

Caractérisation perceptive de bruits

**Patrick Susini, Nicolas Misdariis,
Suzanne Winsberg,**
Institut de Recherche et de Coordination
Acoustique/Musique (IRCAM),
1 place Igor-Stravinsky,
75004 Paris,
Tél. : 01 44 78 43 18,
Fax : 01 44 78 15 40,
Email : {susini,misdariis,smc,winsberg}@ircam.fr

Stephen McAdams,
Laboratoire de Psychologie Expérimentale (CNRS),
Université René Descartes, EPHE,
28 rue Serpente,
75006 Paris,
Tél. : 01 40 51 98 56,
Fax : 01 40 51 70 85

Les enjeux économiques et sociaux dans le domaine de la qualité sonore nous obligent à réfléchir sérieusement sur le problème de la caractérisation subjective des bruits produits par les machines qui peuplent notre environnement quotidien. Il ne s'agit pas simplement de réduire le niveau du bruit, car pour opérer efficacement une machine d'usine ou une voiture, par exemple, nous avons besoin de certaines des informations contenues dans les sons qu'ils génèrent pour nous informer sur le fonctionnement de l'appareil : le changement de timbre de notre aspirateur nous informe lorsque le sac est plein, le bon fonctionnement d'un ordinateur est confirmé par le son du disque dur. En outre certains aspects des bruits contribuent de façon significative à l'image (auditive!) projetée par l'objet qui les produit et sont importants pour souligner son identité face à une clientèle de plus en plus exigeante. L'analyse perceptive des bruits devient donc cruciale. Le problème est que les bruits sont complexes et multiformes à la fois du point de vue acoustique et du point de vue de la perception auditive humaine. Nous avons alors besoin de techniques appropriées pour les caractériser de ces deux points de vue et d'établir une relation étroite entre eux. Les recherches dans ce sens sont menées au sein de l'équipe Psychoacoustique à l'IRCAM qui réunit des compétences en psychologie expérimentale, acoustique, traitement de signal et analyse de données subjectives.

Des premiers pas par l'équipe dans ce sens ont été faits dans le domaine des sons d'instruments de musique. Si on présente deux sons produits par deux instruments différents à la même hauteur et avec la même intensité perçue (la sonie), un auditeur peut toujours les distinguer sur le plan de leur timbre et même évaluer le degré de différence entre eux sur une échelle numérique, effectuant ainsi un jugement de dissemblance. Le seul instrument de mesure approprié pour mettre en évidence la structure de ces différences est un être humain entendant.

Depuis les travaux de Plomp [1], plusieurs études psychoacoustiques ont montré clairement que le timbre est un attribut multidimensionnel de la perception auditive en utilisant les techniques d'analyse multidimensionnelle de dissemblances.

Grey [2] a dégagé trois dimensions perceptives saillantes relatives à ce facteur qui sont partagées par tous les timbres testés. Utilisant une autre technique d'analyse, EXSCAL [3], Krumhansl [4] a également trouvé un espace en trois dimensions partagées, mais avec un poids (appelé "spécificité") plus ou moins fort sur certains timbres qui indique un degré de différence individuel, dont l'espace des dimensions communes ne rend pas compte, et qui éloigne ce timbre de tous les autres selon un critère perceptif supplémentaire.

Nous avons réussi à établir une relation quantitative entre des paramètres acoustiques et les coordonnées des timbres dans cet espace perceptif, certaines dimensions ayant un caractère spectral, d'autres un caractère temporel, et d'autres encore un caractère spectro-temporel [5].

Plus récemment à l'IRCAM, nous avons obtenu une fois de plus un espace perceptif tridimensionnel qui s'accorde avec les deux précédents, mais en employant un plus grand éventail d'auditeurs [6]. Outre les dimensions perceptives communes à l'ensemble des stimuli et les spécificités sur certains timbres, la technique d'analyse multidimensionnelle, CLASCAL [7], utilisée dans cette dernière étude permet de faire apparaître les différences dans la perception des auditeurs qui s'organisent en classes. A chaque classe correspond une configuration d'importances perceptives accordées aux différentes dimensions et à l'ensemble des spécificités.

