



HAL
open science

Comparaison de méthodes d'égalisation pour une restitution en salle d'écoute

Adrien Vidal, Philippe Herzog, Christophe Lambourg

► **To cite this version:**

Adrien Vidal, Philippe Herzog, Christophe Lambourg. Comparaison de méthodes d'égalisation pour une restitution en salle d'écoute . Congrès Français d'Acoustique, Apr 2016, Le Mans, France. pp.201-207. hal-01313821

HAL Id: hal-01313821

<https://hal.science/hal-01313821>

Submitted on 10 May 2016

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

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

CFA/VISHNO 2016

Comparaison de méthodes d'égalisation pour une restitution en salle d'écoute

A. Vidal^a, P. Herzog^b et C. Lambourg^a

^aGENESIS, Domaine du Petit Arbois BP69, 13545 Aix En Provence Cedex 04, France

^bLMA - CNRS, 4 impasse Nikola Tesla CS 40006, 13453 Marseille Cedex 13, France
adrien.vidal@genesis.fr



LE MANS

Lors d'une diffusion sonore, la sensibilité à l'environnement acoustique est indéniable. Cette sensibilité se traduit par une coloration spectrale, qui dépend des propriétés de la salle d'écoute (dimensions, matériaux), mais également des caractéristiques de l'émetteur (directivité, position). Dans le contexte de restitution fidèle d'environnement sonore à partir d'un dispositif de haut-parleurs, cette coloration spectrale n'est pas souhaitée. Il est d'usage de la minimiser, en utilisant notamment une égalisation du système de diffusion. L'égalisation de ce type de systèmes peut être à l'origine d'artefacts audibles, notamment lorsque la position d'écoute n'est pas strictement la même que la position utilisée pour calculer les filtres d'égalisation. L'objectif de ces travaux est de déterminer le meilleur compromis perceptif entre une réponse égalisée et un minimum d'artefacts audible, pour différentes salles d'écoutes, différentes positions de sources et différentes méthodes d'égalisation. Pour cela, des mesures de réponses impulsionnelles ont été réalisées dans plusieurs salles. A partir de ces réponses impulsionnelles, des filtres d'égalisation ont été calculés en utilisant plusieurs méthodes utilisant des filtres à Réponse Impulsionnelle Finie et Infinie (filtres FIR et IIR). Par simulation, le signal est reconstitué au point récepteur ainsi qu'à un point légèrement déplacé en utilisant ou non une égalisation. Des tests perceptifs sont en cours, et nous présenterons les premiers résultats.

1 Introduction

Pour une restitution sonore avec des haut-parleurs, la salle d'écoute joue un rôle important. Son influence peut être souhaitée dans certains cas (en musique notamment), ou non désirée dans d'autres cas (simulation d'environnement sonore par exemple). Cette influence peut être contrôlée dans une certaine mesure en modifiant l'absorption au niveau des parois, et/ou par égalisation du système de diffusion.

L'influence de la salle introduit une certaine coloration spectrale du rendu, qui est dépendante de la position d'écoute. L'égalisation d'un système audio dans une salle d'écoute est un problème complexe, qui est généralement un compromis entre la finesse de l'égalisation et l'étendue de la zone de validité de l'égalisation.

L'égalisation de système audio est un domaine de recherche depuis plusieurs années, et plusieurs méthodes ont été proposées [1]–[4]. Nous proposons ici une évaluation perceptif de plusieurs méthodes d'égalisation, appliquées à cinq salles d'écoute différentes.

Il s'agit de la continuité de travaux visant à adapter un petit système d'écoute transaural à une salle usuelle, dans l'optique d'implémenter un système de restitution léger aux performances comparables aux systèmes plus complexes. Notamment, de récents travaux ont montré la pertinence d'utiliser des sources en champ proches de l'auditeur dans le cas particulier de restitution transaurale [5]. Basés sur des simulations, ces travaux nécessitent d'être approfondis par des tests perceptifs, initiés dans la présente étude. Ainsi, en plus de l'influence de la méthode d'égalisation et de la salle d'écoute, la distance entre la source et le récepteur sera un autre paramètre évalué ici.

L'approche que nous proposons consiste à étudier les dissemblances perceptives pour différentes configurations d'écoute. Nous présentons dans un premier temps les différentes salles d'écoutes évaluées ainsi que les conditions de mesures. Dans un second temps nous décrivons les méthodes d'égalisation employées dans cette étude, et enfin nous présentons le protocole de test d'écoute mis en place.

Dans la suite du document, nous appellerons « configuration » la combinaison d'une salle d'écoute et d'une distance de mesure.

