

**Conservatoire National Supérieur de Musique  
et de Danse de Paris**

**Formation Supérieure aux Métiers du Son**

Mémoire de Fin d'Etudes :

**COMPARAISON DE SYSTEMES DE PRISE DE SON  
MULTICANAUX**

Présenté par :

**Vincent VILLETARD**

Sous la direction de :

**Jean-Marc LYZWA**

**Session : juin 2001**

## Remerciements

Je tiens à remercier particulièrement :

Le Service Audiovisuel du CNSMDP, et plus particulièrement Catherine de Boishéraud, qui sont à l'origine de ce mémoire et qui m'ont apporté le soutien logistique dont j'ai pu avoir besoin au cours de mes expériences.

Jean-Marc Lyzwa, qui m'a guidé de manière souple et éclairée au cours de ce travail.

Jean-Christophe Messonier pour les discussions enrichissantes sur mon sujet de recherche.

Le chœur du Madrigal de Paris, son chef Pierre Calmelet, ainsi que le groupe d'étudiants percussionnistes du CNSMDP (S.David, L.Fraiche, G.Benlolo, P.Gauchon, T.Lecacheux) sous la direction d'Eric Sammut, qui ont participé à la session d'enregistrement et fourni la matière musicale pour ce travail.

Les concepteurs de systèmes de prise de son qui se sont prêtés à l'expérience.

Les étudiants de la FSMS pour leur participation amicale aux tests d'écoute individuels.

Tous les participants aux séances d'écoutes de la Convention de l'AES qui ont contribué, par leurs commentaires, aux premiers résultats de ce mémoire.

## Table des matières

Introduction .....	5
1 <sup>e</sup> partie : la session d'enregistrement .....	7
1.1. Présentation	
1.2. Intérêt	
2 <sup>e</sup> partie : les systèmes de prise de son .....	12
2.1. L'ASM5 de Brauner	
2.2. Le Multichannel Microphone Array de M.Williams et G.Le Dû	
2.3. Le Compact Surround System de N.Jacques et A.Mora	
2.4. Le système de J-M.Porcher (Radio France)	
2.5. Le système de J-M.Lyzwa et J-C.Messonier (CNSMDP)	
2.6. Le KFM360 de J.Bruck	
2.7. Le système du Grenier à Son	
3 <sup>e</sup> partie : l'analyse des commentaires d'écoute de la Convention de l'AES .....	23
3.1. La matière	
3.2. L'analyse qualitative	
3.3. L'analyse quantitative	
3.4. Conclusions sur l'analyse qualitative	
3.5. Conclusions sur l'analyse quantitative	
3.6. Conclusion	
4 <sup>e</sup> partie : les tests individuels d'analyse d'écoute .....	34
4.1. L'élaboration et le déroulement	
4.2. Les résultats et la comparaison avec les écoutes collectives	
4.3. Influence de l'ordre d'écoute	
4.4. Synthèse des résultats	
5 <sup>e</sup> partie : méthodologie des tests de localisation .....	41
5.1. Méthodes de réponse pour les tests de localisation	
5.2. Test de comparaison des méthodes	
5.3. Les résultats	
5.4. Conclusion	

6° partie : les tests individuels de localisation -----	47
6.1. Principe	
6.2. Les résultats	
6.3. Commentaires	
7° partie : synthèse des résultats -----	58
7.1. Le système de J-M.Porcher (Radio France)	
7.2. Le KFM360 de J.Bruck	
7.3. Le système de J-M.Lyzwa et J-C.Messonier (CNSMDP)	
7.4. L'ASM5 de Brauner	
7.5. Le Compact Surround System de N.Jacques et A.Mora	
7.6. Le Multichannel Microphone Array de M.Williams et G.Le Dû	
7.7. Le système du Grenier à Son	
Conclusion -----	74
Annexe 1 : recommandation ITU-R BS 775-1 -----	75
Annexe 2 : plan du grand plateau d'orchestre -----	80
Annexe 3 : feuilles de commentaires de la Convention -----	82
Annexe 4 : critères et feuille de notation des tests individuels d'analyse -----	85
Annexe 5 : résultats complets des tests individuels d'analyse d'écoute -----	88
Bibliographie -----	93

## Introduction

Ce mémoire est parti d'une expérience organisée au Conservatoire National Supérieur de Musique et de Danse de Paris (CNSMDP) en novembre 1999, qui a consisté à enregistrer des œuvres musicales simultanément grâce à 7 systèmes de prise de son multicanale. Cette session d'enregistrement, décrite dans la 1<sup>e</sup> partie de ce mémoire, réunissait plusieurs concepteurs parmi les plus engagés dans la recherche sur l'enregistrement multicanal et allait nous donner l'occasion de pouvoir comparer les systèmes qu'ils ont mis au point. Ces systèmes, décrits dans la 2<sup>e</sup> partie, constituaient un panel intéressant des techniques utilisables pour l'enregistrement multicanal. Ainsi, cette expérience allait permettre, plus que de comparer les systèmes entre eux, d'extrapoler les résultats envisageables avec telle ou telle technique. Ainsi, lorsque le Service Audiovisuel m'a proposé de m'associer à l'analyse de cette expérience, j'y ai vu un intérêt dépassant la simple comparaison de ces 7 systèmes de prise de son multicanale, et cela s'est confirmé tout au long des expériences et des analyses qui ont succédé à la session d'enregistrement.

La démarche du Service Audiovisuel avait pour aboutissement la diffusion des enregistrements réalisés par les différents systèmes de prise de son au cours de séances d'écoute en petit groupe durant la 108<sup>e</sup> Convention de l'Audio Engineering Society (AES) qui eut lieu à Paris en février 2000. Pour moi, il s'agissait du point de départ de mon travail. Ces séances m'ont permis de récolter une quantité importante - par sa valeur comme par son nombre - de commentaires d'écoute libres. La 3<sup>e</sup> partie est consacrée à l'analyse, tant qualitative que quantitative, de ces données. Cette analyse s'est révélée très riche sur le plan qualitatif : elle nous a permis d'établir de façon exhaustive une grille d'analyse utilisée par l'ensemble des auditeurs pour décrire les images sonores multicanales qu'ils écoutaient. L'analyse quantitative a permis de dégager des tendances, mais une étude statistique rigoureuse n'était pas possible du fait qu'on avait volontairement demandé aux auditeurs des commentaires libres.