Ces recherches mettent donc en évidence :

- 1) une représentation mentale de l'ensemble des timbres selon des dimensions perceptives multiples,
- 2) leurs relations étroites avec des propriétés acoustiques et psychoacoustiques des sons,
- 3) des propriétés spécifiques à certains timbres qui affectent leur perception et,
- 4) des différences entre auditeurs dans l'importance perceptive accordée à ces divers attributs auditifs.

La cohérence des résultats à travers ces diverses recherches indique une grande stabilité des processus perceptifs liés à la perception des sons d'instruments de musique. Cette cohérence conforte notre modèle [8] qui postule que la reconnaissance des sources sonores découle d'un processus d'analyse, de calcul et d'extraction d'un certain nombre d'attributs auditifs liés aux paramètres physiques des signaux et aux paramètres intrinsèques des

sources acoustiques. Elle suggère également que cette approche serait valable pour d'autres types d'événements sonores tels les sons intérieurs d'habitacles de voitures et les sons extérieurs générés par les différentes sources de l'environnement urbain. Ces deux cas seront examinés et comparés ci-dessous et la comparaison révèle des divergences intéressantes d'un point de vue théorique et méthodologique¹.

1. Certains des résultats présentés dans la suite de cet article entrent dans le cadre d'une collaboration contractuelle. Pour des raisons de confidentialité, la description restera donc formelle sans décrire explicitement les résultats obtenus (noms des paramètres et des stimuli, valeurs numériques). Cette contrainte n'affecte cependant en rien l'objectif de l'article : la présentation des méthodes expérimentales et d'analyse de données employées. En outre, les programmes d'analyses utilisés ne peuvent pas être détaillés suffisamment dans l'espace qui nous est accordé pour cet article et des informations complémentaires à leur sujet pourront être obtenues dans les articles cités.

Le premier cas donne un espace perceptif homogène et continu à partir duquel les paramètres acoustiques peuvent être déterminés. Dans le second cas, l'espace est de nature catégorielle et l'extraction de paramètres acoustiques correspondant aux dimensions est exclue. Une autre approche est donc envisagée pour ce cas de figure. En plus de l'approche d'analyse des dissemblances où une distance perceptive sépa-

rant deux sons est exprimée par l'auditeur, nous utilisons l'analyse uni- ou multidimensionnelle des données dites "de dominance" où l'auditeur doit décider si l'un ou l'autre d'une paire de sons est plus "préférable", "gênant", "désagréable", etc. Le lien entre ces analyses de données subjectives et l'analyse objective des sons nous permet d'établir soit un espace psychophysique quantitatif, soit une échelle psychophysique quantitative qui deviennent des outils de réflexion pour les ingénieurs et urbanistes dont la tâche est d'améliorer la qualité sonore.

Caractérisation des son intérieurs de voitures : un espace continu

Comme premier pas vers le but d'améliorer la qualité sonore des habitacles de voitures, une série d'expériences a été réalisée à l'IRCAM [9] en association avec les constructeurs automobiles Renault et PSA Peugeot-Citroën, en partant de l'hypothèse que les structures perceptives qui sous-tendent la comparaison des sons de véhicules peuvent être dégagées par les mêmes techniques que celles appliquées au timbre musical.

L'étude comprend différentes étapes :

- Un test de dissemblance permettant de comparer les sons les uns par rapport aux autres. Aucun critère n'est imposé. Les auditeurs effectuent des jugements directs en fonction d'un certain nombre de critères perceptifs propres à chacun. Les données obtenues sont analysées par la technique CLASCAL décrite précédemment afin de déterminer les attributs auditifs communs à l'ensemble des sujets.
- Une série d'analyses acoustiques permettant de quantifier les indices perceptifs dégagés dans l'analyse multidimensionnelle en trouvant les paramètres acoustiques ou psychoacoustiques qui sont fortement corrélés aux coordonnées des sons dans l'espace perceptif.

- Un test de préférence qui permet de construire une cartographie en fonction des paramètres acoustiques perceptivement significatifs révélés dans l'étape précédente qui indique les combinaisons de ces paramètres qui sont les plus ou les moins préférées.