2 Protocole de mesures

2.1 Salles évaluées

Les cinq salles choisies pour cette étude sont les suivantes :

- Salle de réunion rectangulaire de 19 m² (« Reu »)

Cette salle ne bénéficie d'aucun traitement acoustique. Ses dimensions sont d'environ 3.2 x 6 m.

- Studio d'écoute de 18 m² (« Stu »)

Cette salle est destinée à l'écoute, et bénéficie d'un traitement acoustique important : mousse absorbante alvéolé de 5 cm sur tous les murs, faux plafond en laine minérale et moquette au sol. Cette salle est la seule à ne pas avoir des dimensions symétrique. De plus, le plenum entre le faux plafond et le plafond communique avec d'autres locaux, ce qui peut favoriser la dispersion d'énergie. Sa surface est approximativement de 18 m².

- Bureau moyen carré de 16 m² (« BuM »)

Ce bureau n'a pas de traitement acoustique spécifique, à l'exception du faux plafond qui est identique au studio d'écoute de 18 m² : plenum communiquant avec d'autres locaux et faux plafond en laine minérale.

- Petit bureau carré de 7.5m² (« BuP »)

Ce bureau ne bénéficie d'aucun traitement acoustique. Ses dimensions sont d'environ 2.7 x 2.8 m.

- Cabine audiométrique de 4 m² (« Cab »)

Cette salle a les plus petites dimensions de notre panel de salles, et bénéficie d'un important traitement acoustique. Ses dimensions sont d'environ 1.8 x 2.3 m.

On notera que ces cinq salles sont de dimensions inférieures de celles préconisées pour l'écoute : l'ITU recommande en effet d'utiliser une salle d'écoute dont la surface est comprise entre 20 m² et 60 m² pour l'évaluation des systèmes de reproduction stéréophonique [6]. L'AES recommande également d'utiliser une salle d'écoute d'au moins 20 m² [7].

Les salles « Réunion », « Petit Bureau » et « Cabine » sont situées dans le même bâtiment, et les salles « Studio » et « Bureau moyen » sont situées dans un autre bâtiment

Pour chacune des salles, le temps de réverbération estimé pour les bandes d'octaves allant de 63 Hz à 8 kHz est représenté dans la Figure 1.

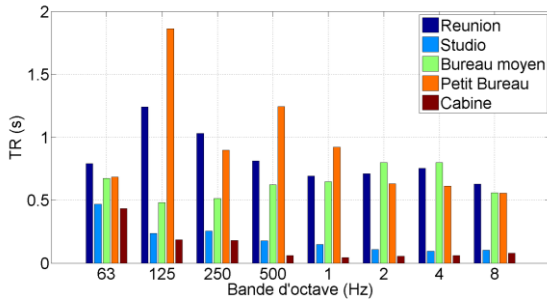


Figure 1 : Temps de réverbération des différentes salles d'écoutes

Les salles d'écoutes ont des comportements acoustiques différents : le studio et la cabine sont très mats : leur temps de réverbération est inférieur à 300 ms au-dessus de 63 Hz. La salle de réunion et le petit bureau sont les plus réverbérants, notamment en basses fréquences. Le bureau moyen a un gabarit de temps de réverbération assez atypique comparativement aux autres, il est plus élevé pour les hautes fréquences que les basses fréquences.

2.2 Positions mesurées

Pour toutes les salles, la source et le récepteur ont été positionnés sur la plus grande diagonale de la salle d'écoute, à une hauteur de 120 cm du sol. La source a été positionnée au tiers de cette diagonale, et le microphone à 40 cm et 80 cm de distance, sur cette même diagonale en direction de l'angle le plus éloigné. Pour la cabine, la mesure à 80 cm n'a pas été réalisée.

Pour toutes les configurations, une mesure a été systématiquement effectuée en décalant la position du microphone de 10 cm par rapport à sa position initiale, à distance à la source constante. A noter que l'orientation de la source n'est pas modifiée entre la mesure dans l'axe et décalée (la mesure décalée est donc influencée par la directivité de la source).

Toutes les mesures de réponses impulsionnelles sont réalisées en utilisant un sweep exponentiel entre 50 Hz et 15 kHz à la fréquence d'échantillonnage de 48 kHz. La réponse impulsionnelle associée à cette mesure est calculés par la méthode présentée dans [8].

Suite à des écoutes informelles, il semblerait que les configurations Cabine 40 cm et Studio 40 cm n'introduisent que peu de coloration du rendu. Cela signifie que la correction à apporter par égalisation est tenue pour ces cas-là.

