



Equipe Lutheries - Acoustique - Musique, Institut Jean Le Rond d'Alembert
Université Pierre et Marie Curie

Mémoire de stage de Master 2 Sciences et technologie

Mention SDI, spécialité Acoustique, parcours ATIAM

Caractérisation acoustique de salles simulées au format Ambisonics

Juan Pablo Espitia Hurtado

Encadré par

Jean Dominique Polack



A mis padres

Remerciements

Tout d'abord, je tiens à remercier mon encadrant Jean-Dominique Polack de m'avoir accueilli dans son équipe de recherche, pour la confiance qu'il m'a accordée dès le départ, pour ses enseignements et ses précieux conseils.

Mes remerciements s'adressent aussi à toute l'équipe du LAM pour son accueil chaleureux et amical.

Merci à mes collègues de stage : Eric, Augustin, Delphine, Rémi, Hélène, Leonard, Quentin, Matilde, François et Hugo pour avoir créé une si agréable ambiance de travail, pour leur amitié et aussi pour les cours de français.

Mes remerciements sincères aux professeurs du Master ATIAM pour leurs enseignements et aussi à mes camarades pour les bons moments partagés et leur amitié.

Je remercie Carlos Agon pour sa disponibilité et ses encouragements tout au long de l'année.

Table de Matières

1. Présentation du Laboratoire	6
2. Introduction	7
3. Notions Théoriques	8
3.1 Le Système Ambisonics	8
3.2 Ambisonics d'ordre 1 (B-Format)	10
3.3 Qualité Acoustique d'une salle	12
3.4 Réponse Impulsionnelle d'une salle.....	13
3.5 Paramètres Acoustiques	15
4. Protocole Expérimental	19
4.1 Egalisation du système source en salles réelles	19
4.2 Système de simulation des salles	21
4.3 Mesure de la réponse impulsionnelle des salles synthétisées	24
4.4 Calcul des indices acoustiques	27
5. Analyses et Résultats	30
5.1 Analyses Objectives	30
5.1.1 Etude des valeurs centrales et dispersions.....	30
5.1.2 Analyse en composantes principales.....	34
5.2 Analyse du Seuil Différentiel des paramètres acoustiques	37
5.3 Test Subjectif.....	41
5.4 Réponses impulsionnelles des salles réelles et synthétisées	45
6. Conclusion	46
Bibliographie	47
Annexes	49

Table des illustrations

Figure 1. Harmoniques Sphériques pour $m=0, 1, 2$ et 3	9
Figure 2. Schéma du microphone SoundField ST250	11
Figure 3. Spectrogramme d'un sinus glissant logarithmique et son filtre inverse.	14
Figure 4. Etapes du protocole expérimental.	19
Figure 5. Réponse en fréquence de la source.	20
Figure 6. Salle sèche du LAM. À droite, schéma du placement des haut-parleurs de la salle.....	21
Figure 7. Plugin VST Voxengo Pristine Space et décodeur Ambisoncis DecoPro.	23
Figure 8. Partie linéaire de la RI et ses produits de distorsion.....	24
Figure 9. Spectrogramme 3D de la RI d'une salle avec des erreurs.	25
Figure 10. Clics sur la RI pendant l'enregistrement.....	26
Figure 11. Spectrogramme 3D de une RI nette correspondant à une salle simulée.....	26
Figure 12. Fenêtre Options de l'application OpenMIDAS.....	28
Figure 13. Réponse en fréquences des enceintes Studer A1 numéro 2 et 5.....	29
Figure 14. Boîtes à moustaches de l'indice T30.....	31
Figure 15. Boîtes à moustaches de l'indice EDT.	31
Figure 16. Boîtes à moustaches de l'indice C80.	32
Figure 17. Boîtes à moustaches de l'indice Ts.	32
Figure 18. Boîtes à moustaches de l'indice G.	33
Figure 19. Boîtes à moustaches de l'indice LF.	33
Figure 20. Analyse en composantes principales 1 et 2 pour 6 indices acoustiques.	35
Figure 21. Analyse en composantes principales 1 et 3 pour 6 indices acoustiques.	36
Figure 22. Analyse en composantes principales 2 et 3 pour 6 indices acoustiques.	36
Figure 23. Différence en JND entre les salles réelles et synthétisées pour les indices T30 et EDT.	38
Figure 24. Différence en JND entre les salles réelles et synthétisées pour les indices C80 et Ts.....	38
Figure 25. Différence en JND entre les salles réelles et synthétisées pour l'indice LF	39
Figure 26. Différence en JND entre les salles réelles et synthétisées pour l'indice G.....	40
Figure 27. Différence en JND entre les salles réelles et synthétisées pour l'indice G par fréquence. ...	40
Figure 28. Différence en JND entre les salles réelles et synthétisées pour l'indice LF par fréquence. .	41
Figure 29. Arbre des sujets initial	42
Figure 30. Arbre des objets obtenu après avoir fait la suppression du sujet 2.....	43
Figure 31. Réflectogrammes correspondants à la Cité de la Musique.....	44

1. Présentation du Laboratoire

L'équipe Lutheries - Acoustique – Musique (LAM) est une équipe de l'Institut Jean le Rond d'Alembert qui est une Unité Mixte de Recherche de l'Université Pierre et Marie Curie et du CNRS, également soutenue par le Ministère de la Culture.

Le LAM rassemble des chercheurs animés par l'étude de l'objet musical avec une approche pluridisciplinaire, à travers les sciences pour l'ingénieur (physique, acoustique, traitement du signal) et des sciences humaines (psychologie cognitive, linguistique). Cette approche, pluridisciplinaire et originale, permet au LAM d'obtenir des résultats scientifiquement fondés et des développements applicatifs trouvant des débouchés industriels et commerciaux.

Créé en 1963 par Emile Leipp dans le Département de Mécanique de la Faculté des Sciences, le LAM reçoit en 1982 le soutien du département SPI du CNRS et du Ministère de la Culture, sous la direction de Michèle Castellengo. Ces liens institutionnels se formalisent par la création d'une Unité Mixte de Recherche avec l'UPMC en 1993. Depuis cette année le LAM a commencé à développer un axe de recherche relatif à l'environnement sonore (perception des bruits, acoustique architecturale et urbaine), notamment grâce à l'arrivée de Jean Dominique Polack et à la collaboration de Danièle Dubois (LCPE, département SHS du CNRS). En 1994, le LAM joue un rôle important dans la création du DEA ATIAM qui réunit autour de l'IRCAM plusieurs laboratoires opérant en Acoustique, Traitement du signal et Informatique Appliqués à la Musique. En 1997, la Bibliothèque Nationale de France installe au LAM ses études sur la conservation et la restauration des enregistrements sonores, et audio-visuels. En 2003 D. Dubois et son équipe sont accueillis au laboratoire. À présent, le LAM est dirigé par Hugues Genevois et les activités de recherche sont regroupées en trois grands thèmes : les Instruments de Musique et la Voix, la Perception et la Cognition, et les Techniques Audio.

Actuellement le laboratoire est installé au 11 Rue de Lourmel où il dispose de 800 m² où sont installés bureaux, bibliothèque, locaux d'écoute, salles d'expérimentations physiques, ainsi que plusieurs équipements spécifiques : une salle d'écoute sourde (simulations d'ambiances sonores écologiquement valides), une salle d'écoute claire (évaluation de la qualité des instruments de musique), une cabine psycho-acoustique, trois studios analogiques et numériques de traitement des sons, un atelier de lutherie traditionnelle et de matériaux composites, un laboratoire de conservation des documents sonores, un laboratoire photographique et un atelier d'électronique.

2. Introduction

L'équipe Lutheries – Acoustique – Musique de l'institut Jean Le Rond d'Alembert a réalisé une campagne de mesures acoustiques dans une quinzaine de salles de concert et de théâtres de la région parisienne. Les salles ont été sélectionnées pour leur intérêt historique, architectural, ou acoustique. Cette campagne a fait l'objet de la thèse de Fábio Leao Figureiredo, soutenue en juillet 2011 [Fig11].

L'une des particularités de ce travail a été l'utilisation systématique d'un microphone Ambisonics pour mesurer les réponses impulsionnelles. Cela a permis à Fabio de simuler toutes ces salles dans la chambre d'écoute sèche du LAM, et d'effectuer des tests d'écoute pour approfondir l'étude des rapports objectifs et subjectifs des indices acoustiques des salles. Lors de la mise en relation des résultats des tests d'écoute avec les mesures acoustiques in situ, il est apparu des divergences dont l'origine n'a pas pu être déterminée faute de mesures systématiques des caractéristiques acoustiques des champs sonores synthétisés.

Par conséquent, le but de notre travail a été de remettre en place la simulation de toutes les salles mesurées (salles réelles) et de mener à bien la caractérisation systématique des champs sonores (salles synthétisées). Pour cela, dans une première phase, une partie représentative des réponses impulsionnelles (RIs) obtenues lors de la campagne de mesures a été utilisée pour faire la simulation dans la salle d'écoute. Les champs sonores ont été mesurés à l'aide du même protocole que celui utilisé par Fabio dans les salles. À partir des RIs des salles réelles et synthétisées, la deuxième phase a consisté à calculer un ensemble d'indices acoustiques. Dans une troisième phase, des analyses statistiques ont été faites afin de comparer tous les indices acoustiques obtenus dans les salles réelles avec ceux des salles synthétisées. De plus, le seuil différentiel a été utilisé pour analyser leurs différences sur le plan perceptif, conformément à ce qui est décrit dans la littérature [Iso09] [Mar10]. Enfin, un test d'écoute a été effectué pour évaluer la similitude entre les salles réelles et synthétisées.

Ce document se divise en quatre parties. La première présente les notions théoriques concernant l'auralisation Ambisonics, la réponse impulsionnelle d'une salle et les paramètres objectifs et subjectifs utilisés pour l'évaluation de la qualité acoustique d'une salle. La seconde partie détaille le protocole expérimental utilisé pour faire la simulation, ainsi que les mesures des réponses impulsionnelles et les calculs d'indices acoustiques. Les analyses effectuées sur ces indices telles que boîtes à moustache, analyse en composantes principales (ACP) et le critère de seuil différentiel, sont exposés dans la troisième partie. Enfin, les résultats obtenus à partir du test d'écoute sont décrits en utilisant une représentation arborée.

3. Notions Théoriques

L'objectif de ce chapitre est de présenter le système d'auralisation Ambisonics, ainsi que les concepts liés à l'analyse acoustique de salles de concert.

3.1 Le Système Ambisonics

Les techniques de restitution sonore spatialisée ont pour objet la reproduction d'un espace sonore. On peut regrouper ces procédés en trois classes:

1. La stéréophonie traditionnelle sur deux haut-parleurs.
2. Les techniques binaural et transaural qui ont pour objectif la reconstruction du champ sonore au niveau des oreilles.
3. Et, des procédés tels qu'Ambisonics, caractérisés par la restitution de l'image sonore au centre du dispositif des haut-parleurs.

La technologie Ambisonics est basée sur la solution de l'équation d'onde en coordonnées sphériques. En tout point de l'espace, la pression acoustique peut être décrite par une décomposition de Fourier-Bessel (1) où apparaissent les fonctions harmoniques sphériques Y_{mn}^σ .

$$p(kr, \theta, \delta) = \sum_{m=0}^{\infty} i^m j_m(kr) \sum_{n=0}^m \sum_{\sigma=\pm 1} B_{mn}^\sigma Y_{mn}^\sigma(\theta, \delta) \quad (1)$$

où kr est le nombre d'onde, θ est l'angle d'azimut, δ est l'angle de site et B_{mn}^σ correspond aux coefficients de pondération.

Les fonctions harmoniques sphériques Y_{mn}^σ représentent une décomposition spatiale de l'espace acoustique en fonction de l'angle d'azimut et de l'angle de site. Elles sont définies par :

$$Y_{mn}^\sigma(\theta, \delta) = \tilde{P}_{mn}(\sin \delta) \begin{cases} \cos n\theta & \text{si } \sigma = +1 \\ \sin n\theta & \text{si } \sigma = -1 \end{cases} \quad (2)$$

où \tilde{P}_{mn} représente les versions semi-normalisées des fonctions de Legendre associées d'ordre m et degré $n = 0, 1, \dots, m$. Les fonctions nulles Y_{m0}^{-1} sont ignorées. Pour chaque ordre m il y a donc $2m + 1$ composantes. La figure 1 montre les fonctions harmoniques sphériques pour $m = 0, 1, 2$ et 3 .

La décomposition de Fourier-Bessel (1) doit être tronquée à un ordre fini M par suite de limitations pratiques. La précision de la reproduction et les dimensions du champ sonore reconstruit (zone d'écoute) dépendent de l'ordre des fonctions harmoniques sphériques. Donc, le champ sonore est décrit à partir d'un nombre limité de coefficients B_{mn}^σ ($m=0, 1, \dots, M$) nommés aussi composantes Ambisonics. Dans le cas d'une onde plane ceux-ci sont déterminés par [Dan00]:

$$B_{mn}^{\sigma} = Y_{mn}^{\sigma}(\theta_S, \delta_S)S \quad (3)$$

L'équation précédente définit le processus de codage d'une source sonore S moyennant une représentation Ambisonics.

L'information codifiée est le signal S multiplié par la valeur de la fonction harmonique sphérique Y_{mn}^{σ} associée évaluée dans l'orientation (θ_S, δ_S) . Le nombre de composantes K d'une représentation Ambisonics 3D est calculé à partir de l'ordre M :

$$K = (M + 1)^2 \quad (4)$$

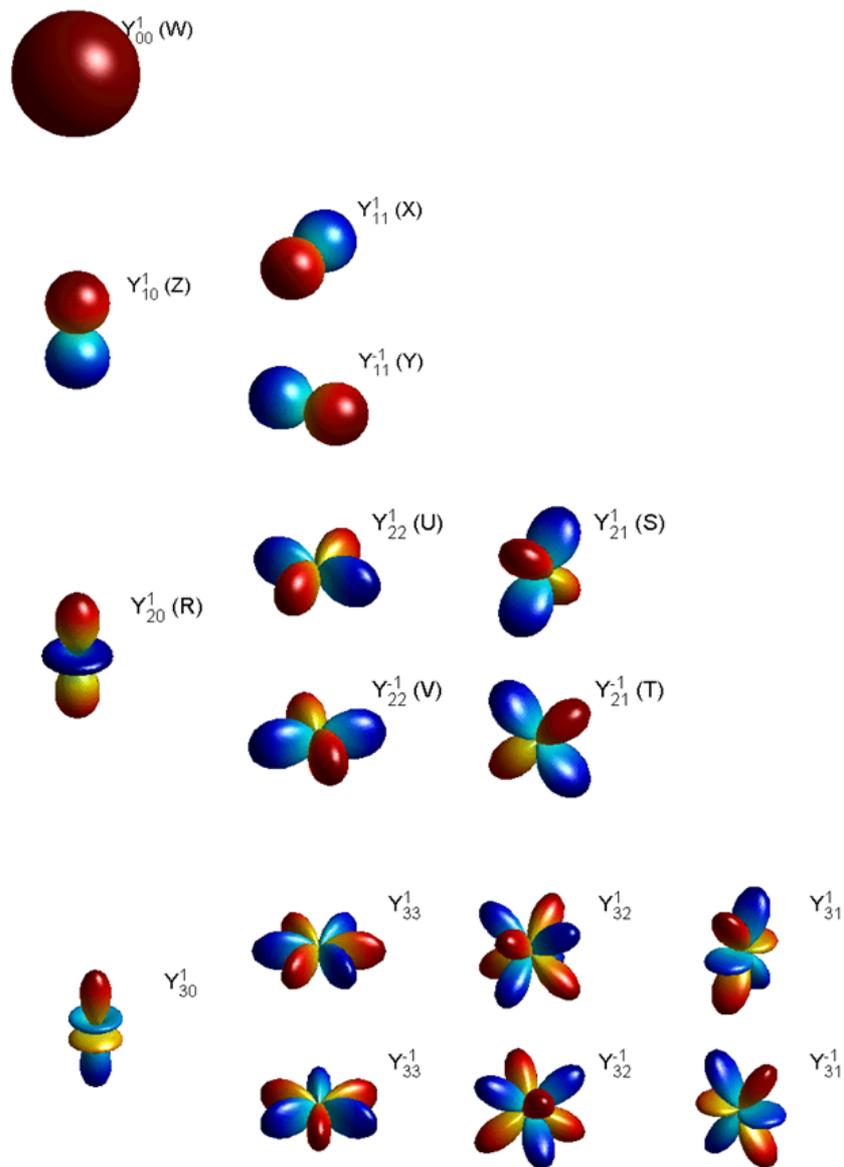


Figure 1. Illustration tirée de la thèse de J. Daniel [Dan00]. Harmoniques Sphériques pour $m=0, 1, 2$ et 3

3.2 Ambisonics d'ordre 1 (B-Format)

Avec $M=1$, le système a quatre canaux, ou composantes, qui contiennent l'information directionnelle de l'image sonore. Dans les années soixante-dix, Michael Gerzon a développé un système de codage et de décodage pour le premier ordre d'Ambisonics nommé B-Format. À partir de l'équation (3), les canaux sont exprimés par :

$$\begin{aligned} W &= B_{00}^1 = Y_{00}^1(\theta_S, \delta_S)S \\ X &= B_{11}^1 = Y_{11}^1(\theta_S, \delta_S)S \\ Y &= B_{11}^{-1} = Y_{11}^{-1}(\theta_S, \delta_S)S \\ Z &= B_{10}^1 = Y_{10}^1(\theta_S, \delta_S)S \end{aligned} \quad (4)$$

où les fonctions harmoniques sphériques sont définies par [Dan00]:

$$\begin{aligned} Y_{00}^1 &= \alpha_{00} \\ Y_{11}^1 &= \alpha_{11} \cos \theta \cos \delta \\ Y_{11}^{-1} &= \alpha_{11} \sin \theta \cos \delta \\ Y_{10}^1 &= \alpha_{10} \sin \delta \end{aligned} \quad (5)$$

Les quatre canaux W , X , Y et Z composent le B-Format. La composante W est omnidirectionnelle (avec $m=0$) et les composantes X , Y et Z sont bidirectionnelles (avec $m=1$). Les facteurs α_{mn} sont les coefficients de normalisation. Différentes conventions de normalisation ont été définies, chacune d'elles avec des propriétés spécifiques [Dan00, Sal09]. Par exemple, dans le cas d'un champ 3D parfaitement diffus, les valeurs $\alpha_{00}=1$, $\alpha_{10}=\sqrt{3}$ et $\alpha_{11}=\sqrt{3}$ assurent une puissance égale aux composantes encodées.

Une méthode a été développée pour permettre la prise de son en B-Format en utilisant quatre capsules cardioïdes placées de façon quasi-coïncidente au centre d'un tétraèdre régulier. Les signaux LF (left-front), LB (left-back), RF (right-front) et RB (right-back), correspondants aux sorties de chaque capsule sont nommés A-Format. Chaque canal Ambisonics (B-Format) est construit à partir de la combinaison linéaire de ces signaux par :

$$\begin{aligned} W &= LF + LB + RF + RB \\ X &= LF - LB + RF - RB \\ Y &= LF + LB - RF - RB \\ Z &= LF - LB - RF + RB \end{aligned} \quad (6)$$

Une compensation d'amplitude et de phase est appliquée aux canaux Ambisonics afin de corriger les effets introduits par la non-coïncidence entre capsules. La figure 2 illustre un schéma du microphone SoundField qui implémente ce principe pour faire le codage Ambisonics.