Les sons sont choisis de manière à représenter une gamme moyenne de 16 voitures. Ils sont enregistrés à l'aide d'une tête artificielle à l'intérieur de l'habitacle. La voiture roule sur piste en régime stable. Deux régimes moteur sont envisagés : R1 (3e vitesse à 4 000 tours/mn) et R2 (5e vitesse, à 3 500 tours/mn). Dans un premier temps, pour chacun des régimes, le paramètre psychoacoustique ϕ_1 (sonie), est variable d'un stimulus à l'autre. Pour une deuxième série d'expériences, ce paramètre est égalisé sur l'ensemble des stimuli. La composition des 4 groupes de stimuli utilisés est résumée dans le tableau 1.

	Régime R1	Régime R2
Paramètre ϕ_1 variable	Groupe 1	Groupe 2
Paramètre ϕ_1 constant	Groupe 3	Groupe 4

Tableau 1 : Composition des 4 groupes de stimuli utilisés.

De plus, afin de restituer le champ binaural, les sons numérisés sont traités par filtrage inverse en temps réel, sur une carte de traitement de signal. Ce filtrage est déterminé par la mesure des réponses impulsionnelles des têtes artificielles à partir desquelles on modélise un filtre inverse pour éliminer les effets dus à la chaîne d'enregistrement. Enfin, la chaîne de diffusion est constituée d'un convertisseur numérique-analogique, d'un amplificateur stéréo et d'un casque.

Les échantillons sonores sont présentés par couple à chaque auditeur, placé dans une cabine audiométrique. On demande à l'auditeur d'effectuer un jugement de dissemblance pour chacune des paires en déplaçant un curseur sur une échelle continue de jugement, allant de "très similaire" à "très dissemblable". A partir de ces données, l'analyse multidimensionnelle permet d'obtenir une représentation géométrique dans laquelle les distances d'un modèle Euclidien d'un nombre de dimensions donné s'accorde le mieux avec les jugements de dissemblance. Le nombre de dimensions k choisi est optimal par rapport à des critères statistiques [voir 7]. Dans notre cas, l'analyse aboutit à un espace perceptif tridimensionnel avec spécificités pour le Groupe 3 et bidimensionnel avec spécificités pour les autres groupes. Une représentation graphique des données correspondant aux sons du Groupe 3 est donnée en exemple sur la figure 1a.

Il nous faut donner ensuite une interprétation psychophysique aux axes, c'est-à-dire une relation systématique entre les caractéristiques acoustiques des stimuli et leurs positions dans l'espace. Pour ce faire, une boucle empirique d'écoute face à une représentation de l'espace perceptif dégagé permet dans un premier temps d'identifier la nature auditive des dimensions. Dans le cas des sons de voiture, leur répartition dans l'espace est homogène et leur évolution sur chaque axe est bien décrite par un attribut auditif. L'analyse des signaux, guidée par cette écoute informelle, a permis de dégager un ensemble de paramètres acoustiques et psychoacoustiques — ϕ_1 , ϕ_2 , ϕ_3 , ϕ_4 et ϕ_5 — bien corrélés avec les coordonnées des stimuli sur une dimension donnée (voir tableau 2).

