérarchiques sur les sujets : • représentation en arbre ultramétrique

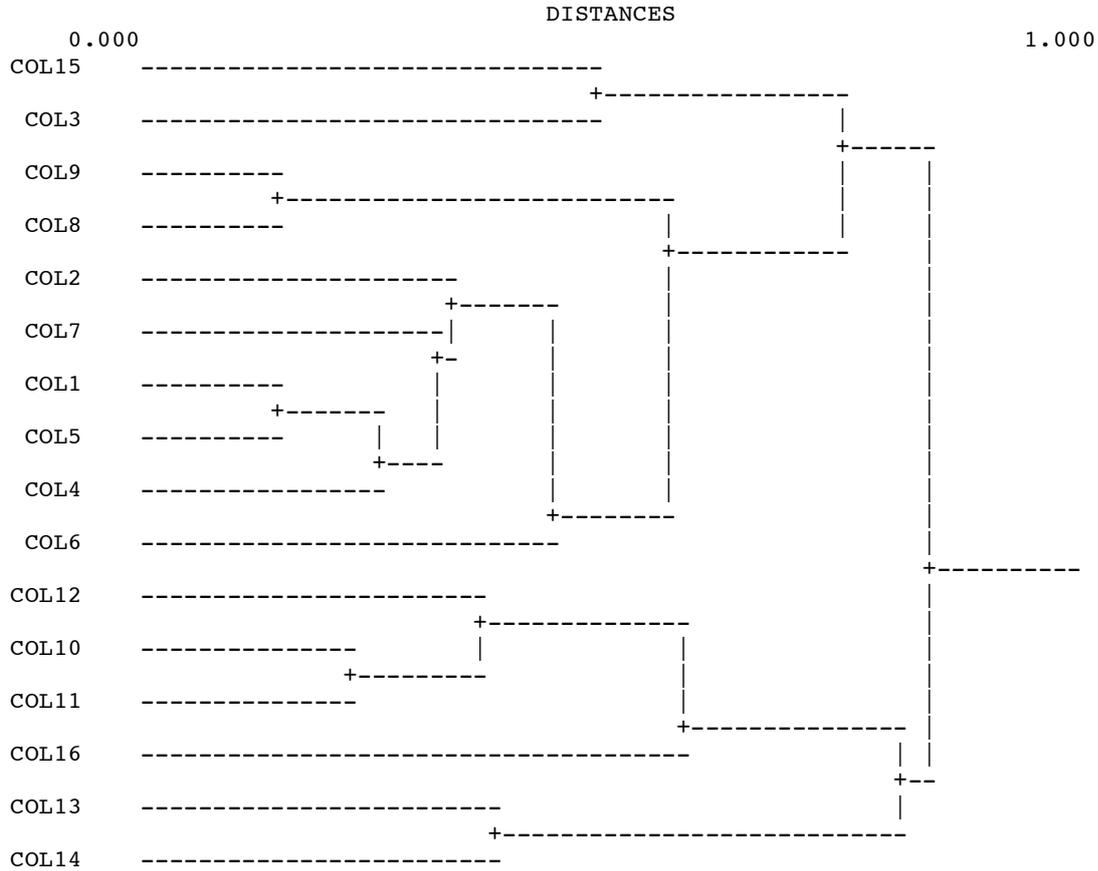
COMPLETE LINKAGE METHOD (FARTHEST NEIGHBOR)
TREE DIAGRAM



ANNEXE H

- Analyse de clusters hiérarchiques sur les stimuli :
- représentation en arbre ultramétrique (résultat avec 30 sujets)

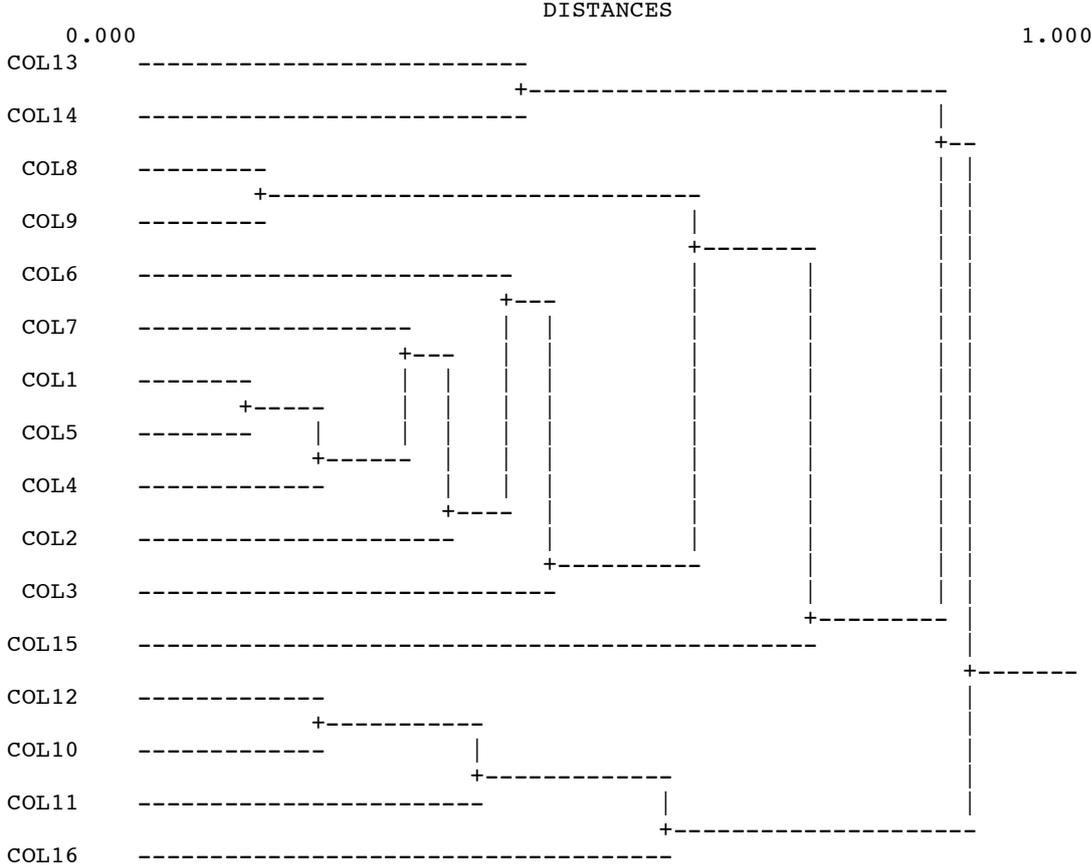
COMPLETE LINKAGE METHOD (FARTHEST NEIGHBOR)
TREE DIAGRAM



ANNEXE H

Analyse de clusters hiérarchiques sur les stimuli : • représentation en arbre ultramétrique (résultat avec 17 sujets)

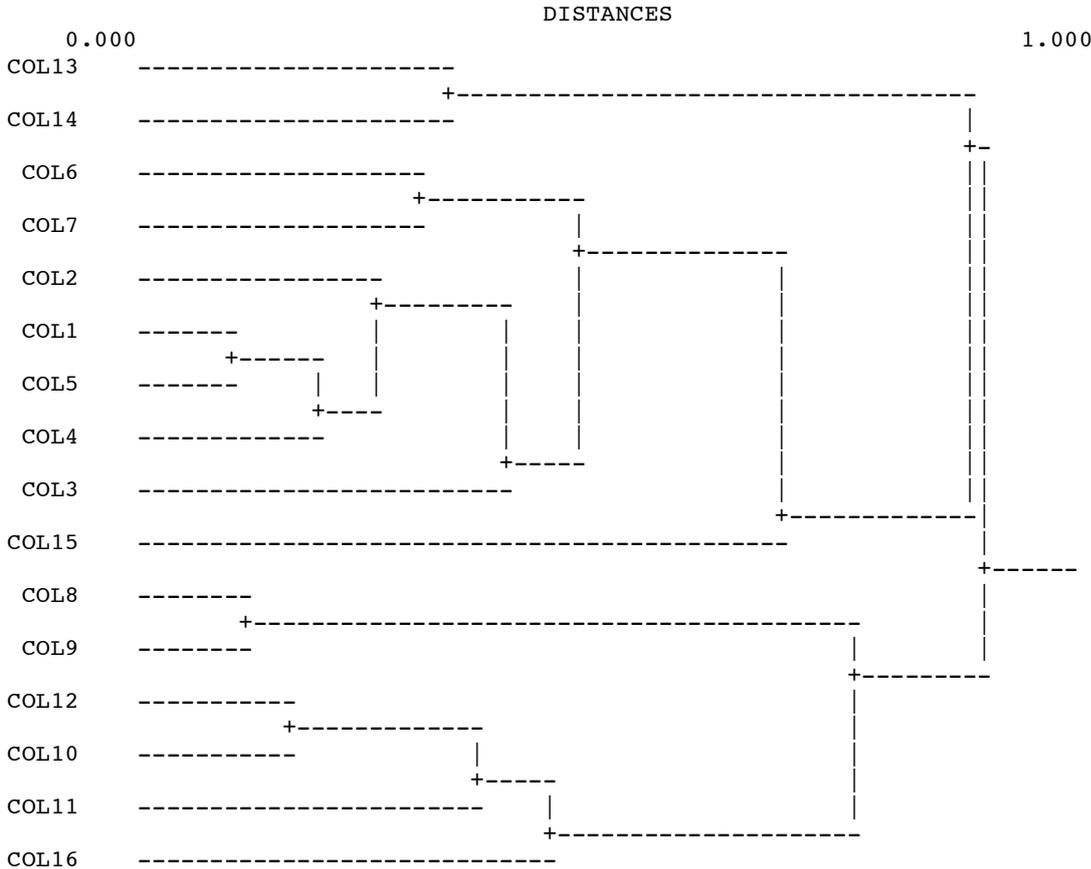
COMPLETE LINKAGE METHOD (FARTHEST NEIGHBOR)
TREE DIAGRAM



ANNEXE H

Analyse de clusters hiérarchiques sur les stimuli : • représentation en arbre ultramétrique (résultat avec 11 sujets)

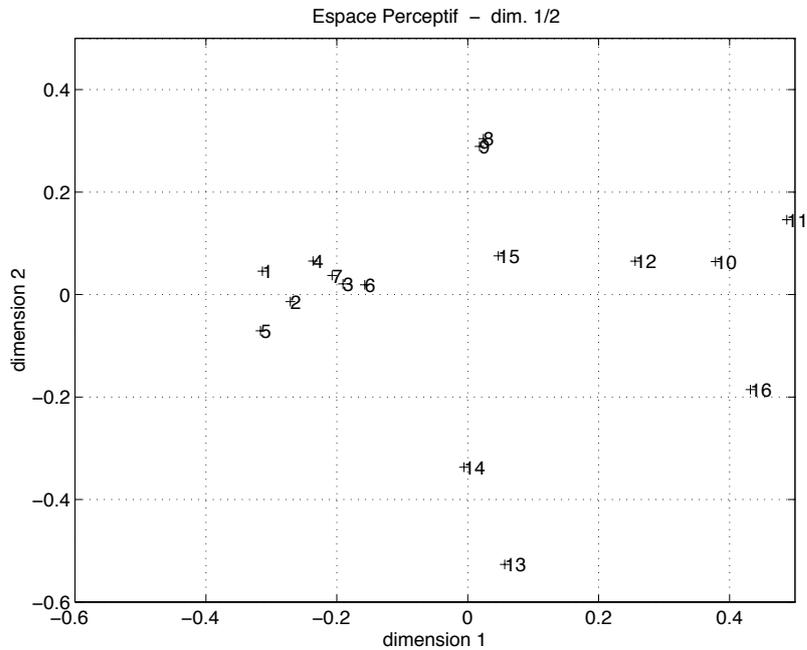
COMPLETE LINKAGE METHOD (FARTHEST NEIGHBOR)
TREE DIAGRAM



ANNEXE I

Analyse multidimensionnelle 'partielle' :

- résultat obtenu avec le 1^{er} sous-groupe de 17 sujets



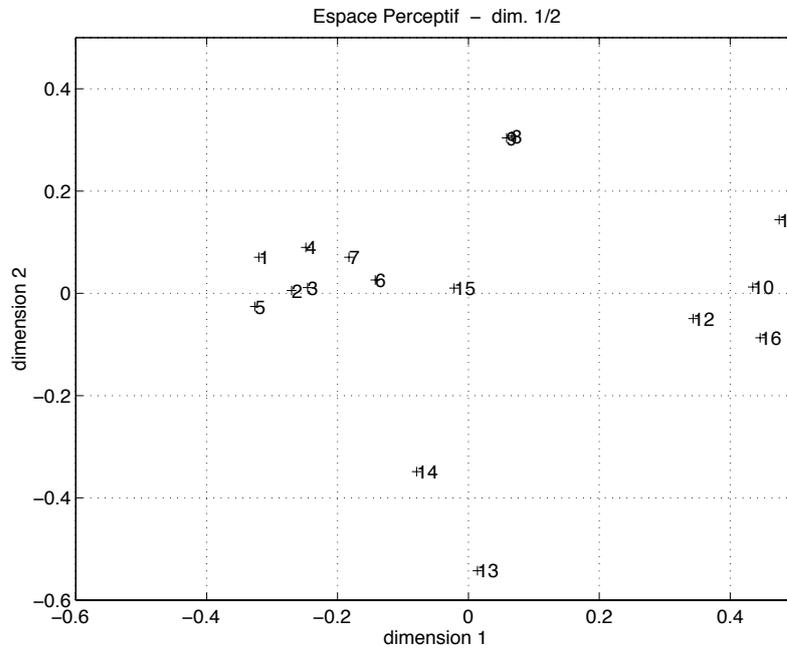
Valeurs des spécificités (racine carré) :

son 1	son 2	son 3	son 4	son 5	son 6	son 7	son 8
0.03	0.21	0.28	0.14	0.00	0.29	0.23	0.21

son 9	son 10	son 11	son 12	son 13	son 14	son 15	son 16
0.14	0.11	0.08	0.21	0.00	0.38	0.51	0.30