L'étape suivante de mon travail a donc été l'organisation de séances d'écoutes individuelles avec une évaluation chiffrée d'un certain nombre de critères. La grille d'analyse élaborée suite à l'analyse des commentaires de la Convention avait justement permis de dégager les critères prépondérants utilisés par les auditeurs. L'élaboration et l'analyse de cette nouvelle série de test sont décrites dans la 4<sup>e</sup> partie. Ces tests débouchent sur un classement, critère par critère, des différents systèmes de prise de son. Ces résultats concernant l'analyse des images sonores rendues par chaque système constituent la première partie des données qui vont servir à une analyse approfondie de chaque système, dans la dernière partie du mémoire.

La deuxième partie des données qui allaient servir à cette analyse concernait les propriétés de localisation de chacun des systèmes. En effet, au cours de la session d'enregistrements, outre les œuvres musicales, nous avons enregistré un certain nombre de signaux tests, dans différentes directions, qui allaient me servir à l'élaboration de tests individuels de localisation. L'élaboration de ces tests a été pour moi l'occasion de me questionner sur les méthodes de réponses possibles pour les auditeurs au cours de tels tests de localisation sur 360°. Ainsi, préalablement aux tests à proprement parler, j'ai organisé une petite expérience afin de comparer la validité et le degré de confiance de plusieurs méthodes possibles, afin de choisir la plus appropriée pour les tests de localisation que j'allais conduire. Cette expérience annexe est décrite et analysée dans la 5<sup>e</sup> partie.

Les tests individuels de localisation eux-même font l'objet de la 6<sup>e</sup> partie. Ils ont conduit à l'élaboration des courbes de localisation de chacun des système, qui allaient servir à alimenter l'analyse approfondie de chaque système.

Cette analyse, qui synthétise l'ensemble des résultats collectés au cours de ce travail de recherche, est présentée dans la 7<sup>e</sup> partie. Plus que de rendre compte des résultats de chacun des systèmes dans la situation d'enregistrement donnée, j'ai tenté du mieux possible d'interpréter ces résultats au vu des principes à l'origine des systèmes, de la situation d'enregistrement et de diffusion et des acquis de la recherche sur la stéréophonie à 2 canaux. Ainsi, le but de ce travail n'est pas de "classer" de façon définitive les systèmes, mais plutôt de dégager les potentialités de chacun et, plus largement, des principes qui les sous-tendent.

## **Première partie**

### **La session d'enregistrement**

## 1.1. Présentation

En novembre 1999, le Service Audiovisuel du Conservatoire National Supérieur de Musique et de Danse de Paris (CNSMDP) organisait sur 3 jours une session d'enregistrement particulière. Il s'agissait d'enregistrer avec différents systèmes de prise de son destinés à la restitution multicanale 5.1 (recommandation ITU-R BS 775-1, Annexe 1 et Figure 1) deux œuvres "spatialisées" sur 360 degrés, c'est-à-dire pour lesquelles l'auditeur est censé être au centre des exécutants. Ces pièces étaient, d'une part, le motet "Spem in alium" de Thomas Tallis pour chœur mixte de 40 chanteurs répartis sur un cercle (Figure 2) et, d'autre part, une improvisation à 5 percussions (batterie, 2 marimbas, xylophone et peaux) également réparties sur un cercle (Figure 3).