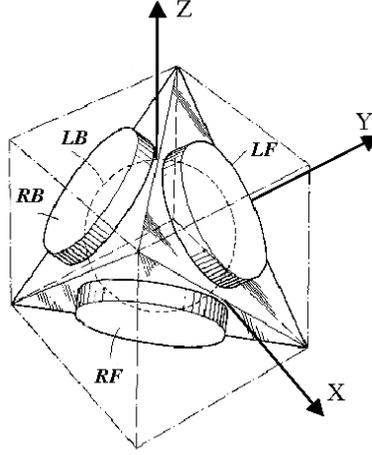


Figure 2. Illustration tirée de la thèse de J. Daniel [Dan00]. Schéma du microphone SoundField ST250

Dans le processus de décodage basique d'Ambisonics, les haut-parleurs ont été considérés comme uniformément distribués autour de la zone d'écoute. Le champ sonore restitué \tilde{B}_{mn}^σ peut être exprimé par [Dan00, Sal09]:

$$\tilde{B}_{mn}^\sigma = CS_N \quad (7)$$

où $S_N = [S_1 \dots S_i \dots S_N]^t$ est la contribution du champ de pression de chacun des N haut-parleurs et C est la matrice des harmoniques sphériques Y_{mn}^σ . Celles-ci sont fonction des angles (θ_i, δ_i) qui indiquent la position de chaque haut-parleur. En égalisant le champ sonore encodé et le champ sonore restitué ($B_{mn}^\sigma = \tilde{B}_{mn}^\sigma$), S_N peut être exprimé par:

$$S_N = DB_{mn}^\sigma \quad (8)$$

où D est la matrice de décodage. Dans le cas d'une distribution régulière des haut-parleurs, D est défini par [Dan00]:

$$D = \frac{1}{N} C^t \quad (9)$$

En combinant les équations (8) et (9), on obtient le signal d'alimentation de chaque haut-parleur en fonction des composantes encodées W, X, Y et Z et de leurs coordonnées (θ_i, δ_i) :

$$\begin{bmatrix} S_1 \\ \vdots \\ S_i \\ \vdots \\ S_N \end{bmatrix} = \frac{1}{N} \begin{bmatrix} Y_{00_1}^1 & Y_{11_1}^1 & Y_{11_1}^{-1} & Y_{10_1}^1 \\ \vdots & \vdots & \vdots & \vdots \\ Y_{00_i}^1 & Y_{11_i}^1 & Y_{11_i}^{-1} & Y_{10_i}^1 \\ \vdots & \vdots & \vdots & \vdots \\ Y_{00_N}^1 & Y_{11_N}^1 & Y_{11_N}^{-1} & Y_{10_N}^1 \end{bmatrix} \begin{bmatrix} W \\ X \\ Y \\ Z \end{bmatrix} \quad (10)$$

où les harmoniques sphériques $Y_{mn_i}^\sigma$ sont définis par :

$$\begin{aligned}
 Y_{00_i}^1 &= \alpha_{00_i} \\
 Y_{11_i}^1 &= \alpha_{11_i} \cos \theta_i \cos \delta_i \\
 Y_{11_i}^{-1} &= \alpha_{11_i} \sin \theta_i \cos \delta_i \\
 Y_{10_i}^1 &= \alpha_{10_i} \sin \delta_i
 \end{aligned}
 \tag{11}$$

Ainsi, il s'agissait dans cette partie d'introduire la méthode d'auralisation qui a été utilisée dans le présent travail. Pour un développement détaillé du système Ambisonics d'ordre 1 ou Ambisonics d'ordres supérieurs (HOA), on peut consulter [Dan00, Ger85, Sal09].

3.3 Qualité Acoustique d'une salle

Pour décrire la qualité acoustique d'une salle de concert ou d'un théâtre, notamment à travers leur faculté à restituer le son émis par un instrument de musique ou la voix, plusieurs dimensions subjectives peuvent être envisagées [Bar93, Ber96]. Les paramètres subjectifs les plus importants sont : la réverbération (*Liveness*), la chaleur (*warmth*), la brillance, l'intimité ou la présence, la précision (*clarity*), le niveau sonore (*loudness*) et l'impression spatiale (*spaciousness*).

La *réverbération* fait référence à la persistance du son dans une salle, suite à son arrêt soudain. Si la durée de la réverbération est perçue comme longue, la salle est dite *vivante*. À l'inverse, si la durée est perçue comme courte la salle est dite *sèche*.

Les facteurs de *chaleur* et de *brillance* sont aussi liés au concept de réverbération. La chaleur en musique est définie comme la richesse de basses fréquences (entre 75Hz et 350Hz), par rapport aux fréquences moyennes (350Hz à 1400Hz). En revanche, la brillance dénote une forte présence de hautes fréquences.

Une salle est « *intime* » si la musique restituée donne l'impression d'être jouée dans un petit espace. Une salle intime est une salle qui a de la *présence*. La *précision* fait référence au degré de distinction entre les différents sons d'une pièce musicale. Deux types de *précision* sont définis : l'horizontal qui mesure le degré de distinction entre sons consécutifs et le vertical qui mesure le degré de distinction entre sons simultanés. La valeur de précision désirée dépend du type de musique [Ber96, Tis09].

Le *niveau sonore* fait référence à la sensation de force sonore. En comparant deux salles avec la même réverbération, celle qui a moins d'auditeurs (sièges) aura un niveau sonore plus fort. De la même manière, un son est perçu plus fort dans une salle très réverbérante en comparaison d'une salle moins réverbérante, même si elles ont une taille identique. De plus, il est possible de différencier le niveau sonore précoce du niveau sonore réverbérant. Le premier est déterminé par l'énergie du son direct et des premières réflexions

tandis que le deuxième dépend de l'énergie des réflexions tardives (généralement après une période de 80ms).

L'impression spatiale est expliquée à partir de deux concepts. La sensation d'enveloppement (LEV – *Listener envelopment*) et la largeur apparente de la source (ASW – *Apparent source width*) [Ber96, Tis09]. Le LEV décrit la sensation d'être entouré par le son. Une place déterminée où le son réverbéré est perçu comme arrivant de toutes les directions avec le même niveau est une place avec un indice LEV optimal. Le LEV dépend de la direction d'incidence des réflexions sur l'auditeur, de la séparation entre les premières réflexions et les réflexions tardives et du niveau d'énergie perçu des réflexions. L'ASW décrit la sensation que la musique provient d'une source plus grande que la largeur visuelle de la source réelle. Plus grande est la largeur apparente, plus solide sera l'impression spatiale.

3.4 Réponse Impulsionnelle d'une salle.

Dans le domaine de l'acoustique des salles, la mesure de la réponse impulsionnelle (RI) est primordiale : c'est à partir de la RI que les paramètres qui décrivent le comportement acoustique d'une salle peuvent être calculés. Plusieurs méthodes ont été définies afin de mesurer la RI d'un système acoustique. Pour toutes les méthodes, un signal d'excitation qui couvre toute la plage de fréquences audibles avec un niveau d'énergie suffisamment fort est nécessaire afin de garantir un rapport signal sur bruit minimal pour calculer des différents indices acoustiques. Si on considère la salle comme un système linéaire et invariable dans le temps, la réponse de la salle $y(t)$ produit par le signal d'excitation $x(t)$, peut être exprimée par [Far00]:

$$y(t) = x(t) * h(t) + n(t) \quad (12)$$

où $h(t)$ est la RI de la salle et $n(t)$ est le bruit du système assumé comme bruit blanc gaussien. Une impulsion acoustique utilisée en tant que signal d'excitation constitue la forme la plus directe pour mesurer la RI. Des dispositifs comme les pistolets sont utilisés pour générer ce type d'onde. Néanmoins, la directivité, la coloration du spectre fréquentielle et la non répétabilité de la mesure font de ces techniques des mécanismes peu fiables pour l'obtention de résultats précis. C'est pourquoi, les techniques les plus utilisées sont celles qui se servent de signaux d'excitation déterministes tels que les séquences MLS et les sinus glissants.

Les signaux MLS sont des séquences binaires pseudo-aléatoires de longueur $N = 2^k - 1$ qui ont des propriétés favorables permettant de mesurer la fonction de transfert d'un système. La fonction d'autocorrélation des signaux MLS est très proche de la fonction Dirac. Les méthodes basées sur le sinus glissant utilisent un signal $x(t) = \sin 2\pi f(t)$, où $f(t)$ est la variation fréquentielle, linéaire ou logarithmique, et dont le spectre est celui d'un bruit blanc et rose respectivement. L'équation (13) montre l'expression de $f(t)$ pour un balayage linéaire :

$$f(t) = f_1 t + (f_2 - f_1)t^2/2T \quad (13)$$

où f_1 est la fréquence initiale, f_2 est la fréquence finale et T est la durée du balayage sinusoïdal. Dans ces types de méthodes, un filtre inverse $x^{-1}(t)$ doit être utilisé afin d'obtenir la RI de la salle. Ce filtre est calculé pour que la convolution avec le signal d'excitation $x(t)$ soit égale à la fonction Dirac.

$$x(t) * x^{-1}(t) = \delta(t) \quad (14)$$

Si on considère les signaux d'excitation comme des séquences MLS, un filtre inverse peut être réalisé très efficacement en utilisant la transformée Hadamard [Car98, Sch79]. Par ailleurs, en utilisant un balayage linéaire ou logarithmique, le filtre inverse est simplement le signal inversé temporellement. Dans le cas d'un balayage logarithmique, une modulation d'amplitude est ajoutée pour compenser l'énergie générée en basses et hautes fréquences. La figure 3 illustre la densité spectrale de puissance du signal d'excitation par la méthode d'un sinus glissant logarithmique et son filtre inverse.

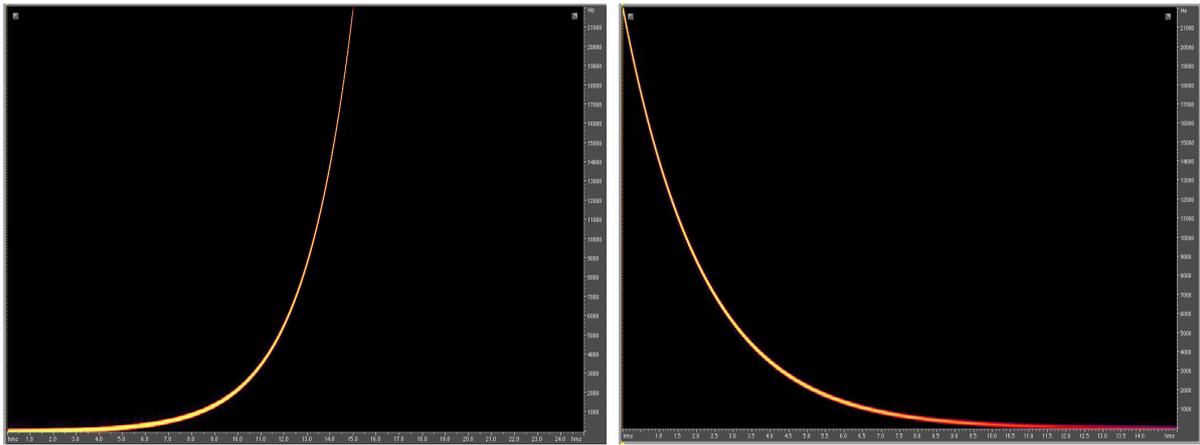


Figure 3. À gauche le spectrogramme d'un sinus glissant logarithmique entre 22Hz à 22kHz. À droite le spectrogramme du filtre inverse associé.

En combinant (12) et (14), la RI de la salle peut être exprimée à partir de la déconvolution par :

$$h(t) = y(t) * x^{-1}(t) + n(t) \quad (15)$$

Pourtant, la réponse de la salle n'est pas complètement linéaire. La plupart des non-linéarités sont générées par le système d'émission du signal d'excitation (les haut-parleurs). La méthode MLS donne une mesure de la RI d'une salle de façon précise et avec une haute tolérance au bruit. Toutefois, un désavantage de cette technique est sa sensibilité au comportement non linéaire de la source, ce qui peut donner des signaux faux dans la RI [Gri96]. Dans le cas du balayage linéaire, les produits de distorsion générés par la source génèrent un type de bruit qui est corrélé avec le signal d'excitation.

Par contre, Angelo Farina [Far00] [Far07] a démontré qu'en utilisant un balayage logarithmique, les produits de distorsion apparaissent clairement séparés de la partie linéaire de la RI et peuvent être facilement supprimés. L'équation (16) montre le signal d'excitation utilisant un sinus à fréquence glissante logarithmique. L'équation (17) exprime le délai Δt entre la réponse linéaire de la RI et ses harmoniques avec f_1 la fréquence initiale, f_2 la fréquence finale, T la durée du balayage sinusoïdale et N l'ordre de la composante de distorsion.

$$x(t) = \sin 2\pi \left[\frac{f_1 T}{\ln(f_2 / f_1)} (e^{\frac{t}{T} \ln(f_2 / f_1)} - 1) \right] \quad (16)$$

$$\Delta t = T \cdot \frac{\ln(N)}{\ln\left(\frac{f_2}{f_1}\right)} \quad (17)$$

3.5 Paramètres Acoustiques

Dans une salle, la mesure de la réponse impulsionnelle, dans une certaine position, permet la description des phénomènes acoustiques correspondants à cet endroit particulier. La norme ISO 3382 définit les différents paramètres acoustiques objectifs associés à ces phénomènes, qui sont utilisés pour estimer les facteurs subjectifs. Ceux-ci sont généralement classés comme des paramètres temporels ou spatio-temporels [Tis09]. Dans le premier groupe se trouvent le Temps de Réverbération (RT60), le Temps de Décroissance Précoce (EDT), le Temps de Délai Critique (ITDG), la Définition (D50), la Clarté (C80), le Temps Central (Ts), le Facteur de Force (G) et d'autres comme le *Bass ratio* (BR) et le *Treble ratio* (TR). Dans le deuxième groupe, se trouvent le Coefficient d'intercorrélacion binaurale (IACC), la Fraction Latérale (LF), le Facteur de Force d'ondes Latérales Tardives (LG) et d'autres indices comme le rapport gauche/droite (LRR), le rapport avant/arrière (FBR).

Le RT60 est à la fois inversement proportionnel à l'absorption totale de la salle et directement proportionnel à son volume. Le calcul du RT60 se fait à partir de la décroissance de l'énergie exprimée par l'intégrale de Schroeder (18) où $h(\tau)$ est la RI de la salle. Dans un champ diffus parfait, la décroissance est une pente en ligne droite. Le temps pris par la pente pour diminuer de 60dB est défini comme le RT60. Pourtant et à cause des restrictions de la dynamique, la mesure se fait dans l'intervalle de temps correspondant à -5dB et à -35dB. Cet indice est nommé T30 (19). L'EDT (20) dépend des réflexions précoces générées par la salle et sa valeur peut différer du T30 selon les caractéristiques du point de mesure. Si le champ sonore est hautement diffus en un point de la salle, leurs valeurs vont être presque de la même magnitude. L'EDT se calcule dans l'intervalle de temps correspondant à 0dB et à -10dB.

$$E(t) = \int_t^{\infty} h^2(\tau) d\tau \quad (18)$$

$$T30 = 2[t_{-5} - t_{-35}] \text{ en secondes} \quad (19)$$

$$EDT = 6[t_0 - t_{-10}] \text{ en secondes} \quad (20)$$

L'EDT et le T30 sont utilisés pour estimer la *réverbération* d'une salle. Du rapport entre eux résulte l'indication de la visibilité de la source. Il a été noté que l'impression de la réverbération est principalement reliée au paramètre EDT [Bar93]. Les indices BR et TR sont utilisés pour évaluer la sensation de *chaleur* et la *brillance* d'une salle respectivement. Ils sont définis par :

$$BR = \frac{T30_{125Hz} + T30_{250Hz}}{T30_{500Hz} + T30_{1000Hz}} \quad (21)$$

$$TR = \frac{T30_{2000Hz} + T30_{4000Hz}}{T30_{500Hz} + T30_{1000Hz}} \quad (22)$$

où $T30_f$ est le temps de réverbération de la salle mesuré à la fréquence f .

Le C80 est le rapport de l'énergie précoce à l'énergie tardive des réflexions. Cet indice est utilisé pour l'étude de la précision, c'est-à-dire, le degré de distinction entre deux notes successives. Le T_s est le barycentre énergétique de la réponse impulsionnelle de la salle. Le temps central est l'indicatif de la *précision* du son en un point spécifique de la salle et il est fortement lié à l'EDT. Le C80 et le T_s sont définis par :

$$C_{80} = 10 \log \left(\frac{\int_0^{0.08} h^2(t) dt}{\int_{0.08}^{\infty} h^2(t) dt} \right) \text{ en dB} \quad (23)$$

$$t_s = \left(\frac{\int_0^{\infty} t h^2(t) dt}{\int_0^{\infty} h^2(t) dt} \right) \text{ en ms} \quad (24)$$

où $h(t)$ est la RI de la salle. Une valeur faible de T_s indique qu'une grande quantité d'énergie arrive dans les premières réflexions, ce qui donne une sensation de précision. Plus haute est la valeur de T_s , plus faible est la précision obtenue. À l'inverse, plus haute est C80, plus ample sera la sensation de précision.

Le facteur de force G (25) est le rapport de l'énergie sonore produite en un siège de la salle par une source non directionnelle à l'énergie sonore de la même source mesurée à 10m en champ libre (en chambre anéchoïque). L'indice G est appelé aussi le gain de la salle.

$$G = \left(\frac{\int_0^{\infty} h^2(t) dt}{\int_0^{\infty} h_{anech,10m}^2(t) dt} \right) \text{ en dB} \quad (25)$$

Avec un sonomètre, G peut être calculé directement grâce à l'équation suivante :

$$G = L_p - L_{p_{anech,10m}} \text{ en dB} \quad (26)$$

où L_p est la pression du son en un point de la salle et $L_{p_{anech,10m}}$ est la pression du son mesuré à une distance de 10m en chambre anéchoïque. Sont également utilisés les indices G_{mid} et G_{low} qui ne sont que la valeur moyenne de G calculée dans les moyennes fréquences (500Hz et 1000Hz) et dans les basses fréquences (125Hz et 250Hz) respectivement. Pour estimer le *niveau sonore*, le paramètre G est utilisé, ce niveau est déterminé principalement par G_{mid} . L'indice G est dépendante de l'inverse de l'absorption totale et du volume de la salle [Ber96].

Le LF (27) est le rapport entre l'énergie précoce des premières réflexions reçues avec un microphone figure-8 (sans la composante du son direct), et l'énergie des premières réflexions reçues par un microphone omnidirectionnel. Barron et Marshall ont démontré que les mesures subjectives de la largeur apparente de la source (ASW - « Apparent Source Width ») sont hautement corrélées avec l'indice LF. Une autre alternative pour estimer l'ASW est l'indice LFC (28). La différence entre les deux indices est que le LFC varie seulement avec le cosinus de l'angle tandis que le LF varie avec le carré du cosinus de l'angle. On considère que le LFC a une meilleure corrélation avec l'impression subjective.

$$LF = \left(\frac{\int_0^{80ms} h_{micro\infty}^2(t) dt}{\int_0^{80ms} h_{omni}^2(t) dt} \right) \text{ en \%} \quad (27)$$

$$LFC = \left(\frac{\int_0^{80ms} |h_{micro\infty}(t)h_{omni}(t)| dt}{\int_0^{80ms} h_{omni}^2(t) dt} \right) \text{ en \%} \quad (28)$$

Pour évaluer le LEV (enveloppement), l'indice GLL a été proposé, celui-ci prend en compte le rapport entre l'énergie sonore des réflexions latérales en un siège de la salle, et l'énergie sonore de la même source mesurée à 10m en champ libre.

Le tableau suivant résume les facteurs subjectifs représentant la qualité acoustique d'une salle avec leurs paramètres acoustiques objectifs associés.

Facteur Subjectif	Paramètre Objectif
Niveau Sonore	G, G _{mid}
Réverbération	EDT, T30
Précision	C80, Ts, D50
Largeur Apparente de Source (AWS)	LF, LFC
Enveloppement (LEV)	GLL

Table 1. Paramètres acoustiques subjectifs et objectifs

Les paramètres acoustiques décrits dans cette partie correspondent aux paramètres d'analyse du présent travail. Pour plus d'information sur d'autres paramètres, on peut se référer, entre autres à [Bar93, Ber96, Car98, Tis09].

4. Protocole Expérimental

Pour étudier la qualité acoustique des différentes salles et afin de réaliser des comparaisons et des groupements, par rapport aux facteurs subjectifs, il est nécessaire de créer l'impression d'être immergé dans l'espace sonore des salles et de le faire dans un endroit contrôlé.