ANNEXE I

- Analyse multidimensionnelle 'partielle' :**
- résultat obtenu avec le 2^{ème} sous-groupe de 11 sujets



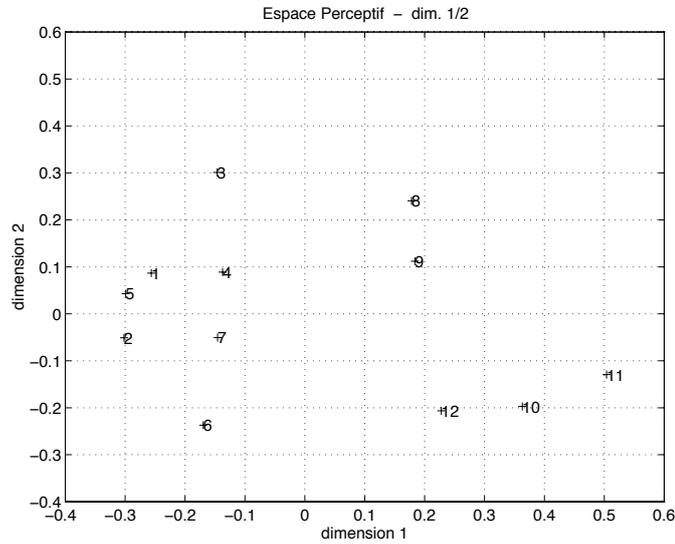
Valeurs des spécificités (racine carré) :

son 1	son 2	son 3	son 4	son 5	son 6	son 7	son 8
0.00	0.21	0.24	0.12	0.00	0.31	0.22	0.23

son 9	son 10	son 11	son 12	son 13	son 14	son 15	son 16
0.00	0.00	0.14	0.18	0.07	0.31	0.52	0.35

ANNEXE J

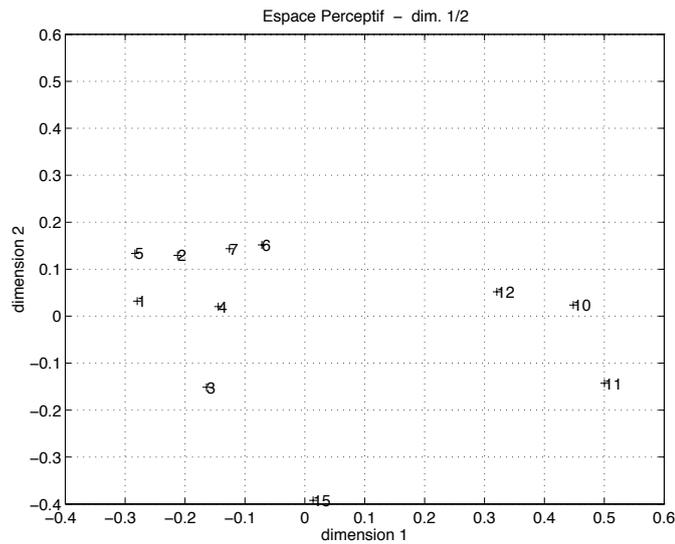
- Analyse multidimensionnelle 'partielle' :**
- **résultat obtenu avec l'ensemble à 12 stimuli**



Valeurs des spécificités (racine carré) :

son 1	son 2	son 3	son 4	son 5	son 6	son 7	son 8	son 9	son 10	son 11	son 12
0.12	0.13	0.12	0.20	0.12	0.21	0.17	0.00	0.18	0.13	0.00	0.24

- **résultat obtenu avec l'ensemble à 10 stimuli**

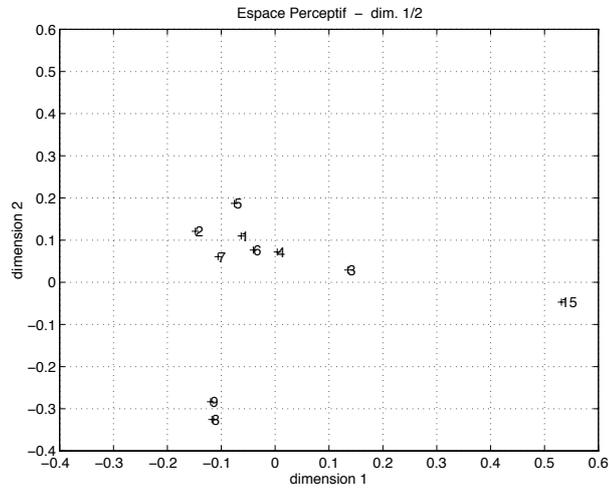


Valeurs des spécificités (racine carré) :

son 1	son 2	son 3	son 4	son 5	son 6	son 7	son 10	son 11	son 12	son 15
0.00	0.27	0.27	0.25	0.03	0.35	0.22	0.00	0.15	0.26	0.39

ANNEXE J

Analyse multidimensionnelle 'partielle' : • résultat obtenu avec l'ensemble à 9 stimuli



Valeurs des spécificités (racine carré) :

son 1	son 2	son 3	son 4	son 5	son 6	son 7	son 8	son 9	son 15
0.10	0.20	0.25	0.20	0.00	0.33	0.20	0.00	0.15	0.00

ANNEXE K

Documents concernant le test de dominance : • feuille de présentation

I.R.C.A.M. - EQUIPE PSYCHOACOUSTIQUE

EXPERIENCE SUR LA PERCEPTION DES SONS DE TRAFIC

Nous nous intéressons à la façon dont les personnes perçoivent les sons de trafic des moyens de transport actuels. Dans cette expérience, les sons ont fait l'objet d'un enregistrement sur site; ils sont au nombre de 16.

- Pour vous donner une idée d'ensemble, nous vous présentons, au début de l'expérience, tous les sons dans un ordre quelconque.
- L'expérience est un test de dominance dans lequel les échantillons sonores sont présentés par paire. Le but du test n'est pas d'essayer d'identifier les sons mais plutôt d'utiliser votre impression globale pour répondre à la question suivante : " **Entre les deux sons de la paire, lequel trouvez-vous le plus désagréable ?** ".
Après l'écoute des deux sons, vous validez votre réponse en cliquant avec la souris soit sur le bouton de gauche [*le son 1 est plus désagréable*], soit sur le bouton de droite [*le son 2 est plus désagréable*].
Une nouvelle paire de sons se présente alors automatiquement à vos oreilles et il vous est impossible de revenir sur la paire précédente.
- Au préalable, une phase d'entraînement est proposée ; elle s'effectue sur 10 paires choisies aléatoirement dans l'ensemble des échantillons de l'expérience. Profitez de ce moment pour vous familiariser avec l'outil informatique et la tâche à effectuer.
- Le test de dominance globale comprend 120 paires. Il s'effectue, à votre convenance, en une seule fois ou bien avec des pauses (1 ou 2 maximum, pour ne pas rallonger démesurément la durée du test).

Remarque : si vous n'êtes pas habitué au maniement de la souris, il est possible de faire votre choix avec les touches '1' et '2' du clavier de l'ordinateur.

ANNEXE K

Documents concernant le test de dominance :
• premières réactions

I.R.C.A.M. - EQUIPE PSYCHOACOUSTIQUE

EXPERIENCE SUR LA PERCEPTION DES SONS DE TRAFIC
- TEST DE DOMINANCE : COMMENTAIRES -

Date : _____

Nom du sujet : _____

N° de l'expérience : _____

Commentaires

• longueur du test : _____

• difficulté du test : _____

• remarques sur le test {pertinence de la question, nature des réponses (instinctive, affective, analytique, etc. ...)} :

ANNEXE L

Test de dominance :

• Remarques concernant les critères utilisés dans la stratégie de jugement

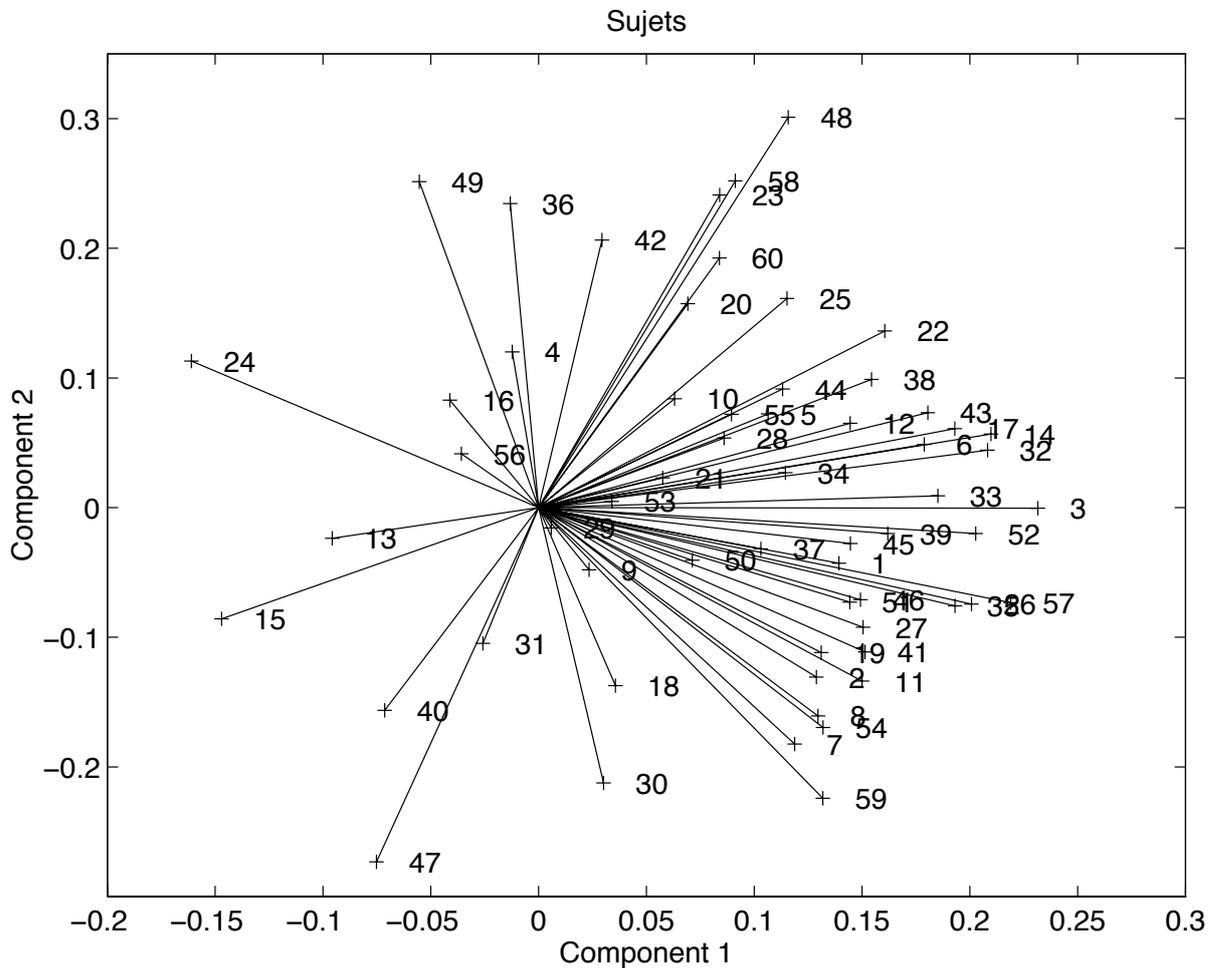
- **Sujet 1** : "problème de critère de jugement : on ne sait pas si l'on doit juger en termes de bruits de trafic (un avion est plus désagréable qu'une voiture) ou en termes de mélodies, musiques, etc. ..."
- **Sujet 2** : "jugement par rapport aux ambiances que représentent les sons"
- **Sujet 3** : "jugement assez analytique, pas très cohérent par moment"
- **Sujet 4** : "jugement plutôt affectif
les fréquences aigus sont plus désagréables que les graves"
- **Sujet 5** : "au départ, jugement affectif
l'ordre de présentation de la paire est important : le premier son prend toute l'impression"
- **Sujet 6** : "jugement plutôt affectif, part de psychologie dans le jugement
avion - train - voiture : va du plus désagréable au moins désagréable
effet de contexte"
- **Sujet 7** : "jugement, en général, affectif et des fois analytique"
- **Sujet 8** : "jugement plutôt instinctif"
- **Sujet 9** : "jugement plutôt instinctif"
- **Sujet 10** : "jugement instinctif"
- **Sujet 11** : "jugement instinctif, certains sons dérangent tout le temps
le son d'avion est beau"
- **Sujet 12** : "on associe le son à un contexte,
jugement parfois instinctif, parfois analytique"
- **Sujet 13** : "jugement plutôt analytique
les modes de transport préférés sont les plus désagréables"
- **Sujet 14** : "jugement plutôt analytique sur la vitesse du bruit, son amplitude"
- **Sujet 15** : "premier critère : sensation d'agression"
- **Sujet 16** : "jugement en fonctions des sons plus aigus et plus graves"
- **Sujet 17** : "jugement davantage analytique
manque de critères précis pour la notion de désagréable"
- **Sujet 18** : "jugement plutôt affectif par rapport à l'environnement quotidien"