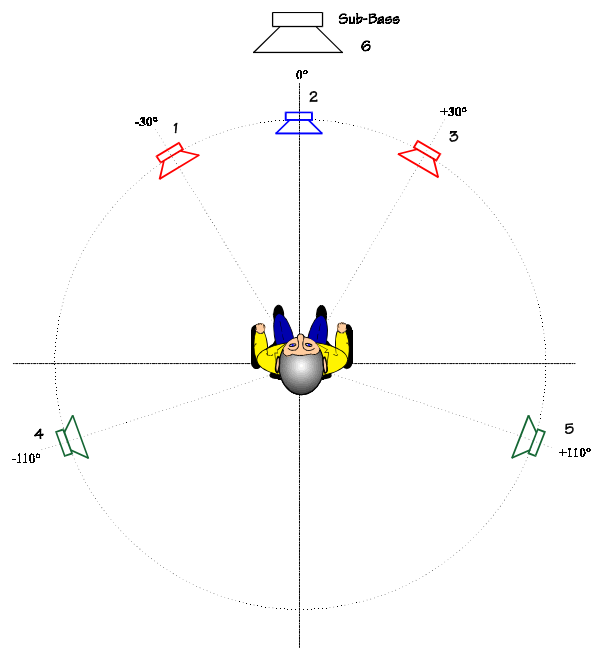


Figure 1

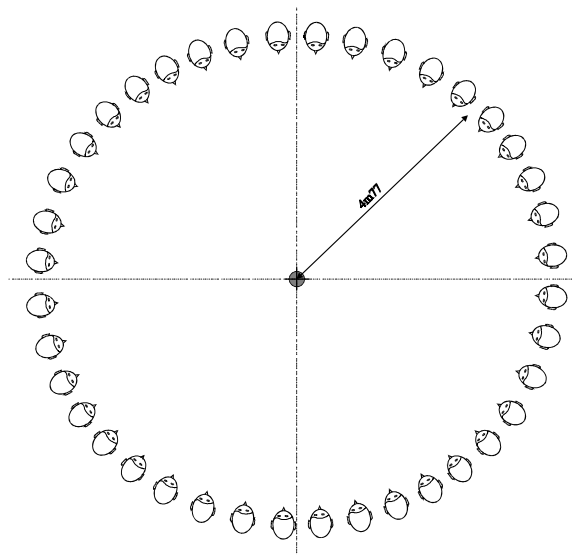


Figure 2



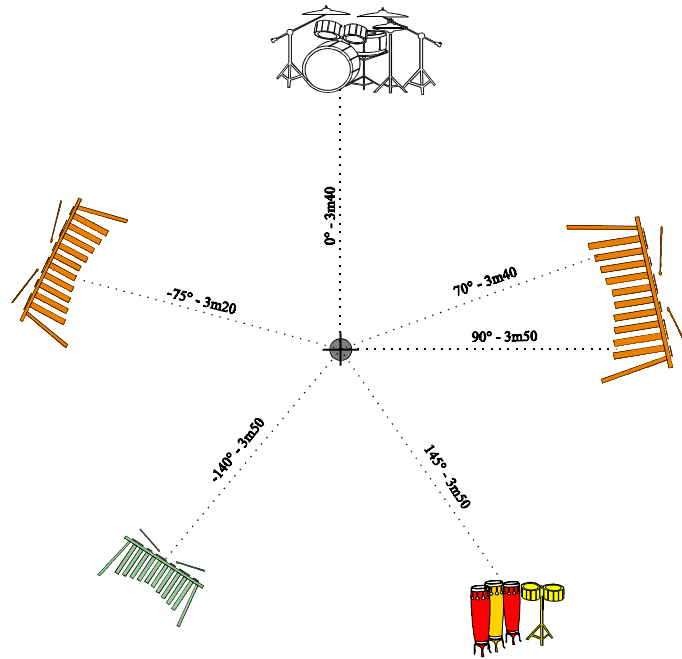


Figure 3

A cette occasion, plusieurs concepteurs de systèmes de prise de son multicanale ont été contactés et invités à participer à l'expérience. Au final, 7 systèmes furent associés à l'expérience : celui de Jean-Marie Porcher de Radio France, le système KFM360 de Jerry Bruck distribué par Schöeps, un système développé par Jean-Marc Lyzwa et Jean-Christophe Messonier du CNSMDP (en association avec l'IRCAM), le système ASM5 de Brauner, le Compact Surround System de Nicolas Jacques et Arnaud Mora développé à partir de leur mémoire de fin d'études de la Formation Supérieure au Métiers du Son (FSMS) au CNSMDP, le Multichannel Microphone Array de Michael Williams et Guillaume Le Dû ainsi qu'un système de Pierre-Antoine Signoret et Jacques Warnier du Grenier à Son (LGS).

L'enregistrement, auquel participèrent le chœur du Madrigal de Paris ainsi qu'un groupe d'étudiants de la classe de percussions du CNSMDP, fut réalisé au grand plateau d'orchestre (plan en Annexe 2). Les pièces furent enregistrées simultanément par les différents systèmes de prise de son, installés au centre du cercle formé par les interprètes (choristes ou percussionnistes) et disposés les uns sous les autres - les systèmes non ponctuels étant centrés autour de ce point. Le positionnement en hauteur des différents systèmes s'est fait par entente mutuelle entre les participants ; comme on peut le voir sur la Figure 4, les systèmes se situaient à une hauteur comprise entre 2,32m et 3,99m (sachant que la tête des chanteurs, disposés sur des estrades, se situait à environ 2,40m du sol et les percussions à environ 1m du sol). Le rayon du cercle était de 4,77m pour le chœur et de 3,40m pour les percussions. Le rayon critique du grand plateau mesuré pour une source omnidirectionnelle est de 2m. Les ingénieurs du son attachés à chaque système disposaient de régies d'écoute multicanales afin d'effectuer leur balance. Les enregistrements s'effectuaient en 16 bit/48 kHz sur bandes multipistes Hi8 et étaient synchronisés par un distributeur général de time-code. Les enregistrements nécessitant une post-production étaient confiés aux ingénieurs du son afin qu'ils l'effectuent dans leurs propres studios.

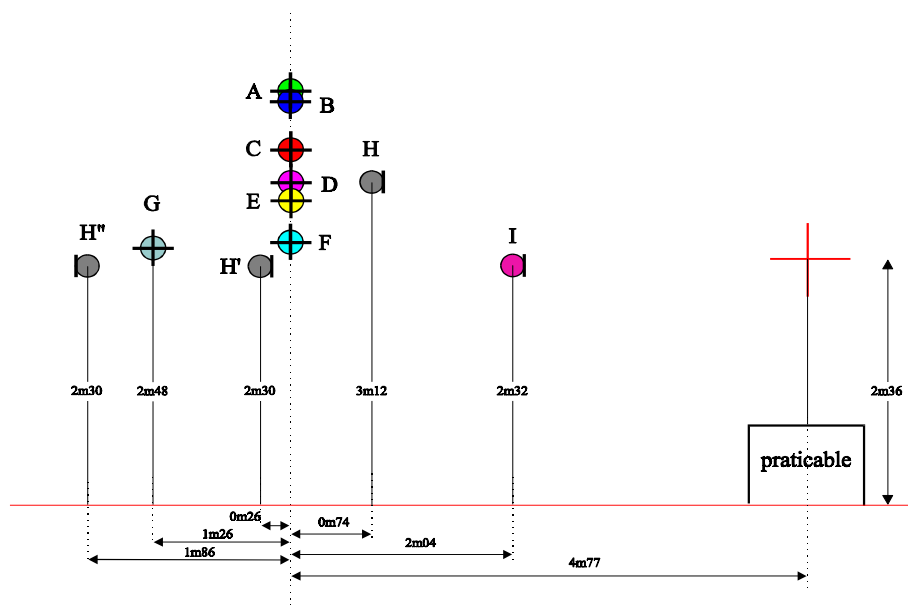


Figure 4

A : Brauner	hauteur 3m99
B : M. Williams / G. le Dû	hauteur 3m90
C : Nicolas Jacques - Arnaud Morah	hauteur 3m43
D : Soundfield	hauteur 3m12
E : Tête artificielle - le grenier à son	hauteur 2m95
F : Sphère Shoeps	hauteur 2m54
H, H', H'' : Radio-France	hauteur 2m30
I : Conservatoire de Paris	hauteur 2m32

Cette expérience était menée en vue d'une démonstration lors de la 108<sup>e</sup> Convention de l'Audio Engineering Society (AES) qui eut lieu à Paris en février 2000. Un auditorium avec diffusion multicanale, pour un auditoire de 20 personnes, fut installé au Palais des Congrès ; tout au long de la Convention se sont déroulées des séances d'écoute avec la diffusion d'un extrait de chacune des deux pièces dans l'enregistrement des 7 systèmes.