Pour ce faire, il est nécessaire de synthétiser l'acoustique de ces salles dans une salle très absorbante. Pour valider la synthèse, il faut donc mesurer la réponse impulsionnelle (RI) de la salle simulée et la comparer à celle de la salle réelle. Cela permet de vérifier la pertinence des tests subjectifs.

Le protocole qui est expliqué par la suite développe les mêmes procédures que celles utilisées dans la thèse de Fabio Figueiredo [Fig11] pour mesurer la RI dans les salles réelles, ainsi que pour restituer le champ sonore de ces mêmes salles dans une chambre sèche. La figure suivante illustre les différentes étapes du protocole.

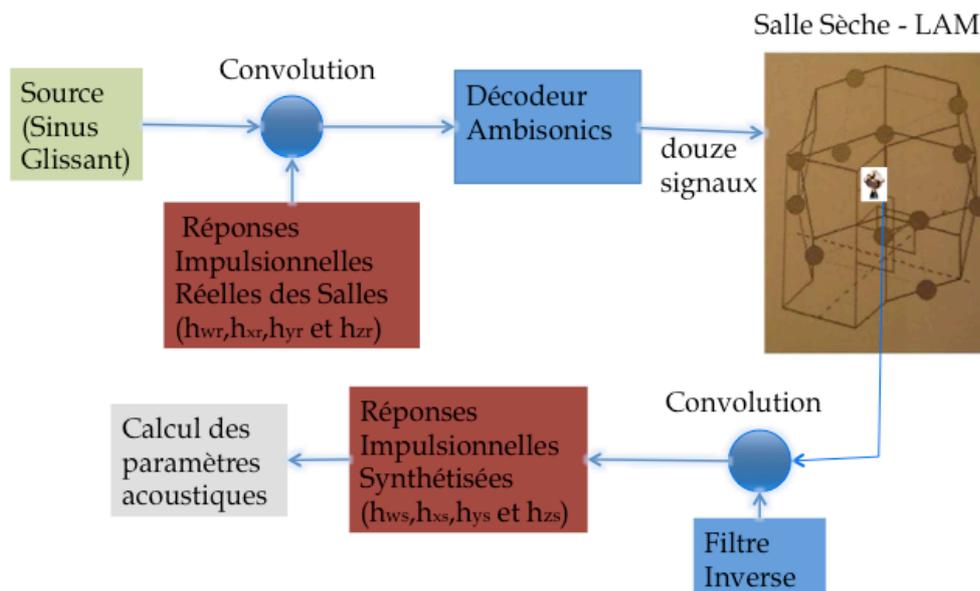


Figure 4. Etapes du protocole expérimental.

4.1 Egalisation du système source en salles réelles

Les réponses impulsionnelles (RIs), prises dans des salles réelles (*in-situ*), tiennent compte de la réponse en fréquence du système source. Autrement dit, le signal d'excitation généré dans les salles pour mesurer la RI a été filtré afin d'égaliser la réponse en fréquence de l'enceinte omnidirectionnelle utilisée pendant les enregistrements.

Suivant l'équation (29), à partir de la RI de l'enceinte $e(t)$ qui a été mesurée dans la chambre anéchoïque du LNE¹, deux processus de filtrage ont été établis, $f_{pre}(t)$ et $f_{pos}(t)$, l'un avant et l'autre après l'enregistrement dans la salle.

$$f_{pre}(t) * e(t) * f_{pos}(t) = \delta(t) \quad (29)$$

La méthode de filtrage en deux étapes a été réalisée pour appliquer le processus d'égalisation sur un large intervalle de fréquences sans dégrader le rapport signal sur bruit [Fig11]. La pré-égalisation est limitée, entre 100Hz et 4000Hz et la post-égalisation entre 40Hz et 18000Hz. La figure 5 illustre la réponse en fréquence de la source, la réponse après le processus de pré-égalisation et la réponse après le processus de post-égalisation. Ainsi, les mesures des RIs ont été faites en utilisant un signal d'excitation filtré x_{EQ} , défini par:

$$x_{EQ}(t) = x(t) * f_{pre}(t) \quad (30)$$

où $x(t)$ est le signal d'excitation et $f_{pre}(t)$ est le filtre de pré-égalisation. De cette façon, la réponse d'une salle $y(t)$ est exprimée par :

$$y(t) = x_{EQ}(t) * e(t) * h(t) + n(t) \quad (31)$$

où $e(t)$ est la RI de l'enceinte, $h(t)$ est la RI de la salle et $n(t)$ est le bruit du système assumé comme bruit blanc gaussien. Enfin, en appliquant l'équation (15) et le filtre de post-égalisation $f_{pos}(t)$, $h(t)$ est égal à :

$$h(t) = y(t) * x^{-1}(t) * f_{pos}(t) + n(t) \quad (32)$$

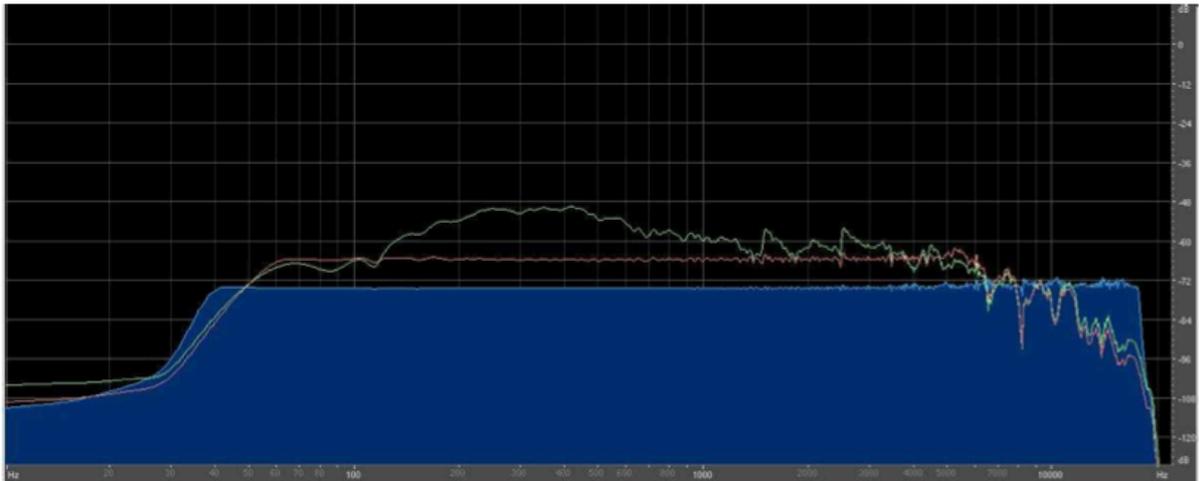


Figure 5. Illustration tirée de la thèse de F.L. Figueiredo [Fig11]. Réponse en fréquence de la source (vert). Réponse après le processus de pré-égalisation (rouge). Réponse après le processus de post-égalisation (bleu).

¹ Laboratoire National de Métrologie et d'Essais

² Le microphone SoundField ST250 permet la prise du son dans le format B (canaux W, X, Y et Z) du système

Comme il a été démontré par Figueiredo [Fig11], le processus d'égalisation de la source est important pour la réalisation d'auralisation. La coloration que donne la source (l'enceinte) génère une perception du son assez différente parmi les auditeurs. C'est pour cette raison que les RIs utilisées dans la simulation des salles ont été calculées en se basant sur cette procédure.

4.2 Système de simulation des salles

La simulation ou synthèse de l'acoustique des différentes salles de concert, étudiées par F.L. Figueiredo pendant les tests subjectifs [Fig11], a été effectuée avec le système Ambisonics installé dans la salle sèche du LAM/IJLDRA. Ce système a été sélectionné car la restitution Ambisonics se caractérise par une bonne sensation d'immersion donnant l'impression aux auditeurs d'être au milieu de la scène sonore [Gua03].

Le système de reproduction se sert d'une carte audio RME. Il est composé de douze enceintes Studer A1 disposées autour du centre de la chambre, ainsi que d'un caisson de basses fréquences JBL 4645C. La figure 6 illustre la disposition des haut-parleurs dans la salle sèche du LAM.

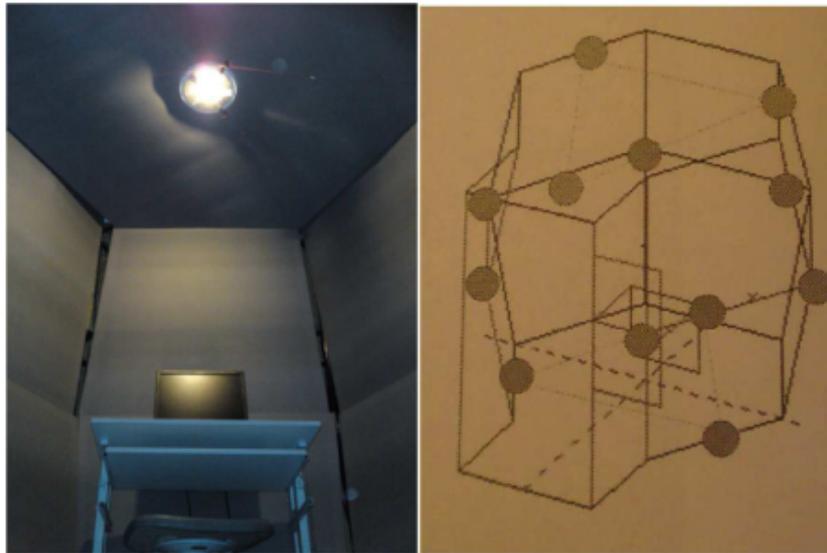


Figure 6. Illustration tirée de la thèse de F.L. Figueiredo [Fig11]. À gauche, la photographie de la salle sèche du LAM. À droite, schéma du placement des haut-parleurs de la salle.

L'auralisation Ambisonics a été implémentée à l'ordre 1 (B-Format). Comme il a été expliqué en 3.2, le signal d'alimentation de chaque haut-parleur est fonction des composantes encodées (W, X, Y et Z) et de leurs coordonnées (angle d'azimut, angle du site), dont l'origine du repère est le point central de la zone d'écoute.

L'auralisation Ambisonics se fait en deux étapes :

- La convolution entre les réponses impulsionnelles de la salle et la source.
- Le décodage Ambisonics à l'ordre 1 sur les douze enceintes.

Les RIs des salles ont été mesurées en utilisant le microphone SoundField ST250², qui donne donc quatre RI pour chaque salle sous B-Format (h_W, h_X, h_Y, h_Z). La source est le signal d'excitation ou un morceau musical enregistré en chambre anéchoïque. De cette manière, et comme défini dans l'équation (30), la composante omnidirectionnelle $x_W(t)$ et les trois composantes bidirectionnelles $x_X(t)$, $x_Y(t)$ et $x_Z(t)$ correspondent à la convolution de la source $x(t)$ avec les quatre réponses impulsionnelles mesurées au B-Format.

$$\begin{aligned}x_W(t) &= x(t) * h_W(t) \\x_X(t) &= x(t) * h_X(t) \\x_Y(t) &= x(t) * h_Y(t) \\x_Z(t) &= x(t) * h_Z(t)\end{aligned}\tag{30}$$

L'enregistrement des RIs *in-situ* a été réalisé en plusieurs positions d'écoute en plaçant la source au centre, à gauche et à droite sur la scène. Il est possible donc de synthétiser le champ sonore avec une source mono ou une source stéréo. Pour une source monophonique, les RIs ont été mesurées en situant la source d'excitation au centre de la scène de la salle. Dans le cas d'une source stéréo, les RIs ont été mesurées en fixant la source d'excitation dans les positions gauche et droite de la scène. De cette façon, il existe quatre composantes Ambisonics par canal (donc, huit en total). Néanmoins, les contributions des composantes gauche et droite sont mixées afin d'obtenir les quatre composantes Ambisonics avec les contributions spatiales correctement codées [Fig11].

Pour obtenir les douze signaux qui alimentent les haut-parleurs, le décodage Ambisonics utilise les composantes x_W, x_X, x_Y et x_Z . Tel qu'il a été exprimé dans les équations (10) et (11), à partir des coordonnées des enceintes, le décodeur délivre un signal précis à chaque enceinte afin de restituer le champ sonore d'une salle déterminée. Ces signaux sont envoyés directement au système de reproduction mais aussi au système d'enregistrement pour créer un fichier de 13 canaux (12 canaux *full band* + 1 canal de basses fréquences) qui peut être utilisé par une autre application.

Cette méthode de décodage considère une distribution régulière des haut-parleurs, c'est-à-dire, le système ne prend pas en compte la distance des enceintes au centre de la salle. Etant donné que le placement des haut-parleurs dans la salle sèche du LAM n'est pas uniforme autour de la zone d'écoute, un réglage du mécanisme de reproduction/enregistrement a été effectué afin de garantir au centre de la salle le même niveau de pression sonore (hors simulation) pour chaque enceinte. De plus, tous les haut-parleurs ont été orientés vers le point central de la zone d'écoute.

² Le microphone SoundField ST250 permet la prise du son dans le format B (canaux W, X, Y et Z) du système d'auralisation Ambisonics. Voir section 3.2

Les processus décrits dans les paragraphes précédents sont contrôlés par le logiciel Audiomulch³. L'auralisation Ambisonics est réalisée avec les plugins VST Voxengo Pristine Space⁴ et Deccopro⁵. Voxengo permet de faire la convolution en temps réel de plusieurs canaux et Deccopro fait le décodage Ambisonics de premier ordre. La figure 7 illustre les applications mentionnées.

La simulation des salles a été faite en utilisant la base de données des réponses impulsionnelles des salles réelles ainsi que les fichiers de 13 canaux enregistrés après le décodage Ambisonics. Ces deux méthodes ont été testées afin de réaliser la synthèse du champ sonore. Une méthode à partir de la convolution en temps réel avec la réponse impulsionnelle en utilisant Voxengo et l'autre avec les fichiers de 13 canaux décodés en faisant le routage avec MAX/MSP. Cette dernière méthode a été suivie afin de maintenir la cohérence avec la méthode utilisée dans la thèse de Figueiredo pour les tests d'écoute [Fig11], et de vérifier la conformité de la méthode en temps réel.

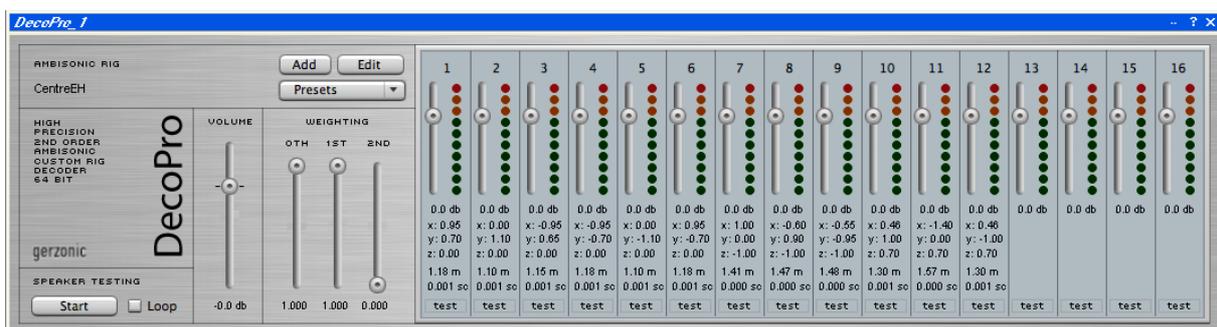
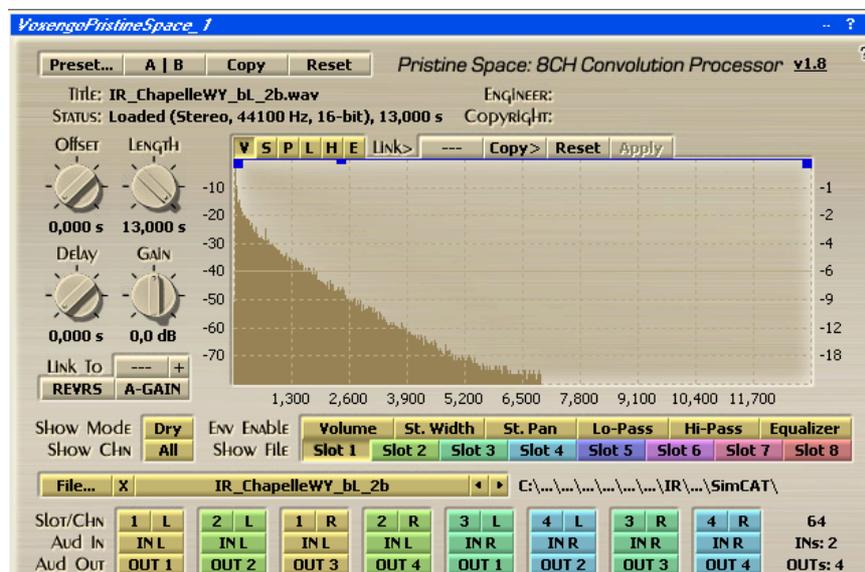


Figure 7. En haut le plugin VST Voxengo Pristine Space. En bas le décodeur Ambisonic DecoPro.

³ <http://www.audiomulch.com>. AudioMulch est un logiciel pour le traitement d'audio, le design sonore et la composition musicale.

⁴ <http://www.voxengo.com/product/pspace>.

⁵ "<http://www.gerzonic.net>"

4.3 Mesure de la réponse impulsionnelle des salles synthétisées

Le sous-ensemble de salles choisies correspond aux salles utilisées dans le test de catégorisation⁶. Celles-ci sont : le Théâtre des Abbesses, le Théâtre de l'Athénée, la Chapelle Royale de Versailles, la Cité de la Musique, la Salle Cortot, la Maison de la Culture du Japon, l'Auditorium du Louvre, le Théâtre de la Porte St. Martin, l'Auditorium du Musée d'Orsay et la Salle Pleyel. La simulation a été faite en utilisant les RIs réelles mesurées dans les trois positions sources (centre, gauche et droite), comme expliqué dans la section précédente. La position de réception varie relativement entre une salle et une autre, mais toutes les positions ont été prises dans un point central près de la scène.

Les RIs des salles synthétisées ont été mesurées en utilisant la méthode de balayage logarithmique qui permet de supprimer les produits de distorsion causés par les non-linéarités du système de reproduction, car ils apparaissent clairement séparés de la partie linéaire de la RI (section 3.4). La figure 8 illustre cette caractéristique.

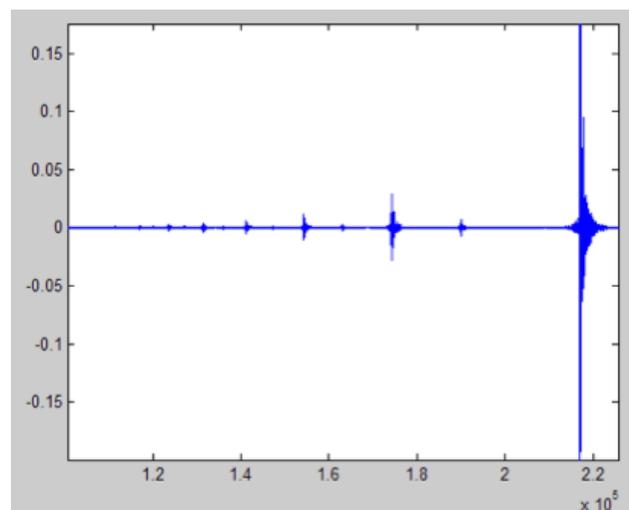


Figure 8. Partie linéaire de la RI (tout à droite) et ses produits de distorsion. Après l'édition seulement la partie linéaire est conservée

Ainsi, l'auralisation Ambisonics (cf. section 3.2 et 4.2) des salles choisies a été faite en prenant comme source un sinus glissant logarithmique sur une plage de fréquences entre 22Hz et 22000Hz et une durée de 30s. Le niveau du signal d'excitation est aussi important au moment de la mesure. Il est nécessaire d'éviter la saturation du signal, mais il est également important de conserver un rapport signal sur bruit suffisant dans toutes les plages de fréquence. Ce niveau a été adapté aux caractéristiques des enceintes. La réponse de chaque salle a été prise à l'aide d'un microphone placé au milieu de la zone d'écoute de l'espace sonore restitué dans la salle sèche du LAM. La RI s'obtient en appliquant le filtre inverse à ce signal (cf. équation 15). Seulement la partie linéaire de celle-ci est retenue.

