HAL
open science

Méthode de test adaptée à l'évaluation perceptive d'un grand nombre de stimuli audio : Application aux enceintes acoustiques

Pierre-Yohan Michaud, Sabine Meunier, Philippe Herzog, Gérard Drouet d'Aubigny, Mathieu Lavandier

To cite this version:

Pierre-Yohan Michaud, Sabine Meunier, Philippe Herzog, Gérard Drouet d'Aubigny, Mathieu Lavandier. Méthode de test adaptée à l'évaluation perceptive d'un grand nombre de stimuli audio : Application aux enceintes acoustiques. 10ème Congrès Français d'Acoustique, Apr 2010, Lyon, France. CD: 390 (6p.). hal-00546825

HAL Id: hal-00546825

https://hal.science/hal-00546825v1

Submitted on 14 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

10ème Congrès Français d'Acoustique

Lyon, 12-16 Avril 2010

Méthode de test adaptée à l'évaluation perceptive d'un grand nombre de stimuli audio : Application aux enceintes acoustiques

Pierre-Yohan Michaud[1], Sabine Meunier[1], Philippe Herzog[1], Gérard Drouet d'Aubigny[2], Mathieu Lavandier[3]

[1] Laboratoire de Mécanique et d'Acoustique, 31 Chemin Joseph Aiguier, 13402 Marseille Cedex 20

[2] Laboratoire Jean Kuntzman, Département de Statistique, 1251 avenue centrale B.P. 47, 38040 Grenoble Cedex 9

[3] École Nationale des Travaux Publics de l'État, Département Génie Civil et Bâtiment, Rue M. Audin, 69518 Vaux-en-Velin Cedex

De nombreux tests perceptifs font intervenir la méthode de comparaison par paires afin d'évaluer des dissemblances perçues entres sons. Cette méthode permet une mesure directe de la dissemblance mais entraîne des tests longs lorsque le nombre de stimuli est grand. Une alternative est d'utiliser la méthode de classification libre qui est souvent mise en oeuvre pour l'étude de panels de sons plus larges. Malgré tout, des travaux ont montré que cette méthode n'est pas toujours adaptée à l'évaluation de la dissemblance. La recherche d'un test permettant l'évaluation d'un grand nombre de stimuli a permis d'adapter une méthode provenant de la perception visuelle. Pour cette méthode, que nous appelons comparaison à une référence, la tâche consiste à évaluer la similarité d'un groupe de stimuli par rapport à une référence. Afin de connaître les limites et biais potentiels de la méthode, nous avons simulé les jugements d'auditeurs sur la base de matrices de dissemblances "artificielles" et ainsi obtenu des indications sur plusieurs paramètres, tels que le nombre de stimuli ou d'auditeurs, nécessaires à une estimation précise des dissemblances. Le principe et les résultats de la simulation seront dans un premier temps décrits. Nous avons ensuite appliqué cette méthode à l'évaluation perceptive des enceintes acoustiques et plus particulièrement à l'évaluation de leur restitution du timbre. La méthode de comparaison à une référence a été utilisée pour analyser les deux corpus de sons testés par Lavandier et al. pour des mesures de dissemblances effectuées sur 12 enceintes en utilisant la comparaison par paires (JASA, 2008) et sur un panel de 37 enceintes avec la classification libre (Forum Acusticum, 2005). Les résultats perceptifs sont comparés à ceux obtenus par Lavandier et al. Cette comparaison des méthodes permet d'évaluer et de valider l'utilisation de la méthode de comparaison à une référence.

1 Introduction

Des tests perceptifs peuvent être mis en oeuvre pour analyser différentes caractéristiques des sons et découvrir les dimensions perceptives utilisées par des auditeurs pour différencier ces stimuli audio. Ces mesures de différences relatives permettent de construire une matrice de dissemblances perçues entre stimuli. A partir de la matrice de dissemblances issue du test, les dimensions sont révélées à l'aide d'une méthode MDS (Multi-Dimensionnal Scaling) [1] qui permet de représenter les dissemblances perçues par les auditeurs entre les stimuli dans un espace multidimensionnel.