Lors de la session d'enregistrement, outre les 2 œuvres, nous avons enregistré avec Jean-Marc Lyzwa une série de signaux (que nous nommerons signaux tests) en vue d'élaborer des tests de localisation. Il s'agissait de coups de wood-blocks et de maracas ainsi que d'une voix chantée diffusée par un haut-parleur mobile, dans différentes directions régulières autour des systèmes.

## 1.2. Intérêt

Cette expérience a été l'occasion de réunir un nombre important de systèmes de prise de son multicanale, commercialisés ou en cours de développement, couvrant une grande partie des méthodes actuellement utilisées. L'enregistrement simultané par ces différents systèmes allait permettre une comparaison valide de leurs propriétés. De plus, la diffusion des enregistrements au cours de la Convention de l'AES était l'occasion de réunir l'appréciation d'un nombre très important d'auditeurs appartenant au milieu de l'audio et allait constituer une base de données de grande valeur.

Les œuvres choisies présentaient cela de particulier qu'elles étaient spatialisées sur 360°. Cela créait une situation particulière, digne d'intérêt. En effet, une des possibilités du multicanal est de pouvoir restituer des œuvres composées dans le but d'être spatialisées de façon particulière autour de l'auditeur ; en ce sens, la restitution multicanale a un intérêt musicologique en respectant l'esthétique des œuvres et constitue même une motivation à la création musicale, comme on peut le voir actuellement dans les classes de composition du CNSMDP. Il faut d'ailleurs souligner que, contrairement à ce que l'on pourrait croire, la spatialisation des œuvres musicales n'est pas apparue seulement au 20<sup>e</sup> siècle (on pense à Stockhausen, Xenakis, Boulez,...) mais a été présente très tôt dans l'histoire de la musique avec la création d'œuvres spécialement composées pour être interprétées dans des lieux à l'architecture particulière. C'est par exemple le cas de Gabrielli, qui a exploité l'architecture de la basilique Saint-Marc à Venise pour disposer des chœurs dans différentes directions. C'est aussi le cas de Thomas Tallis avec le motet "Spem in alium" enregistré au cours de ces séances ; celui-ci a en effet fort probablement été commandé par le duc de Norfolk afin d'être donné dans la salle des banquets de son château, à la forme octogonale. Cela explique la division vocale en 8 chœurs de 5 voix, chaque chœur se situant sur un des côtés de la salle, et l'auditoire au milieu ; cela permet en particulier des effets de réponses, provenant de directions opposées, assez impressionnants.

Si le choix de telles œuvres spatialisées sur 360° constitue un cas un peu particulier, cela est néanmoins un bon moyen d'évaluation des propriétés des systèmes sur l'espace complet. Il est d'autre part bien entendu que l'analyse des systèmes à partir de ces enregistrements ne doit pas conduire à un jugement absolu, voire un classement de ceux-ci. Il s'agit plutôt de dégager les propriétés et les limites des différents systèmes afin d'envisager leur utilisation optimale, voire leur amélioration ou le développement de systèmes dérivés.

Nous allons dans un premier temps présenter les différents systèmes de prise de son associés à l'expérience, avec les théories qui les sous-tendent.

## **Deuxième partie**

### **Les systèmes de prise de son**

Il ne s'agit pas ici de reproduire l'ensemble des documents de présentation des différents systèmes (nombreux pour certains, plus rares pour d'autres), mais plutôt de les synthétiser afin de dégager la théorie qui a servi à leur élaboration. Cela permettra en particulier de mettre en regard les résultats obtenus aux différents tests avec les concepts qui sous-tendent les systèmes. Cette présentation se voulant à la fois concise et la plus complète possible, il est possible que la compréhension de certains principes par le lecteur nécessite une lecture plus approfondie des documents cités.

Trois des systèmes (ceux de Brauner, M.Williams/G.Le Dû et N.Jacques/A.Mora) sont sensiblement basés sur la même théorie, celle des angles de prise de son. L'angle de prise de son d'un couple de microphones en stéréophonie à 2 canaux est la portion angulaire de l'espace d'enregistrement qui sera restituée entre les 2 enceintes dans l'espace d'écoute ; les sources se situant au-delà de l'angle de prise de son seront reproduites sur les enceintes même. M.Williams a étudié, voilà plusieurs années, les angles de prise de son des systèmes stéréophoniques à 2 canaux et fourni les courbes permettant de déterminer l'angle de prise de son d'un couple de microphones en fonction de leur directivité, leur espacement et l'angle qu'ils forment ([1]) ; par exemple, la Figure 5 montre les angles de prise de son pour un couple de microphones cardioïdes. Cette théorie peut être étendue aux systèmes de prise de son multicanaux à 5 microphones en les voyant comme 5 couples adjacents (gauche/centre, centre/droit, droit/arrière droit, arrière droit/arrière gauche, arrière gauche/gauche). Le principe est alors de couvrir l'espace sonore de 360° grâce à leur 5 angles de prise de son.

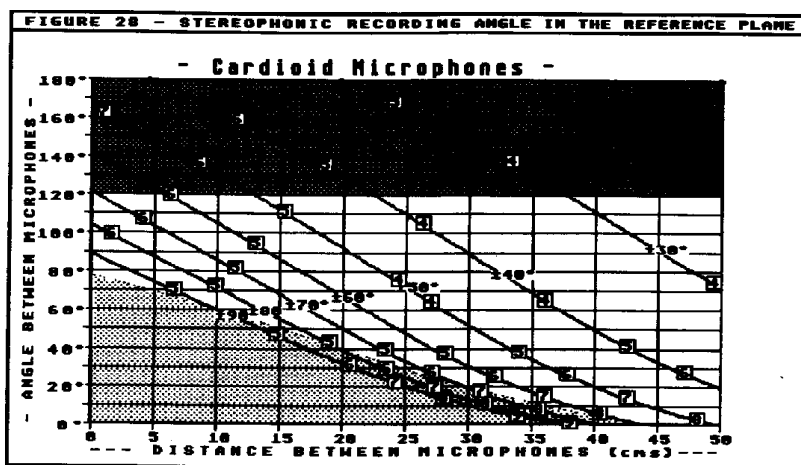


Figure 5

## 2.1. L'ASM5 de Brauner

La configuration de base des microphones pour le système Brauner est le système INA5 (Ideale Nierenanordnung 5: "positionnement idéal de 5 cardioïdes") développé par U.Herrmann et V.Henkels ([2]) à la suite d'une recherche menée sur la comparaison de différentes techniques multicanales. Dans cette configuration, les 3 microphones avant sont disposés au bout de bras de 17,5cm orientés à  $-90^\circ$ ,  $0^\circ$  et  $+90^\circ$  respectivement et les 2 microphones arrière sont au bout de bras de 59,6cm orientés à  $\pm 150^\circ$  (Figure 6).