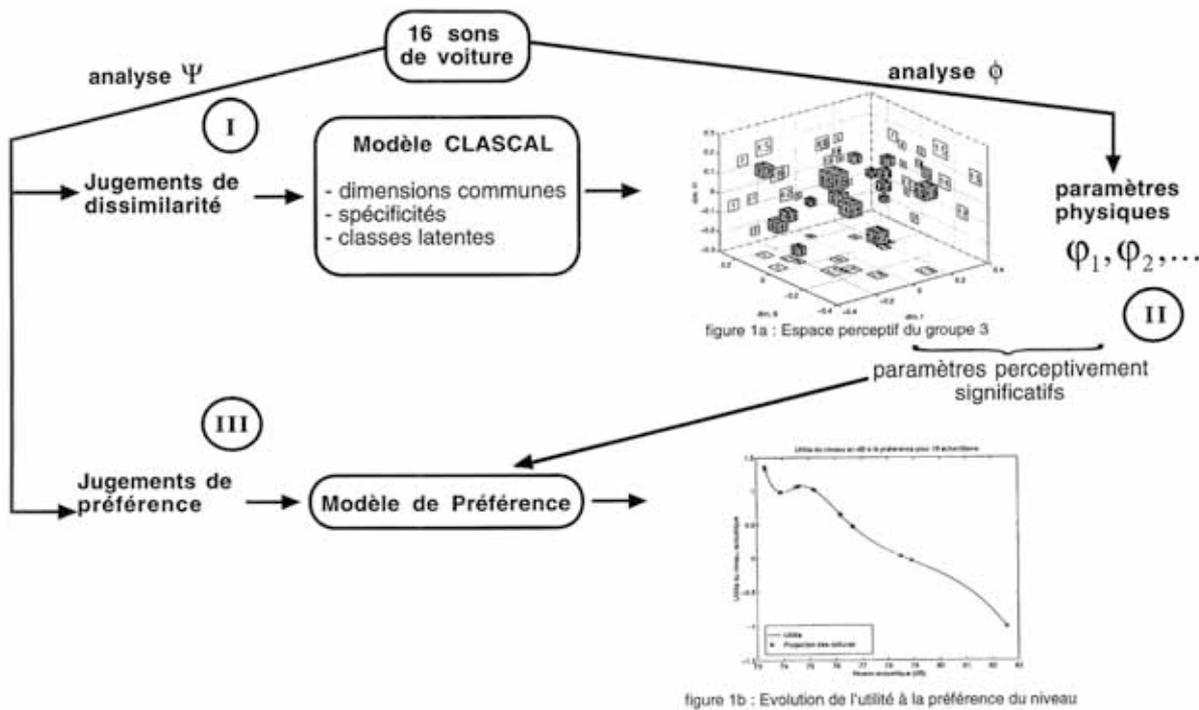


Fig. 1 : Schéma global de l'étude perceptive des véhicules
(a) espace perceptif tridimensionnel obtenu pour les sons du Groupe 3
(b) évolution de l'utilité pour déterminer la préférence en fonction du niveau acoustique des sons

Paramètres	Groupe 1		Groupe 2	
	Dim. 1	Dim.2	Dim.1	Dim.2
φ1	-0,92÷	-0,06	-0,91÷	0,28
φ2	0,09	0,80÷	-	-
φ3	-	-	0,43	-0,88÷

Paramètres	Groupe 3			Groupe 4	
	Dim. 1	Dim. 2	Dim. 3	Dim. 1	Dim. 2
φ2	0,35	-0,7÷	-0,14	-0,93÷	-0,29
φ3	-	-	-	0,51*	0,86÷
φ4	-0,81÷	0,32	-0,33	-	-
φ5	-0,32	0,00	-0,83÷	-	-

Tableau 2 : Coefficients de corrélation entre les données perceptives et les paramètres physiques. La probabilité p telle que les coordonnées perceptives et les valeurs du paramètre acoustique φ soient indépendantes est indiquée : ÷ p < 0,01, *p < 0,05

Il apparaît clairement que la sonie est un facteur prédominant, significativement corrélée avec la première dimension des Groupes 1 et 2. En revanche, lorsque ce paramètre est éliminé — Groupes 3 et 4 —, d'autres paramètres émergent dans le jugement des auditeurs : la structure perceptive écrasée par la sonie se décompose donc en divers attributs lors de la disparition de celle-ci.

Pour la troisième étape de l'étude qui tente de construire la carte de préférence, le protocole expérimental varie légèrement. Comme pour le test de dissemblance, les sons sont diffusés par paires mais l'auditeur doit simplement répondre à la question suivante : "Lequel des deux sons préférez-vous?" Une matrice des proportions de préférence est ensuite calculée sur l'ensemble des sujets, chacune des composantes (i, j) correspondant à la

probabilité que le stimulus i soit préféré au stimulus j. A partir de ces données, le programme d'analyse C5SPLN [10] cherche une fonction qui transforme les valeurs des paramètres physiques en valeurs d'utilité qui dépendent des proportions de préférence empirique. A titre d'exemple, la fonction d'utilité du niveau acoustique est représentée à la figure 1b : plus le niveau augmente plus la valeur d'utilité à la préférence diminue. Ce qui paraît évident pour ce paramètre l'est nettement moins pour les autres paramètres acoustiques déterminés, qui peuvent même avoir des fonctions d'utilité non monotones où une ou deux zones au milieu de la gamme sont préférées aux autres zones.