La méthode de comparaison par paires est principalement recommandée [2, 3] comme étant la méthode standard pour obtenir directement des jugements de dissemblance entre des stimuli audio. Les auditeurs doivent évaluer la dissemblance globale des paires en reportant par exemple leur estimation sur une échelle linéaire variant de "très semblable" à "très dissemblable". Les matrices individuelles de dissemblances sont remplies avec les estimations des auditeurs et conduisent à une matrice moyenne adaptée à la MDS.

Par contre, la méthode de comparaison par paires est seulement adaptée pour l'estimation de dissemblance sur un faible nombre de sons. Le nombre de présentations (paires) du test croît rapidement pour un large panel de stimuli. Pour un test faisant intervenir 30 stimuli, l'évaluation de 435 paires est nécessaire pour remplir la matrice de dissemblances. Dans ce cas, l'influence de la fatigue auditive devient significative pour un test d'une durée importante.

La classification libre est présentée par Bonebright [4] comme étant une méthode alternative permettant l'évaluation de large corpus de sons. Cette méthode indirecte consiste à regrouper des sons, placés aléatoirement dans un espace, en fonction de leur ressemblance. Des sons jugés similaires seront regroupés dans un même groupe, et des sons jugés différents se trouveront dans des groupes différents. Pour cette méthode la matrice de dissemblances est obtenue en effectuant la moyenne de chaque matrice individuelle remplie avec des valeurs de dissemblances de 0 pour deux stimuli apparaissant dans un même groupe, et 1 pour deux stimuli placés dans des groupes différents.

Malgré tout, Alrdrich et al. [5] ont comparé la méthode de comparaison par paires et celle de classification libre sur des corpus de sons identiques. Ils ont démontré que les participants à la classification libre avaient tendance à grouper par catégories plutôt que par similarité.

Aldrich et al. ont conclu que pour un corpus supérieur à 20 stimuli, la classification libre est valable mais tend à fournir des informations catégorielles.

Cette étude vise à présenter une méthode permettant l'évaluation perceptive d'un large panel de stimuli sonores. Adaptée d'une expérience de perception visuelle menée par Rogowitz et al. [6], la méthode proposée, que nous appelons méthode de comparaison à une référence, fait intervenir l'évaluation de la similarité entre un stimulus de référence et d'autres sons de comparaison. La méthode demande autant de temps que la comparaison par paires mais présente l'avantage de pouvoir être divisée en plusieurs parties grâce à la présentation d'une référence à chaque essai. Afin d'explorer cette méthode pour en connaître les limites et obtenir des indications sur ses principaux paramètres, nous avons simulé des tests.

La méthode a ensuite été appliquée à des tests d'écoute pour l'évaluation des enceintes acoustiques. Ces travaux reprennent ceux initiés par Lavandier et al. dont le but étaient de découvrir les dimensions de la restitution du timbre d'un extrait musical à travers les enceintes. Deux panels différents comprenant respectivement 12 et 37 enceintes ont été évalués par Lavandier et al. en utilisant la comparaison par paires pour l'évaluation du premier panel [7] et la classification libre pour le panel de 37 enceintes [8].

Deux tests d'écoute ont été effectués sur les deux panels avec la méthode de comparaison à une référence. Les résultats perceptifs ont été comparés avec ceux obtenus par Lavandier et al. dans le but de valider l'utilisation de cette méthode pour les tests perceptifs.

2 Présentation de la méthode

La méthode développée dans cette étude provient d'une expérience menée par Rogowitz et al. [6] dans le but d'évaluer la similarité sur un large panel d'images. Dans ce test, chaque sujet comparait 9 images choisies aléatoirement parmi un total de 97. Une d'entre elle était considérée comme la référence et le sujet devait évaluer laquelle des 8 images restantes paraissait la plus similaire à la référence. Toutes les images apparaissaient une fois en tant que référence, et était comparée aux 96 autres distribués aléatoirement en groupes composés de 8 images.

A partir d'une matrice de dissemblances initialisée à 1, lorsque un sujet choisissait le stimulus j comme étant le plus similaire à la référence i, alors l'élément d_{ij} de la matrice de dissemblances était décrémentée de 1. Pour deux stimuli, i et j, il existe deux essais dans lesquels l'un apparaît en tant que référence et l'autre parmi les stimuli de comparaison. Ainsi, pour chaque sujet, les éléments d_{ij} ou d_{ji} peuvent avoir une valeur de dissemblance différente. Pour obtenir une valeur moyenne, Rogowitz et al. proposent de faire la moyenne de chaque paire d_{ij} et d_{ji}. La matrice totale moyenne issue du test correspond à la moyenne arithmétique de chaque matrice individuelle.