2.3 Source

La source utilisée pour l'ensemble des mesures est une Tannoy system 600. Il s'agit d'une enceinte de monitoring coaxiale, pour laquelle nous avons bouché les événements. Des mesures de réponses en fréquence ont été réalisées dans la chambre anéchoïque du Laboratoire de Mécanique et d'Acoustique (Marseille). Elles sont représentées dans la Figure 2.

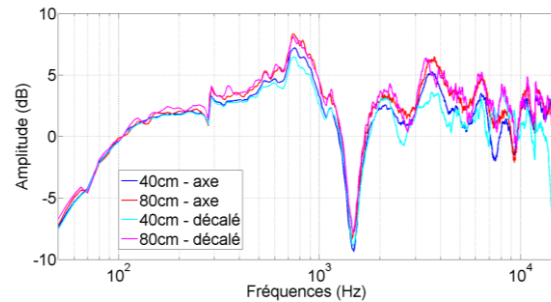


Figure 2: Réponses en Fréquences à 40 cm et 80 cm en champ libre, dans l'axe et en position décalée

On peut observer un accident significatif autour de 1500 Hz pour l'ensemble des mesures. Cet accident intervient autour de la fréquence de crossover du haut-parleur, et avait déjà été identifié dans de précédents travaux utilisant cette enceinte [9].

On constate que les réponses mesurées à 80 cm dans l'axe et décalée sont très proches. Les écarts entre les réponses mesurées à 40 cm dans l'axe et décalées sont plus importants, pouvant atteindre 2 dB entre 3 et 4 kHz.

Un filtre destiné à égaliser la réponse de la source en champ libre a été calculé. Ce filtre est un FIR d'ordre 4096 (pour $F_c = 48$ kHz) calculé à partir des mesures anéchoïques dans l'axe, dont la transformée de Fourier H est obtenue en suivant la méthode présentée dans [10] :

$$H = [C^*C + \beta]^{-1} \cdot C^*A$$

C est la réponse en fréquence de l'enceinte, β le paramètre de régularisation (dépendant de la fréquence, voir Figure 3), A la transformée de Fourier d'un retard pur de 2048 échantillons, et $*$ correspond à l'opération de transconjugaison. Le filtre de compensation de la source est obtenu par transformée de Fourier inverse de H . Dans notre cas, seule la composante à phase minimale du filtre est utilisée pour compenser l'effet de la source.

Deux filtres ont ainsi été calculés, l'un à partir de la réponse de l'enceinte dans l'axe à 40 cm, et l'autre à partir de la réponse de l'enceinte dans l'axe à 80 cm. Ces filtres ont été appliqués à l'ensemble des mesures correspondantes.

La Figure 3 représente les caractéristiques du filtre d'égalisation obtenu.

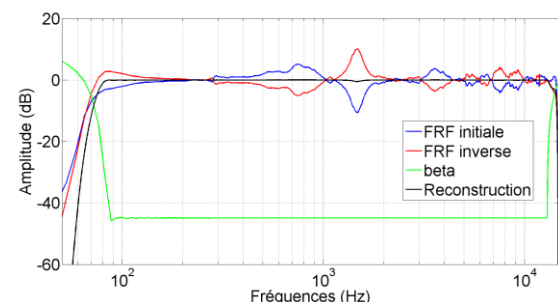


Figure 3 : FRF anéchoïque à 40 cm, FRF du filtre de déconvolution, paramètre de régularisation et reconstruction

La source étant compensée, les différences évaluées porteront uniquement sur les différences introduites par les salles d'écoutes. Dans le cas d'une égalisation d'un système complet, une partie du problème de l'égalisation est déjà résolu. La question dont nous cherchons la réponse est de savoir si l'égalisation de la salle représente un apport significatif.

3 Méthodes d'égalisation

Plusieurs méthodes d'égalisation ont été testées. La source ayant été égalisée par la méthode décrite ci-dessus, cette seconde étape d'égalisation vise uniquement à corriger l'effet de la salle. Pour cela, quatre méthodes différentes ont été évaluées. D'autres méthodes ont été implémentées, mais elles ont été éliminées lors d'écoutes informelles car elles introduisaient des artefacts significatifs.

3.1 Egalisation IIR

Cette méthode d'égalisation repose sur la publication de Ramos et Lopez [11]. Le principe consiste à mettre en cascade plusieurs filtres IIR d'ordre 2, où chacun d'entre eux vise à corriger un défaut particulier de la réponse fréquentielle à égaliser. Ces filtres IIR sont des filtres biquadratiques (« biquads » en anglais) de type « peak » ou « notch ».