- **Sujet 19** : "au début, on juge avec les références (affectif)
le jugement est parfois lié à des images
évolution de la sensation (p. ex. pour l'avion)"
- **Sujet 20** : "les sons les moins désagréables sont les sons de trafic urbain"
- **Sujet 21** : "certains jugements liés à l'affectif et démarche analytique pour d'autres"
- **Sujet 22** : "essayer de juger spontanément mais beaucoup d'affectif
on aurait envie d'avoir une alternative identique - carrément désagréable"
- **Sujet 23** : "l'affectif passe avant"
- **Sujet 24** : "il y a des sons qui parlent plus, vibrations dans les avions crispantes"
- **Sujet 25** : "jugement instinctif"
- **Sujet 26** : "jugement sur la première impression"
- **Sujet 27** : "jugement analytique"
- **Sujet 28** : "jugement affectif, relation image/son"
- **Sujet 29** : "problème de l'occurrence des sons (on est plutôt influencé par le second)
en général, jugement affectivo-analytique"
- **Sujet 30** : "jugement physique et analytique"
- **Sujet 31** : "jugement instinctif"
- **Sujet 32** : "je me mets dans la peau des gens qui subissent tous les jours ces sons"
- **Sujet 33** : "jugement plutôt instinctif"
- **Sujet 34** : "au début, plutôt affectif et ensuite, avec la répétition, c'est la sensation de
nuisance qui influe sur la réponse (-> contradictions éventuelles)"
- **Sujet 35** : "jugement variable : instinctif ou analytique"
- **Sujet 36** : "jugement par rapport à l'environnement sonore quotidien
critère sur la fréquence (les basses sont moins dérangeantes) et sur l'hétérogénéité
(les sons homogènes sont moins dérangeants)"
- **Sujet 37** : "jugement dépendant des sons : mélange (instinctif, affectif et analytique)
par exemple, pour le bus -> jugement affectif"
- **Sujet 38** : "séparation sons connus/sons inconnus, longueur des sons, sons saccadés"
- **Sujet 39** : "bruit aigu et heurté -> désagréable"
- **Sujet 40** : "jugement analytique : harmoniques graves et sons continus -> moins désagréable"

- **Sujet 41** : "jugement instinctif (en théorie)"
- **Sujet 42** : "jugement mélangé : aspect culturel, projection d'un environnement, sensible à certaines fréquences; par exemple, le train fait penser à la fraise du dentiste"
- **Sujet 43** : "jugement affectif, image associée au son sons aigus, saccadés -> dérangeant"
- **Sujet 44** : "jugement affectif et analytique : acuité et rugosité"
- **Sujet 45** : "remise des sons dans le contexte par exemple, bruit d'avion répété -> désagréable trafic routier pendant 2 heures -> désagréable"
- **Sujet 46** : "mélange de critères, désagréable implique de la subjectivité pas stable dans les jugements"
- **Sujet 47** : "jugement sur les bruits qui dérangent (p. ex., l'avion plus fort que les autres) jugements analytiques (bruits stridents à éviter)"
- **Sujet 48** : "jugement instinctif"
- **Sujet 49** : "cliché par rapport aux sons (association du son et de la source)"
- **Sujet 50** : "jugement instinctif, sons stridents -> désagréable"
- **Sujet 51** : "critères : niveau sonore et longueur du son"
- **Sujet 52** : "sons aigus -> désagréable"
- **Sujet 53** : "jugement analytique"
- **Sujet 54** : "au début, les sons ne sont pas si désagréable, à la fin c'est l'inverse jugement sur le moment (au moment où le son est joué) et après le silence (donc après le deuxième son de la paire)"
- **Sujet 55** : "jugement analytique ; harmoniques élevées et effet Doppler -> désagréable jugement affectif : ras le bol des avions " (a habité près de Roissy)
- **Sujet 56** : "l'affectif rejoint l'analytique : analyse des sons préférés (pas trop de différence de niveau, pas de variation de fréquence importante et un spectre homogène)"
- **Sujet 57** : "jugement instinctif mais sons aigus et de forte intensité -> désagréable"
- **Sujet 58** : "sons lisses, sans transitoires -> bonne intégration trafic routier : plusieurs petits sons -> désagréable globalement, trafic routier -> association avec la ville : désagréable"
- **Sujet 59** : "jugement affectif (p. ex. le bruit de bus); agression des sons aigus"
- **Sujet 60** : "association avec la couleur du son (froide, chaude)"

ANNEXE M

- Analyse en composantes principales :
- position des sujets sur les 2 premières composantes



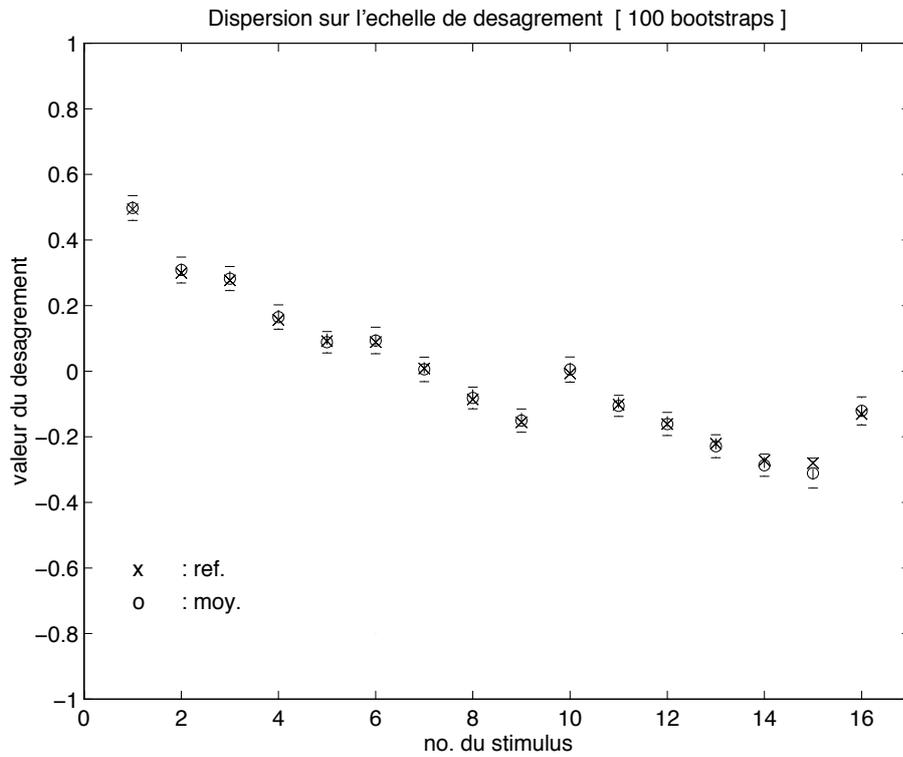
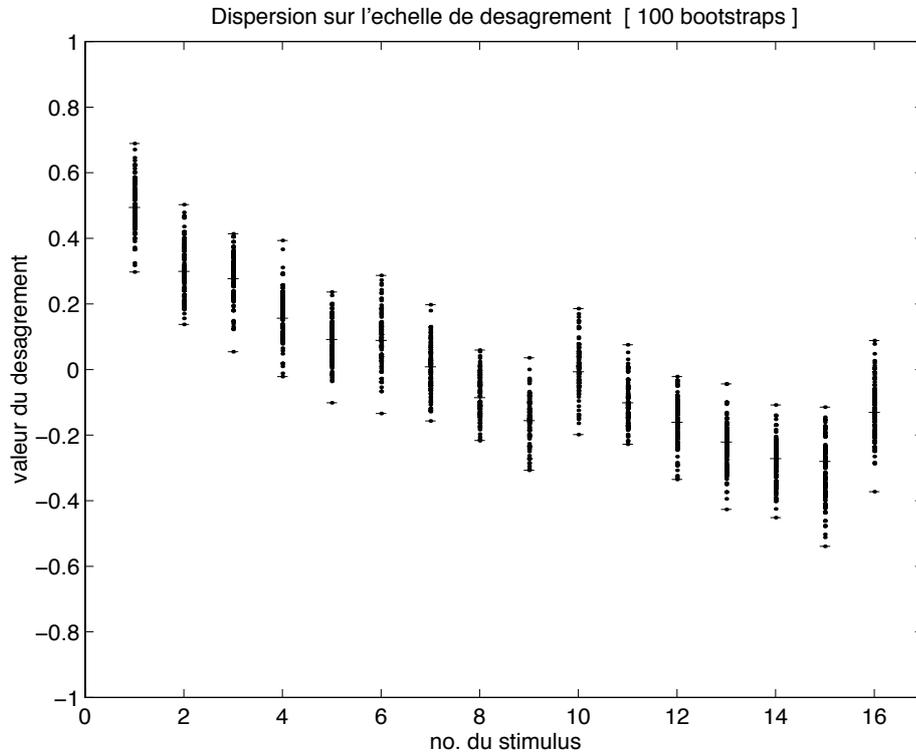
- pourcentage de variance expliquée par les 2 premières composantes

- composante 1 : 24 %
- composante 2 : 11 %

ANNEXE N

Technique de bootstrap :

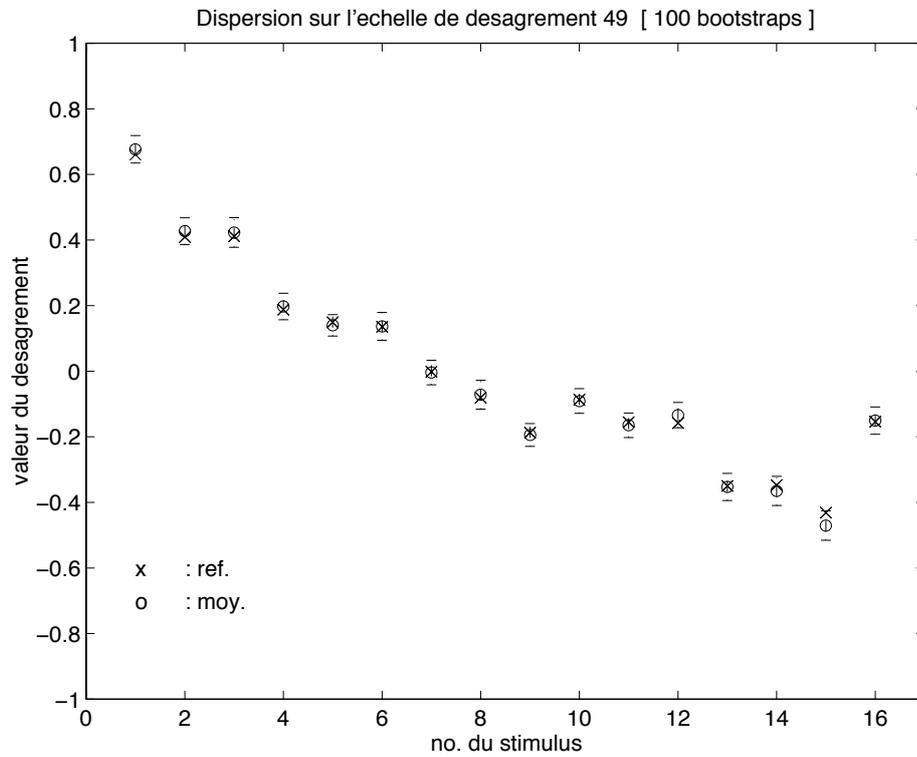
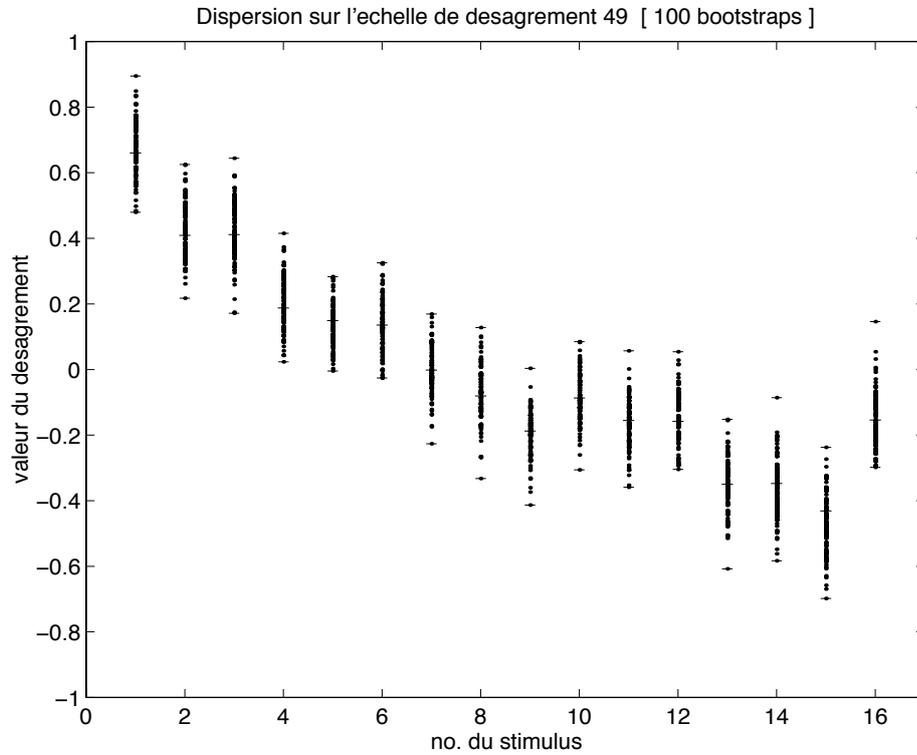
- résultat du bootstrap sur la population de 60 sujets, avec 100 réplifications