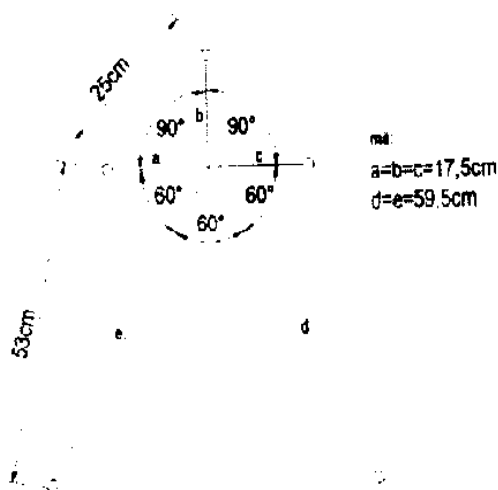


Figure 6

La disposition des 3 microphones avant est calculée de sorte qu'il y ait juxtaposition et non recouvrement des angles de prise de son gauche/centre et centre/droit ; ceux-ci sont égaux à  $90^\circ$ . Le positionnement des microphones arrière est fait de sorte que l'on ait environ  $60^\circ$  d'angles de prise de son arrière et latéraux (soit un total de  $360^\circ$ ).

Le système ASM5 de Brauner utilise des microphones VM1 à double membrane et présente une certaine flexibilité par rapport à la configuration de base dans la mesure où les longueurs des bras arrière sont ajustables sur une plage relativement importante ; de plus, l'orientation des capsules est elle aussi modulable entre  $\pm 45^\circ$ . D'autre part, le système microphonique est associé au processeur dédié Atmos 5.1 de SPL. Celui-ci permet de gérer la directivité des microphones continuellement d'omnidirectionnels à bidirectionnels ; il permet également un panoramique de chacun des microphones dans l'espace multicanal ainsi qu'un élargissement des zones couvertes avant et arrière (par une simulation électronique de modification de distance entre les microphones gauche/droit et arrière gauche/arrière droit).

Il faut noter que toute modification des paramètres de directivité, de distance, d'angle ou de panoramique des microphones ne permettra plus de vérifier la propriété de la configuration de base, à savoir un total de  $360^\circ$  pour les 5 angles de prise de son ; néanmoins cette flexibilité permet d'aborder de façon plus pragmatique l'enregistrement en fonction de l'acoustique du lieu, la distance aux sources, leur positionnement... D'ailleurs, au cours de la session d'enregistrement, la directivité adoptée a été l'hypercardioïde, certainement du fait de la réverbération importante du lieu et de la distance aux sources (le système Brauner était celui positionné le plus haut) ; les microphones étaient orientés dans le prolongement des bras.

## 2.2. Le Multichannel Microphone Array de M.Williams et G.Le Dû

M.Williams et G.Le Dû ont étudié de façon approfondie le principe de la juxtaposition des 5 angles de prise de son (appelés angles de couvertures) afin qu'il n'y ait ni trou ni recouvrement entre eux (principe dit du "critical linking"); ils ont d'ailleurs développé un logiciel calculant les paramètres permettant cette juxtaposition ([3]).

Une des caractéristiques des angles de prise de son (ou angles de couverture) est d'être centrés par rapport à l'axe de symétrie du couple de microphones; cela empêche dans bien des cas une juxtaposition précise des 5 angles de couverture. En effet, une fois les microphones positionnés (avec un certain angle et une certaine distance entre eux), l'angle de prise de son est fixé à la fois dans sa valeur et dans sa position. Or, si l'on veut une juxtaposition parfaite des angles de couverture, il ne suffit pas que leur somme fasse  $360^\circ$  (comme pour le système INA5 de Brauner), mais il faut pouvoir les faire pivoter de façon à supprimer leurs recouvrements éventuels (voir exemple de recouvrement Figure 7). Le travail de M.Williams et G.Le Dû est allé dans ce sens : ils ont examiné les méthodes permettant le pivotement des angles de couverture.

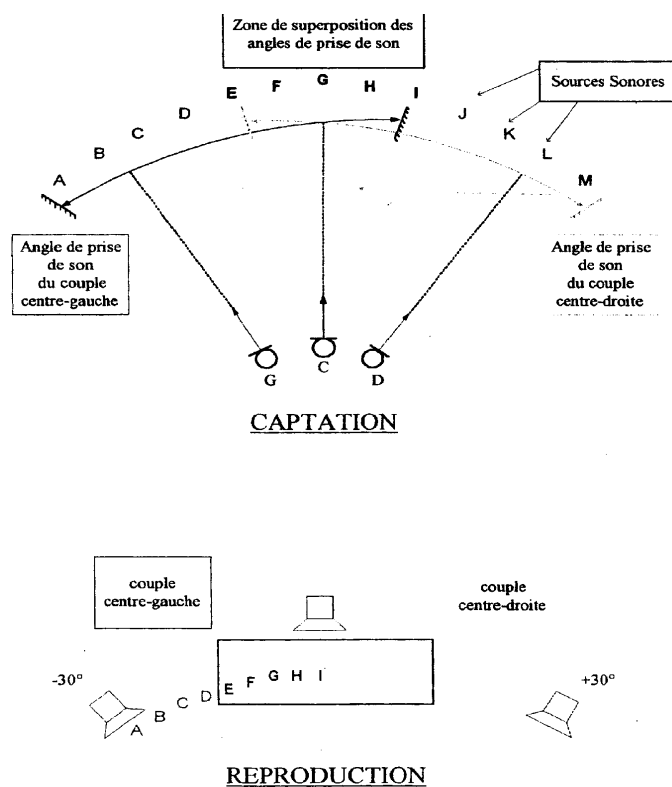


Figure 7

Ces méthodes sont de deux ordres ([4],[5]) : modification physique par pivotement d'un microphone, tout en conservant la distance et l'angle relatif entre les deux microphones du couple considéré, ou bien modification électronique en post-production par ajout d'une différence de temps ou d'intensité constante entre les deux microphones (cette méthode simule la modification physique sans toucher au positionnement des microphones). Le principe de ces méthodes (physique ou électronique) est en fait de créer un écart angulaire entre, d'une part, l'axe des sources provoquant une différence de temps nulle ( $\Delta T=0$ ) et, d'autre part, celui des sources créant une différence d'intensité nulle ( $\Delta I=0$ ). Alors que ces