La procédure est itérative, en corrigeant les défauts par ordre d'importance. Dans notre implémentation, lorsque le plus grand défaut observé est inférieur à un certain seuil, l'algorithme est arrêté. Le filtre d'égalisation est composé d'au plus 31 biquads dans notre implémentation.

Pour chaque défaut à corriger, les paramètres des biquads (fréquence centrale, amplitude et facteur de qualité) sont déterminés par un algorithme de Monte-Carlo.

Dans notre cas, la réponse fréquentielle à égaliser est préalablement lissée par bandes de $1/12^{\text{ème}}$ d'octave. Deux alternatives sont comparées ici, correspondant à deux valeurs du seuil d'arrêt : 0.1 et 0.5 dB octave (1 dB octave correspondant à un défaut de 1dB sur une bande d'octave).

Le nombre de filtres permettant d'atteindre cet objectif est un indicateur de la quantité de correction appliquée : plus le nombre de biquads est élevé, et plus le filtre d'égalisation est complexe. Le nombre de biquads utilisés en fonction de la configuration est représenté dans le Tableau 1.

Un exemple de réalisation de ce filtrage est présenté en Figure 4 avec un seuil fixé à 0.1 dB octave, et en Figure 5 avec un seuil fixé à 0.5 dB octave.

Tableau 1 : nombre de biquads utilisés

	0.5dB octave	0.1dB octave
Reu_40	7	20
Reu_80	9	31
Stu_40	4	25
Stu_80	6	31
BuM_40	5	31
BuM_80	11	31
BuP_40	11	31
BuP_80	17	24
Cab_40	4	29

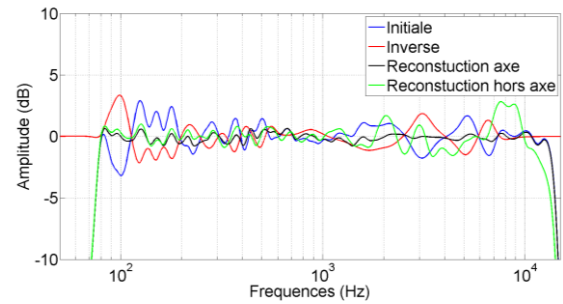


Figure 4 : filtre d'égalisation pour le bureau moyen avec la méthode IIR. Seuil à 0.1 dB octave

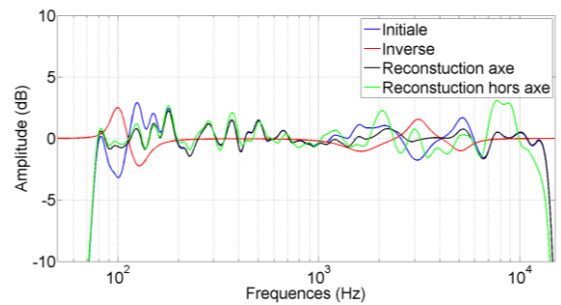


Figure 5 : filtre d'égalisation pour le bureau moyen à 40 cm avec la méthode IIR. Seuil à 0.5 dB octave

3.2 Egalisation FIR à partir d'un gabarit spectral

Dans un premier temps, les gains en bandes fines de la cible et de la FRF à égaliser sont lissés en $1/6^{\text{ème}}$ d'octave. Un gabarit fréquentiel de filtre inverse est déterminé à partir du rapport entre les gains lissés de la de cible et de la réponse à égaliser. On en déduit un filtre FIR inverse à phase linéaire en appliquant la méthode de la fenêtre [12], implémentée dans la fonction *fir2* de Matlab. La composante à phase minimale est ensuite extraite ([13], fonction *rceps* de Matlab). Le filtre FIR ainsi obtenu est d'ordre 2048 échantillons (soit 43 ms pour une fréquence d'échantillonnage de 48 kHz).

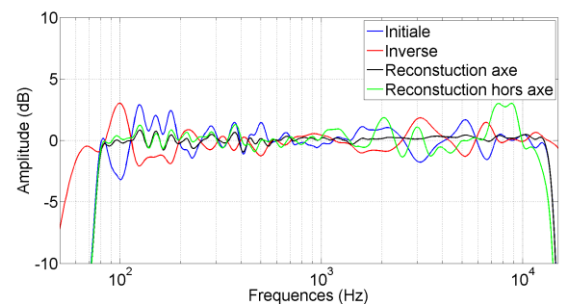


Figure 6 : filtre d'égalisation pour le bureau moyen à 40 cm avec la méthode FIR par gabarit.