Afin d'adapter cette méthode à l'évaluation de la dissemblance entre stimuli sonores, certaines précautions ont été prises. Dans l'expérience menée par Rogowitz, les sujets avaient une vision globale de la référence ainsi que des 8 images à comparer. Pour des tests d'écoutes, les stimuli sont nécessairement présentés successivement, et les auditeurs doivent mémoriser les sons afin d'établir des comparaisons. La méthode a donc été adaptée pour que les auditeurs entendent dans un ordre chronologique la référence puis les différents autres sons à comparer les uns après les autres.

Il semblait aussi évident de réduire le nombre de stimuli présenté à chaque essai afin de prendre en compte les effets de mémoires auditives. Pour chaque essai, la tâche a donc été adaptée pour que chaque référence soit comparée à 3 sons de comparaisons. Comme chaque stimulus apparaît comme référence, un test entier sera composé de $n.(n-1)/3$ essais, n étant le nombre de stimuli.

Pour évaluer la stabilité de la méthode, ainsi que sa capacité à fournir des valeurs de dissemblance précises et fiables, des simulations ont été menées et sont décrites dans la partie suivante.

3 Simulation

Pour étudier l'applicabilité de la méthode aux tests d'écoute, la simulation des jugements d'un échantillon d'auditeurs "parfaits" permet d'évaluer l'influence de plusieurs paramètres tels que le nombre de stimuli ou le nombre d'auditeurs simulés.

3.1 Principe

Les jugements de dissemblances ont été simulés sur la base de valeurs de dissemblances "artificielles". Pour obtenir ces valeurs, un certain nombre de points ont été aléatoirement disposés dans un espace, et la distance entre chaque paire de points a été calculée pour remplir une matrice de dissemblances. Le nombre de dimensions de l'espace dans lequel les points sont placés, a été fixé arbitrairement à 3 dimensions.

L'auditeur simulé est considéré comme "parfait" car chaque élément de la matrice de dissemblances est interprété comme la distance perceptive simulée entre deux stimuli. Au cours d'un essai simulé, la plus petite valeur de dissemblance entre les 3 stimuli de comparaison et la référence indique quel stimulus parmi les 3 est choisi comme le plus semblable à la référence. En simulant les jugements d'auditeurs "parfaits", la matrice de dissemblances issue de la simulation du test devrait correspondre à la matrice initiale. Ainsi, en comparant ces deux matrices, l'efficacité de la méthode de comparaison à une référence peut être évaluée.

Les tests ont été simulés en faisant varier le nombre de stimuli puis le nombre d'auditeurs simulés. Comme décrit ci-dessus, un stimulus parmi n servant de référence est comparé aux $(n-1)$ autres stimuli aléatoirement distribués en $(n-1)/3$ trios. Pour chaque simulation, le coefficient de corrélation linéaire de Pearson r a été calculé entre la matrice de dissemblances de départ et la matrice de dissemblances issue de la simulation du test, afin d'évaluer si les informations contenues dans la matrice initiale ont été retrouvées.

3.2 Résultats

Deux paramètres ont été évalués grâce à la simulation : le nombre de stimuli impliqués dans le test puis le nombre d'auditeurs.

3.2.1 Nombre de stimuli

Pour s'assurer que la méthode est stable pour différents nombres de stimuli impliqués dans le test, le coefficient de corrélation a tout d'abord été calculé pour un nombre de stimuli allant de 10 à 50 pour un nombre d'auditeurs constant fixé à 100. Il en résulte que la corrélation reste constante avec une valeur moyenne de $r = 0.92$ et un écart-type de 0.01. Ces valeurs montrent donc que la méthode semble applicable pour des tests impliquant n'importe quel nombre de stimuli.

Ayant des informations sur le nombre de stimuli, il s'avérait intéressant de connaître, pour un nombre de stimuli fixé, combien d'auditeurs simulés sont nécessaires à l'obtention d'une valeur de corrélation stable.