deux axes sont confondus au départ (il s'agit de l'axe de symétrie du couple), ils sont dissociés par l'utilisation d'une de ces méthodes, ce qui fait pivoter l'angle de prise de son (ainsi, une source située sur l'axe de symétrie du couple ne sera plus localisée au centre, mais décalée, car elle ne créera plus à la fois une différence d'intensité et de temps nulles ; un exemple est donné Figures 8 et 9).

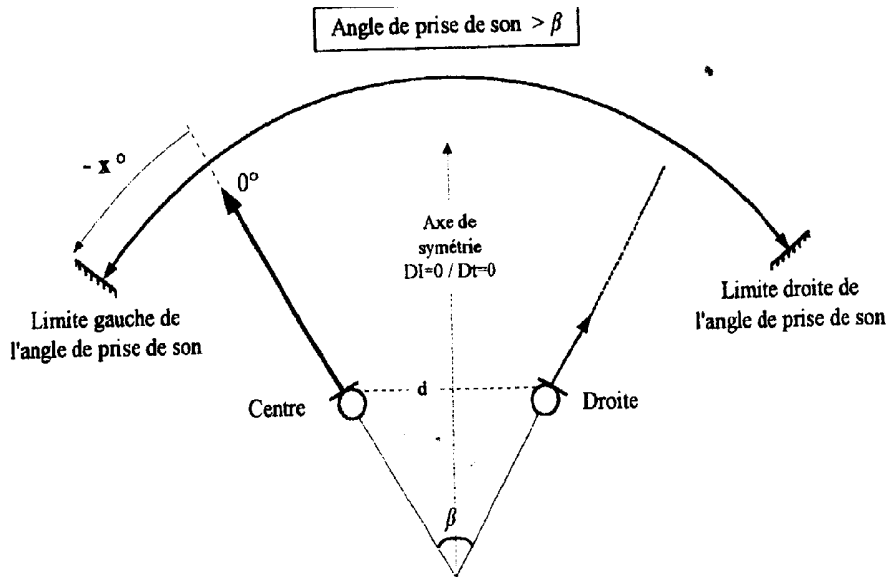


Figure 8

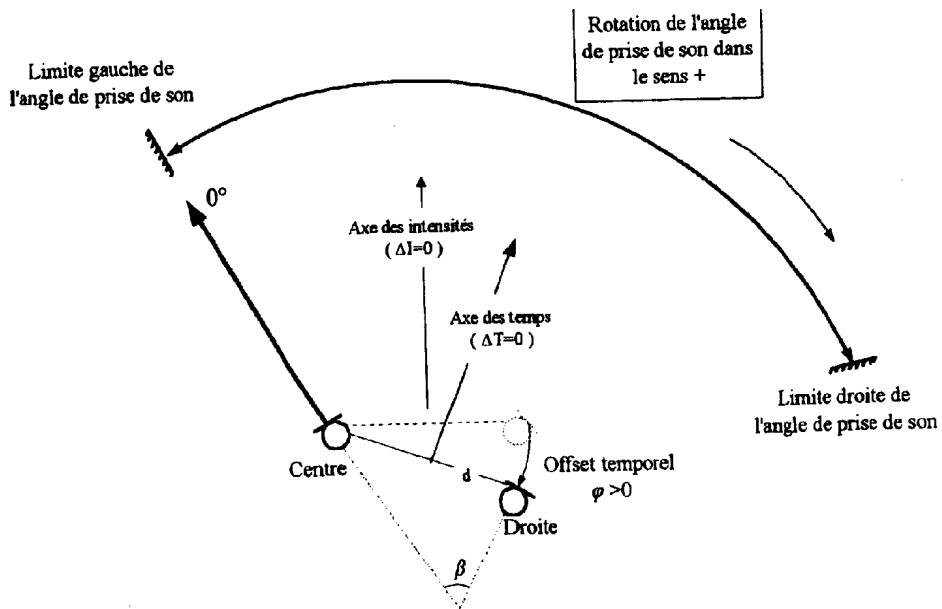


Figure 9



L'idée de base de M.Williams et G.Le Dû est de travailler avec 5 microphones cardioïdes Schoeps MK4. Cependant, pour l'expérience, ils n'ont utilisé que 4 microphones (sans utilisation du canal central), à savoir deux microphones cardioïdes MK4 à l'avant et deux hypocarديوïdes MK21 à l'arrière. La méthode électronique a été utilisée pour jouxter les angles de couverture. Ceux-ci ont été choisis égaux aux angles physiques entre les haut-parleurs :  $60^\circ$  à l'avant et  $140^\circ$  à l'arrière (donc  $90^\circ$  sur les côtés) ;. Le choix d'égalité entre les angles de couverture et les angles physiques entre les haut-parleurs a pour but d'obtenir une restitution linéaire de l'espace, en évitant les distorsions angulaires ; il semblerait que pour réaliser cela, la configuration à 5 microphones était difficilement gérable, d'où le choix d'utiliser 4 microphones.

### 2.3. Le Compact Surround System de N.Jacques et A.Mora

N.Jacques et A.Mora ont développé leur propre système (Compact Surround System) dans le cadre de leur mémoire ([6]), prolongé par une recherche personnelle. Les concepteurs étant peu disposés à fournir des informations techniques précises sur leur système (des raisons commerciales ont été invoquées), il ne m'a pas été possible d'obtenir plus de détails que ceux que je fournis ici.