ANNEXE N

Technique de 'bootstrap' :

- résultat du bootstrap sur la population des 49 sujets, avec 100 répliquions



ANNEXE O

Corrélation entre paramètres physiques et échelle de désagrément : • récapitulatif des paramètres avec leur descriptif

Nom du paramètre	Descriptif
rmsdB	niveau RMS sur toute la durée du signal
rmsdBcut2TO4	niveau RMS du signal compris entre 2 et 4 secondes
rmsdBcut4TO6	niveau RMS du signal compris entre 4 et 6 secondes
rmsdBcut6TO8	niveau RMS du signal compris entre 6 et 8 secondes
mixrmsdB	niveau RMS interaural moyen
cgs_r15	centre de gravité spectral sur un spectre de résolution 1 Hz
cgs_r11	centre de gravité spectral sur un spectre de résolution 20 Hz
cgs_r8	centre de gravité spectral sur un spectre de résolution 170 Hz
cgsA	centre de gravité spectral sur un spectre de résolution 1 Hz, en dBA
cgsB	centre de gravité spectral sur un spectre de résolution 1 Hz, en dBB
cgsC	centre de gravité spectral sur un spectre de résolution 1 Hz, en dBC
cgsERB	centre de gravité spectral physiologique (banc de 38 filtres auditifs)
cgs_SR16	centre de gravité spectral sur un spectre sous-échantillonné à 16 kHz
cgscut2TO4	centre de gravité spectral sur le spectre du signal compris entre 2 et 4 sec.
cgscut4TO6	centre de gravité spectral sur le spectre du signal compris entre 4 et 6 sec.
cgscut6TO8	centre de gravité spectral sur le spectre du signal compris entre 6 et 8 sec.
spectypdBdev	déviations spectrale du spectre par rapport à un profil-type en dB
spectypdBAdev	déviations spectrale du spectre par rapport à un profil-type en dBA
variatspect	variation spectrale
specfluct	flux spectral
specfluct_SR16	flux spectral calculé sur un signal sous-échantillonné à 16 kHz
energmean2500_5000Hz	énergie moyenne comprise entre 2500 et 5000 Hz
energmean1000_5000Hz	énergie moyenne comprise entre 1000 et 5000 Hz
energmeanERB5bands	énergie moyenne comprise dans 5 filtres ERB => [2632 Hz - 4739 Hz]
energmeanERB14bands	énergie moyenne comprise dans 14 filtres ERB => [949 Hz - 5322 Hz]
energmeanBF0_4kHz	énergie moyenne 'basse fréquence' comprise entre 0 et 4 kHz
energmeanHF4_8kHz	énergie moyenne 'haute fréquence' comprise entre 4 et 8 kHz
energmeanRAP_HFBF	rapport de l'énergie 'haute fréquence' et de l'énergie 'basse fréquence'

ANNEXE O1

Corrélation entre paramètres physiques et échelle de désagrément :
• valeurs des paramètres relatifs au niveau RMS

	rms dB		rmsdB cut2TO4		rmsdB cut4TO6		rmsdB cut6TO8		mixrms
	voix G	voix D	voix G	voix D	voix G	voix D	voix G	voix D	
son 1	64.3	65.7	67.1	67.4	65.9	68.1	64.1	67.3	62.0
son 2	65.8	68.0	69.2	68.2	68.1	71.2	64.2	70.1	64.3
son 3	61.2	65.1	62.2	66.4	63.5	67.2	54.7	61.7	60.8
son 4	60.3	64.4	63.4	67.2	61.8	66.4	50.1	59.1	60.1
son 5	64.7	65.7	67.9	68.5	66.1	67.4	57.4	62.1	62.1
son 6	63.6	65.7	65.6	66.4	65.4	67.6	63.6	66.1	61.5
son 7	67.1	61.5	69.5	64.1	68.8	63.1	66.8	61.4	61.9
son 8	62.9	61.2	65.4	63.6	62.6	63.1	64.4	60.2	59.2
son 9	66.0	63.0	66.0	62.4	68.7	63.7	68.3	66.9	62.1
son 10	61.3	59.8	61.1	60.0	63.3	61.7	58.3	58.0	58.6
son 11	61.6	60.2	60.7	58.4	63.9	63.4	64.2	62.1	59.4
son 12	60.5	55.6	61.9	56.7	61.4	56.8	63.3	58.1	54.4
son 13	63.5	65.4	66.5	66.4	65.2	66.4	63.3	67.2	64.3
son 14	59.5	64.3	57.1	56.5	61.2	65.7	65.0	70.3	61.5
son 15	63.2	63.4	63.4	63.8	65.8	65.4	64.9	64.6	61.8
son 16	65.1	64.0	70.3	68.3	60.8	64.4	51.2	61.1	64.3

• corrélations avec les valeurs de désagrément

Nom du paramètre	Coeff. de corrélation
rmsdB (voix G)	0.19
rmsdB (voix D)	0.44
rmsdBcut2TO4 (voix G)	0.33
rmsdBcut2TO4 (voix D)	0.56
rmsdBcut4TO6 (voix G)	0.23
rmsdBcut4TO6 (voix D)	0.53
rmsdBcut6TO8 (voix G)	-0.25
rmsdBcut6TO8 (voix D)	0.06
mixrmsdB	0.10

ANNEXE O2

Corrélation entre paramètres physiques et échelle de désagrément : • valeurs des paramètres relatifs au CGS

	cgs _r15		cgs _r11		cgs _r8		cgs A	
	voix G	voix D	voix G	voix D	voix G	voix D	voix G	voix D
son 1	2367.9	2279.0	2381.9	2296.8	2489.5	2414.3	2376.2	2288.8
son 2	2432.2	2305.8	2447.6	2323.1	2570.4	2452.5	2487.8	2371.7
son 3	2600.6	2682.8	2604.0	2682.8	2719.6	2822.7	2599.3	2746.2
son 4	2293.8	2385.5	2306.1	2392.6	2452.6	2561.0	2383.8	2505.8
son 5	2436.2	2357.0	2463.7	2381.0	2564.8	2489.4	2400.2	2322.5
son 6	2001.6	1715.3	1986.4	1696.7	2141.9	1856.2	2112.7	1894.0
son 7	2021.9	1959.0	2033.4	1971.4	2149.1	2073.3	2168.7	2061.3
son 8	1833.6	1958.8	1845.1	1968.5	1968.7	2082.7	1940.2	2058.8
son 9	1650.0	1878.1	1652.1	1882.2	1754.7	1959.4	1809.9	2028.9
son 10	2907.5	3344.6	2887.2	3335.5	2949.6	3351.8	2973.4	3332.2
son 11	3309.8	3605.5	3258.4	3554.7	3483.0	3698.3	3440.4	3648.6
son 12	3212.6	3476.6	3200.3	3473.0	3405.5	3621.0	3276.2	3469.2
son 13	1610.9	1537.6	1622.6	1543.6	1762.6	1740.4	2142.5	1978.1
son 14	3940.8	4166.1	3945.2	4166.6	4020.4	4335.0	4155.5	4126.2
son 15	2991.9	2928.4	2958.7	2878.0	2957.2	2871.8	2943.5	2929.8
son 16	2907.8	3071.8	2753.3	2967.6	3081.3	3323.2	3687.1	3528.2

	cgs B		cgs C		cgs ERB		cgs _SR16	
	voix G	voix D	voix G	voix D	voix G	voix D	voix G	voix D
son 1	2179.0	2097.7	2193.3	2105.8	3384.5	3416.2	2095.9	1985.4
son 2	2334.2	2215.4	2342.1	2217.2	2947.8	3060.2	2332.4	2193.2
son 3	2318.6	2430.1	2310.4	2395.7	3616.1	3754.4	2087.6	2160.8
son 4	2146.9	2251.8	2133.6	2225.6	3024.5	3226.7	2038.2	2135.2
son 5	2208.7	2128.1	2224.9	2136.9	3583.2	3625.2	2081.6	1962.3
son 6	1842.9	1635.8	1819.2	1598.8	3252.3	2951.9	1701.3	1555.7
son 7	1876.7	1802.5	1861.4	1799.2	3219.8	2933.9	1769.8	1710.3
son 8	1699.7	1798.9	1672.8	1777.8	3169.3	3158.1	1560.0	1654.3
son 9	1526.0	1715.1	1491.0	1687.9	3233.2	3219.9	1375.5	1555.3
son 10	2423.1	2781.5	2311.4	2694.4	4120.0	4200.7	1722.8	2056.1
son 11	2747.8	3002.2	2568.9	2872.1	4288.3	4309.1	1787.0	2106.6
son 12	2676.0	2890.0	2522.4	2770.4	4221.3	4173.6	1836.1	2069.1
son 13	1626.7	1558.6	1482.8	1429.9	2985.4	3004.6	1459.4	1398.1
son 14	3464.5	3401.5	3299.0	3288.5	4384.4	4535.5	2603.8	2208.1
son 15	2422.6	2382.9	2364.1	2314.2	4140.4	4119.2	1778.7	1766.1
son 16	2539.9	2587.2	2133.4	2263.3	4449.7	4380.3	1378.3	1447.8

ANNEXE O2

Corrélation entre paramètres physiques et échelle de désagrément :
• valeurs des paramètres relatifs au CGS (suite)

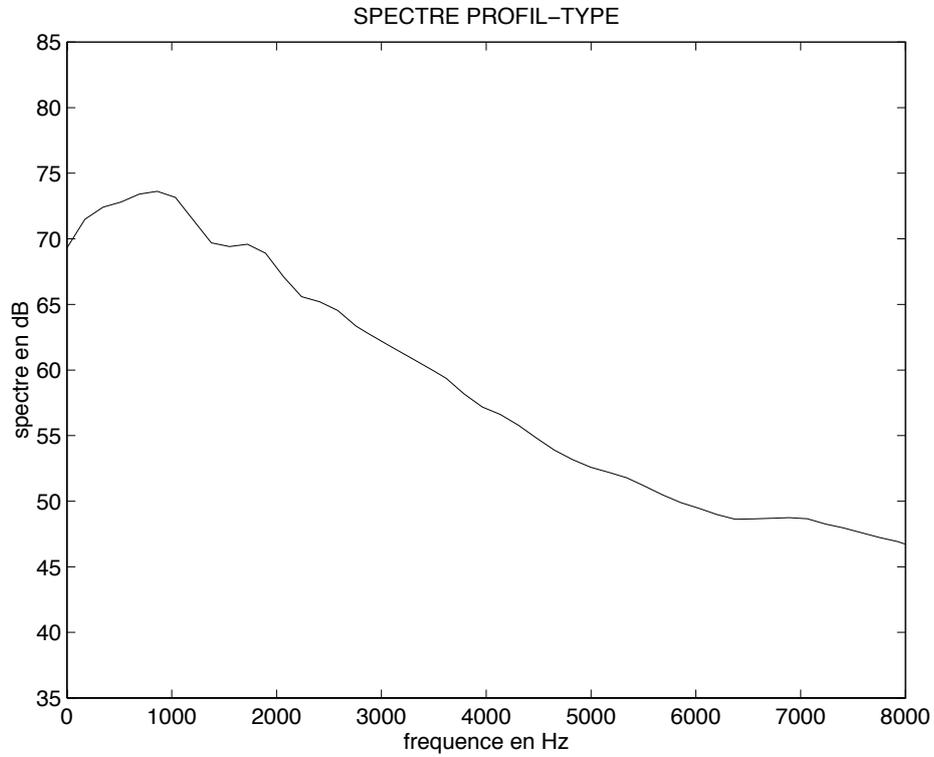
	cgscut 2TO4		cgscut 4TO6		cgscut 6TO8	
	voix G	voix D	voix G	voix D	voix G	voix D
son 1	2385.1	2272.3	2363.9	2268.5	2404.5	2298.6
son 2	2392.8	2149.7	2496.8	2352.0	2467.7	2352.9
son 3	2461.4	2548.8	2660.2	2686.2	2792.4	2835.3
son 4	2290.8	2401.1	2324.3	2430.0	2020.6	2237.9
son 5	2461.6	2454.9	2294.6	2213.6	2246.7	2139.8
son 6	2081.7	1766.8	1942.4	1765.9	1948.4	1761.4
son 7	2030.6	1965.5	1956.1	1932.6	2152.6	2065.8
son 8	1909.0	2007.1	1640.5	1975.2	1881.4	1898.4
son 9	1571.0	2099.8	1657.8	1729.8	1638.0	1801.9
son 10	2955.1	3537.3	2794.6	3353.8	3136.9	3340.6
son 11	3200.5	3412.9	3451.1	3785.5	3392.5	3642.0
son 12	3125.4	3339.0	3144.8	3264.3	3432.0	3761.1
son 13	1683.0	1630.8	1584.0	1529.4	1669.8	1582.7
son 14	2218.4	1908.0	3472.5	3406.8	4296.9	4477.5
son 15	2811.2	2811.5	3096.7	2980.9	2976.7	2933.2
son 16	2649.9	2812.3	3276.5	3426.0	3307.5	3289.3