3.2.2 Nombre d'auditeurs simulés

Sachant que le nombre de stimuli n'influence pas la méthode, les simulations suivantes ont été effectuées avec un nombre de stimuli fixé à 37 afin de simuler un test impliquant un panel de la taille de celui évalué par Lavandier et al. [8]. Ainsi, 36 stimuli de comparaison, distribués en 12 trios, ont été comparés à chaque référence. Le test entier était finalement composé de 444 essais.

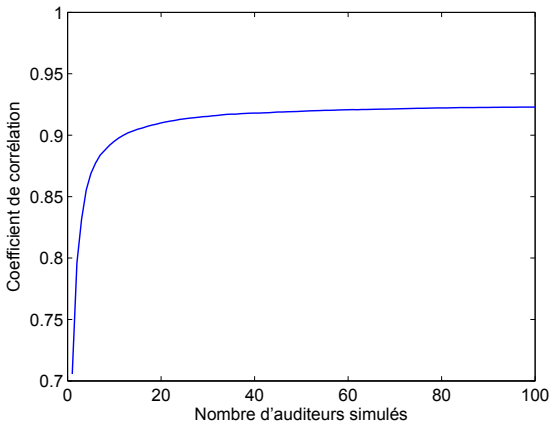

FIGURE 1 – Corrélation entres les matrices de dissemblances initiale et celle issue de la simulation pour différents nombres d'auditeurs simulés

La corrélation entre les matrices de dissemblances croît rapidement entre 1 et 10 auditeurs puis se stabilise. Au delà de 20 auditeurs simulés, la majorité des informations de dissemblances est obtenue. La méthode atteint rapidement une valeur stable, par contre elle ne dépasse pas $r = 0.92$ même pour 100 auditeurs. Ce seuil montre que la méthode n'est pas capable de retrouver exactement la matrice de dissemblances initiale. Ce biais reste faible et plus de 90% des informations de dissemblance sont retrouvées avec la méthode de comparaison à une référence. Pour information, la simulation de la méthode de comparaison par paires aurait donné une corrélation égale à 1.

Le biais a été identifié comme étant intrinsèque à la méthode et provient du fait de présenter $(n-1)/3$ trios pour chaque référence. Pour chaque référence, les auditeurs ne peuvent seulement apporter que $(n-1)/3$ réponses sur $(n-1)$ sons de comparaison, ce qui entraîne que les matrices individuelles ne sont remplies qu'avec un tiers des valeurs de dissemblance.

3.3 Discussion

L'adaptation de la méthode proposée par Rogowitz et al. aux tests d'écoute semble apporter certains avantages. Premièrement, la tâche proposée aux auditeurs est simple. En écoutant quatre extraits sonores, ils doivent seulement choisir un son parmi trois comme étant le plus semblable à une référence (choix forcé), sans faire intervenir l'estimation du niveau de dissemblance sur une échelle. Le test entier comprend $n.(n-1)/3$ essais par auditeurs, contre les $n.(n-1)/2$ essais nécessaires pour la comparaison par paires. Cependant, pour l'évaluation de large corpus de sons, la durée des tests avec les deux méthodes reste conséquente. Néanmoins, pour la méthode de comparaison à une référence, du fait de la présence d'un stimulus de référence, il devient possible de diviser les tests en plusieurs sessions afin d'éviter la fatigue auditive (sans craindre que l'auditeur change d'échelle d'évaluation entre deux sessions).

Finalement, la simulation a permis l'exploration de la méthode afin d'en découvrir les limites et biais et de caractériser les paramètres essentiels à l'obtention de valeurs de dissemblances fiables. Même si la méthode ne permet pas de retrouver 100% des informations de dissemblance, la majorité de ces informations est obtenue avec peu d'auditeurs simulés. Les résultats de la simulation montrent que la méthode peut être appliquée pour des tests d'écoute pour l'évaluation de larges panels de sons.

4 Application aux enceintes acoustiques

Pour étudier l'applicabilité de la méthode de comparaison à une référence, cette étude s'inscrit dans le prolongement des travaux de Lavandier et al. [7, 8, 9] concernant l'évaluation perceptive des enceintes acoustiques. A travers plusieurs séries de tests, ils se sont intéressés à déterminer les dimensions perceptives sur lesquelles les auditeurs s'appuyaient pour différencier des extraits sonores correspondant à l'enregistrement de différentes enceintes acoustiques.