⁶ Thèse de Fabio Figueiredo [Fig11]. Chapitre 5.

Afin de donner l'impression d'être dans une salle particulière, le niveau sonore du système de simulation (l'auralisation Ambisonics) doit être soigneusement réglé par rapport au niveau sonore réel. Dans les salles réelles, ce niveau a été mesuré en utilisant un bruit rose⁷. Ces valeurs ont été conservées pour chacune des positions où la RI a été mesurée. Ainsi, l'ensemble des enceintes du système de reproduction de la salle sèche a été étalonné à l'aide du sonomètre CEL 573. Cet étalonnage a été effectué suivant le niveau mesuré par Figueiredo dans chacune des salles réelles et en utilisant le même signal de calibration, ainsi que la même synthèse du champ sonore.

Lors des premières expériences et après avoir fait l'analyse des indices acoustiques, quelques valeurs non cohérentes ont été identifiées. Comme expliqué en 3.5, plusieurs paramètres acoustiques sont calculés à partir de la décroissance de l'énergie. Cela a permis l'identification d'une erreur liée à la prise de la réponse impulsionnelle. La figure 9 illustre le spectrogramme 3D de la RI d'une salle et l'erreur en question. Elle montre une décroissance très particulière de l'énergie, parcourant toute la plage des hautes fréquences jusqu'aux basses fréquences.

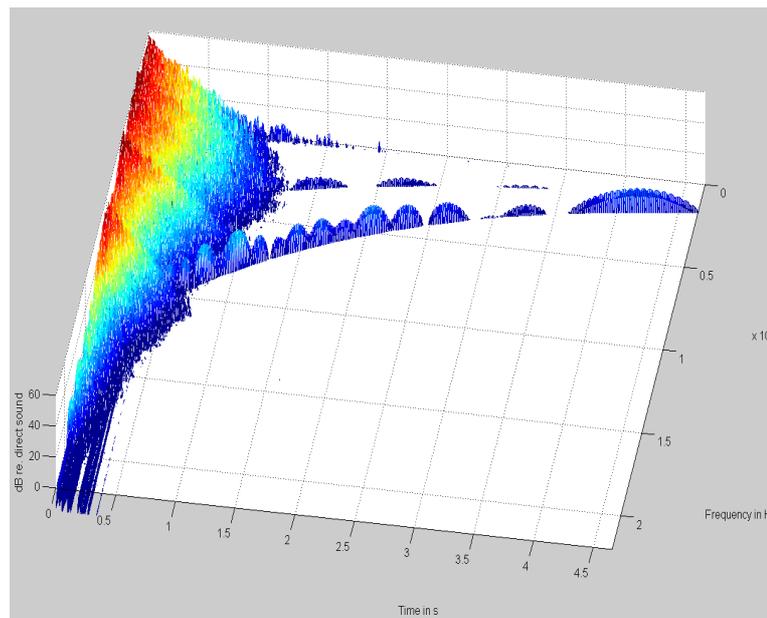


Figure 9. Spectrogramme 3D (Intensité-temps-fréquence) de la RI d'une salle quelconque. La partie en bleu, en forme d'arc, représente une erreur dans la RI résultant du processus de déconvolution, conséquence des discontinuités temporelles du signal pendant l'enregistrement. Sa forme correspond au spectrogramme du filtre inverse du signal sinus glissant logarithmique

⁷ Ce même signal a été utilisé pendant l'étalonnage de la source omnidirectionnelle au LNE (Laboratoire National de Métrologie et d'Essais). [Fig11]

Après avoir vérifié toutes les étapes de mesure, de minuscules discontinuités (clics) sont observées en zoomant le signal dans le domaine temporel et spectral. La figure 10 illustre ces discontinuités générées pendant l'enregistrement et qui ne sont pas visibles à simple vue.

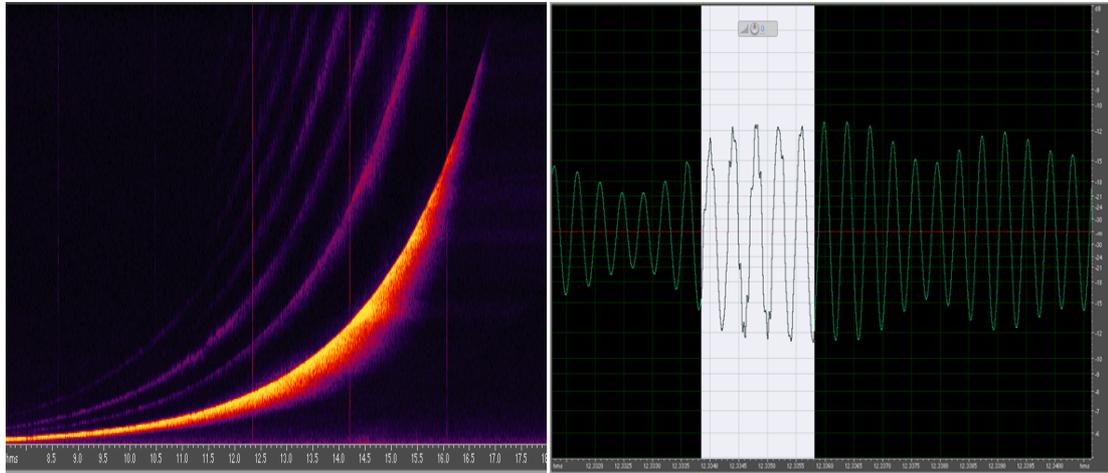


Figure 10. À gauche, le spectre de la RI. Les lignes verticales en rouge sont les clics. À droite, le signal dans le domaine temporel zoomé sur un des clics.

Les raisons de ces erreurs sont diverses et, en général, attachées à la performance de l'ordinateur dans sa gestion de la carte audio. Une première erreur est liée à l'utilisation de deux logiciels d'audio (par exemple, le logiciel Adobe Audition pour faire des enregistrements et le logiciel Audiomulch pour reproduire le signal d'excitation). La seconde est due à l'utilisation de clés USB, ou d'autres dispositifs, avec enregistrement simultané. Bien que ce problème soit propre au système utilisé dans la salle sèche, il a permis de dévoiler d'autres erreurs éventuelles pendant la mesure de la RI. La figure suivante illustre un exemple des RIs obtenues sans aucun type d'erreur ou de distorsion.

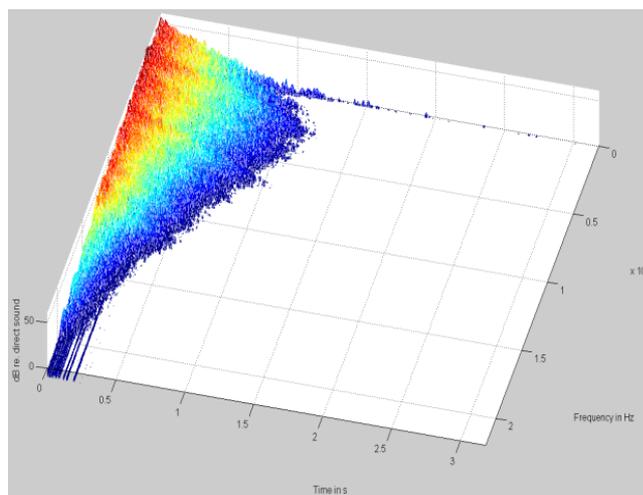


Figure 11. Spectrogramme 3D (Intensité-temps-fréquence) d'une RI nette correspondant à une salle simulée.

Après vérification du protocole de mesure, le microphone SoundField ST250 a été utilisé pour faire la prise des RIs en format Ambisonics, avec une fréquence d'échantillonnage de 44.1kHz. Ainsi, les quatre signaux Ambisonics B-Format $\tilde{x}_W, \tilde{x}_X, \tilde{x}_Y$ et \tilde{x}_Z ont été pris par le microphone pendant l'auralisation de la salle. Comme mentionné déjà dans cette section, les RIs $\tilde{h}_W, \tilde{h}_X, \tilde{h}_Y$ et \tilde{h}_Z de la salle simulée sont obtenues par la méthode de balayage logarithmique (voir section 3.2 et 3.4). La déconvolution des signaux Ambisonics B-Format $\tilde{x}_W, \tilde{x}_X, \tilde{x}_Y$ et \tilde{x}_Z (c'est-à-dire la convolution avec le filtre inverse) a été faite avec la suite d'Aurora⁸ développée par Farina [Far00, Far07]

De cette façon, pour chacune des 10 salles choisies, 3 ensembles de 4 RIs en B-Format ont été mesurés (un pour chaque position source: centre, gauche et droite). C'est-à-dire un total de 120 réponses impulsionnelles. Il a été nécessaire de suivre un processus d'édition pour le début et la fin de chaque RI, avant de faire le calcul des indices acoustiques.

4.4 Calcul des indices acoustiques

Les RIs des salles réelles et des salles synthétisées ont été analysées. Tous les indices sont calculés en utilisant la RI omnidirectionnelle h_W ou \tilde{h}_W , sauf l'indice de spatialisation LF/LFC qui a besoin aussi de la RI directionnelle h_Y ou \tilde{h}_Y , pour être calculé. Ainsi, pour chaque mesure, une paire de RIs (W, Y) par position de la source (centre, gauche et droite) a été analysée. Donc, un total de 30 paires de RIs analysées pour les salles réelles et de 30 paires de RIs aussi pour les salles synthétisées.

Les indices retenus sont les descripteurs acoustiques classiques⁹: le temps de réverbération (T30), le temps de décroissance précoce (EDT), la clarté (C80), le temps central (Ts), l'amplification sonore (G) et aussi l'efficacité latérale (LF). En plus, nous calculons les rapports BR (« bass ratio ») et TR (« treble ratio »). Les calculs ont été faits avec Aurora et OpenMIDAS¹⁰ dans les octaves 125Hz, 250Hz, 500Hz, 1000Hz, 2000Hz et 4000Hz d'après la norme ISO 3382. Procéder aux mêmes calculs à l'aide de ces deux logiciels permet d'augmenter la fiabilité des résultats obtenus.

Aussi, deux facteurs ont été considérés afin de minimiser des imprécisions pendant les calculs des indices acoustiques :

- Le temps du début des réponses impulsionnelles.
- Les caractéristiques des sources (enceintes) utilisées pendant l'enregistrement de ces réponses impulsionnelles.

⁸ Aurora est un ensemble de plugins VST qui fonctionne sur Adobe Audition développée par Angelo Farina.

⁹ cf section 3.5

¹⁰ OpenMIDAS est la version sur Matlab du logiciel MIDAS développé sous l'impulsion de Jean Dominique Polack.

Le temps de début ($t=0$) de la RI est essentiel pour les calculs des indices acoustiques (comme mentionné en 3.5). Il n'existe pas une méthode établie dans la norme ISO 3382 à ce propos. Il a été démontré par Defrance et al. [DDP08] qu'une variation de seulement 2ms dans le temps de début de la RI peut générer d'importantes variations des indices T_c et C_{80} . Ce sujet n'est pas traité dans le cadre de ce rapport, mais il a été introduit afin d'expliquer que la méthode appliquée, concernant le temps du début, est celle de Figueiredo.

La figure 12 montre la fenêtre d'options de l'application OpenMIDAS. On règle le calcul du temps de début dans la partie « *Acquisition* ». Tant dans OpenMIDAS que dans Aurora, le temps de début a été établi là où l'amplitude de la RI est inférieure de 20dB à son amplitude maximale, et avant l'arrivée de celle-ci.

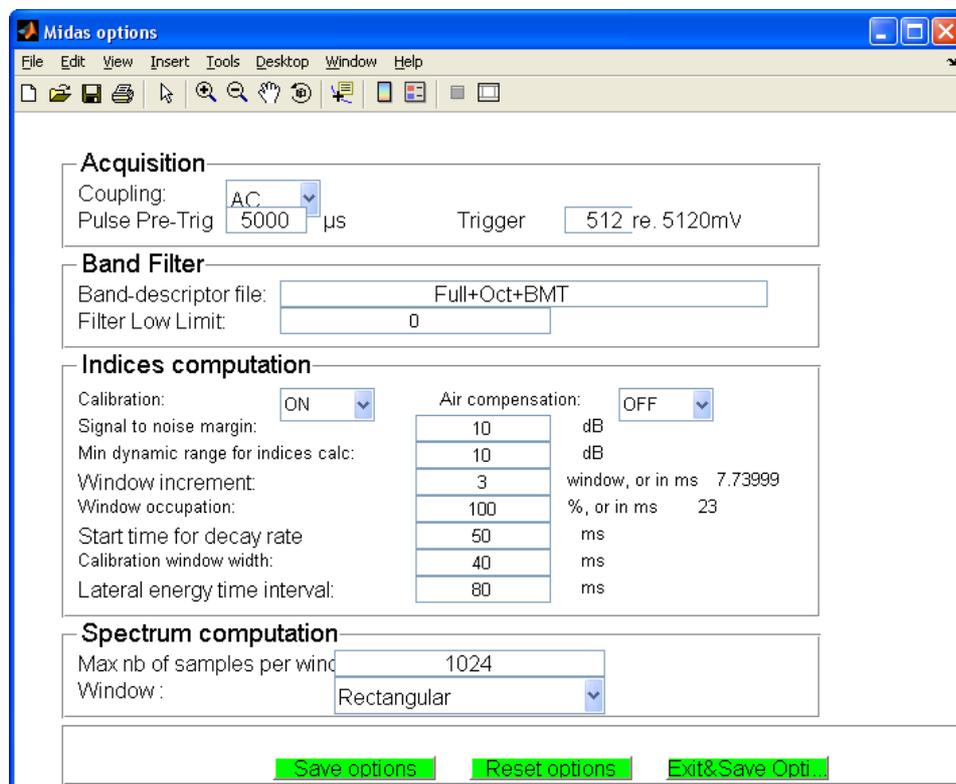


Figure 12. Fenêtre Options de l'application OpenMIDAS

Par ailleurs, il a été expliqué en 3.5, que le facteur de force G est calculé en prenant comme référence l'énergie de la source mesurée en chambre anéchoïque. Par conséquent, un processus de calibration est nécessaire au moment de calculer G . Sur OpenMIDAS dans la partie « *Indices Computation* » (cf. figure 12), on règle tous les paramètres nécessaires concernant le calcul des indices acoustiques. Si l'option de calibration est ON, le logiciel tiendra compte d'un fichier de calibration qui n'est que la RI de la source¹¹.

¹¹ La RI de la source seule mesurée en chambre anéchoïque

Il faut considérer séparément le cas des salles réelles et celui des salles synthétisées. Les RIs des salles réelles ont été mesurées en utilisant une source dodécaphonique omnidirectionnelle tandis que pour les RIs des salles synthétisées les enceintes Studer A1 de la salle sèche du LAM ont été utilisées.

Dans le premier cas, les fichiers correspondants à la RI de la source omnidirectionnelle enregistrés dans la salle anéchoïque du LNE ont été utilisés comme signal de référence. De plus, il a été appliqué le même protocole d'égalisation employé par Fabio : un pré et post filtrage, de façon à obtenir une réponse plate entre les fréquences de 40Hz à 18KHz (cf. section 4.1).

Dans le deuxième cas, il a fallu mesurer la RI des enceintes. Bien que les douze haut-parleurs appartiennent à un même modèle, il existe des différences dans la réponse en fréquence. Après avoir défini le meilleur signal de référence (moyenne des RIs, etc), il a été décidé de choisir ceux qui appartiennent aux enceintes donnant des informations relatives au placement de la source. Les RIs de ces enceintes ont été mesurées dans la salle sèche du LAM. Bien que la salle ne soit pas anéchoïque, elle permet d'avoir une mesure assez fiable dans la plage de fréquences analysées (125Hz à 4kHz). La figure suivante illustre les réponses en fréquences des enceintes choisies.

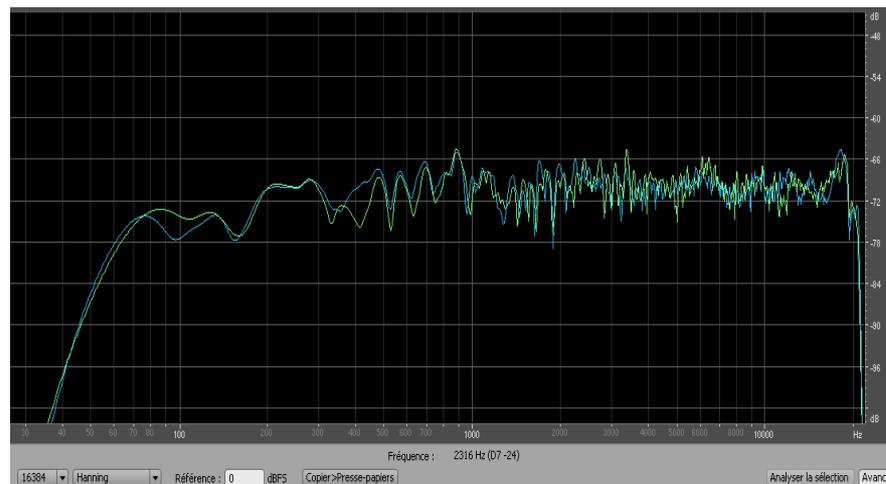


Figure 13. Réponse en fréquences des enceintes Studer A1 numéro 2 et 5. Celles qui correspondent à la direction gauche (vert) et droite (bleu) dans l'émission du son direct dans le système d'auralisation.

On n'a pas suivi un processus d'égalisation des enceintes. Comme expliqué dans l'introduction, l'objectif était d'analyser les salles synthétisées prenant en compte les mêmes conditions que celles que F.L. Figueiredo a observées dans sa thèse, où les tests subjectifs ont été effectués sans modifier la réponse en fréquence des enceintes de la salle du LAM. Néanmoins, dans une deuxième étape de ce travail et afin de réduire le *biais* introduit par le système de reproduction, il sera nécessaire de faire un étalonnage du système de reproduction en mesurant la réponse en fréquence en champ diffus au centre de la salle sèche et en réalisant l'égalisation complète du système.

5. Analyses et Résultats

L'objectif de ce travail était d'analyser des paramètres acoustiques des salles synthétisées et de les comparer à ceux des salles réelles. Il a été décidé aussi de faire un test subjectif de catégorisation avec les mêmes salles que celles étudiées par Fabio Figueiredo mais en ajoutant les salles synthétisées. De cette façon, les analyses se divisent en trois parties : les analyses objectifs, le critère de seuil différentiel, et le test subjectif.

5.1 Analyses Objectives

Six indices acoustiques (T30, EDT, C80, Ts, G et LF) ont été analysés en bande d'octave de 125Hz à 4000Hz dans 20 salles (10 champs sonores réels et 10 champs sonores synthétisés) pour 3 positions de source (centre, gauche et droite). Cet ensemble d'indices acoustiques forme une base de données qui peut être traitée par une approche statistique. Deux méthodes ont été utilisées : les boîtes à moustache, et l'ACP (analyse en composantes principales). L'annexe A.1 contient les tables des valeurs des indices acoustiques utilisés pour les calculs.

5.1.1 Etude des valeurs centrales et dispersions

La boîte à moustache ou « *box plot* » est un graphique ou diagramme qui fournit des informations en 5 chiffres : les valeurs minimale et maximale ; et les quartiles Q1, Q2 ou médiane, et Q3. De plus, ce diagramme illustre l'existence de valeurs aberrantes ou exceptionnelles ainsi que la symétrie de la dispersion [Pot06].

Ces graphiques ont été utilisés par Polack, Figueiredo et Liu afin de comparer un même indice acoustique dans 15 salles Parisiennes, avec plusieurs positions de mesure [PFL12]. Pour la même salle l'indice T30 a présenté une valeur presque constante dans toutes les positions. En revanche, la plupart des indices acoustiques analysés ont présenté une grande dispersion. En général, les résultats ont montré que, pour caractériser les salles, les valeurs des indices ne peuvent pas être considérées distinctement, mais plutôt l'ensemble de valeurs de tous les indices.