Les concepteurs se sont basés sur la théorie des angles de couverture de M.Williams tout en se fixant comme contrainte une portabilité aisée du système, en particulier en vue d'applications télévisuelles. Cela les a conduit à élaborer un système de 5 microphones cardioïdes Schoeps MK4 montés sur une même barrette (voir photo Figure 10). Ils simulent électroniquement, par l'utilisation de délais, l'écartement entre les microphones avant et arrière.

Afin d'avoir une linéarité angulaire pour la restitution, en évitant que certaines zones angulaires soient contractées ou au contraire dilatées, le choix des angles de couverture a été celui des angles physiques entre les haut-parleurs.



Figure 10

Pour la détermination des angles de couverture sur les côtés et l'arrière, N.Jacques et A.Mora ne se sont pas basés sur les courbes établies pour la stéréophonie à 2 canaux (qui concernent la perception sur les  $60^\circ$  frontaux) mais ont effectué des tests de perception afin d'établir les relations entre, d'une part, les différences de temps et d'intensité et, d'autre part, la localisation dans les zones latérales et arrière.

## **2.4. Le système de J-M.Porcher (Radio France)**

Dans la lignée des systèmes à 5 microphones, se trouve celui de J-M.Porcher, de Radio France. Son élaboration a été plus empirique que les trois systèmes précédents, et il était originellement destiné à des scènes sonores frontales. Néanmoins, J-M.Porcher a été intéressé par l'expérience d'enregistrement d'une source sonore circulaire.

L'idée de départ pour ce système est d'utiliser pour l'avant une configuration élaborée par la firme DECCA ("triangle DECCA"); cette configuration est constituée de 3 microphones omnidirectionnels largement espacés, disposés en triangle équilatéral (1,40m de côtés), dont la pointe est en avant, dirigée à environ 45° vers le haut. Cette disposition, élaborée dans le cadre de la stéréophonie à 2 canaux (le canal central était alors mixé aux canaux gauche et droite), avait pour but de "remplir" le centre de l'image sonore ; en effet, dans le cadre de l'utilisation d'un couple omnidirectionnel, les sources ont tendance à être rapidement latéralisées (du fait de la non linéarité de la localisation en différence de temps). Les microphones utilisés par J-M.Porcher sont des TLM50 de Neumann, omnidirectionnels, mais présentant une certaine directivité dans l'aigu. Celle-ci a été exploitée en orientant les microphones vers l'extérieur d'un angle de l'ordre de 60°.

L'utilisation prévue pour ce système étant de capter les scènes frontales, les canaux arrières étaient utilisés pour apporter de la réverbération et de l'enveloppement ; la disposition classique est donc constituée de 2 microphones identiques à ceux de l'avant mais disposés suffisamment loin de ceux-ci et suffisamment espacés afin de fournir des signaux décorrélés permettant une meilleure sensation d'enveloppement. Dans le cadre de l'expérience, la configuration a été adaptée à la situation d'une source circulaire en rapprochant les microphones arrières de l'avant (à une distance de 1,60m) et en les rapprochant l'un de l'autre (espacement de 48cm).

## **2.5. Le système de J-M.Lyzwa et J-C.Messonier (CNSMDP)**

Dans la catégorie des systèmes à microphones non coïncidents se trouve également celui du CNSMDP, élaboré par J-M.Lyzwa et J-C.Messonier du Service Audiovisuel. Contrairement aux précédents, il ne s'agit plus d'un système à 5 canaux, mais utilisant 10 microphones omnidirectionnels BK4006 régulièrement disposés sur un cercle de 2,04m de rayon (à 0°, +/-36°, +/-72°, +/-108°, +/-144° et 180°, comme présenté sur la Figure 11), à l'intérieur du cercle formé par les chanteurs ou les percussionnistes. Ce système met en œuvre une post-production importante, qui exploite différents principes de localisation en écoute binaurale.

La perception par différence de temps est exploitée sur la base avant où les microphones à -36°, 0° et +36° sont directement routés vers les haut-parleurs gauche (à -30°), centre (à 0°) et droit (à +30°). Il s'agit donc de considérer 2 couples de microphones omnidirectionnels sur l'avant ; leur angle de prise de son est environ de 40°.

Sur l'arrière, les concepteurs utilisent une perception par différence d'intensité en panoramiquant les microphones entre les haut-parleurs arrière gauche et arrière droit : ceux à +/-108° alimentent directement les haut-parleurs (puisque'ils correspondent à leur direction), celui à 180° est panoramiqué exactement au centre, ceux à +/-144° sont dirigés en utilisant les signaux tests enregistrés à 140° afin de faire correspondre au mieux la direction perçue avec la direction réelle.