D'une part, lors d'une première expérience, la méthode de comparaison par paires a été mise en oeuvre avec 27 auditeurs sur un premier panel de 12 enceintes [7]. Ensuite, un test perceptif a été mené sur un deuxième panel plus important comprenant 37 enceintes avec la méthode de classification libre, avec un total de 56 auditeurs [8].

A partir des matrices de dissemblances, Lavandier et al. ont dévoilés, grâce à la MDS, des dimensions percep-

tives représentant les critères utilisés par les auditeurs pour effectuer leurs jugements. En reprenant cette procédure, deux séries de tests ont été effectuées sur les deux panels de 12 et 37 enceintes afin de comparer les deux méthodes utilisées par Lavandier et al. et la méthode de comparaison à une référence.

4.1 Stimuli

Pour chacun des deux panels de 12 et 37 enceintes, les stimuli correspondaient à l'enregistrement de la restitution d'un extrait musical à travers différentes enceintes dans une même pièce. Les enregistrements effectués par Lavandier et al. ont été réalisés avec une prise de son stéréophonique pour une restitution par une seule enceinte d'un extrait de 3 secondes (McCoy Tyner). Les chaînes de restitution et d'enregistrement restaient identiques tout au long des sessions d'enregistrement et la seule variable restante était, par conséquent, l'enceinte acoustique.

4.2 Auditeurs

Pour le test d'écoute sur les 12 enceintes du panel 1, 27 auditeurs ont participé et 18 auditeurs ont évalué le panel 2. Tous les auditeurs qui ont participé aux deux tests ne présentaient aucun trouble auditif.

4.3 Procédure

Pour le panel 1, chaque référence était comparée aux autres 11 sons distribués aléatoirement en 4 trios pour chaque auditeurs. Les 11 enregistrements ne permettaient pas de former exactement 4 trios et seuls les trois premiers trios étaient complets avec 9 sons. Le dernier trio était composé des deux sons restants et compléter avec un son choisi aléatoirement parmi les 9 sons déjà tirés. Le test entier comptait 48 essais. Pour le deuxième panel de 37 enceintes, chaque référence était comparée aux 36 sons restants distribués aléatoirement en 12 trios. Le test entier était donc composé de 12x37 essais, ce qui fait un total de 444 essais. Les tests ont été divisés en plusieurs sessions grâces à la présence d'une référence à chaque essai.

Les tests perceptifs ont été menés avec une interface MATLAB permettant automatiquement d'entendre successivement les 4 sons (référence + 3 sons de comparaison) représentés graphiquement par des carrés. Les auditeurs avaient la possibilité de réécouter les sons dans l'ordre qu'ils souhaitaient avant de répondre, en cliquant sur le carré correspondant. Tout comme les tests effectués par Lavandier et al., les enregistrements étaient reproduits sous casque.

4.4 Résultats

4.4.1 Comparaison des matrices de dissemblances

Dans un premier temps les matrices de dissemblances résultant des tests d'écoute sur chaque panel ont été comparées à celles obtenues par Lavandier et al.

Pour chaque panel, la comparaison entre les matrices de dissemblances consiste à calculer la corrélation entre la matrice moyenne de la méthode utilisée par Lavandier et al. et les matrices individuelles des auditeur issues de la méthode de comparaison à une référence, ajoutées une à une, aléatoirement. Les résultats des corrélations correspondent à la moyenne de 50 tirages aléatoires de l'ordre d'addition des matrices individuelles.

La figure 2 présente les résultats des corrélation entre la méthode de compraison à une référence et la comparaison par paires pour le panel 1, puis la classification libre pour le panel 2, en fonction du nombre d'auditeurs.

Pour le panel 2, Les tests avec la méthode de comparaison à une référence étant en cours de réalisation, la figure 2 ne montre que la corrélation pour un total de 18 auditeurs. Les résultats définitifs avec un nombre d'auditeurs plus important seront présentés ultérieurement.