3.3 Egalisation FIR temporel

Cette méthode d'égalisation est la seule qui vise à corriger la phase de la réponse impulsionnelle. En d'autres termes, c'est la seule méthode étudiée ici qui vise à corriger la réverbération. Cette méthode repose sur le calcul d'un filtre inverse avec régularisation de Tikhonov, comme explicité dans [1]. Cette méthode est analogue à celle présentée dans la section 2.3, mais implémentée dans le domaine temporel. Cela permet de s'affranchir de

problèmes de recouvrement temporel lié à l'hypothèse implicite de périodicité lors du calcul de transformée de Fourier. Le filtre inverse h est déterminé de la manière suivante :

$$h = [C^T C + \beta B^T B]^{-1} \cdot C^T a_m$$

Avec C la matrice de covariance de forme Toeplitz déterminée à partir de la réponse impulsionnelle à égaliser (préalablement tronquée et fenêtrée), β le paramètre de régularisation, et a_m un vecteur correspondant à un retard pur. La réponse impulsionnelle à égaliser est tronquée à la demi-longueur du filtre et la seconde moitié de cette réponse impulsionnelle est apodisée en utilisant une demi-fenêtre de Hann. Cette réponse impulsionnelle est complétée avec des échantillons nuls de telle sorte que sa durée soit la même que la longueur du filtre.

Dans chacun des cas, le retard introduit correspond à la demi-longueur du filtre. Compte tenu de la taille maximale des matrices pouvant être inversées sur notre plateforme (Matlab 32 bits sur un i7 avec 8 Go de RAM), l'ordre maximum des filtres qui peuvent être calculés par cette méthode est limité à 6000. Pour maximiser la durée de réponse impulsionnelle prise en compte dans l'égalisation, celle-ci a été sous-échantillonnée à 30 kHz avant le calcul du filtre.

La valeur du paramètre β ainsi que la taille du filtre correspondant à chacune des configurations sont données dans le Tableau 2. Ces paramètres ont été fixés par approches successives, dans le cadre d'écoutes informelles, en tentant de maximiser l'égalisation tout en limitant les artefacts de type pré-échos.

Tableau 2 : paramètres des filtres FIR

	β	Longueur de filtre (ms)
Reu_40	10^{-5}	200
Reu_80	10^{-2}	133
Stu_40	10^{-5}	200
Stu_80	10^{-3}	200
BuM_40	10^{-5}	200
BuM_80	$3 \cdot 10^{-2}$	200
BuP_40	10^{-5}	200
BuP_80	$5 \cdot 10^{-2}$	133
Cab_40	10^{-5}	200

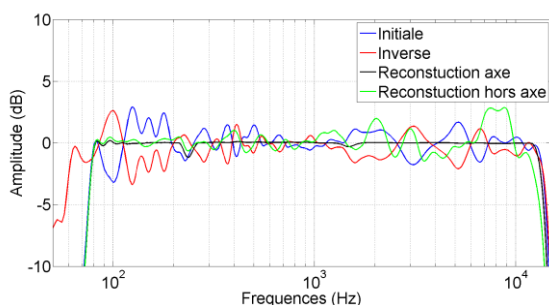


Figure 7 : filtre d'égalisation pour le bureau moyen à 40 cm avec la méthode FIR temporel.

Les méthodes que nous évaluons ici ont un support temporel qui n'est pas suffisant pour traiter la réponse de la salle, qui dans certains cas est extrêmement long. La source

sonore étant préalablement égalisée, les corrections que nous proposons sont très tenues.

4 Protocole de test

Une interface de test de type MUSHRA (Multiple Stimuli with Hidden Reference and Anchor) a été mise en place [14]. Il est demandé au sujet de juger la dissemblance entre une référence et différents signaux. Pour cela, le sujet dispose d'un curseur associé à chaque son à évaluer, à positionner entre « le plus proche » (note 100) et « le plus différent » (note 0). L'interface est représentée dans la Figure 8.



Figure 8 : exemple de réponses pour l'interface de test

La référence dans notre cas est le signal anéchoïque à 40 cm, pour lequel le filtre de déconvolution a été appliqué. Le filtrage de l'enceinte est donc exclu de l'évaluation réalisée ici.

Une référence cachée fait partie des signaux à évaluer, ainsi qu'une ancre basse et une ancre moyenne.

L'ancre basse correspond au signal non égalisé dans les conditions les plus réverbérantes : le petit bureau à 80 cm. Pour cette ancre, le récepteur est décalé de 20 cm par rapport à la configuration dans l'axe.

L'ancre moyenne correspond au signal non égalisé dans le bureau moyen à 80 cm de distance.