A la différence des analyses réalisées dans [PFL12], nous cherchions à savoir si les valeurs des indices acoustiques analysés dans les salles réelles étaient semblables à celles des salles synthétisées. Ainsi, une première comparaison entre les salles réelles et les salles synthétisées a été faite en observant les valeurs moyennes et les dispersions des indices acoustiques pour une même position de source.

Pour calculer les boîtes à moustaches, la valeur moyenne de chaque indice a été prise pour toutes les salles sur les octaves entre 250Hz et 2000Hz. Par conséquent, les diagrammes « *box*

plot » montrent la dispersion de ces indices dans les 10 salles choisies¹² (champ sonore réel et synthétisé) pour les trois positions source. La convention utilisée a été la suivante:

- C : Source au Centre / Salles Réelles
- CS : Source au Centre / Salles Synthétisées
- G : Source à Gauche / Salles Réelles
- GS : Source à Gauche / Salles Synthétisées
- D : Source à Droite / Salles Réelles
- DS : Source à Droite / Salles Synthétisées

Les boîtes à moustaches ont été faites sous MATLAB en utilisant la fonction `boxplot`. Les figures 14 et 15 illustrent les diagrammes des indices T30 et EDT. Afin de mieux analyser la dispersion, leurs graphiques ont été tracés avec le logarithme de leurs valeurs.

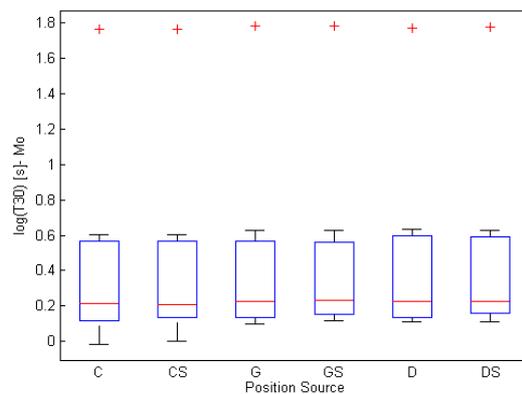


Figure 14. Boîtes à moustaches de l'indice T30.

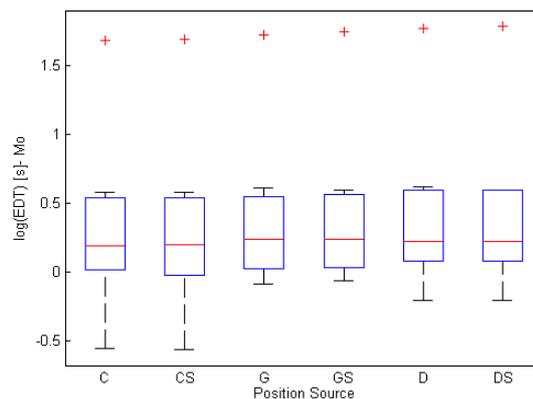


Figure 15. Boîtes à moustaches de l'indice EDT.

En faisant la comparaison de la valeur moyenne et de la dispersion de l'indice T30, nous observons que les diagrammes sont presque identiques entre les salles réelles et synthétisées pour une même position source. De plus, les valeurs sont presque constantes dans les trois

¹² cf. section 4.3

positions. La croix rouge qui marque les valeurs exceptionnelles appartient à la Chapelle Royal de Versailles qui a un T30 beaucoup plus grand que les autres salles. Dans le cas de l'indice EDT, quelques variations entre les trois positions sources sont observées. Ces différences sont prévisibles car l'EDT dépend des réflexions précoces qui varient par rapport à la position de la source. Néanmoins, de la même façon que le T30, la valeur moyenne et la dispersion de l'indice EDT, sont conservées entre les salles réelles et les salles synthétisées pour la même position source. A partir de ces résultats, nous pourrions dire (à priori) que les réflexions tardives du champ sonore réel sont bien reconstruites dans le champ sonore synthétisé.

Les figures 16, 17, 18 et 19 illustrent les boîtes à moustaches des indices C80, Ts, G et LF. Elles ont été calculées directement sur les valeurs originales.

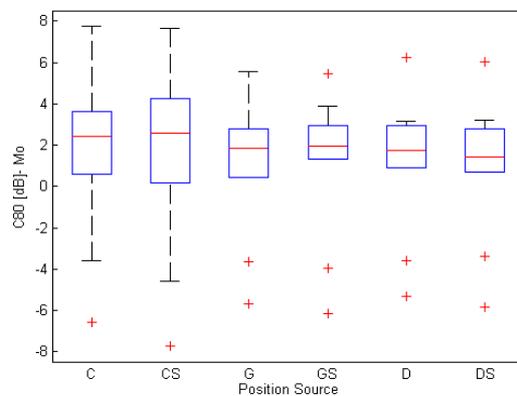


Figure 16. Boîtes à moustaches de l'indice C80.

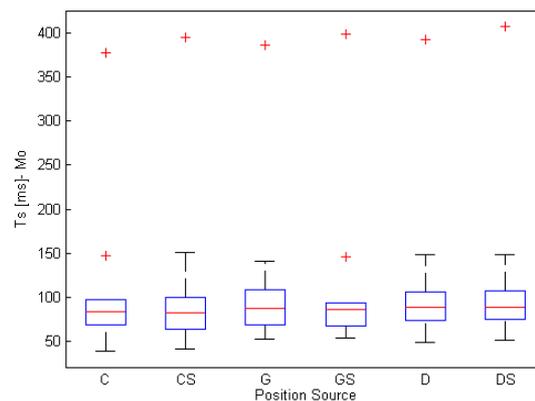


Figure 17. Boîtes à moustaches de l'indice Ts.

Les indices C80 et Ts présentent beaucoup plus des variations entre les trois positions source. Comme pour l'indice EDT, les différences sont explicables, car ces deux indices sont très influencés par l'énergie des réflexions précoces. Bien que la valeur moyenne soit maintenue dans l'ensemble de salles ainsi que quelques valeurs maximales et minimales, la dispersion présente des différences entre salles réelles et salles synthétisées pour la même position. Ces

variations sont plus grandes pour C80 que pour Ts. Les croix rouges appartiennent à la Chapelle Royal de Versailles¹³, La Cité de la Musique¹⁴ et le Théâtre de l'Athénée¹⁵

Pour les indices G et LF, les analyses ont été faites seulement pour la position source au centre¹⁶. La lettre « R » dans la figure, fait référence au calcul de G qui a été fait à partir de la mesure directe (avec un sonomètre) de la pression acoustique générée dans les salles¹⁷. En revanche, les boîtes à moustache en « C » et « CS » ont été faites à partir des RIs réelles et synthétisées respectivement¹⁸ (cf. section 3.5). Nous pouvons observer que G en R présente une dispersion en peu plus grande qu'en C et CS. De plus, G (CS) est décalé par rapport à G (C). Pour l'indice LF, le diagramme montre clairement une dispersion plus faible dans les salles synthétisées que dans les salles réelles.

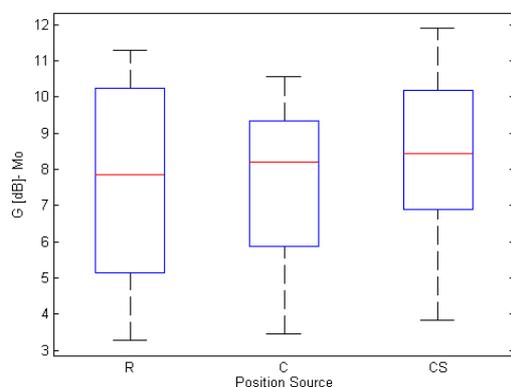


Figure 18. Boîtes à moustaches de l'indice G.

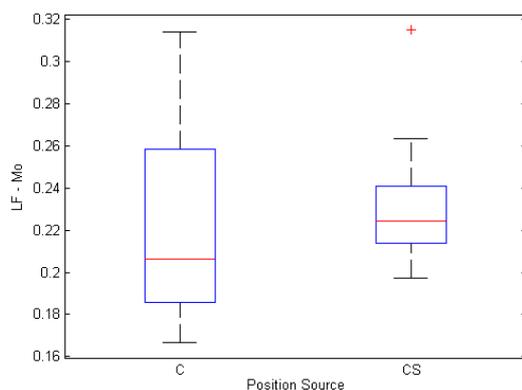


Figure 19. Boîtes à moustaches de l'indice LF.

¹³ Les valeurs plus basses pour C80 et les plus élevées pour Ts

¹⁴ Les croix au milieu pour C80 et les deux croix autour de 150ms pour Ts

¹⁵ Seulement pour l'indice C80, les valeurs plus élevées

¹⁶ La mesure directe (avec sonomètre) de G a été réalisée seulement en position centrale. Les analyses ont été donc effectuées dans cette même position. Pour l'indice LF, la position centrale donne la meilleure réponse.

¹⁷ Cf. F.L Figueiredo [Fig11] p.25

¹⁸ Il a fallu appliquer un gain afin d'approcher les valeurs de celles mesurées par la méthode directe.

5.1.2 Analyse en composantes principales

L'ACP est une méthode non-paramétrique utilisée pour réduire le nombre des variables (dimensions) d'un ensemble de données. A partir d'un ensemble d'observations décrites par un nombre de variables initiales, l'ACP cherche à décrire ce même ensemble de données par de nouvelles variables en nombre réduit nommées *composantes principales*. Cette technique se base sur les assumptions suivantes [Shl09]:

- Linéarité : les composantes principales peuvent être écrites comme des combinaisons linéaires des variables initiales.
- Les grandes variances ont une structure importante : les composantes principales avec des grandes variances représentent une structure importante tandis que celles avec des petites variances représentent un bruit.
- Les composantes principales sont orthogonales.

Dans d'autres termes, l'ACP cherche de nouvelles variables (les composantes principales) non corrélées entre elles, et les plus corrélées avec l'ensemble des variables initiales. Un des avantages majeur de l'APC découle de la quantification de l'importance de chaque variable pour décrire la variabilité d'un ensemble de données. Par conséquent, cette technique est utilisée pour trouver les causes de la variabilité dans un ensemble de données (système) et de les ranger par ordre d'importance.

Ainsi, cette méthode a été utilisée pour vérifier si la variabilité des indices acoustiques calculés à partir du champ sonore synthétisé correspondait à la variabilité trouvée par F.L Figueiredo pour les indices acoustiques calculés à partir du champ sonore réel. Une explication plus détaillée de la méthode et son application à ce sujet est exposée dans [Liu11] et [Fig11].

La table 2 montre l'ensemble de données analysées. Les colonnes représentent les variables (indices acoustiques) et les lignes représentent les individus (salles synthétisées).

	T30 [s]	EDT [s]	C80 [dB]	Ts [ms]	G [dB]	LF
Abesses	1,14	1,03	3,70	69,33	9,40	0,22
Athenée	1,16	0,79	6,42	54,00	6,55	0,20
Chapelle	5,36	4,81	-6,45	346,83	10,47	0,24
Cité	1,73	1,70	-3,77	143,83	6,30	0,26
Cortot	1,13	1,04	3,38	72,00	9,57	0,22
Japan	1,27	1,28	0,15	99,67	4,93	0,23
Louvre	1,46	1,21	1,18	93,83	9,72	0,20
Martin	1,09	0,65	7,12	46,83	3,02	0,22
Orsay	1,23	1,13	1,95	88,17	11,35	0,32
Pleyel	1,71	1,68	1,80	92,17	6,97	0,21

Table 2. Valeurs moyennes des 6 indices acoustiques des salles synthétisées

L'ACP a été faite en prenant la valeur moyenne des indices acoustiques T30, EDT, C80, Ts, G et LF de chaque salle synthétisée, correspondant à la mesure de la RI en position source au centre. Les graphiques ont été faits sous MATLAB en utilisant la fonction `princomp`. De plus, la rotation orthogonale *varimax*¹⁹ a été utilisée pour avoir une interprétation plus aisée des résultats. La figure 20 illustre l'analyse pour les deux premières composantes. Chaque indice acoustique est représenté par un point sur le graphique. La direction et la distance de ce point à l'origine indiquent la contribution de l'indice aux composantes principales [PFL12]. Il est observé que la première composante est fortement corrélée avec les indices Ts, EDT, T30 et C80, cette composante représente 65.87% de la variance des données, tandis que la deuxième composante est fortement corrélée à l'indice LF et représente 19.25% de la variance. C'est-à-dire, ces deux composantes expliquent le 85.12% de la variance totale.

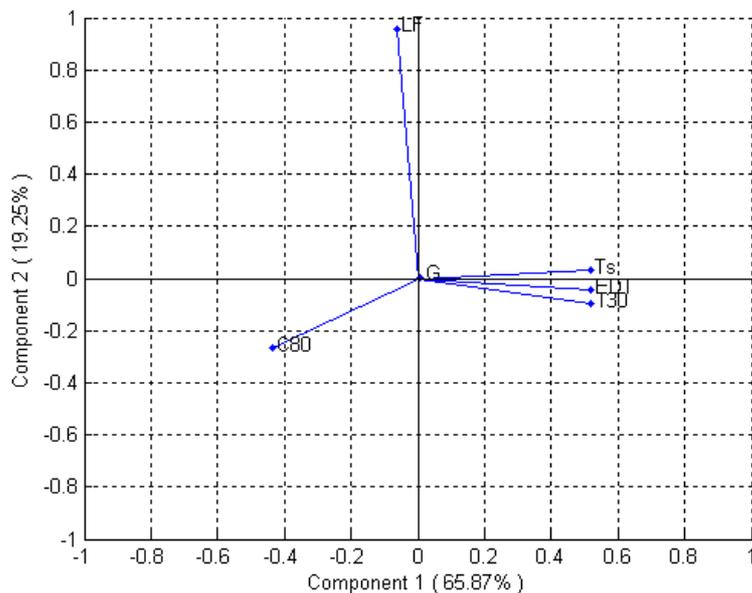


Figure 20. Analyse en composantes principales 1 et 2 pour 6 indices acoustiques.

La figure 21 illustre l'analyse pour la première et la troisième composante. On y observe la même corrélation de la première composante avec les indices T30, Ts, EDT et C80. Par ailleurs, la troisième composante principale est fortement corrélée avec l'indice G et représente 11.1% de la variance. Ainsi, les trois premières composantes expliquent 96.22% de la variance de l'ensemble de données analysées. La figure 22 constate la forte contribution de LF à la seconde composante et de G à la troisième composante.

¹⁹ Le critère mathématique *varimax* cherche à maximiser la variance de la série des corrélations au carré avec les variables initiales, ce qui fait que chaque composante principale est plus fortement corrélée à certaines variables et beaucoup moins corrélée aux autres.

Dans le plan ACP, il est possible de visualiser la corrélation entre deux variables (deux indices acoustiques) en évaluant le cosinus de l'angle entre elles. Nous pouvons observer sur les figures 20 et 21 que les indices T30, EDT et Ts sont bien corrélés. En revanche, l'indice C80 présente une forte corrélation négative avec ceux-ci. Ces mêmes résultats ont été observés par J.D. Polack, J.F Figueiredo et S. Liu [PFL12].

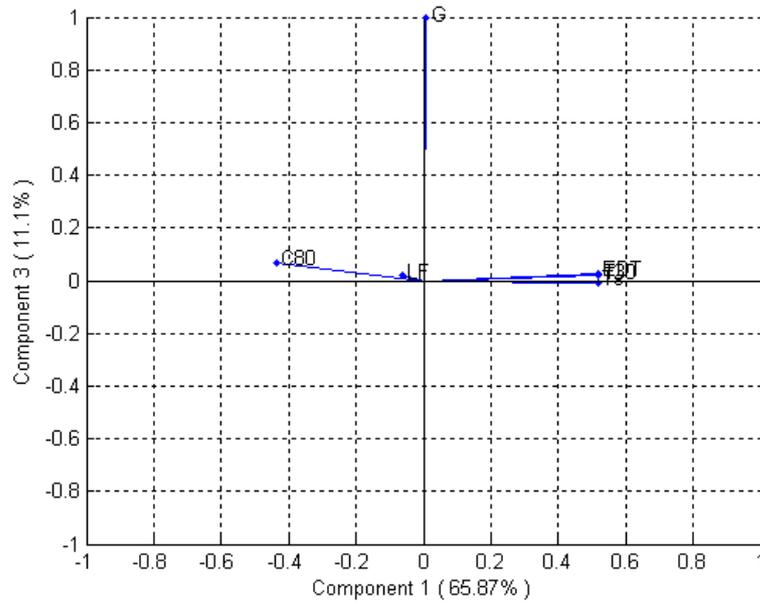


Figure 21. Analyse en composantes principales 1 et 3 pour 6 indices acoustiques.

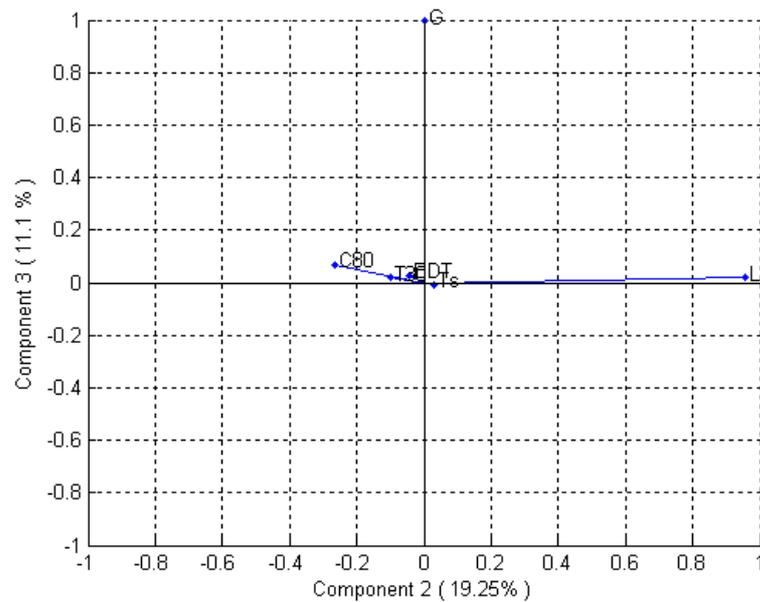


Figure 22. Analyse en composantes principales 2 et 3 pour 6 indices acoustiques.

Ces analyses ont été réalisées aussi pour l'ensemble des indices acoustiques calculés dans les salles réelles. Les graphiques de l'annexe A.2 montre que les résultats observés sont similaires à ceux des salles synthétisées.

Des résultats similaires à ceux de S. Liu [Liu11] ont été obtenus en réalisant l'analyse ACP avec 8 indices acoustiques, à savoir les 6 indices déjà analysés plus les indices D50²⁰ et I²¹. La première composante représente 59.92% de la variance des données, fortement corrélée aux indices T30, EDT, Ts, C80 et D50. La seconde représente 20.30%, corrélée aux indices G et I et la troisième représente un 12.59%, corrélée à l'indice LF. Ces résultats sont exposés dans l'annexe A.2.

5.2 Analyse du Seuil Différentiel des paramètres acoustiques

Les méthodes précédentes (boite à moustache et ACP) ont permis d'analyser les salles réelles et synthétisées par rapport aux valeurs moyennes des paramètres acoustiques ainsi que leur dispersion. En plus, nous avons observé la variance de l'ensemble de données et nous avons pu constater l'existence de corrélation entre certains indices.

Par ailleurs et afin d'évaluer si la salle synthétisée crée l'illusion à un auditeur d'être immergé dans le champ sonore de la salle réelle correspondante, le critère de seuil différentiel ou JND (*Just Noticeable Difference*) est utilisé. La méthode consiste à observer la différence entre les valeurs du champ sonore réel et synthétisé pour chaque paramètre acoustique objectif. Si cette différence est au-dessous du seuil différentiel, l'auditeur ne percevra pas un changement entre la salle synthétisée et la salle réelle pour cet indice particulier. Les valeurs de seuil différentiel utilisées pour ces analyses ont été tirées de la norme ISO 3382-1²². La table 3 illustre le JND pour les paramètres acoustiques étudiés.