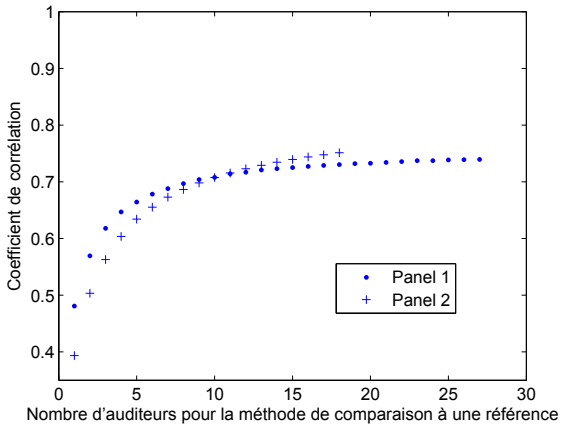

FIGURE 2 – Corrélation entre les matrices de dissemblances pour le panel 1 (.) et le panel 2 (+) en fonction du nombre d'auditeurs impliqués dans la méthode de comparaison à une référence.

Panel 1 :

La corrélation croît rapidement jusqu'à 10 auditeurs puis se stabilise pour un nombre d'auditeurs supérieur. Cette tendance est semblable à celle obtenue grâce à la simulation mais le biais mis en évidence sur la figure 1 est plus important pour les données de dissemblances issues du vrai test. Lors de la simulation, le biais était uniquement du au fait de remplir un tiers de la matrice pour chaque auditeur dont les jugements étaient considérés comme "parfaits". Dans cette comparaisons de données expérimentales les différences observées pourraient provenir d'un bruit de mesure et d'un biais systématique présent dans la méthode.

Les deux méthodes fournissent quand même des données de dissemblance proches alors que pour chaque auditeur les manières de remplir les matrices de dissemblances sont différentes.

Panel 2 :

Les données actuelles semblent converger vers une tendance similaire à celle obtenue pour le panel 1 mais avec des valeurs de corrélation légèrement supérieures. La corrélation entre les valeurs de dissemblance obtenues par les deux méthodes croît jusqu'à 0.70 pour 18

auditeurs. Cette valeur indique que la matrice de dissemblances actuelle issue de la méthode proposée ressemble à celle obtenue avec la méthode de classification libre.

4.4.2 Comparaison des espaces perceptifs

Les dimensions perceptives obtenues par les différents tests d'écoute ont été comparées. Sans chercher à expliquer les dimensions révélées par les travaux de Lavandier et al., cette comparaison permet se focaliser sur la capacité de la méthode proposée à fournir des résultats perceptifs cohérents.

Panel 1 :

L'analyse MDS des données de dissemblance issues de comparaison par paires fournit un espace à deux dimensions [7]. Pour comparer les deux espaces perceptifs, la corrélation entre les coordonnées des points représentant les enregistrements a été calculée suivant chaque dimension. La corrélation entre les deux espaces perceptifs s'élève à 0.96 pour la dimension 1, et 0.88 suivant la dimension 2.

Ces valeurs de corrélation montrent que les dimensions résultant de l'analyse MDS des valeurs de dissemblances pour le test avec la méthode de comparaison à une référence sont similaires aux dimensions trouvées avec la comparaison par paires.

Afin de comparer visuellement les dimensions obtenues grâce aux deux méthodes, la figure 3 présente les deux espaces perceptifs.

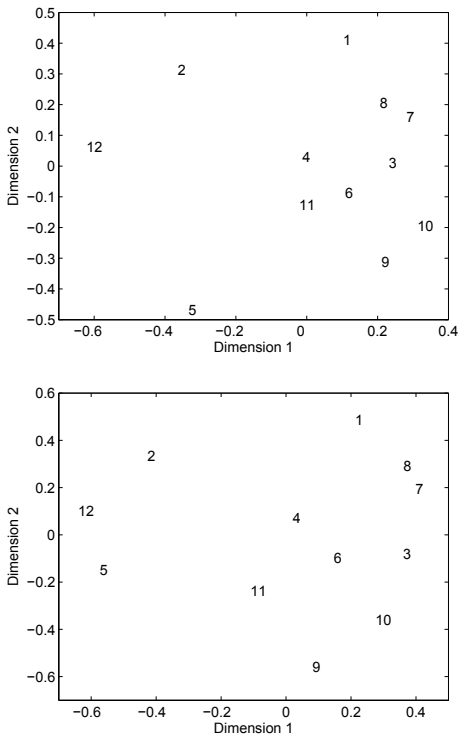

FIGURE 3 – Espaces perceptifs obtenus avec la méthode de comparaison par paires (en haut) et la méthode de comparaison à une référence (en bas). Chaque numéro correspond à l'enregistrement d'une enceinte.