• corrélations avec les valeurs de désagrément

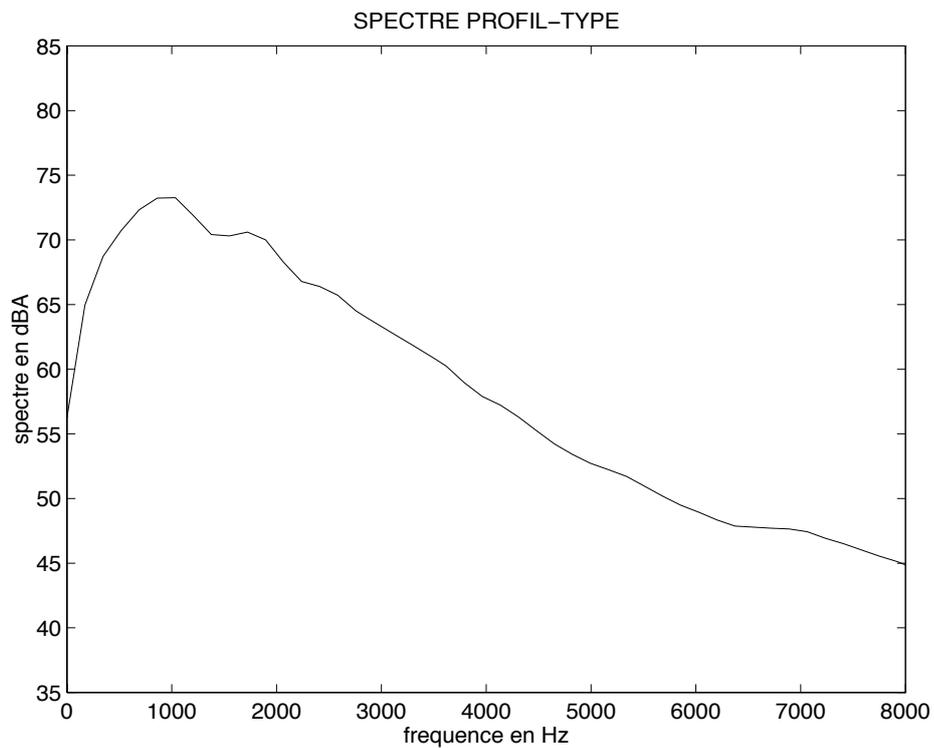
Nom du paramètre	Coefficient de corrélation	
	voix G	voix D
cgs_r15	-0.24	-0.31
cgs_r11	-0.22	-0.29
cgs_r8	-0.22	-0.28
cgsA	-0.35	-0.37
cgsB	-0.16	-0.22
cgsC	-0.05	-0.15
cgsERB	-0.45	-0.40
cgs_SR16	0.39	0.38
cgscut2TO4	-0.02	-0.12
cgscut4TO6	-0.20	-0.25
cgscut6TO8	-0.27	-0.31

ANNEXE O3

- Corrélation entre paramètres physiques et échelle de désagrément :
- profil spectral 'type' (dB) obtenu par analyse en composantes principales



- profil spectral 'type' (dBA) obtenu par analyse en composantes principales



ANNEXE O3

Corrélation entre paramètres physiques et échelle de désagrément :
 • valeurs du paramètre relatif à la Déviation Spectrale

	spectyp dBdev		spectyp dBAddev	
	voix G	voix D	voix G	voix D
son 1	515.3	/	505.4	/
son 2	1288.5	/	1275.9	/
son 3	195.8	/	182.8	/
son 4	984.8	/	959.1	/
son 5	412.3	/	415.6	/
son 6	799.8	/	775.1	/
son 7	702.4	/	707.6	/
son 8	668.0	/	640.4	/
son 9	529.5	/	504.8	/
son 10	817.3	/	824.5	/
son 11	1021.3	/	1032.2	/
son 12	946.4	/	955.1	/
son 13	1029.7	/	1003.1	/
son 14	841.9	/	859.6	/
son 15	982.1	/	998.3	/
son 16	1088.8	/	1092.5	/

• corrélations avec les valeurs de désagrément

Nom du paramètre	Coefficient de corrélation	
	voix G	voix D
spectypdBdev	-0.36	/
spectypdBAddev	-0.38	/

ANNEXE O4

Corrélation entre paramètres physiques et échelle de désagrément :
• valeurs des paramètres relatifs aux variations spectro-temporelles

	variat spect		spect fluct		spectluct _SR16	
	voix G	voix D	voix G	voix D	voix G	voix D
son 1	0.9609	0.9621	48.0	85.3	33.9	56.6
son 2	0.9636	0.9655	37.3	72.4	42.4	64.3
son 3	0.9420	0.9382	36.7	67.0	31.7	36.9
son 4	0.9565	0.9553	84.5	52.9	86.8	33.9
son 5	0.9327	0.9350	106.8	109.1	38.1	37.4
son 6	0.9567	0.9605	17.4	11.3	9.0	8.5
son 7	0.9484	0.9387	19.7	18.5	15.1	12.9
son 8	0.9647	0.9616	48.4	36.8	33.4	14.5
son 9	0.9666	0.9623	21.7	29.0	18.9	41.1
son 10	0.9508	0.9370	26.4	19.0	8.4	35.9
son 11	0.9500	0.9367	97.9	169.6	6.5	60.9
son 12	0.9538	0.9419	89.7	88.1	26.9	30.2
son 13	0.9540	0.9604	103.9	71.1	110.4	72.9
son 14	0.9181	0.9436	1104.8	1507.9	423.2	443.3
son 15	0.9500	0.9505	39.6	39.2	20.0	44.2
son 16	0.9343	0.9441	321.2	218.3	457.2	163.8

• corrélations avec les valeurs de désagrément

Nom du paramètre	Coefficient de corrélation	
	voix G	voix D
variatspect	0.27	0.20
specfluct	-0.35	-0.30
specfluct_SR16	-0.31	-0.33

ANNEXE O5

Corrélation entre paramètres physiques et échelle de désagrément :
• valeurs des paramètres relatifs à la répartition de l'énergie en fréquence

	energmea 2500_5000		energmea 1000_5000		energmea ERB5band		energmea ERB14ban	
	n		n		n		n	
	voix G	voix D	voix G	voix D	voix G	voix D	voix G	voix D
son 1	62.7	62.3	67.0	67.4	45.1	44.8	49.1	49.9
son 2	68.3	68.6	71.1	72.1	51.4	52.1	52.4	53.9
son 3	58.8	63.7	62.8	67.1	41.6	46.4	45.1	49.2
son 4	59.8	65.4	63.5	68.5	42.5	47.9	45.3	49.9
son 5	62.6	61.3	67.2	66.8	45.1	43.9	49.3	49.3
son 6	57.0	57.4	61.5	62.1	39.4	39.9	44.4	45.4
son 7	63.2	55.8	66.8	60.2	45.5	38.4	49.1	42.9
son 8	54.5	54.5	60.8	60.2	37.0	36.7	44.1	43.1
son 9	55.4	55.3	61.2	60.4	37.6	37.4	44.4	43.1
son 10	55.7	57.6	59.9	60.7	37.7	39.4	42.1	42.2
son 11	55.8	57.2	59.1	59.8	37.7	38.9	40.7	41.2
son 12	56.2	53.6	59.7	56.5	38.3	35.5	41.4	37.8
son 13	58.1	59.1	59.8	62.2	41.5	42.5	41.3	44.3
son 14	58.8	58.3	59.9	61.6	40.5	39.9	40.5	43.8
son 15	57.2	56.6	61.2	60.4	39.3	38.6	43.6	42.9
son 16	53.3	53.5	54.9	56.2	35.2	35.3	35.6	37.7

	energmea BF0_4kHz		energmea HF8_4kHz		energmea RAP_HFBF	
	n		n		n	
	voix G	voix D	voix G	voix D	voix G	voix D
son 1	67.6	68.6	51.1	51.2	0.76	0.75
son 2	70.6	72.6	51.0	51.4	0.72	0.71
son 3	64.3	68.4	50.2	55.3	0.78	0.81
son 4	64.7	69.4	47.5	52.9	0.73	0.76
son 5	67.9	68.4	50.9	50.5	0.75	0.74
son 6	64.2	65.5	48.2	47.2	0.75	0.72
son 7	69.2	62.9	53.1	46.2	0.77	0.74
son 8	64.2	62.8	44.2	45.1	0.69	0.72
son 9	65.4	63.6	46.3	46.8	0.71	0.74
son 10	62.3	62.0	49.2	51.9	0.79	0.84
son 11	61.5	61.3	50.7	53.0	0.83	0.86
son 12	62.0	58.1	49.8	47.9	0.80	0.83
son 13	62.8	65.9	44.9	45.0	0.71	0.68
son 14	60.6	64.3	56.5	59.1	0.93	0.92
son 15	63.7	63.0	51.4	51.6	0.81	0.82
son 16	57.4	59.2	48.3	49.1	0.84	0.83

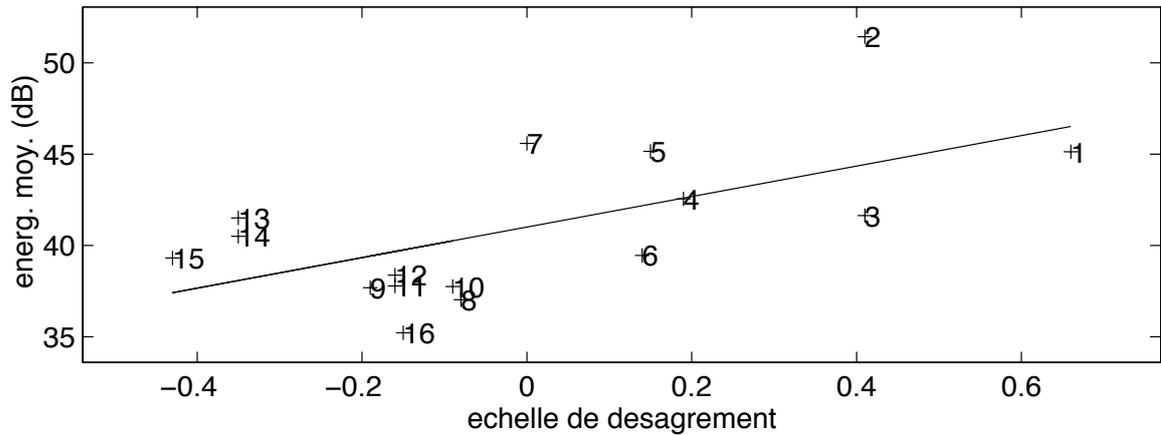
ANNEXE O5

Corrélation entre paramètres physiques et échelle de désagrément :
 • corrélations avec les valeurs de désagrément

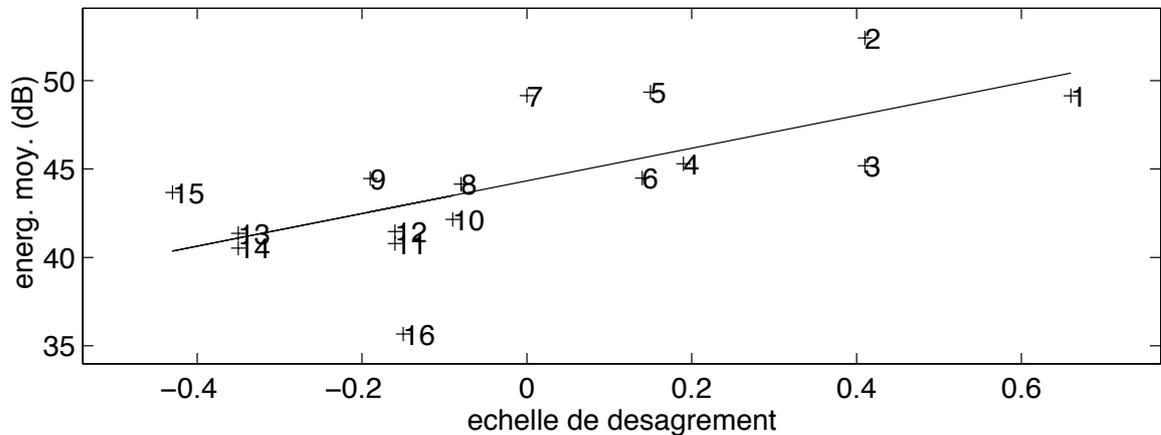
Nom du paramètre	Coefficient de corrélation	
	voix G	voix D
energmean2500_5000Hz	0.61	0.69
energmean1000_5000Hz	0.68	0.74
energmeanERB5bands	0.61	0.69
energmeanERB14bands	0.67	0.74
energmeanBF0_4kHz	0.60	0.68
energmeanHF4_8kHz	0.07	0.14
energmeanRAP_HFBF	-0.35	-0.35