Paramètre Objectif	JND
T30 [s]	5 %
EDT [s]	5 %
C80 [dB]	1 dB
Ts [ms]	10 ms
G [dB]	1 dB
LF [%]	5 %

Table 3. Seuil différentiel des Paramètres acoustiques

Pour vérifier le JND la valeur moyenne des indices doit être calculée. D'après la recommandation ISO 3382-1-2009 pour tous les paramètres acoustiques cette valeur correspond à la moyenne entre les octaves de 500Hz et 1000Hz sauf pour LF dont la valeur

²⁰ La Définition D50 est le rapport de l'énergie précoce à l'énergie totale. La limite d'intégration temporelle de l'énergie précoce est de 50ms. Cet indice est dédié à la parole.

²¹ L'indice I = G/T30 a été utilisé par F. Figueiredo afin d'évaluer l'impression d'intimité d'une salle.

²² Annexe A du standard ISO 3382-1-2009 "Acoustics — Measurement of room acoustic parameters"

moyenne est calculée entre les octaves de 125Hz et 1000Hz. Les analyses du JND ont été faites pour chaque une des 10 salles choisies (cf. section 4.2). Les figures 23, 24, 25 et 26 illustrent la différence en JND entre les salles réelles et les salles synthétisées pour les indices de la table 3.

Concernant l'indice T30 on observe dans toutes les salles, que la différence est en dessous de 1 JND. La valeur la plus haute appartient au Théâtre de la Porte St. Martin qui a moins de 0.5 JND. Par rapport à l'EDT, les différences en JND sont plus élevées que pour le T30 sauf pour le Théâtre de la Porte St. Martin. Néanmoins, toutes les valeurs sont aussi inférieures à 1 JND. C'est-à-dire que, théoriquement, il n'y a pas de différences perceptibles entre salles réelles et salles synthétisées quant à la *réverbération*.

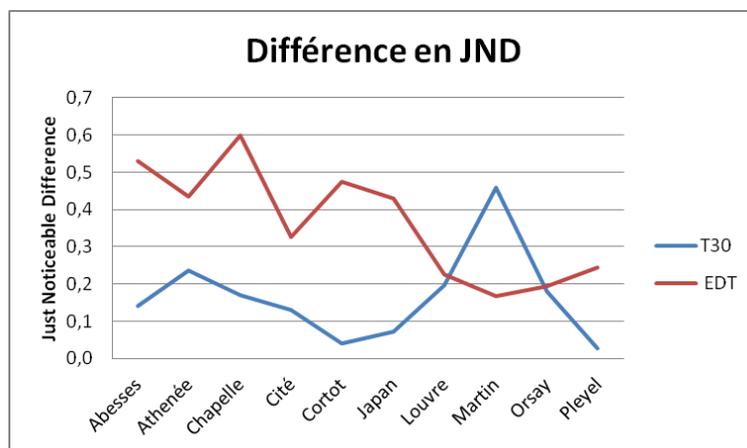


Figure 23. Différence en JND entre les salles réelles et les salles synthétisées pour les indices T30 et EDT.

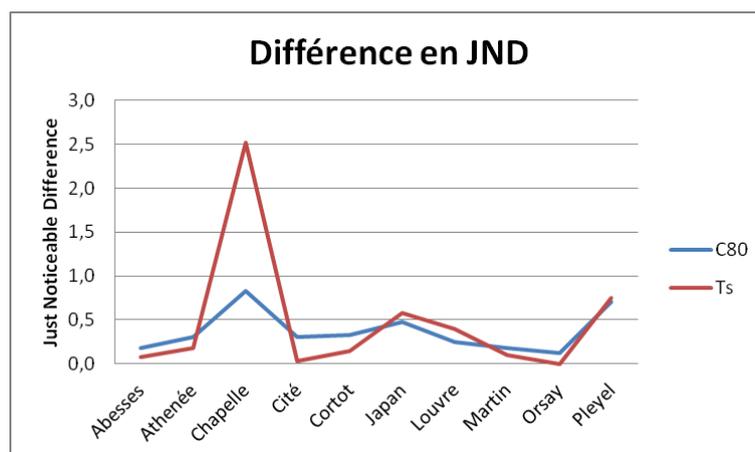


Figure 24. Différence en JND entre les salles réelles et les salles synthétisées pour les indices C80 et Ts.

Les indices C80 et Ts montrent aussi des différences en dessous de 1 JND, sauf dans le cas de la Chapelle Royale de Versailles qui a une valeur de 2.5 JND pour le Ts. Il faut remarquer que les valeurs de JND recommandées par le standard ISO 3382-1-2009 ont été évaluées pour les auditoriums. Martelotta [Mar10] a signalé que, dans les salles présentant un long temps de réverbération, le seuil différentiel de l'indice Ts est dépendant du temps de réverbération. Dans ces salles, le JND du Ts peut être autour de 8.5% de la valeur de base du Ts.

Dans le cas de la Chapelle Royale (avec un T30 d'environ 6 secondes), la valeur de base pour le Ts est 410ms. Par conséquent, le JND est autour de 34ms. C'est-à-dire, le JND est 3,4 fois plus grande que le JND recommandé pour les auditoriums et en conséquence la différence de la valeur Ts de la Chapelle se trouve aussi au dessous du seuil différentiel. Ainsi, de la même façon que pour la *réverbération*, théoriquement il n'y a pas de différences perceptibles entre salles réelles et salles synthétisées quant à la *précision*.

Quant à l'indice LF, la figure 25 montre une grande variabilité du JND avec la salle. Le fait d'avoir quelques salles dont la différence est au-dessus de 1 JND, nous indique que la *largeur apparente de la source (ASW)* n'est pas bien simulée dans les salles synthétisées.

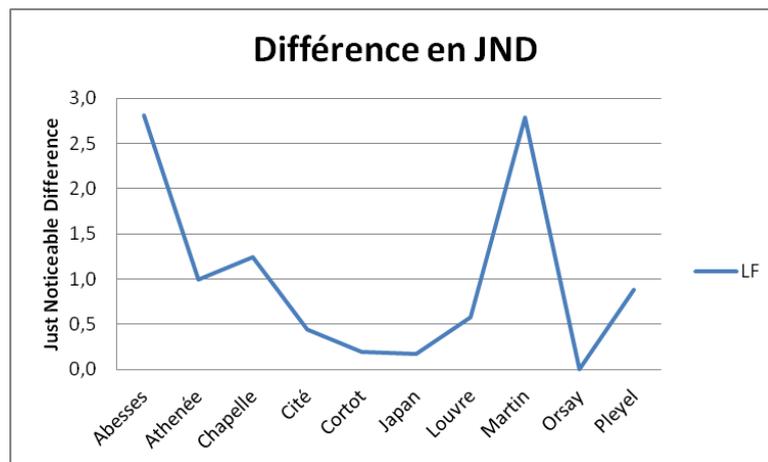


Figure 25. Différence en JND entre les salles réelles et les salles synthétisées pour l'indice LF

Les analyses de l'indice G ont été faites en ajustant le gain des mesures G(C) et G(CS)²³ pour s'approcher du niveau qui a été mesuré par la méthode direct avec le sonomètre dans les salles réelles. De cette façon, on peut avoir des différences de l'indice G au dessous de 1 JND comme illustre la figure 25. Ainsi, à priori il n'y a pas de différences perceptibles entre salles réelles et salles synthétisées quant au *niveau sonore*.

²³ G(C) et G(CS) sont les valeurs de l'indice G calculés à partir des réponses impulsionnelles de la salle réelle et synthétisée respectivement et les fichiers de référence comme expliqué en la section 4.1

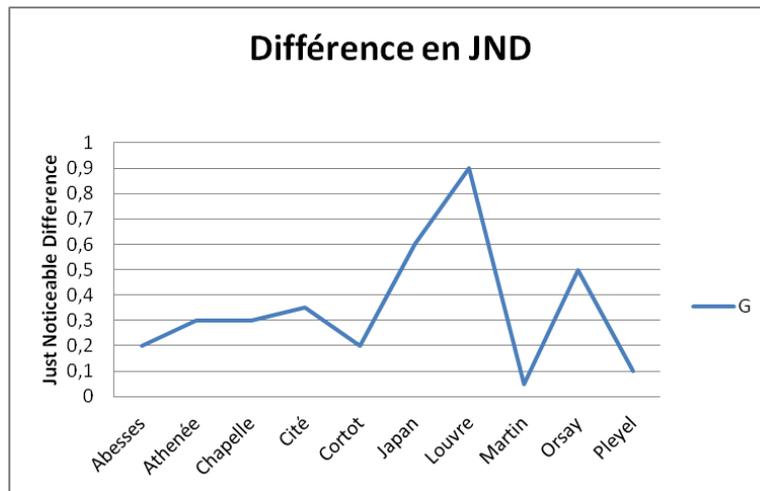


Figure 26. Différence en JND entre les salles réelles et les salles synthétisées pour l'indice G.

Il est importante de remarquer que ces analyses ont été réalisées avec la fréquence moyenne, comme recommandé par la norme ISO 3382-1-2009. Néanmoins, Bradley [Bra10] a signalé qu'il n'y pas beaucoup de justification à soutenir que les valeurs à la fréquence moyenne soient les plus représentatives des changements perceptifs pour tous les indices.

En général, les différences des valeurs pour les indices T30, EDT, C80 et Ts restent en dessous de 1 JND pour toutes les octaves entre 125Hz à 4kHz. Par contre, si on regarde les résultats de G dans toute la plage des fréquences, les différences ne sont pas négligeables. Dans le cas de l'indice LF, les différences sont plus grandes, surtout en hautes fréquences. Les figures 27 et 28 illustrent ces deux cas.

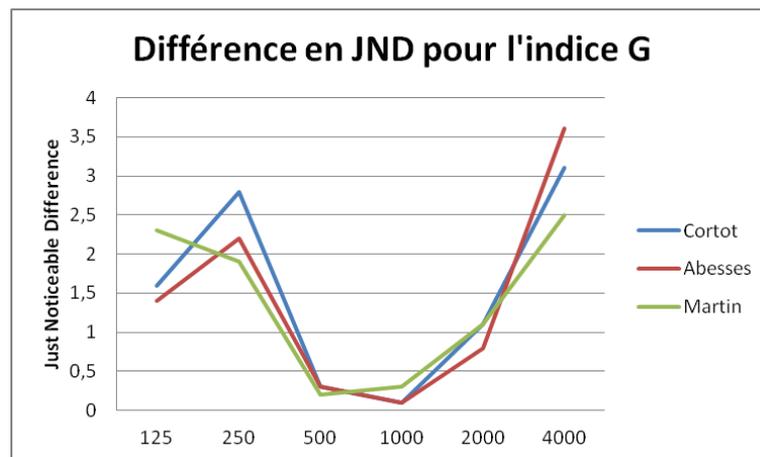


Figure 27. Différence en JND entre les salles réelles et les salles synthétisées pour l'indice G en fonction de la fréquence.

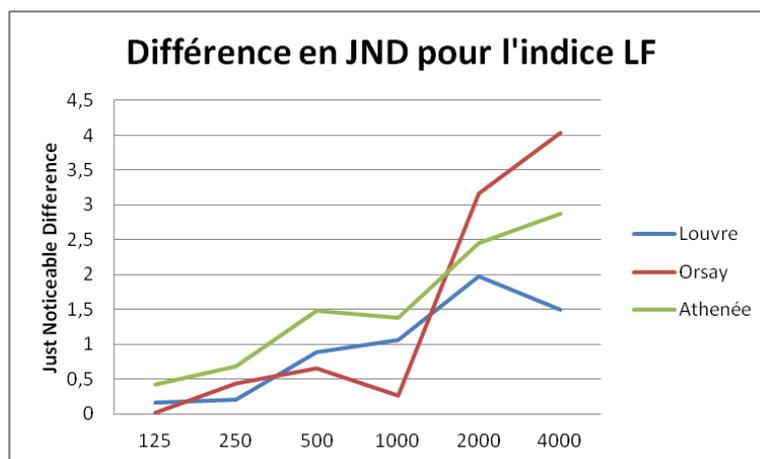


Figure 28. Différence en JND entre les salles réelles et les salles synthétisées pour l'indice LF en fonction de la fréquence.

5.3 Test Subjectif

Les analyses faites avec le critère du seuil différentiel ont montré que les indices acoustiques T30, EDT, C80 et Ts sont bien récréés dans les salles synthétisées. On a constaté avec l'analyse en composantes principales que ces indices représentent 65.87% de la variance des données analysées (cf. 5.1). L'objectif de cette section est d'évaluer les similarités perceptives entre la simulation des champs sonores réels (salles réelles) et celle des champs sonores synthétisés (salles synthétisées).

Pour cela, un test de catégorisation a été fait à partir du protocole subjectif conçu par Jean-Dominique Polack et Fabio Figueiredo [PF12, Fig11]. L'expérience a été réalisée dans la salle sèche du LAM en utilisant le système d'auralisation Ambisonics (cf. 3.2 et 4.2). A différence du test de Fabio, cette procédure a été faite en utilisant les 10 RI des salles réelles ainsi que les 10 RI des salles synthétisées²⁴, c'est-à-dire qu'il a été réalisé la simulation des 10 salles réelles et la simulation de la « simulation » des ces salles.

L'application développée en MAX/MSP pendant la thèse de Figueiredo pour les tests subjectifs a été réutilisée pour cette expérience. Cette application permet de faire l'auralisation Ambisonics à partir des fichiers de 13 canaux qui ont été enregistrés auparavant (cf. 4.2). La source utilisée pendant l'auralisation a été un morceau anéchoïque²⁵ de l'œuvre musicale d'Haendel *Water Music*. Au total, 16 sujets ont réalisé le test d'écoute. Les participants ne savaient pas que la moitié des salles qu'ils devaient écouter constituait la simulation des champs sonores synthétisés. La question posée a été la suivante :

²⁴ Celles qui correspondent à la mesure de la RI des salles réelles choisies pour le test de Catégorisation (cf. section 4.3)

²⁵ Enregistrements anéchoïques de la collection Denon. [Fig11]

« Écoutez les séquences A à T et groupez ensemble des séquences sonores qui se ressemblent. Il s'agit des enregistrements d'un même morceau de 30s, joué dans différentes salles. Vous pouvez les écouter autant de fois que vous voulez et faire autant de groupes que vous voulez. Indiquez vos groupes et caractérisez chaque groupe avec quelques mots »

Ainsi, les données à étudier ont été l'ensemble des groupes de salles produits par les sujets. A partir de cet ensemble, la matrice de distances ou dissimilarités a été construite (cf. A.3). Les analyses ont été faites en utilisant la méthode des arbres de similarité additifs de Sattath et Tversky [Poi02].

Les représentations arborées permettent d'identifier des relations entre objets. Les feuilles de l'arbre sont les objets (dans notre cas les salles analysées). Les arêtes indiquent la distance entre deux objets. La pertinence et la qualité de la représentation arborée s'évaluent à partir d'un certain nombre d'indices ou de critères qui peuvent être métriques ou topologiques. On a utilisé le Stress de Kruskal²⁶ et le taux de quadruples bien représentés²⁷ comme critère métrique et topologique respectivement. Ainsi, plus le Stress de Kruskal s'approche à zéro, mieux les distances sont reproduites. Une valeur de ce critère inférieure à 0,05 révèle une bonne qualité de représentation. En revanche, plus le taux de quadruples bien représentés est élevé, plus la représentation est considérée comme fiable.

La représentation graphique des résultats a été faite à l'aide du logiciel *AddTree*²⁸. D'abord, l'arbre des sujets (cf. figure 29) a été analysé. Si le Stress de Kruskal est inférieur à 10% (0,0764), le taux de quadruplets bien représentés n'atteint que 66.76%. C'est à dire la représentation arborée est faible.

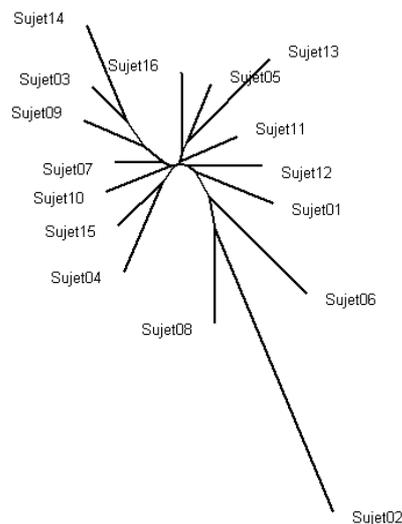


Figure 29. Arbre des sujets initial

²⁶ Le stress de Kruskal mesure l'écart entre la matrice initiale des distances (entre les objets) et la matrice de distances ou dissimilarité observée (matrice de la représentation produite).

²⁷ Le taux de quadruples bien représentés mesure le pourcentage de quadruples qui présentent la même topologie dans la matrice initiale des distances (entre les objets) et la matrice de distances de l'arbre obtenu.

²⁸ Version de Barthélemy et Guénoche, 1991

Les indices de la qualité de représentation de cet arbre sont 0,088 pour le Stress de Kruskal et 82% pour le taux de quadruplets bien représentés. En se référant aux études de simulation effectuées par Guénoche et Garreta (2001), les valeurs obtenues permettent de considérer que ce modèle d'arbre apparaît raisonnablement ajusté aux données [Poi02]. On établit comme « groupements fiables », ceux qui s'opposent aux autres objets de la topologie par une arête interne supportée par un taux de quadruplets supérieur et égal à 0.90 [DR02].

Ainsi, trois « groupements fiables » sont identifiés :

- CHPr et CHPs correspondant à la Chapelle Royale. Quant à la verbalisation on trouve: « *très réverbérant* », « *salle grande* », « *beaucoup de réverbération* » et « *son d'église* ».
- CORr et CORs correspondant à la salle Cortot. Quant à la verbalisation on trouve: « *salle clair* », « *Absences de graves* », « *froid* », « *Basses fréquences faibles* » et « *sec* ».
- MARr et MARs correspondant à la Porte St. Martin. Quant à la verbalisation on trouve: « *lointain* », « *sans relief quoique net* », « *distant* », « *étouffé* », « *un peu voilé* », « *l'orchestre joue très éloigné de l'auditeur* »

Ainsi, pour les trois salles précédentes, on observe que les sujets ont été capables de percevoir des caractéristiques similaires entre la simulation du champ réel et synthétisé. En revanche, les autres salles ont été perçues différemment et avec « des groupements non fiables ». C'est-à-dire que, pour une salle, la plupart des sujets n'ont pas groupé dans une même catégorie le champ réel et le champ synthétisé. Néanmoins, au-delà de la distance entre les objets, on voit aussi qu'une partie importante des sujets a perçu les salles ATHr et ATHs plus similaires entre elles qu'avec une autre salle.

D'ailleurs, LOUr et ABEs ont été groupées avec un taux de quadruplets égal à 0.84, et ABEr et ORSr avec un taux de quadruplets égal à 0.78. On constate que ce trinôme (Abbesses, Louvre et Orsay) a été aussi groupé en [Liu11] comme un groupement fiable. De même, ce groupe correspond à la classe A de *clustering* dans les analyses développées en [PFL12, Fig11].

Dans l'étude réalisée par Liu, les salles CHP, CIT et PLE ont été groupées dans une même catégorie, de la même façon que les salles COR et JAP. On voit que l'utilisation des salles synthétisées dans notre test a permis aux sujets de classer facilement la Chapelle pour son temps de réverbération et la salle Cortot pour sa couleur particulière. En conséquence, les salles Pleyel, Cité de la Musique et la Maison de la Culture de Japon n'ont pas été groupées dans une catégorie définie, contrairement à ce que Liu a trouvé.

Une autre analyse importante consiste à chercher les indices acoustiques implicitement utilisés par les sujets pour cette catégorisation subjective. Plus concrètement, on se demande si les sujets tiennent compte de l'indice LF car celui-ci a marqué la plus grande variabilité entre salles réelles et synthétisées. Pour ce faire, il faudrait appliquer la méthode MDS³⁰ (Multidimensional Scaling) à la matrice de dissimilarités (cf. A.4) pour obtenir l'espace de représentation. Afin de tester si les dimensions de cet espace peuvent être expliquées par les indices objectifs, une régression multiple doit être faite entre chaque indice objectif et ces dimensions. Faute de temps, cette dernière analyse n'a pas été faite.