A titre d'information, Lavandier et al. [9] ont trouvé que la première dimension était liée à l'équilibre grave/aigu, alors que la deuxième correspondait à une notion de clarté liée à l'émergence des mediums.

Panel 2 :

Pour la classification libre, 4 dimensions ont été dévoilées après analyse MDS [8]. La méthode de comparaison à une référence a permis de mettre en évidence 3 dimensions perceptives. Afin d'obtenir une comparaison avec les dimensions issues de la classification libre, les coordonnées des enregistrements de chaque espace perceptif ont été analysées. Les résultats des corrélations sont de 0.80 pour la dimension 1, 0.66 pour la dimension 2 et 0.85 suivant la dimension 3. La méthode de classification libre fait ressortir une quatrième dimension qui n'apparaît pas dans les résultats pour la méthode de comparaison à une référence.

Ces valeurs montrent que les espaces perceptifs obtenus avec les deux méthodes sont ressemblants. Pour 18 auditeurs impliqués pour la méthode présentée, les corrélations pour les dimensions 1 et 3 indiquent que ces dimensions sont proches pour les deux tests. Par contre la dimension 2 est elle aussi commune aux deux méthodes mais présente une ressemblance plus faible.

D'après Lavandier et al. [8], la première dimension correspond à la balance spectrale. Cette dimension coïncide avec la dimension 1 issu des tests sur le panel 1. La dimension 3 est liée l'émergence des mediums, dimension qui correspond à la dimension 2 obtenu pour le premier panel. La dimension 2 est une nouvelle dimension, définie par Lavandier et al. [8] comme étant liée à une sensation d'espace ou autrement dit de réverbération qui traduirait l'interaction de l'enceinte enregistrée avec la pièce.

4.5 Discussion

La comparaison des matrices de dissemblances permet d'analyser les données issues directement des tests perceptifs. La méthode de comparaison à une référence offre la possibilité de construire une matrice de dissemblances, adaptée à la MDS, similaire de celles obtenues avec les autres méthodes. Pour les deux panels, une différence apparaît et pourrait provenir des différentes manières de remplir les matrices de dissemblances pour chaque méthode. Ce biais révélé grâce à la simulation doit être mélé à un bruit de mesures sachant que, pour chaque panel, sont confrontées deux séries de données issues de deux méthodes expérimentales différentes.

L'analyse MDS de chaque matrice de dissemblances a permis de comparer les espaces perceptifs obtenus avec les différentes méthodes. La méthode de comparaison à une référence permet de retrouver des espaces perceptifs proches de ceux obtenus par Lavandier et al. Il est à noter que les valeurs de corrélation entre les dimensions des différents espaces perceptifs issus de la MDS sont, dans l'ensemble, supérieures aux corrélations entre matrices de dissemblances. Cette différence pourrait provenir du fait que l'analyse MDS représente au mieux les dissemblances dans un espace euclidien à plusieurs dimensions et donc ajuste les valeurs des coordonnées suivant les différentes dimensions.

Les espaces perceptifs issus des deux tests effectués sur le panel 1 sont très ressemblants. En revanche, pour

le panel 2, d'après les résultats intermédiaires, la méthode testée permet de mettre en évidence 3 dimensions communes parmi les 4 obtenues avec la classification libre. Parmi ces trois dimensions, les dimensions 1 et 3 sont très semblables et la deuxième diffère légèrement, d'une méthode à l'autre. Il faudrait attendre d'avoir effectuer le test avec la méthode de comparaison à une référence sur plus d'auditeurs pour pouvoir conclure sur la comparaison des espaces perceptifs.

Il est intéressant de voir que les deux dimensions obtenues pour 12 enceintes se retrouvent dans l'espace perceptif des 37 enceintes. L'apparition des nouvelles dimensions provient du fait d'évaluer un plus grand nombre de sons permettant aux auditeurs d'effectuer plus de comparaisons et ainsi élaborer plus de critères de différenciation entre les sons.