• représentation des droites de régression linéaire

Correlation desagscale49/energmeanERBbands (5 bands) – R = 0.613

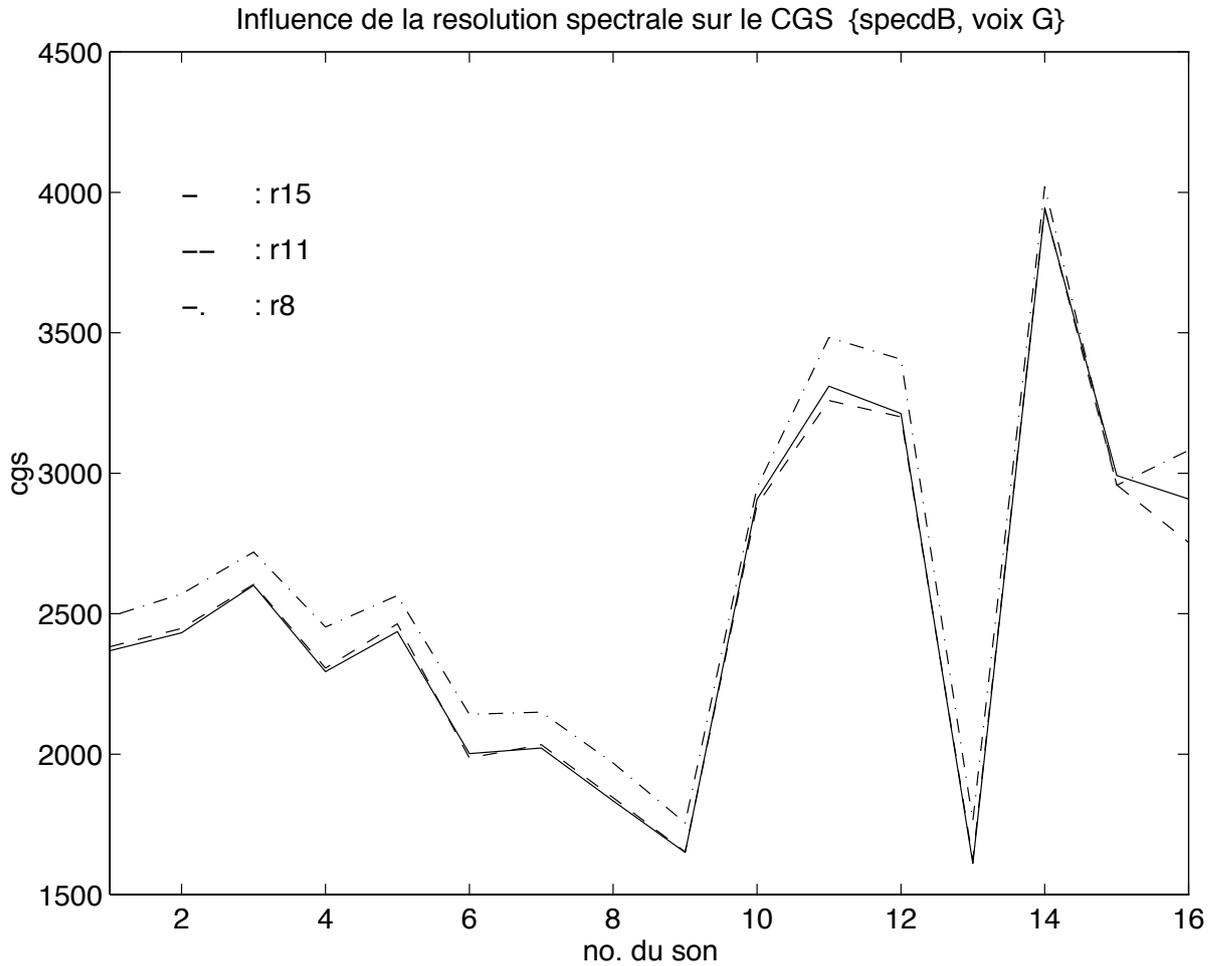


Correlation desagscale49/energmeanERBbands (14 bands) – R = 0.6747



ANNEXE P

- Influence de la résolution spectrale sur la valeur du CGS :**
• comparaison des CGS sur un spectre calculé avec $r = 15, 11$ et 8



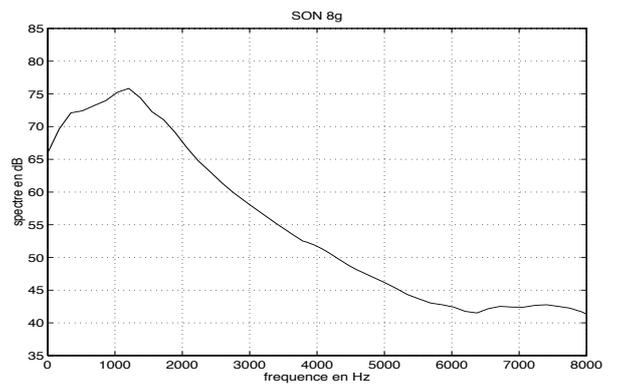
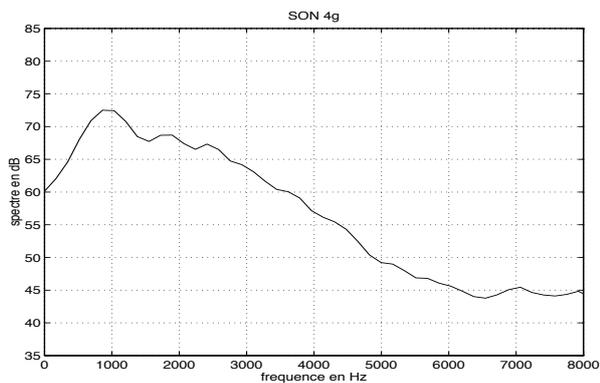
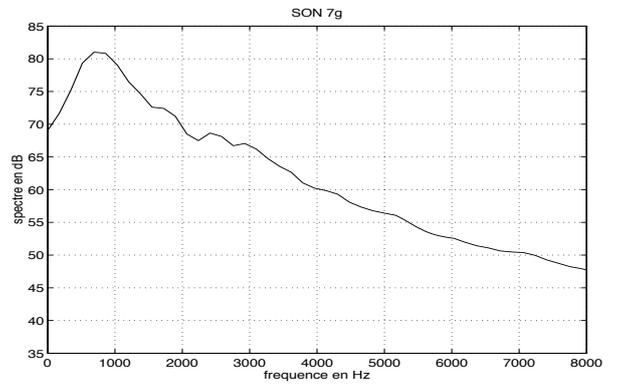
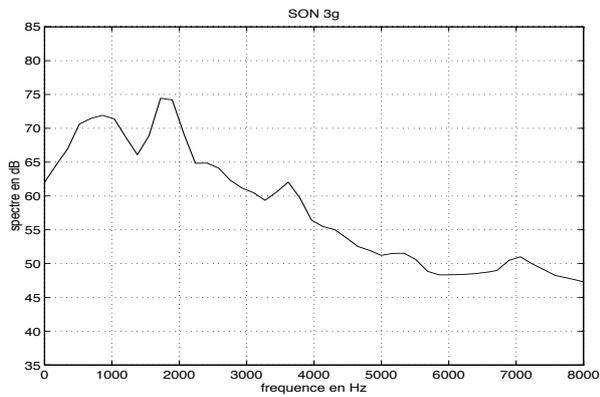
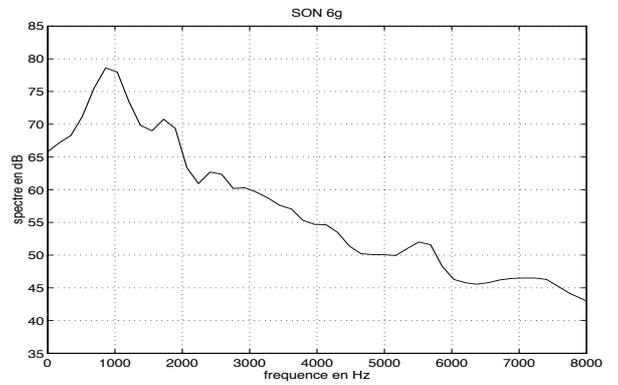
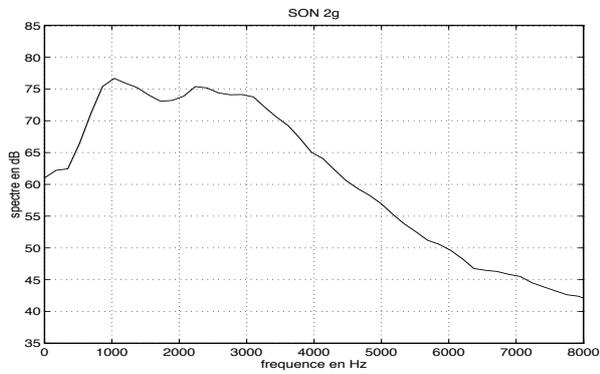
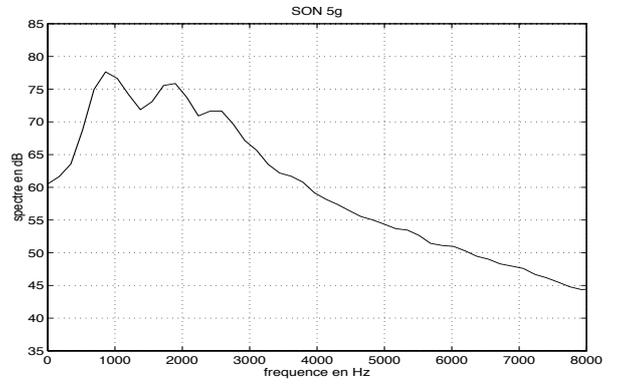
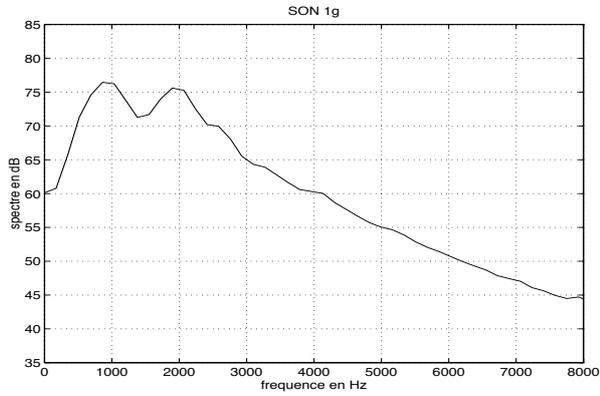
ANNEXE Q

Caractéristiques du banc de filtres auditifs :
• fréquence centrale, largeur de bandes et gain des filtres ERB

N° du filtre	Fréquence centrale	Larg. de bande (ERB)	gain
38	19710.490	2152.231	-332.179
37	17616.653	1926.224	-328.585
36	15742.691	1723.951	-325.197
35	14065.514	1542.918	-321.778
34	12564.458	1380.895	-318.317
33	11221.029	1235.887	-314.816
32	10018.674	1106.106	-311.280
31	8942.579	989.953	-307.711
30	7979.485	885.998	-304.113
29	7117.526	792.959	-300.488
28	6346.081	709.690	-296.840
27	5655.647	635.165	-293.170
26	5037.715	568.466	-289.482
25	4484.672	508.771	-285.776
24	3989.705	455.345	-282.054
23	3546.715	407.529	-278.319
22	3150.243	364.734	-274.571
21	2795.405	326.433	-270.812
20	2477.828	292.154	-267.044
19	2193.600	261.475	-263.266
18	1939.220	234.017	-259.480
17	1711.551	209.443	-255.687
16	1507.791	187.449	-251.888
15	1325.427	167.765	-248.083
14	1162.214	150.148	-244.272
13	1016.139	134.381	-240.457
12	885.404	120.270	-236.638
11	768.398	107.640	-232.816
10	663.678	96.337	-228.989
9	569.955	86.220	-225.160
8	486.074	77.166	-221.328
7	411.001	69.063	-217.494
6	343.812	61.811	-213.658
5	283.678	55.320	-209.819
4	229.859	49.511	-205.979
3	181.692	44.312	-202.137
2	138.582	39.658	-198.293
1	100.000	35.494	-194.446