Les signaux à comparer correspondent aux 4 égalisations, pour laquelle la reconstruction a été effectuée au point initial, ainsi qu'au point décalé. Dans ce test, les sujets étaient contraints d'utiliser la dynamique complète de l'échelle de notation.

Le stimulus choisi est une salve de quatre bouffées de bruit rose. Chacune des bouffées dure 200 ms, et sont séparées de 200 ms de silence. Ce type de stimulus a été comparé à différents extraits d'enregistrement musicaux et vocaux dans le cadre d'écoutes informelles, et s'est avéré être le plus discriminant pour différencier les méthodes d'égalisation.

Il y a au total 12 stimuli à comparer : la référence cachée, les ancrs basse et moyenne, le signal non égalisé et les 4 égalisations reconstruite dans l'axe et décalées. La Figure 9 représente un schéma résumant la confection des différents stimuli. Au cours du test, un stimulus considéré comme manquant a été rajouté au panel. Il s'agit du signal décalé et non égalisé. L'analyse des résultats sera effectuée en considérant l'ensemble des sujets, en excluant ce stimulus ajouté. Nous considérons que la présence ou l'absence de ce stimulus n'influe que peu la réponse des sujets.

Le test d'écoute est divisé par configuration : le sujet n'évalue qu'une seule salle et distance à la fois. Nous évaluons 9 configurations, le sujet effectue donc 9 tests.

Tous les stimuli sont égalisés en niveau en s'appuyant sur estimation de la sonie définie dans la norme ISO 532B

[15]. Les signaux sonores sont diffusés à l'aide d'un casque Beyerdynamic DT990 Pro et d'une carte audionumérique ESI U24 XL.

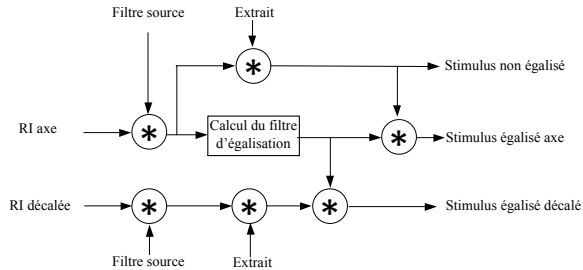


Figure 9 : Schéma de la confection des stimuli. * représente l'opération de convolution

Pour ce test préliminaire, notre objectif était d'estimer l'importance de l'égalisation de la salle pour un auditeur moyen, sans a priori sur le processus de reproduction. Nous avons donc fait appel à des auditeurs non entraînés spécifiquement pour ce type de stimulus, sans critère de sélection particulier. Aucun sujet n'a reporté de problèmes d'écoute, mais leur audition n'a pas non plus été évaluée. La fiabilité de leur réponse a pu être contrôlée indirectement via la discrimination de la référence cachée, conformément à la recommandation [14].

A ce jour, 16 sujets ont passé le test d'écoute, dont 9 sont des sujets rémunérés, sans liens avec la reproduction sonore. Nous présentons dans la section suivante des premiers résultats d'analyse, qui méritent d'être approfondis, et enrichis de réponses de plus de sujets.

5 Résultats préliminaires

Deux types de réactions ont été relevés suite à une discussion informelle postérieure au test. Certains sujets ont qualifié le test de relativement facile, tandis que d'autres l'ont trouvé particulièrement difficile. Ces difficultés énoncées portent notamment sur le fait qu'il y ait beaucoup de stimuli à évaluer en même temps, et qu'il faille classer des différences qui ne sont pas toujours du même type. Les sujets qui ont qualifié le test d'assez facile ne se sont probablement pas attardés sur les détails qui font pourtant la subtilité du test.

Dans un premier temps, nous observons la matrice de corrélation des réponses présentée dans la Figure 10. On constate que les réponses du sujet n°9 sont moyennement corrélées avec celles des autres sujets. Il s'avère que ce sujet a attribué une note intermédiaire à la référence cachée pour la majorité des configurations. Nous préférons éliminer l'ensemble de ses réponses. Les réponses des autres sujets semblent assez corrélées entre elles, et ne permettent pas à première vue d'identifier d'éventuels groupes de sujets.

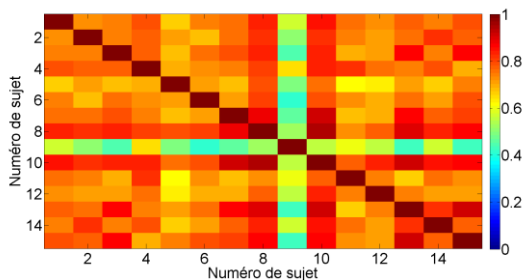


Figure 10 : Matrice de corrélation des réponses des sujets

Il s'avère que plusieurs sujets n'ont pas su identifier la référence cachée pour certaines configurations. Un second tri a été effectué, en supprimant les séries de réponses pour lesquelles la référence cachée n'a pas été identifiée (réponse inférieure à 90). Le nombre de sujets retenus par configuration est présenté dans le Tableau 3.