En conclusion, pour les deux régimes, lorsque la sonie est un facteur de comparaison, elle domine largement dans les jugements de préférence. Par contre, le rôle des paramètres secondaires varie selon le régime et ont globalement moins d'importance dans les jugements de préférence.

Néanmoins, l'égalisation de la sonie a permis de faire émerger d'autres facteurs contribuant à la préférence. Dans ce cas, aucune famille de paramètres ne domine les jugements de préférence pour les deux régimes. De manière générale, la saillance perceptive des dimensions n'est pas la même pour les deux régimes R1 et R2, ce qui indique que l'importance relative des indices perceptifs évolue avec le régime et le contexte (sonie égalisée ou non). Le modèle de représentation perceptive ainsi obtenu procure aux constructeurs un outil à partir duquel l'espace sonore des habitacles peut être amélioré selon des critères associés à la qualité auditive.

Caractérisation des sons de l'environnement urbain : un espace catégoriel

Dans un souci de caractériser les critères perceptifs liés au désagrément occasionné par diverses situations urbaines, l'IRCAM mène des études psychoacoustiques sur la perception des sons de l'environnement urbain. Les hypothèses et la méthodologie retenues a priori pour cette application sont les mêmes que pour l'étude précédente. Après une présentation générale du contexte expérimental, nous nous intéresserons donc plus particulièrement aux divergences possibles des résultats par rapport aux conclusions précédentes. Dans ce cas, nous présenterons des résultats pour illustrer la démarche scientifique et la problématique psychoacoustique en question. Ils représentent une idéalisation ayant des caractéristiques d'études réelles sur des bruits extraits de l'environnement urbain qui nous intéressent pour notre propos, notamment l'organisation catégorielle des bruits.

Dans ce second cas, le corpus d'échantillons sonores comprend 20 séquences représentatives d'un grand nombre de contextes urbains : rue, autoroute, aéroport... L'ensemble peut être segmenté en 7 familles distinctes ($f1$, $f2$, $f3$, $f4$, $f5$, $f6$ et $f7$). Les sons ont été enregistrés en extérieur, et les signaux sont restitués aux auditeurs, sans filtrage, dans les mêmes conditions que pour les sons de voitures. Par ailleurs, il a été choisi de faire égaliser globalement les sons en sonie par un panel d'auditeurs.

Une tâche d'identification des sons est incluse à la fin du test de dissemblance. Cette tâche complémentaire est motivée par la nature des sons proposés et l'intuition d'un rôle influent de la reconnaissance des sources. De fait, l'analyse CLASCAL des données de dissemblance révèle une structure perceptive fortement catégorielle (Fig. 2) liée à un taux élevé de reconnaissance des types de sources sonores.

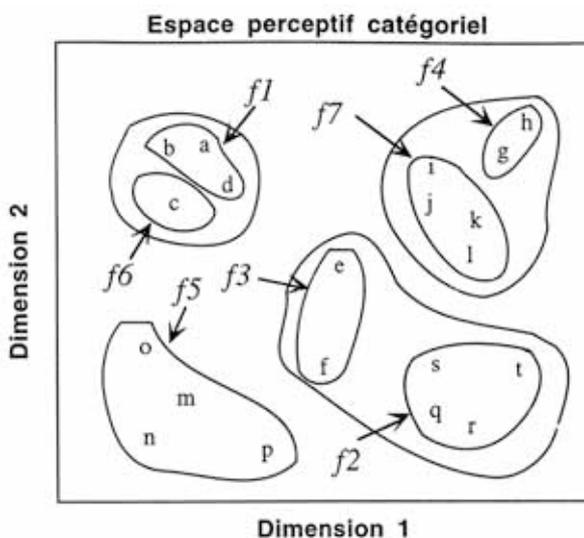


Fig. 2 : Solution illustrative en deux dimensions d'une analyse multidimensionnelle des jugements de dissemblance pour un ensemble hétérogène de sons. Le résultat d'une analyse de cluster est indiquée par les contours qui regroupe de façon hiérarchique les sons appartenant à une même classe