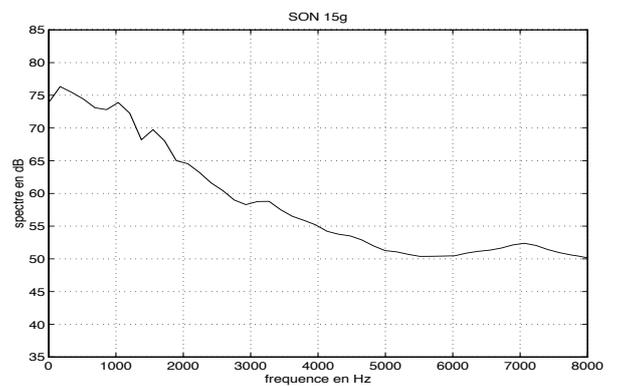
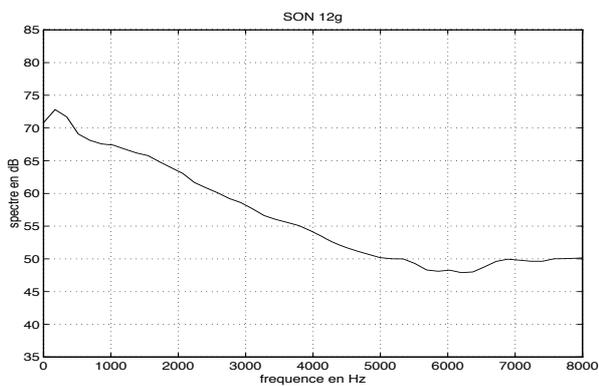
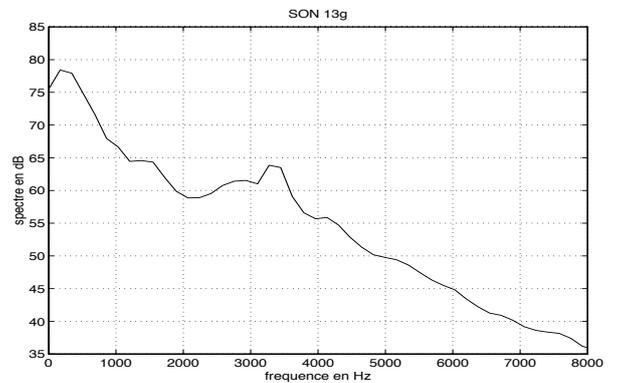
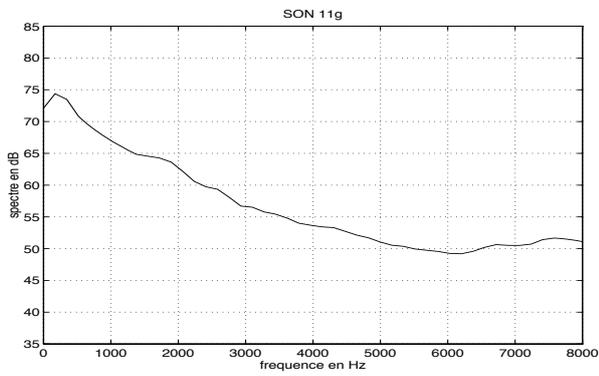
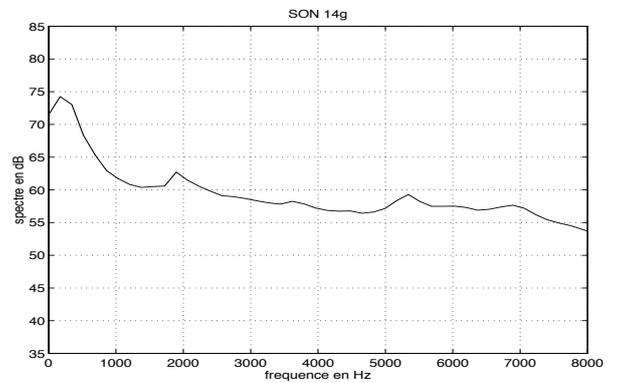
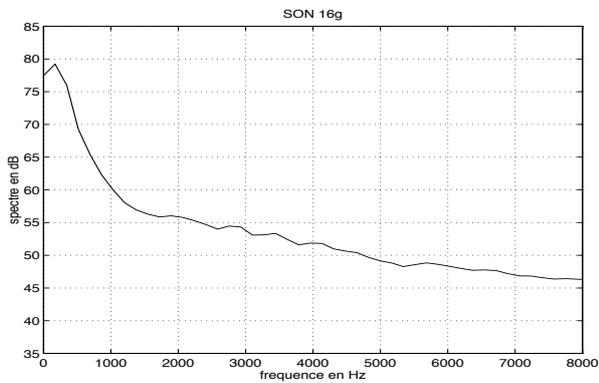
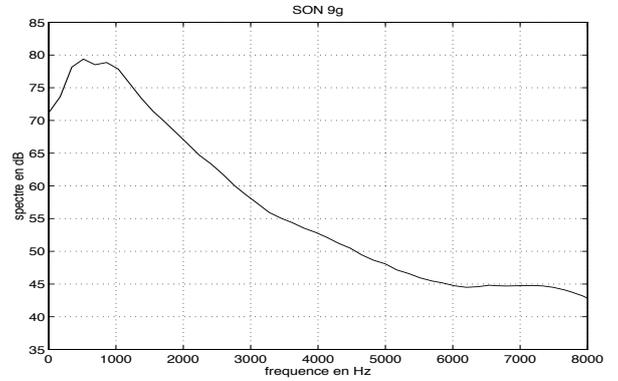
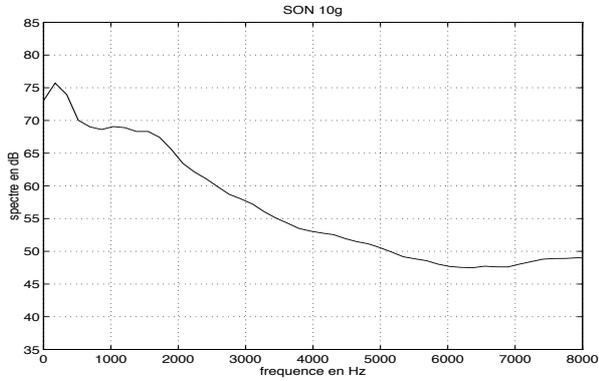
ANNEXE R

Profils spectraux des échantillons sonores : • spectres (dB, voix G) dans l'ordre décroissant du désagrément



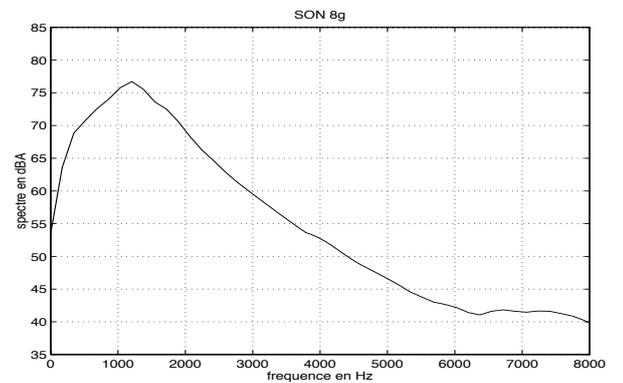
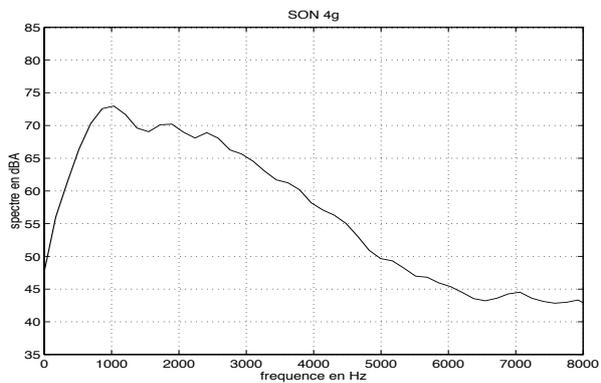
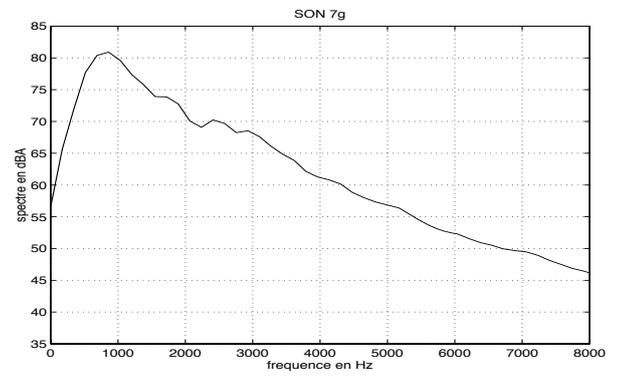
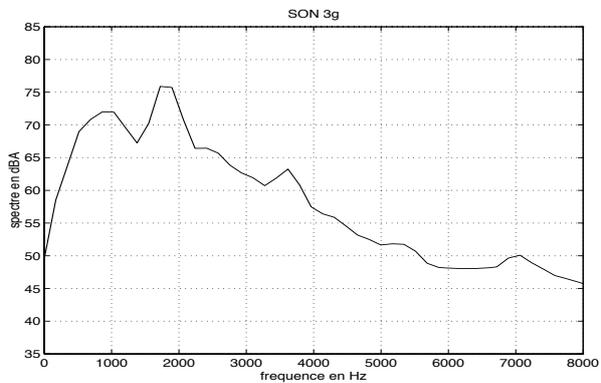
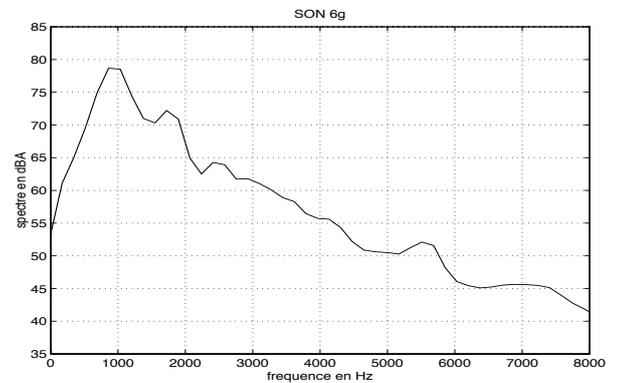
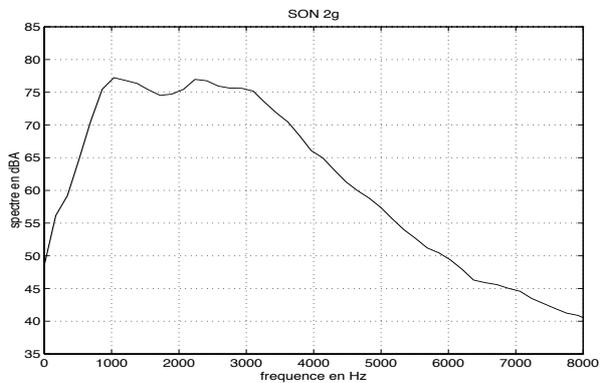
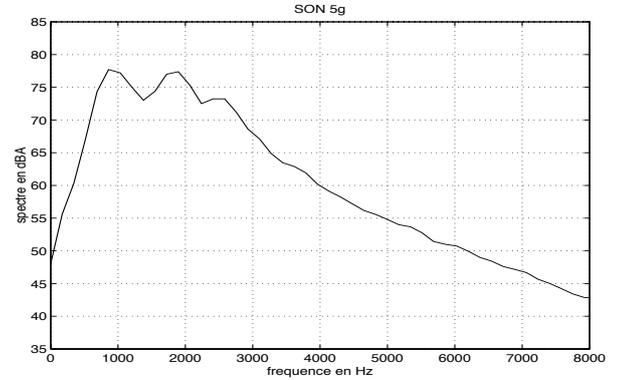
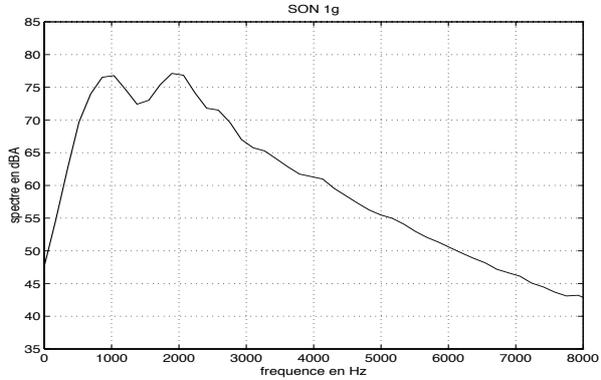
ANNEXE R

Profils spectraux des échantillons sonores : • spectres (dB, voix G) dans l'ordre décroissant du désagrément (suite)