5.4 Réponses impulsionnelles des salles réelles et synthétisées

Bien que les indices acoustiques des champs sonores synthétisés soient bien récréés dans la plupart des salles (comme montrent les analyses de seuil différentiel), il a été constaté dans le test subjectif que les sujets n'ont pas perçu une ressemblance entre le champ réel et le champ synthétisé. Une explication peut être trouvée en observant les réflectogrammes³¹ réelles et synthétisées de chaque salle. La figure 31 illustre les deux réflectogrammes correspondant à la Cité de la Musique. L'annexe A.5 montre d'autres exemples.

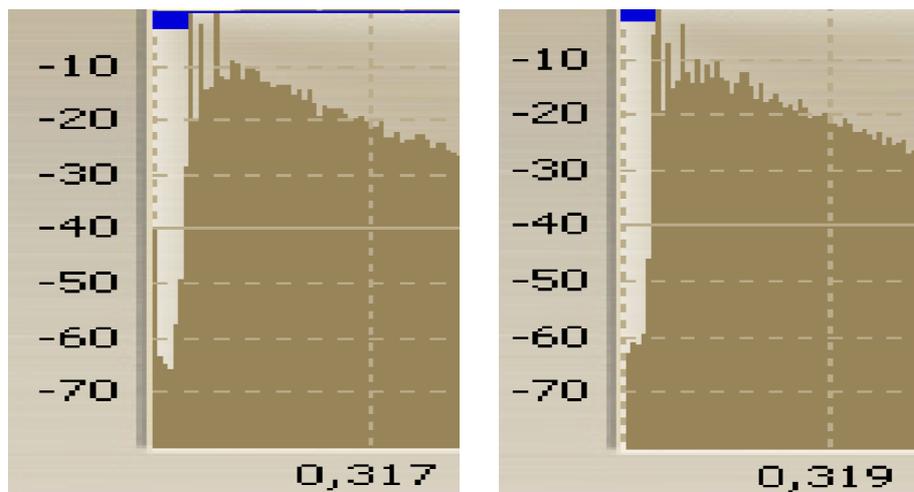


Figure 31. Réflectogrammes correspondant à la Cité de la Musique.
Champ sonore réel (à gauche) et synthétisé (à droite)

On peut voir que les courbes varient significativement dans les premières dizaines de millisecondes, c'est-à-dire pour les premières réflexions. Cela est plus évidente si on observe les réflectogrammes correspondant à la composante Ambisonics Y (composante latérale), ce qu'on peut relier à une faible restitution de l'indice LF comme indiqué par avant.

³⁰ Les méthodes MDS (Kruskal, 1964) permettent de déterminer un espace de représentation de faible dimension dans lequel les distances entre objets reflètent au mieux les dissimilarités originales.

³¹ Le réflectogramme est la représentation logarithmique de la réponse impulsionnelle.

6. Conclusion

La mesure de la réponse impulsionnelle de la simulation d'un ensemble de salles réelles a été faite dans la salle d'écoute du LAM, en reprenant la méthode utilisée par Fabio Figueiredo au cours de sa thèse. Ainsi, la caractérisation des champs sonores simulés a été réalisée à partir des indices acoustiques calculés de ces réponses impulsionnelles. Les analyses montrent qu'à l'exception de T30 et EDT, la dispersion des indices acoustiques diffère entre le champ sonore réel et le champ sonore synthétisé. Pour l'indice Ts, cette disparité est beaucoup plus faible que pour l'indice C80. La différence la plus remarquable est celle de l'indice LF.

L'analyse en composantes principales (ACP) réalisée avec les valeurs des indices des champs sonores synthétisés fournit les mêmes résultats que ceux obtenus par Liu et Figueiredo avec les valeurs des indices des champs sonores réels, à savoir que les indices T30, EDT, Ts et C80 sont bien corrélés avec la première composante et représentent la partie la plus importante de la variance de l'ensemble de données. De la même manière, les indices LF et G sont corrélés avec la deuxième et troisième composante.

D'ailleurs, le critère du seuil différentiel a montré que les indices acoustiques T30, EDT, C80, Ts et G sont reproduits correctement par le système de simulation sonore de la salle d'écoute. Néanmoins, les champs sonores simulés de quelques salles ne reproduisent pas bien les variations du LF. C'est-à-dire que le champ sonore synthétisé n'arrive pas à donner les mêmes caractéristiques physiques que le champ sonore réel quant au rapport entre l'énergie précoce des réflexions latérales (sans le son direct) et l'énergie précoce totale. Soit en termes perceptifs, le système ne donne pas bien l'impression de l'effet d'espace ou de la sensation d'être entouré par le son. Il faudrait travailler sur la restitution du champ direct car il semble que la technique utilisée pour cette restitution ait une influence importante sur la perception de l'effet d'espace [Gua03]

Aussi, le test d'écoute nous a permis d'observer les distances perceptives entre les champs sonores réels et simulés. Trois salles (Chapelle, Cortot, Martin) ont été groupées dans une seule catégorie, de cette façon la plupart des sujets ont perçu une grande ressemblance entre les deux champs. Cela n'a pas été le cas pour les autres salles. Il faudrait faire une analyse MDS pour vérifier si la catégorisation a été influencée par l'indice LF ou non.

Finalement et de la même manière que Liu, la catégorie Louvre-Abbesses-Orsay a été observée. Néanmoins, le fait de faire écouter la simulation des champs sonores réels et synthétisés en même temps a changé les autres groupements trouvés par Liu (qui a fait l'expérience seulement avec la simulation des champs sonores réels). Il faudrait faire l'expérience avec la simulation des champs sonores synthétisés indépendamment pour constater si les sujets établissent les mêmes catégories

Bibliographie

- [Bar93] M. Barron. *Auditorium Acoustics and Architectural Design*. E & FN SPON. 1993.
- [Ber96] L. Beranek. *Concert and Opera Hall. How they sound*. Acoustical Society of America, 1996.
- [Bra10] J.S Bradley. *Review of objective room acoustics measures and future need*. In Proceedings of the International Symposium on Room Acoustics, ISRA Melbourne, Australia, pp. 1-8. August 2010
- [Car98] A. Carrión. *Diseño acústico de espacios arquitectónicos*. Universitat Politècnica de Catalunya, 1998.
- [Dan00] J. Daniel. *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia*. Thèse de doctorat. Université Pierre et Marie Curie, CENT, Septembre 2000.
- [DDP08] G. Defrance, L. Daudet, J.D. Polack. *Finding the onset of a room impulse response: Straightforward?* In JASA Express Letters. September. 2008.
- [DR02] S. David, C Rouby. *Groupements d'odeurs ou catégories ? Questions de méthodologie*. Cahiers du LCPE. Décembre 2002.
- [Far00] A. Farina. *Simultaneous measurement of impulse response and distortion with a swept-sine technique*. In 108th Audio Eng. Soc. Convention, Paris, February 2000.
- [Far07] A. Farina. *Advancements in impulse response measurements by sine sweeps*. In 122th Audio Eng. Soc. Convention, Vienna, May 2007.
- [Fig11] F.L. Figueiredo. *Indices acoustiques et leurs rapports statistiques : vérification objective et subjective pour un ensemble de salles de spectacles*. Thèse de Doctorat. Université Pierre et Marie Curie. 2011.
- [Ger85] M. A. Gerzon. *Ambisonics in multichannel broadcasting and video*. Journal Audio Eng. Soc., 33(11): 859–871, November 1985.
- [Gri96] D. Griesinger. *Beyond MLS - Occupied Hall Measurement With FFT Techniques*. In 101th Audio Eng. Soc. Convention. November 1996.

- [Gua03] C. Guastavino. *Etude sémantique et acoustique de la perception des basses fréquences dans l'environnement urbain*. Thèse de Doctorat. Université Pierre et Marie Curie. 2003.
- [Iso09] ISO3382-1:2009. “*Measurement of room acoustic parameters. Part 1: Performance Spaces*”.
- [Liu11] S. Liu. *Analyse statistique des indices acoustiques des salles de concerts parisiennes*. Mémoire de Fin d'études. DEA de Statistique. Université Pierre et Marie Curie. 2011.
- [Mar10] F. Martellotta. *The just noticeable difference of center time and clarity index in large reverberant spaces*. In JASA. Vol. 128, No. 2, pp. 654-663. August 2010.
- [PF12] J.D Polack, F.L Figueiredo. *Room Acoustics auralization with Ambisonics*. In Annual IOA Meeting. Nantes. 2012
- [Poi02] J. Poitevineau. *Méthode des arbres de similarité additifs de Sattah et Tversky : Illustration dans une tâche de catégorisation de situations d'incertitude*. Cahiers du LCPE. Décembre 2002.
- [Pot06] K. Potter. *Visualization of Large and Unstructured Data Sets*. GI-Edition. Lecture Notes in Informatics (LNI). In Hans Hagen, Andreas Kerren, and Peter Dannenmann (Eds.). Vol. S-4, pp. 97-106, 2006
- [PFL12] J.D Polack, F.L Figueiredo, S. Liu. *Statistical analysis of a set of Parisian Concert Halls and Theatres*. In Annual IOA Meeting. Nantes. 2012
- [Sal09] S. Salaün-Bertet. *Formats audio 3d hiérarchiques : caractérisation objective et perceptive des systèmes Ambisonics d'ordres supérieurs*. Thèse de Doctorat. Institut National des Sciences Appliquées de Lyon, Mars 2009.
- [Sch79] M. R. Schroeder, “*Integrated-impulse method for measuring sound decay without using impulses,*” vol. 66, no. 2, pp. 497–500, 1979.
- [Shl09] J. Shlens. *A tutorial on Principal Component Analysis*. Center for Neural Science. New York University. In: <http://www.snl.salk.edu/~shlens/>. April 2009.
- [Tis09] A. Tisseyre. *Concerts Halls and Opera Houses Acoustic Design*. Edition : www.planete-acoustique.com. 2009

Annexes

A.1. Tables des indices acoustiques utilisés dans les analyses statistiques.

Pour calculer les boîtes à moustaches, la valeur moyenne de chaque indice a été prise pour toutes les salles sur les octaves entre 250Hz et 2000Hz

	T30 [s] - Temps de Réverbération					
	C	CS	G	GS	D	DS
abesses	1,13	1,14	1,14	1,12	1,11	1,11
athenee	1,07	1,09	1,16	1,18	1,15	1,18
chapelle	5,83	5,82	5,93	5,93	5,89	5,90
cite	1,83	1,83	1,88	1,88	1,88	1,87
cortot	1,17	1,16	1,18	1,19	1,19	1,17
japan	1,31	1,31	1,34	1,34	1,32	1,33
louvre	1,49	1,48	1,54	1,55	1,54	1,53
martin	0,99	1,00	1,10	1,13	1,13	1,15
orsay	1,16	1,17	1,15	1,17	1,15	1,17
pleyel	1,76	1,76	1,77	1,75	1,81	1,81

	EDT [s] - Temps de Décroissance Précoce					
	C	CS	G	GS	D	DS
abesses	1,01	0,98	0,97	0,95	1,08	1,07
athenee	0,64	0,62	0,91	0,94	0,82	0,81
chapelle	5,36	5,40	5,59	5,72	5,85	5,94
cite	1,78	1,78	1,72	1,76	1,81	1,81
cortot	1,17	1,16	1,18	1,19	1,19	1,17
japan	1,25	1,28	1,36	1,37	1,32	1,31
louvre	1,27	1,31	1,40	1,37	1,32	1,31
martin	0,58	0,57	1,17	1,14	1,17	1,19
orsay	1,02	1,06	1,02	1,03	1,07	1,08
pleyel	1,71	1,72	1,84	1,82	1,85	1,81

	C80 [dB] - Clarté					
	C	CS	G	GS	D	DS
abesses	3,63	4,25	3,55	3,88	3,15	3,20
athenee	7,28	7,68	5,58	5,48	6,23	6,03
chapelle	-6,55	-7,70	-5,65	-6,15	-5,33	-5,83
cite	-3,60	-4,60	-3,63	-3,93	-3,58	-3,35
cortot	2,93	3,43	2,23	2,58	2,18	2,38
japan	0,58	0,18	0,75	1,33	1,18	1,30
louvre	1,78	1,10	1,50	1,73	1,45	1,48
martin	7,75	7,68	2,80	2,95	2,98	2,80
orsay	2,93	2,83	2,28	2,00	2,00	1,40
pleyel	1,93	2,40	0,45	1,93	0,90	0,70

	Ts [ms] - Temps Centrale					
	C	CS	G	GS	D	DS
abesses	68,75	63,75	66,25	62,00	69,75	70,50
athenee	45,25	43,75	52,75	53,75	49,50	51,25
chapelle	377,25	394,75	385,50	398,50	391,75	406,75
cite	147,25	151,00	141,75	146,25	149,25	148,75
cortot	72,50	71,75	82,50	77,75	78,25	76,50
japan	97,50	100,75	95,00	89,00	95,25	94,25
louvre	91,50	97,50	93,50	89,00	95,50	94,50
martin	39,75	41,75	69,25	68,00	74,00	75,50
orsay	78,50	80,50	80,75	83,25	82,25	84,25
pleyel	88,75	85,75	108,50	94,50	107,00	108,00

	G [dB] - Facteur de Force		
	R	C	CS
abesses	10,3	9,4	9,3
athenee	6,5	7,2	7,6
chapelle	9,7	9,6	11,1
cite	4,5	5,1	6,9
cortot	10,2	9,3	10,2
japan	5,1	5,9	5,5
louvre	9,2	9,3	10,2
martin	3,3	3,5	3,8
orsay	11,3	10,6	11,9
pleyel	6,1	6,8	7,2

	LF - Fraction Latérale	
	C	CS
abesses	0,27	0,22
athenee	0,19	0,20
chapelle	0,26	0,24
cite	0,22	0,26
cortot	0,17	0,22
japan	0,19	0,23
louvre	0,18	0,20
martin	0,31	0,22
orsay	0,26	0,32
pleyel	0,19	0,21

L'ACP a été fait en prenant la valeur moyenne des indices acoustiques de chaque salle, correspondantes à la mesure de la RI en position source au centre.

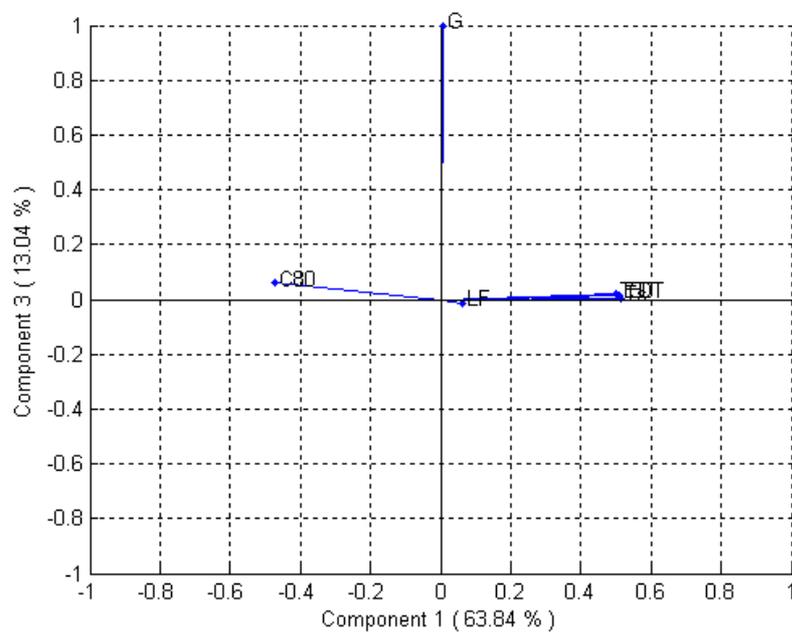
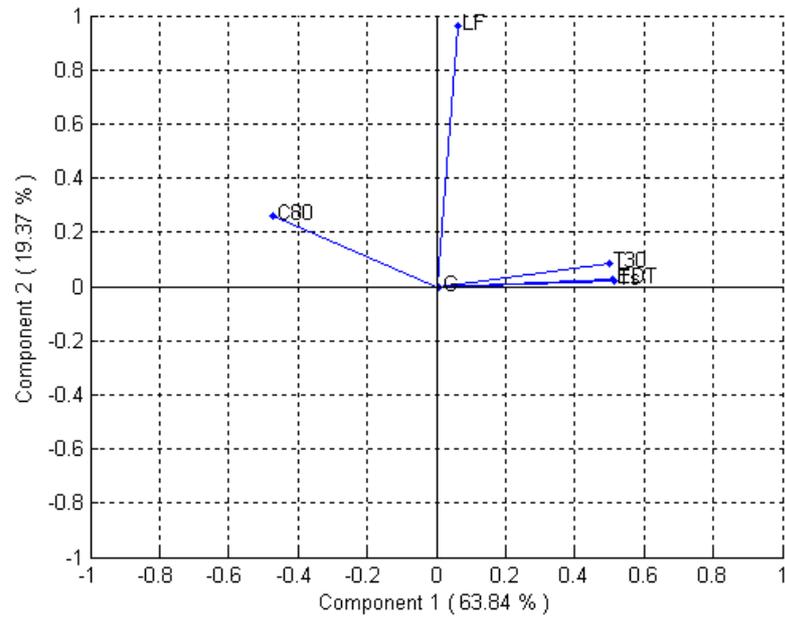
Champs sonores synthétisés (Salles synthétisées)						
	T30 [s]	EDT [s]	C80 [dB]	Ts [ms]	G [dB]	LF
Abesses	1,14	1,03	3,70	69,33	9,40	0,22
Athenée	1,16	0,79	6,42	54,00	6,55	0,20
Chapelle	5,36	4,81	-6,45	346,83	10,47	0,24
Cité	1,73	1,70	-3,77	143,83	6,30	0,26
Cortot	1,13	1,04	3,38	72,00	9,57	0,22
Japan	1,27	1,28	0,15	99,67	4,93	0,23
Louvre	1,46	1,21	1,18	93,83	9,72	0,20
Martin	1,09	0,65	7,12	46,83	3,02	0,22
Orsay	1,23	1,13	1,95	88,17	11,35	0,32
Pleyel	1,71	1,68	1,80	92,17	6,97	0,21

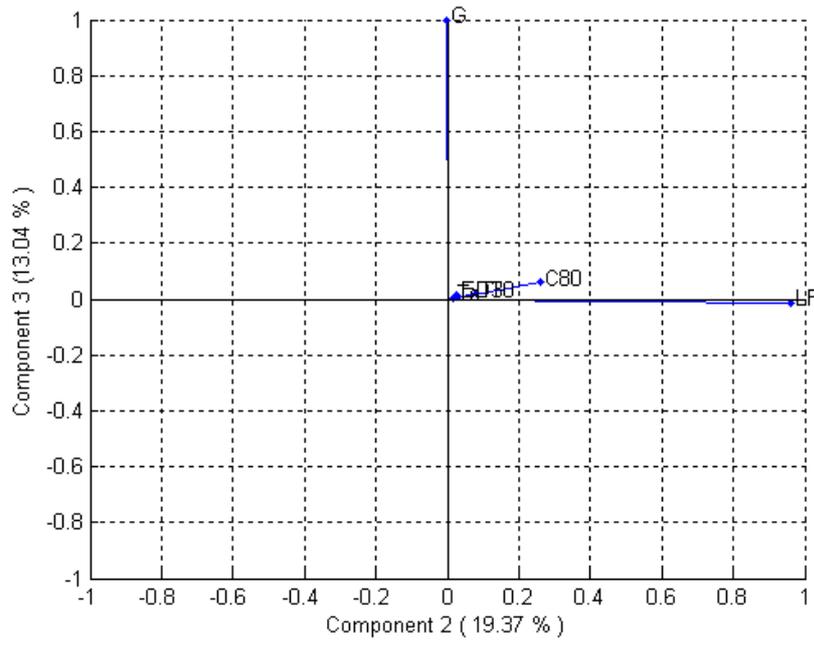
Champs sonores réels (Salles réelles)						
	T30 [s]	EDT [s]	C80 [dB]	Ts [ms]	G [dB]	LF
Abesses	1,15	1,06	3,40	72,17	9,90	0,27
Athenée	1,16	0,79	5,92	56,00	6,52	0,19
Chapelle	5,42	4,79	-5,53	333,17	9,47	0,26
Cité	1,73	1,71	-2,97	140,67	4,73	0,22
Cortot	1,16	1,03	3,20	71,67	9,07	0,17
Japan	1,28	1,25	0,55	96,00	5,47	0,19
Louvre	1,47	1,21	1,57	90,67	8,88	0,18
Martin	1,08	0,67	7,02	45,33	2,73	0,31
Orsay	1,23	1,11	2,02	86,67	10,17	0,26
Pleyel	1,71	1,70	1,58	93,83	6,30	0,19