Le point essentiel est que la méthode de comparaison à une référence permet de retrouver les dimensions découvertes grâce à deux autres méthodes. D'une part, la méthode étudiée semble être équivalente à la méthode de comparaison par paires pour l'évaluation de 12 enceintes. Ces résultats valident l'utilisation de la méthode pour l'évaluation de la dissemblance entre des stimuli audio. D'autre part, en appliquant cette méthode à un panel de sons plus important comprenant 37 enceintes, les résultats perceptifs obtenus à ce jour sont semblables à ceux obtenus par Lavandier et al. Même si les tests d'écoute ne sont pas achevés, ces informations intermédiaires montrent que la méthode semble pouvoir être utilisée pour l'évaluation de la dissemblance pour de larges panels de sons.

5 Conclusion

La méthode de comparaison à une référence, adaptée d'un test de perception visuelle, fournit des données de dissemblances précises et adaptées à une analyse MDS. Tout d'abord la simulation a permis de trouver les paramètres et les limites qui ont montré la possibilité d'appliquer cette méthode à de vrais tests d'écoutes. Pour le déroulement des tests, la méthode offre l'avantage de pouvoir sectionner le test en plusieurs sessions afin d'éviter les tests trop longs et les problèmes de fatigue auditive. De plus, la tâche proposée est simplifiée. L'auditeur doit seulement juger parmi 3 sons celui qui ressemble le plus à un son de référence.

Afin de valider expérimentalement l'utilisation de cette méthode, deux tests perceptifs sur l'évaluation des enceintes acoustiques ont été menés sur deux panels différents déjà évalués par Lavandier et al. avec deux autres méthodes. La méthode de comparaison à une référence fournit des espaces perceptifs similaires à ceux obtenus avec la méthode de comparaison par paires et proches de ceux obtenus avec la classification libre.

Les dimensions retrouvées avec la méthode de comparaison à une référence valident l'utilisation de cette méthode pour l'évaluation de la dissemblance pour des panels de sons plus ou moins importants.

La prochaine étape consistera à comparer les espaces perceptifs avec des espaces issus de mesures physiques afin d'attribuer des critères objectifs aux dimensions perceptives obtenues grâce aux tests d'écoute. Cette démarche permettra d'analyser en détails les dimensions, d'expliquer les différences entre les méthodes de tests et de conclure sur les dimensions de la restitution du timbre par les enceintes acoustiques.

Remerciements

Nous tenons à remercier tous les auditeurs qui ont accepté de participer aux deux tests d'écoute.

Références

[1] Borg I., Groenen, P. "Modern Multidimensional Scaling, Theory and Applications" *Springer* (1997)

[2] International Electrotechnical Commission, "Rapport technique IEC 60268 - Part 13 : Listening tests on loudspeakers" (1998)

[3] Lipshitz S.P., Vanderkooy J., "The great debate : Subjective evaluation" *J. Audio Eng. Soc.* 29(7/8) 482-491 (1981)

[4] Bonebright, T.L. "An investigation of data collection methods for auditory stimuli : Paired comparison versus a computer sorting task" *Behavior research Methods, Instruments, and Computers* 28(2), (1996)

[5] Aldrich, K.M., Hellier E.J., Edworthy J., "What determines auditory similarity ? The effect of stimulus group and methodology" *The quartely journal of experimental psychology* 62(1), 63-83 (2008)

[6] Rogowitz B.E., Frese T., Smith J.R., Bouman C.A., Kalin, E., "Perceptual image similarity experiment" *Human Vision and Electronics Imaging III* (1998)

[7] Lavandier M., Herzog P., Meunier S., "Comparative measurements of loudspeakers in a listenning situation", *J. Acoust. Soc. Am.* 123(1), 77-87 (2008)

[8] Lavandier M., Meunier S., Herzog P., "Perceptual and physical evaluation of differences among a large panel of loudspeakers", *Forum Acusticum*, Budapest, Hungary (2005)

[9] Lavandier M., Meunier S., Herzog P., "Identification of some perceptual dimensions underlying loudspeaker dissimilarities", *J. Acoust. Soc. Am.* 123(6), 4186-4198 (2008)