Ainsi, ce facteur cognitif préminent — l'identification et la relation directe du son à l'objet — empêche de concevoir dans ce cas précis les mécanismes de perception selon un ensemble de traits auditifs abstraits liés à des paramètres qui varient de manière continue sur une dimension donnée. Cependant, d'un point de vue théorique, la notion de catégorisation est pertinente dans la mesure où elle fait référence à certains modèles perceptifs. Ils postulent que la représentation mentale du stimulus sonore passe très tôt d'une nature continue à une nature discrète dans le processus de reconnaissance. L'information sensorielle est ainsi décomposée en un certain nombre de traits discrets qui sont ensuite intégrés avant d'être comparés à des catégories mémorisées [8]. Néanmoins, ces modèles sont en contradiction avec les hypothèses issues de la première étude qui ont malgré tout été posées a priori pour la seconde étude qui paraissait en beaucoup de points semblable. Il est vraisemblable que les différences acoustiques entre les différents types de sources (avions, camions, voitures...), sont très grandes par rapport aux différences entre les diverses marques de voitures de l'étude précédente.

Ce résultat préliminaire nous a amené à considérer d'autres types d'analyses qui formalisent la similarité selon un appariement de propriétés discrètes partagées par les stimuli : il s'agit essentiellement des analyses dites "de cluster hiérarchique" [11] dont les résultats confirment la structure catégorielle discontinue relative à la perception des sons proposés (cf. contours sur fig. 2). Par ailleurs, on observe généralement dans ce genre d'étude que les catégories obtenues à l'issue de ces dernières analyses correspondent globalement aux familles de sons définies précédemment (illustrées par $f1$ à $f7$ sur la figure 2).

Dans le cadre du second cas de figure, la dominance s'exprime ici en termes de désagrément, la question étant : "Parmi ces deux sons lequel trouvez-vous le plus désagréable?". Cependant, étant donnée la nature des premiers résultats, les données de dominance sont traitées directement et non plus par le programme d'analyse C5SPLN utilisé dans la première étude ; ce programme nécessite l'existence de "facteurs objectifs" (paramètres physiques et psychoacoustiques) corrélés avec les dimensions perceptives continues par rapport auxquels le désagrément est analysée.

En fait, dans le cas présent, les matrices de proportion de désagrément sont exploitées grâce à l'algorithme de Bradley-Terry-Luce [12] qui permet de construire une hiérarchie unidimensionnelle entre les sons, lisible sur la seule échelle sémantique choisie comme base de l'expérience, en l'occurrence l'échelle de désagrément. En outre, la dimension ainsi construite devient un axe psychologique pertinent, et la même approche de quantification psychophysique est appropriée pour dégager les indices acoustiques et psychoacoustiques susceptibles de l'expliquer. L'approche pour ce faire est de rechercher, par des analyses de régressions multiples, une combinaison linéaire de paramètres acoustiques dont certains contribuent à l'agrément et d'autres au désagrément. Enfin, on trouve en général que l'ordre des stimuli sur cette échelle de désagrément respecte encore une fois les catégories naturelles des sons ($f1$ à $f7$).

Conclusion

Les hypothèses déduites d'un modèle de représentation perceptive du timbre d'instruments de musique développé par l'équipe Psychoacoustique à l'IRCAM ont été appliquées dans le cadre de deux types d'études s'intéressant aux sons de l'environnement. Le premier type présente un corpus de sons homogènes, lié à une classe unique d'objets sonores. Le champ d'investigation du second type est plus hétérogène et contient des sons de l'environnement issus de différents contextes urbains, tous très caractéristiques.