Champs sonores synthétisés (Salles synthétisées)								
	T30 [s]	EDT [s]	C80 [dB]	Ts [ms]	D50 [%]	G [dB]	LF	I
Abesses	1,14	1,03	3,70	69,33	55,05	9,40	0,22	8,27
Athenée	1,16	0,79	6,42	54,00	60,76	6,55	0,20	5,65
Chapelle	5,36	4,81	-6,45	346,83	20,15	10,47	0,24	1,95
Cité	1,73	1,70	-3,77	143,83	19,07	6,30	0,26	3,65
Cortot	1,13	1,04	3,38	72,00	54,23	9,57	0,22	8,50
Japan	1,27	1,28	0,15	99,67	40,92	4,93	0,23	3,87
Louvre	1,46	1,21	1,18	93,83	49,65	9,72	0,20	6,66
Martin	1,09	0,65	7,12	46,83	58,91	3,02	0,22	2,76
Orsay	1,23	1,13	1,95	88,17	37,31	11,35	0,32	9,23
Pleyel	1,71	1,68	1,80	92,17	43,90	6,97	0,21	4,07

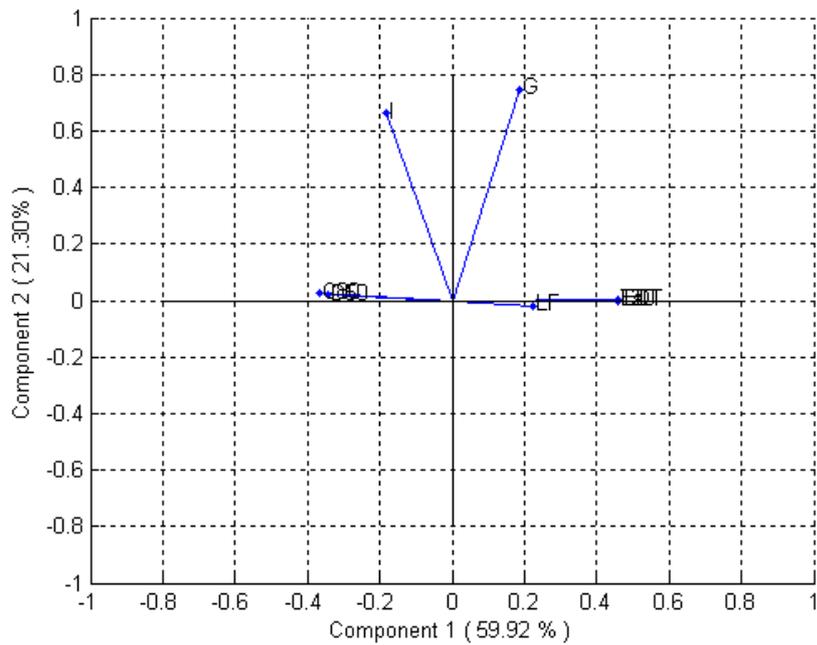
A.2. Résultats de l'ACP

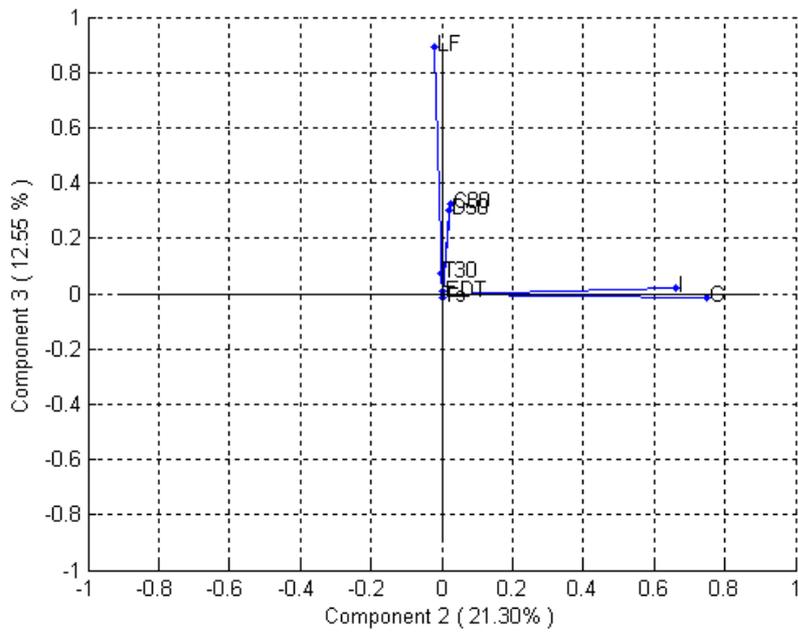
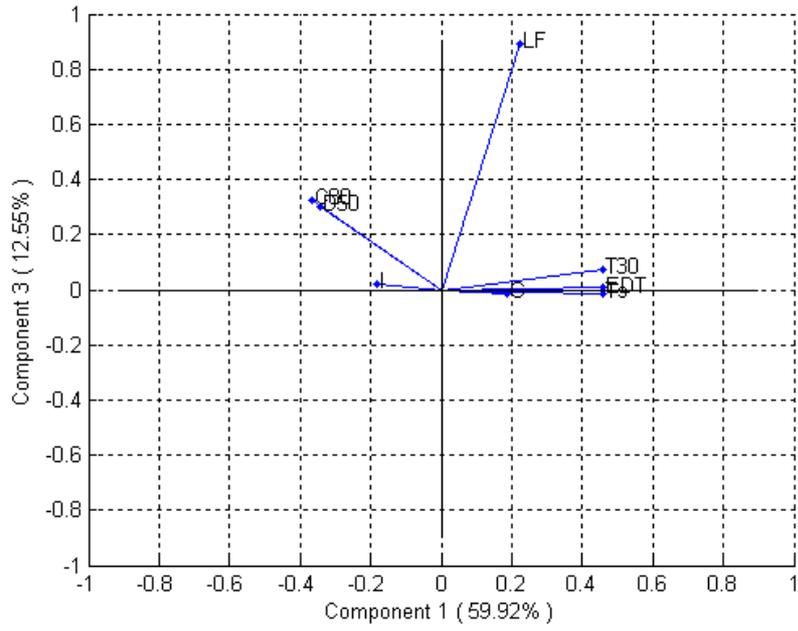
Analyse en composantes principales pour l'ensemble de 6 indices acoustiques des salles réelles.





Analyse en composantes principales pour l'ensemble des 8 indices acoustiques des salles synthétisées.

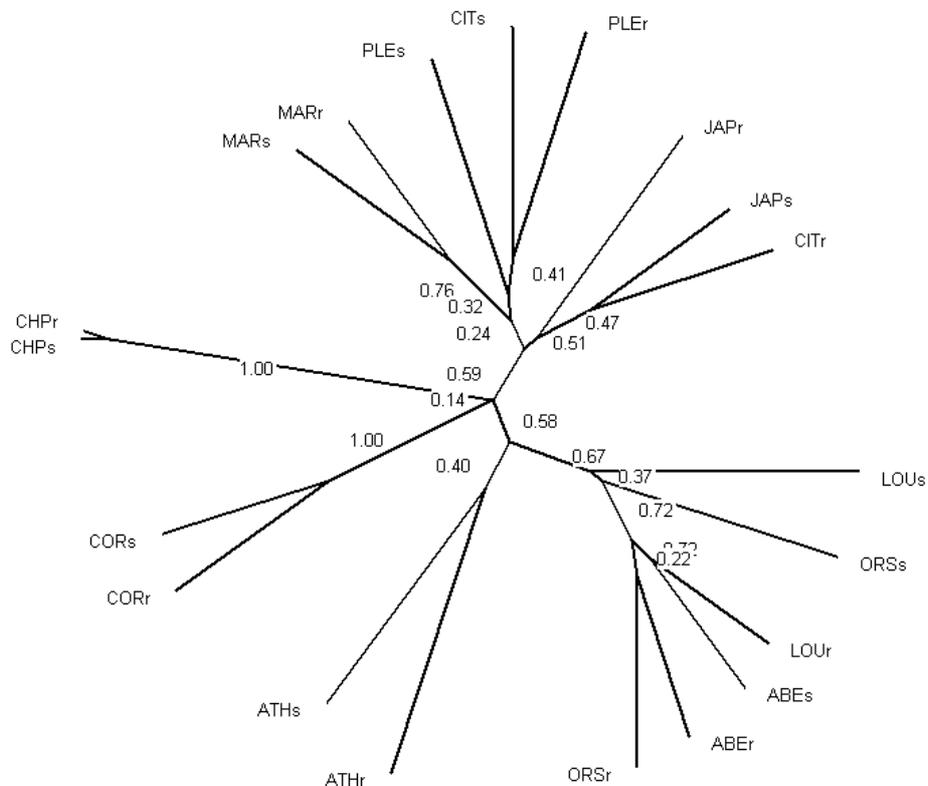




A.3 Matrice de distance entre objets.

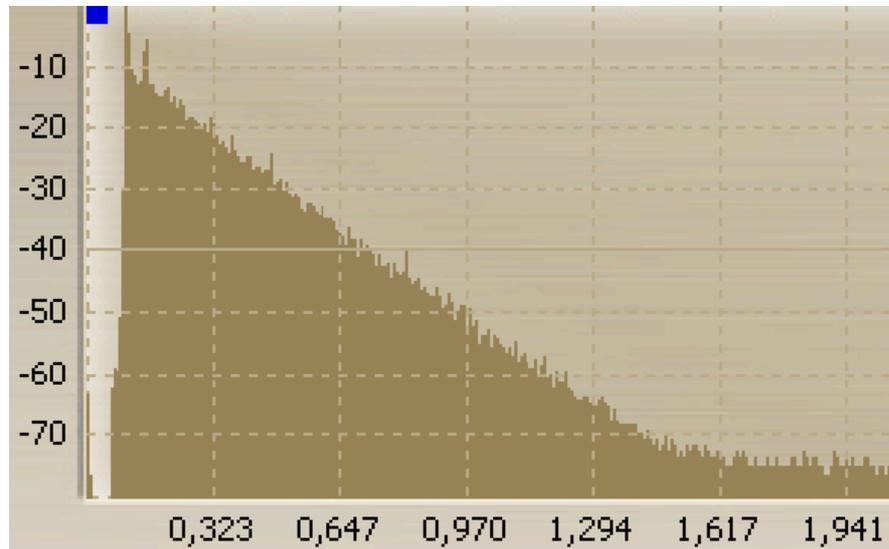
	ABEr	ATHr	CHPs	CITs	CORr	JAPr	LOUs	MARs	ORSr	PLEr	PLEs	ORSs	MARr	LOUr	JAPs	CORs	CITr	CHPr	ATHs	ABEs
ABEr	0	56,25	100	93,75	87,5	100	81,25	100	43,75	87,5	87,5	75	93,75	43,75	93,75	100	93,75	100	93,75	43,75
ATHr	56,25	0	100	87,5	87,5	93,75	81,25	100	87,5	93,75	81,25	87,5	87,5	62,5	93,75	93,75	100	100	68,75	56,25
CHPs	100	100	0	87,5	93,75	87,5	87,5	93,75	100	93,75	100	100	100	100	81,25	93,75	75	6,25	93,75	100
CITs	93,75	87,5	87,5	0	81,25	75	93,75	81,25	93,75	56,25	56,25	81,25	68,75	93,75	81,25	81,25	75	87,5	93,75	93,75
CORr	87,5	87,5	93,75	81,25	0	87,5	93,75	93,75	81,25	100	93,75	93,75	93,75	81,25	81,25	43,75	87,5	93,75	93,75	87,5
JAPr	100	93,75	87,5	75	87,5	0	87,5	62,5	93,75	62,5	81,25	93,75	68,75	87,5	68,75	81,25	50	87,5	87,5	100
LOUs	81,25	81,25	87,5	93,75	93,75	87,5	0	87,5	81,25	100	93,75	62,5	87,5	62,5	87,5	87,5	87,5	93,75	87,5	50
MARs	100	100	93,75	81,25	93,75	62,5	87,5	0	100	62,5	75	93,75	43,75	100	56,25	68,75	62,5	93,75	68,75	100
ORSr	43,75	87,5	100	93,75	81,25	93,75	81,25	100	0	87,5	93,75	56,25	93,75	43,75	100	100	100	100	100	62,5
PLEr	87,5	93,75	93,75	56,25	100	62,5	100	62,5	87,5	0	68,75	87,5	68,75	87,5	81,25	81,25	81,25	93,75	93,75	93,75
PLEs	87,5	81,25	100	56,25	93,75	81,25	93,75	75	93,75	68,75	0	75	50	87,5	50	81,25	75	100	75	93,75
ORSs	75	87,5	100	81,25	93,75	93,75	62,5	93,75	56,25	87,5	75	0	87,5	56,25	87,5	93,75	100	100	87,5	62,5
MARr	93,75	87,5	100	68,75	93,75	68,75	87,5	43,75	93,75	68,75	50	87,5	0	93,75	62,5	87,5	68,75	100	75	93,75
LOUr	43,75	62,5	100	93,75	81,25	87,5	62,5	100	43,75	87,5	87,5	56,25	93,75	0	93,75	93,75	100	100	100	37,5
JAPs	93,75	93,75	81,25	81,25	81,25	68,75	87,5	56,25	100	81,25	50	87,5	62,5	93,75	0	81,25	43,75	81,25	87,5	93,75
CORs	100	93,75	93,75	81,25	43,75	81,25	87,5	68,75	100	81,25	81,25	93,75	87,5	93,75	81,25	0	81,25	93,75	81,25	100
CITr	93,75	100	75	75	87,5	50	87,5	62,5	100	81,25	75	100	68,75	100	43,75	81,25	0	68,75	93,75	100
CHPr	100	100	6,25	87,5	93,75	87,5	93,75	93,75	100	93,75	100	100	100	100	81,25	93,75	68,75	0	93,75	100
ATHs	93,75	68,75	93,75	93,75	93,75	87,5	87,5	68,75	100	93,75	75	87,5	75	100	87,5	81,25	93,75	93,75	0	93,75
ABEs	43,75	56,25	100	93,75	87,5	100	50	100	62,5	93,75	93,75	62,5	93,75	37,5	93,75	100	100	100	93,75	0

A.4. Arbre de objets en tenant compte le sujet 02.

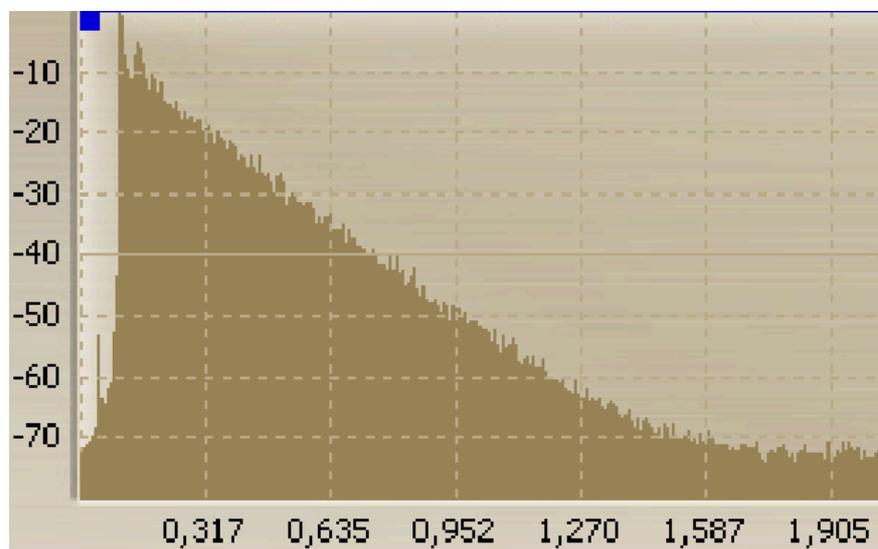


A.5. Réflectogrammes de différentes salles réelles et synthétisés

Réflectogramme - Maison de la Culture du Japon – Composante Ambisonics Y

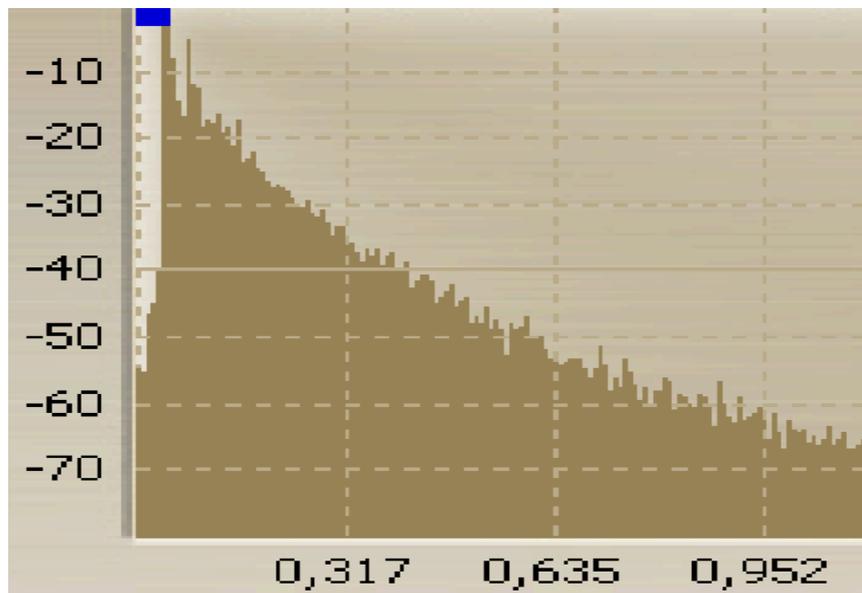


Champ sonore réel

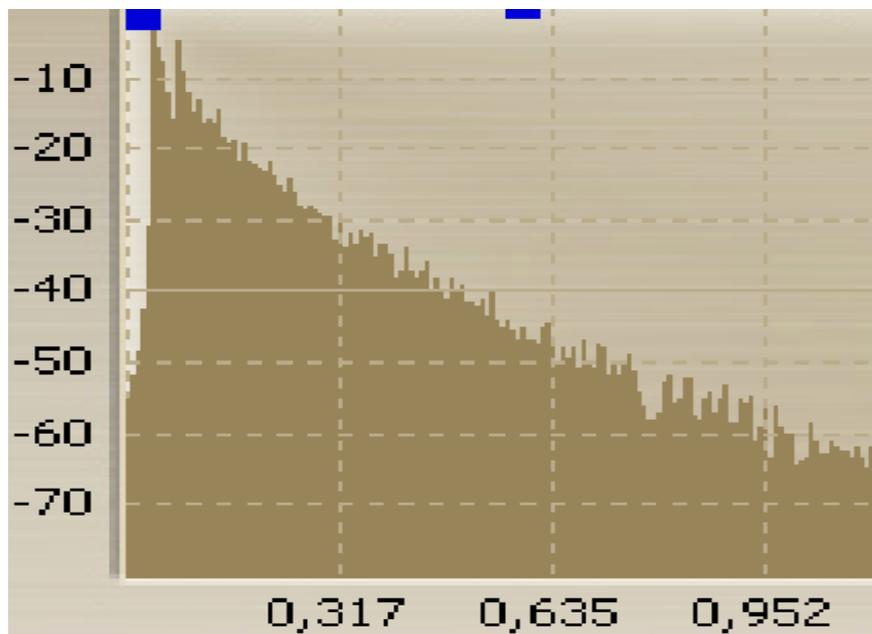


Champ sonore synthétisé

Réfectogramme – Théâtre de l'Athénée - Composante Ambisonics W



Champ sonore réel



Champ sonore synthétisé