Tableau 3: nombre de sujets retenus par configuration

Configuration	Nombre de sujets
Reu_40	15
Reu_80	15
Stu_40	12
Stu_80	15
BuM_40	15
BuM_80	15
BuP_40	15
BuP_80	15
Cab_40	13

Il apparaît qu'il y a plusieurs sujets qui n'ont pas identifié la référence cachée pour les configurations Studio 40 cm et Cabine 40 cm, qui sont les configurations les moins réverbérantes. À l'exception des ancres, la majorité des stimuli présentés pour ces configurations sont très proches de la référence. Pour le sujet peu habitué à l'écoute de systèmes audio, cela peut être perturbant et l'induire en erreur. Il semblerait qu'il soit nécessaire de recruter des sujets qui soient habitués à avoir une écoute critique sur le rendu sonore (professionnels de l'audio, musiciens).

À ce jour, le nombre de participants au test étant réduit, nous n'avons pas mené d'analyse statistique sur les résultats. Il est encore trop tôt pour avancer d'éventuelles conclusions sur les égalisations, nous présentons seulement des résultats préliminaires sur les signaux non égalisés.

Sont représentées sur la Figure 11 les notes moyennes attribuées aux ancres et au signal non égalisé en fonction des différentes configurations. L'ancre basse semble bien remplir son rôle, puisqu'elle semble être significativement notée moins proche que les autres signaux, à l'exception du petit bureau à 80 cm. Cette exception semble cohérente avec le fait que l'ancre basse corresponde à une mesure dans cette même salle, à la même distance.

L'ancre moyenne semble en revanche moins bien choisie, puisqu'elle ne remplit pas son rôle de valeur intermédiaire et est notée relativement bas (moyenne < 40 pour toutes les configurations). L'évaluation de cette ancre n'est pas constante avec les configurations :

- Elle est évaluée du même ordre que l'ancre basse pour une partie des configurations (la cabine, les deux configurations dans le studio, et la configuration à 40 cm dans le bureau moyen)
- Elle est évaluée du même ordre que le signal non égalisé pour une autre partie (les deux configurations dans la salle de réunion, bureau moyen à 80 cm et petit bureau à 40 cm)
- Elle est évaluée plus proche de la référence que le signal non égalisé pour une configuration (petit bureau à 40cm).

Pour les signaux non égalisés, on peut discerner des évaluations différentes pour les configurations. Il semblerait que les configurations Cabine et Studio à 40 cm soit évaluée significativement plus proche de la référence que les autres configurations. Au vu des caractéristiques des salles (voir Figure 1), cela ne semble pas surprenant puisque ces salles sont très mates. Nos écoutes informelles avaient également révélées que ces configurations n'introduisent que peu de colorations spectrales.

Un point intéressant à relever est qu'il semblerait que la configuration Studio à 80 cm soit évaluée d'une manière similaire à la configuration Bureau moyen à 40 cm. Cela signifierait en d'autres termes que le rendu sonore peut être similaire dans une salle destinée à l'écoute à moyenne distance (80 cm) et dans une salle standard à faible distance (40 cm). Cette information nécessite bien entendu d'être confirmée avec un panel plus important de sujet, mais également d'être complétée par l'analyse des différentes égalisations.

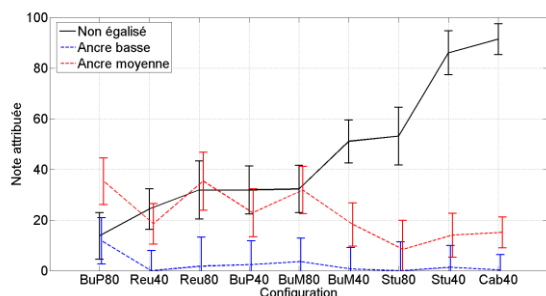


Figure 11 : Notes moyennes attribuées pour les ancres et le signal non égalisé. Intervalles de confiance à 95%

6 Conclusion

Des mesures de réponses impulsionnelles ont été réalisées dans cinq salles d'écoute, pour deux distances source-récepteur différentes (40 et 80 cm), ainsi qu'à une position décalée de 10 cm de l'axe de mesure. Quatre méthodes d'égalisation ont été implémentées, en utilisant les mesures dans l'axe. Les filtres d'égalisation ont été appliqués à l'ensemble des mesures dans l'axe et hors axe.