La comparaison des résultats de ces deux types d'études permet de mesurer en partie l'influence de la nature des sons sur les structures perceptives qui sous-tendent la perception auditive. En effet, il semble que si les stimuli sonores sont assez similaires, donc difficilement séparables en classes distinctes, leur représentation mentale est liée à différents attributs auditifs continus que l'on peut corrélérer à des paramètres physiques des signaux. Par contre, si le corpus de stimuli est de nature extrêmement hétérogène, le facteur cognitif lié à l'identification des sources impose une structure perceptive catégorielle et empêche inévitablement le jugement à partir de critères perceptifs plus fins.

Dans ces deux cas de figure, des approches différentes sont nécessaires pour réussir à franchir l'étape de la quantification psychophysique et donc de pouvoir lier clairement le monde perceptif au monde physique -étape absolument essentielle pour l'ingénieur chargé d'améliorer la qualité acoustique des sources sonores en question. Quoiqu'il en soit, pour atteindre ce but, il est de plus en plus clair que l'instrument de mesure pour caractériser la qualité sonore, dont on ne peut pas se passer au stade actuel de nos connaissances, est l'auditeur humain. La psychoacoustique, menée de façon rigoureuse et avec des compétences nécessaires dans les méthodes de la psychologie expérimentale, a donc une place critique et incontournable dans le domaine de la qualité acoustique. Cependant, il est absolument essentiel de reconnaître que cet instrument de mesure est doté d'un cerveau qui est formé par l'expérience au cours d'une vie ; le domaine de la qualité sonore ne peut donc nullement ignorer des facteurs cognitifs liés aux mesures subjectives.

Remerciements

Nous remercions tout particulièrement Anne Bardot (PSA Peugeot Citroën) et Etienne Parizet (Renault), ainsi que d'autres collaborateurs avec qui nous avons eu l'honneur de travailler, pour leur intérêt et leur collaboration à ces travaux de recherche.

Références bibliographiques

- [1] PLOMP R. (1970) Timbre as a multidimensional attribute of complex tones. In R. Plomp & G.F. Smoorenburg (Eds.), *Frequency Analysis and Periodicity Detection in Hearing*, pp.397- 414, Sijthoff, Leiden.
- [2] GREY J.M. (1997) Multidimensional perceptual scaling of musical timbres, *Journal of the Acoustical Society of America*, 61, 1270-1277.
- [3] WINSBERG S. & CARROLL J. D. (1989) A quasi-nonmetric method for multidimensional scaling via an extended Euclidean model, *Psychometrika*, 54, 217-229.
- [4] KRUMHANSL C.L. (1989) Why is musical timbre so hard to understand? In S. Nielzén & O. Olsson (Eds.), *Structure and Perception of Electroacoustic Sound and Music*, pp. 43-53, Elsevier (Excerpta Medica 846), Amsterdam.
- [5] 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.
- [6] McADAMS S., WINSBERG S., DONNADIEU S., DE SOETE G. & KRIMPHOFF J. (1995) Perceptual scaling of synthesized musical timbres : Common dimensions, specificities, and latent subject classes, *Psychological Research*, 58, 177-192.
- [7] WINSBERG S. & DE SOETE G. (1993) A latent class approach to fitting the weighted Euclidean model, *CLASCAL, Psychometrika*, 58, 315-330.
- [8] McADAMS S. (1994) Reconnaissance de sources et d'événements auditifs. In S. McAdams & E. Bigand (Eds.), *Penser les sons : Psychologie cognitive de l'audition*, pp. 157-214, PUF, Paris.
- [9] SUSINI P., McADAMS S. & WINSBERG S. (1997) Caractérisation perceptive des bruits de véhicules, *Actes du 4e Congrès Français d'Acoustique*, vol. I, 543-546.
- [10] WINSBERG S. & DE SOETE G. (1993) A Thurstonian pairwise choice model with univariate and multivariate spline transformation, *Psychometrika*, 58, 233-256.
- [11] BARTHELEMY J.P., GUENOCHÉ A. (1988) *Les arbres et la représentation des proximités*, Masson, Paris.
- [12] DAVID H. A. (1986) *The method of paired comparisons*, 2nd ed., Oxford University Press, Oxford.