ANNEXE R

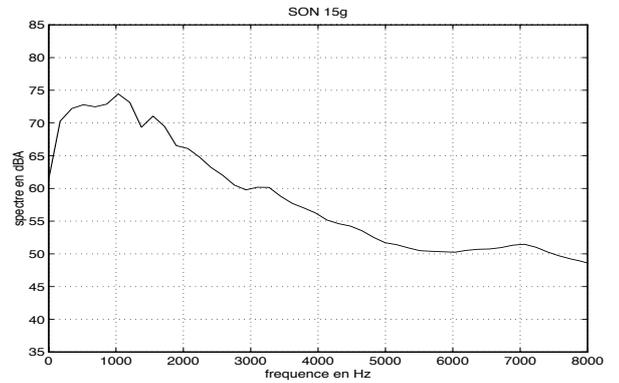
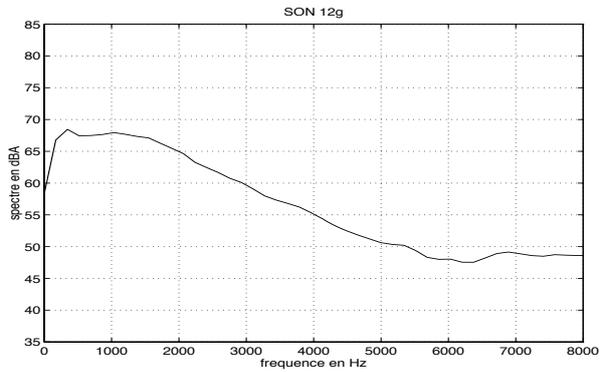
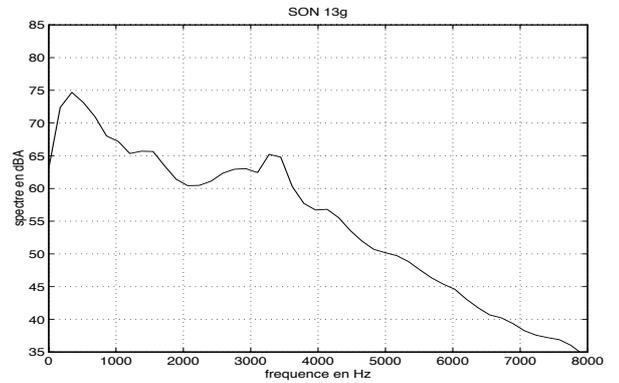
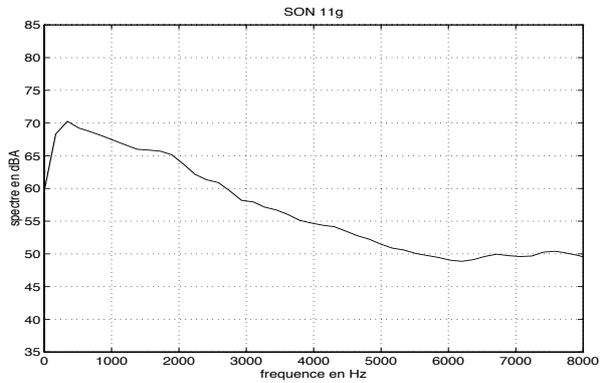
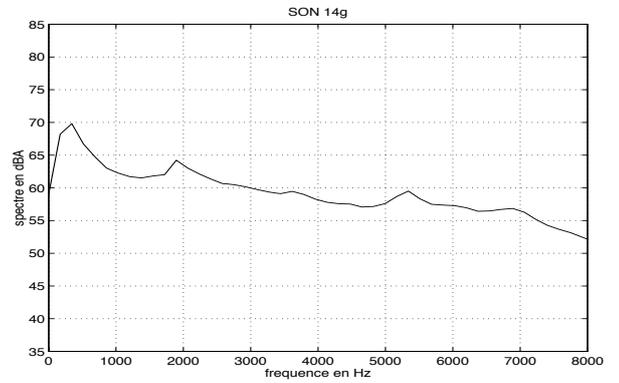
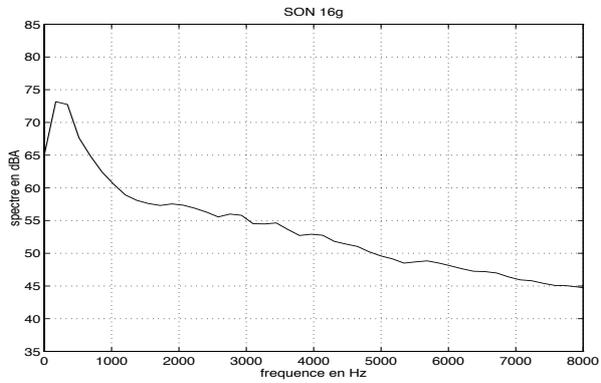
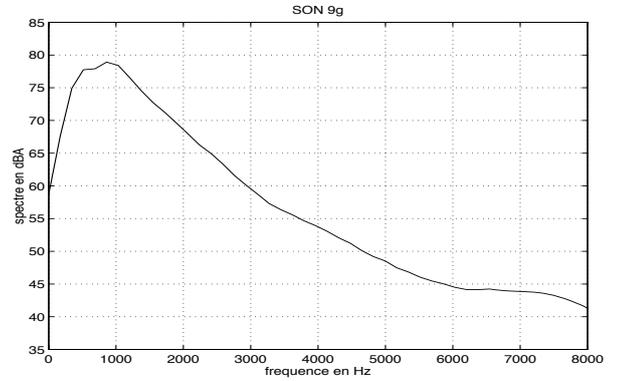
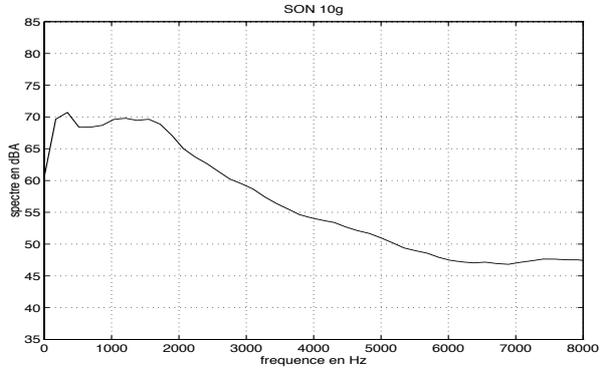
Profils spectraux des échantillons sonores : • spectres (dBA, voix G) dans l'ordre décroissant du désagrément



ANNEXE R

Profils spectraux des échantillons sonores :

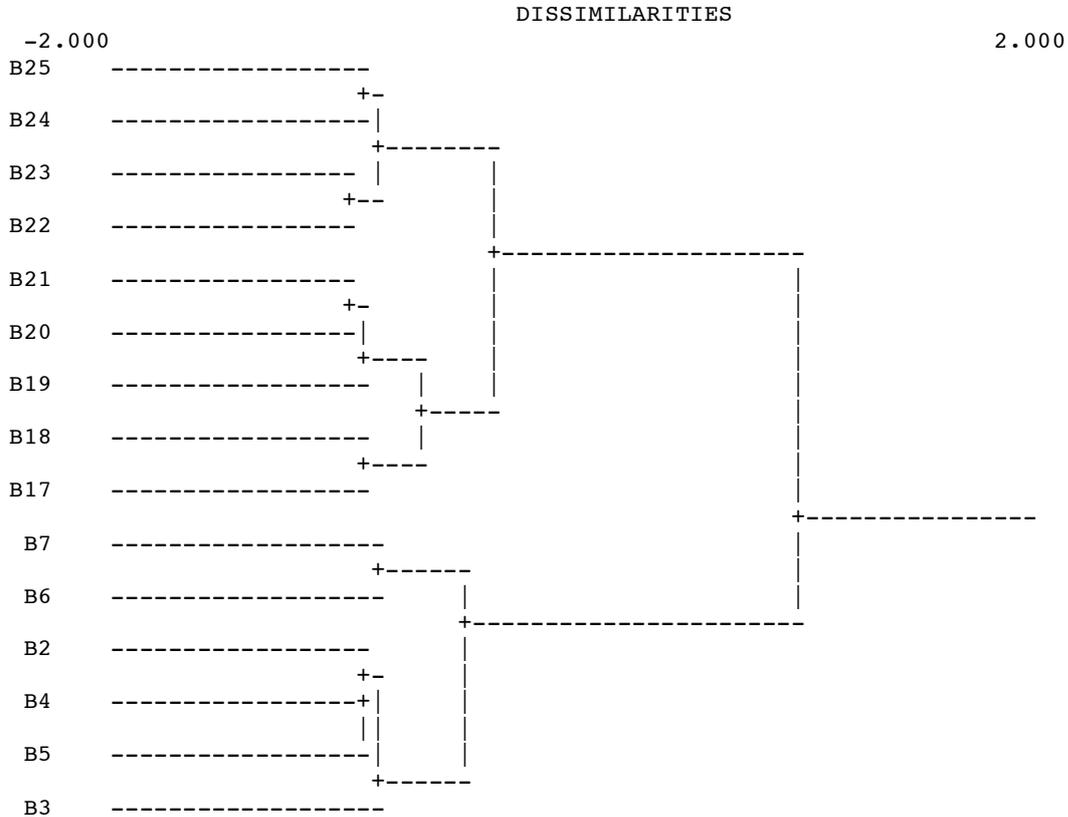
- spectres (dBA, voix G) dans l'ordre décroissant du désagrément (suite)



ANNEXE S1

- Analyse de clusters hiérarchiques des filtres auditifs (ERB)
sur la base de leur intercorrélation à travers les 16 sons - voix GAUCHE :**
- **clustering des filtres dont la corrélation avec l'échelle de désagrément > 0.5**

COMPLETE LINKAGE METHOD (FARTHEST NEIGHBOR)
TREE DIAGRAM



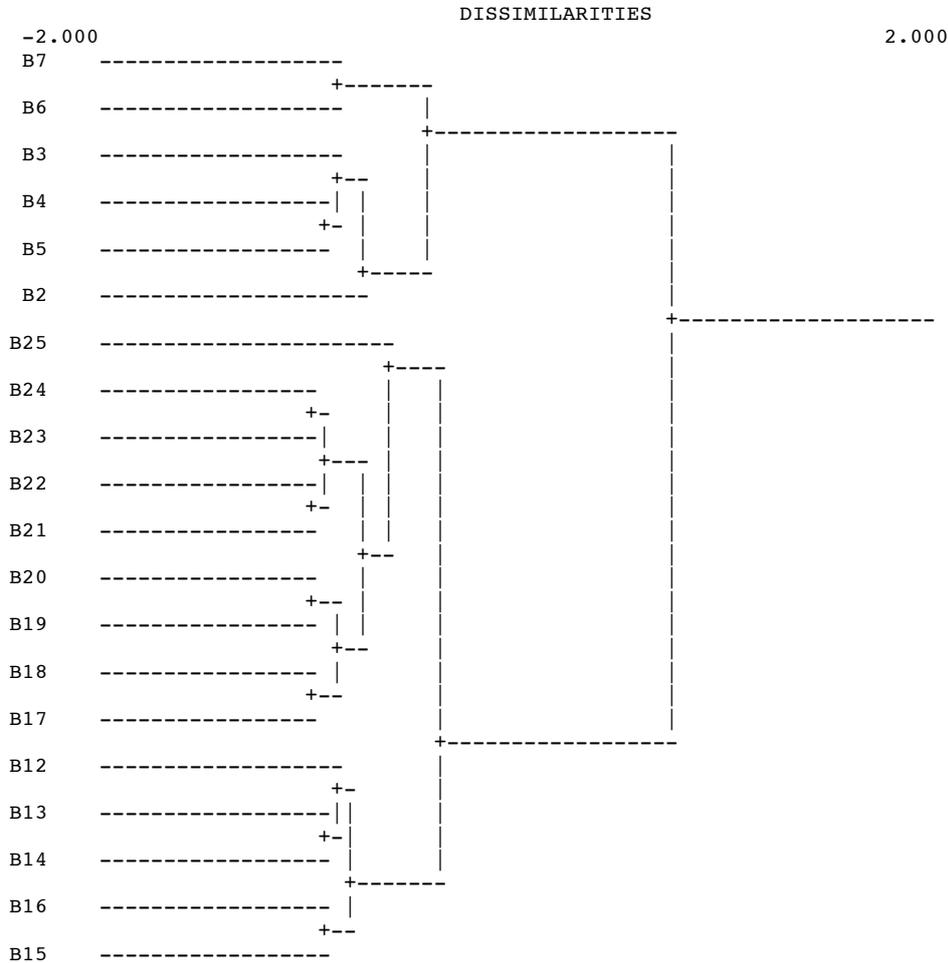
- **Niveau moyen des bandes regroupées pour l'analyse de régression multiple**

Son n°	B2-5	B6-7	B17-21	B22-25
1	36.950	39.900	52.920	44.050
2	39.575	36.700	54.800	50.350
3	42.450	41.150	48.800	41.100
4	39.150	39.500	47.840	41.575
5	39.150	37.750	53.460	43.725
6	42.875	43.450	45.400	38.800
7	49.700	49.700	50.000	44.825
8	45.700	47.500	46.200	35.900
9	52.350	53.150	45.840	36.725
10	56.100	45.400	43.560	37.100
11	52.025	47.200	41.940	37.325
12	51.125	45.250	42.800	37.750
13	57.900	51.300	41.160	41.125
14	53.050	45.850	41.220	40.475
15	55.475	47.400	44.080	38.975
16	57.400	44.750	35.980	35.000

ANNEXE S2

- Analyse de clusters hiérarchiques des filtres auditifs (ERB)
sur la base de leur intercorrélation à travers les 16 sons - voix DROITE :**
- **clustering des filtres dont la corrélation avec l'échelle de désagrément > 0.5**

COMPLETE LINKAGE METHOD (FARTHEST NEIGHBOR)
TREE DIAGRAM



- **Niveau moyen des bandes regroupées pour l'analyse de régression multiple**

Son n°	B3-5	B6-7	B17-24
1	36.100	41.000	50.688
2	38.933	38.500	55.463
3	45.867	45.700	50.550
4	41.433	44.200	50.775
5	38.300	39.000	50.350
6	44.300	46.000	43.950
7	39.267	42.300	41.850
8	43.200	46.200	41.612
9	45.300	51.650	41.588
10	51.467	44.450	41.600
11	49.700	46.900	40.700
12	44.333	40.750	37.525
13	55.500	52.900	44.487
14	52.400	51.750	42.000
15	51.167	47.500	41.350
16	53.100	44.750	36.775

ANNEXE T

Etude de la stabilité des résultats de la régression multiple :
• données issues de la méthode de 'jack-knife'

	voix gauche désag = $i + s \cdot B_{2-5}$			voix droite désag = $s1 \cdot B_{17-24} + s2 \cdot B_{6-7}$			
	i	s	R ²	i	s1	s2	R ²
- son n°1	1,054	-0,022	0,675	-0,301	0,023	-0,016	0,761
- son n°2	1,191	-0,025	0,688	-0,358	0,030	-0,021	0,739
- son n°3	1,185	-0,025	0,712	-0,120	0,023	-0,021	0,742
- son n°4	1,299	-0,027	0,717	-0,332	0,027	-0,019	0,735
- son n°5	1,350	-0,028	0,749	-0,180	0,028	-0,023	0,799
- son n°6	1,265	-0,026	0,719	-0,302	0,026	-0,019	0,761
- son n°7	1,247	-0,026	0,722	-0,342	0,026	-0,018	0,745
- son n°8	1,281	-0,026	0,750	-0,309	0,026	-0,019	0,741
- son n°9	1,234	-0,026	0,712	-0,260	0,026	-0,020	0,736
- son n°10	1,340	-0,028	0,778	-0,351	0,027	-0,018	0,748
- son n°11	1,245	-0,026	0,715	-0,319	0,026	-0,019	0,742
- son n°12	1,235	-0,026	0,719	-0,058	0,024	-0,022	0,743
- son n°13	1,265	-0,026	0,699	-0,517	0,027	-0,015	0,737
- son n°14	1,208	-0,025	0,720	-0,428	0,026	-0,016	0,727
- son n°15	1,205	-0,025	0,697	-0,297	0,025	-0,018	0,752
- son n°16	1,310	-0,027	0,731	-0,438	0,028	-0,018	0,744
global	1,245	-0,026	0,719	-0,308	0,026	-0,019	0,744

N.b. : en grisé dans le tableau, les valeurs minimales et maximales des paramètres