Un corpus de stimuli a alors été créé à partir des réponses impulsionnelles et des filtres d'égalisation. Dans tous les cas, la source a été préalablement égalisée.

Un test perceptif de type MUSHRA a été mis en place afin de comparer les dissemblances de chacun des stimuli avec une référence correspondant à un signal anéchoïque. Le test est encore en cours, et ne permet pas de dégager d'importantes conclusions à l'heure actuelle.

Néanmoins, quelques points peuvent être mentionnés. Tout d'abord, il semblerait que la tâche demandée aux sujets ne soit pas triviale et exige certaines compétences en écoute critique. Pour compléter le panel de participant, il faudrait plutôt faire appel à des professionnels de l'audio.

Au vu des réponses actuelles, il semblerait que l'ancre basse soit bien choisie, puisqu'elle est très souvent évaluée comme très différente de la référence. L'ancre moyenne semble quant à elle plus discutable, puisqu'elle n'est en moyenne pas évaluée au milieu de l'échelle, mais plutôt vers le bas.

Enfin, concernant l'évaluation des signaux non égalisés, il semblerait que deux configurations soit évaluées

significativement plus proches de la référence que les autres. Ces deux configurations correspondent aux deux salles les plus mates, à la plus faible distance. De plus, il semblerait que les résultats obtenus pour les signaux non égalisés dans une salle destinée à l'écoute à 80 cm soit similaires aux résultats obtenus dans une salle standard à 40 cm. Ce point mérite d'être comparé aux résultats des égalisations dans l'axe et en dehors. Ces résultats futurs permettront d'éventuellement conclure sur la possibilité d'avoir un rendu sonore équivalent dans une salle standard à 40 cm que dans une salle d'écoute à 80 cm.

Références

- [1] O. Kirkeby and P. Nelson, "Digital filter design for inversion problems in sound reproduction," *J. Audio Eng. Soc.*, vol. 47, no. 7/8, pp. 583–595, 1999.
- [2] B. Radlovik and R. Kennedy, "Nonminimum Phase Equalization and its subjective importance in room acoustics," *IEEE Trans. Audio Speech Lang. Process.*, vol. 8, no. 6, pp. 728–737, 2000.
- [3] L. Fielder, "Analysis of traditional and reverberation reducing methods of room equalization," *J. Audio Eng. Soc.*, vol. 51, no. 1/2, pp. 3–26, 2003.
- [4] A. Mertins, T. Mei, and M. Kallinger, "Room impulse response shortening/reshaping with infinity- and p-norm optimization," *IEEE Trans. Audio Speech Lang. Process.*, vol. 18, no. 2, pp. 249–259, 2010.
- [5] A. Vidal, P. Herzog, C. Lambourg, P. Boussard, and L. Husnik, "Binaural rendering using near-field loudspeakers," presented at the 3rd International Conference on Spatial Audio, Graz (Austria), 2015.
- [6] "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems - ITU-R BS.1116-1," International Telecommunication Union, 1997.
- [7] "AES recommended practice for professional audio - Subjective evaluation of loudspeakers," Audio Engineering Society, r2007 1996.
- [8] A. Novak, L. Simon, and P. Lotton, "Nonlinear System Identification Using Exponential Swept-Sine Signal," *IEEE Trans. Instrum. Meas.*, vol. 59, no. 8, pp. 2220–2229, 2009.
- [9] P.-Y. Michaud, "Distorsions des systèmes de reproduction musicale : protocole de caractérisation perceptive," Phd thesis, Aix-Marseille université, 2012.
- [10] O. Kirkeby, P. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *IEEE Transaction Speech Audio Process.*, vol. 6, no. 2, pp. 189–194, 1998.
- [11] G. Ramos and J. Lopez, "Filter design method for loudspeaker equalization based on IIR parametric filters," *J. Audio Eng. Soc.*, vol. 54, no. 12, pp. 1162–1178, 2006.
- [12] S. K. Mitra, *Digital Signal Processing A Computer Based Approach*, McGraw-Hill, 1998.
- [13] A. Oppenheim and R. Schaffer, *Digital signal processing*, Prentice-Hall international Editions, 1975.
- [14] "Method for the subjective assessment of intermediate quality level of coding systems - ITU-R BS.1534-1," International Telecommunication Union, 2001.
- [15] "Méthode de calcul du niveau d'isotonie," Norme internationale ISO 532B, 1975.