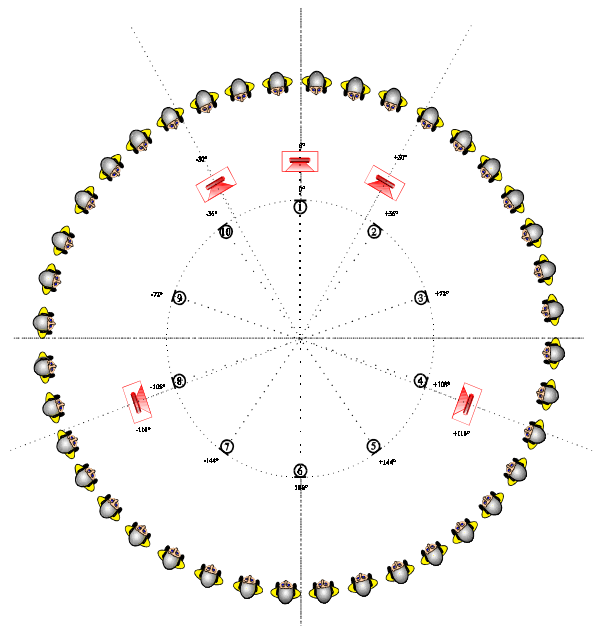


Figure 11

La principale originalité de ce système concerne la localisation latérale : partant du fait que la localisation latérale en différence de temps ou d'intensité a conduit dans de nombreuses expériences à une imprécision importante (extra-largeur de sources, basculement rapide avant-arrière), les concepteurs ont voulu expérimenter une méthode de localisation utilisant le filtrage transaural. La localisation latérale en écoute binaurale étant fortement basée sur les HRTFs (Head Related Transfer Functions), les concepteurs ont utilisé un logiciel de traitement du signal développé par l'IRCAM et Espaces Nouveaux ([7]) : le Spatialisateur. Ce logiciel a (entre autre) pour but une reproduction de sons latéraux, arrières ou en élévation par diffusion sur les haut-parleurs avant gauche et droit, après traitement du signal simulant les filtrages de la tête en écoute naturelle. Cela a tout d'abord été développé pour une écoute au casque (écoute binaurale, où chaque oreille ne perçoit que le signal du côté lui correspondant, ou ipsilatéral) puis adapté à l'écoute aux haut-parleurs ; pour cela, il a fallu s'attacher à annuler les ondes croisées, ou contralatérales (ce que l'oreille gauche - resp. droite - perçoit du signal émis par le haut-parleur droit - resp. gauche) par utilisation d'une matrice de filtrage inverse dite d'annulation croisée (filtrage transaural). L'efficacité du procédé est tributaire de la position de l'auditeur, qui doit se situer précisément entre les deux haut-parleurs.

Le Spatialisateur a été utilisé sur les signaux issus des microphones à  $\pm 72^\circ$  qui, après traitement transaural adapté, ont été diffusés sur la base avant gauche/droite.

Enfin, les concepteurs ont utilisé une autre possibilité du Spatialisateur qui est de générer une réverbération multicanale (ici diffusée sur les 4 haut-parleurs gauche, droit, arrière gauche et arrière droit). Une réverbération précoce (simulation de premières réflexions) a été générée à partir de chaque microphone et diffusée dans le secteur spatial correspondant (gauche/droit, droit/arrière droit, arrière droit/arrière gauche ou arrière gauche/gauche) puis une réverbération tardive, issue de la somme de tous les microphones, a été diffusée sur les 4 haut-parleurs.

Les deux derniers systèmes (KFM360 et Le Grenier à Son), contrairement aux précédents, sont des configurations partiellement coïncidentes (c'est-à-dire dont certains capteurs sont coïncidents).

## 2.6. Le KFM360 de J.Bruck

Le KFM360 ([8]) est constitué d'une sphère Schoeps KFM6 à laquelle sont accolés deux microphones bidirectionnels MK8. La sphère Schoeps est une sphère dure de 18cm de diamètre équipée de deux capteurs omnidirectionnels encastrés, diamétralement opposés ; les deux microphones bidirectionnels sont disposés contre ces capteurs et sont orientés avant/arrière (voir Figure 12). On peut donc voir ce système comme deux systèmes M/S (composés d'un omnidirectionnel et d'un bidirectionnel) dirigés vers les côtés et séparés par une sphère.

Le décodage classique de ces deux systèmes M/S fournit les 4 canaux arrière gauche/gauche et arrière droit/droit. Le canal central est obtenu par un matricage des canaux gauche et droit dépendant de la fréquence ("optimum matrix"), élaboré par M.Gerzon ([9]). Ce système utilise donc à la fois la localisation par différence d'intensité (pour les côtés) et par différence de temps (pour l'avant et l'arrière).

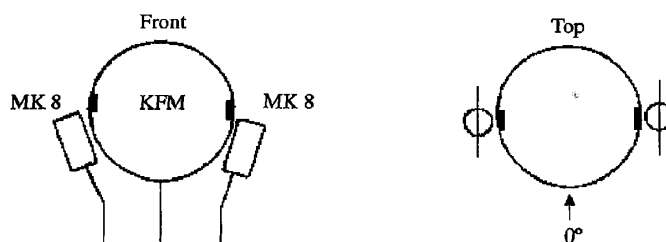


Figure 12

Comme on l'a évoqué plus haut, la localisation latérale en écoute multicanale pose des difficultés. J.Bruck est parti du fait, établi pour la stéréophonie à 2 canaux, que les systèmes coïncidents donnent de meilleurs résultats quant à la localisation (à la fois en précision et en stabilité) et a donc utilisé un tel procédé pour les côtés ; une fois décodés, les M/S correspondent en effet à 2 microphones coïncidents - plus ou moins directifs selon le dosage de l'omnidirectionnel et du bidirectionnel - orientés vers l'avant et l'arrière (Figure 13).

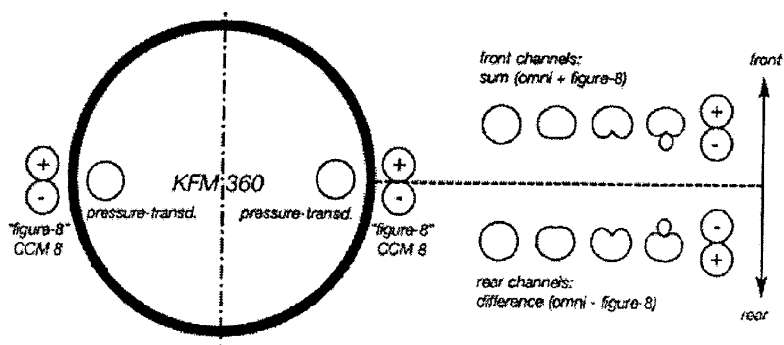


Figure 13

Par contre, pour l'avant et l'arrière, où la localisation pose moins de difficultés, J.Bruck a utilisé les meilleures qualités de rendu spatial des systèmes non coïncidents. Les canaux avant (resp. arrière) se trouvent en effet alimentés par les signaux issus du décodage des deux paires M/S assimilables à deux microphones directifs pointés vers l'avant (resp. vers l'arrière) et écartés de 18cm ; la sphère va, quant à elle, agir comme un obstacle entre ces deux microphones et ajouter une différence d'intensité par diffraction, donc croissante avec la fréquence. Cela simule en fait la création d'un angle entre les microphones "reconstitués", angle d'autant plus important que la fréquence est élevée ; ceci est très proche du principe de la perception binaurale naturelle.

Un des avantages de ce système est d'avoir une grande latitude du décodage en post-production afin de s'adapter au mieux à la situation de prise de son.

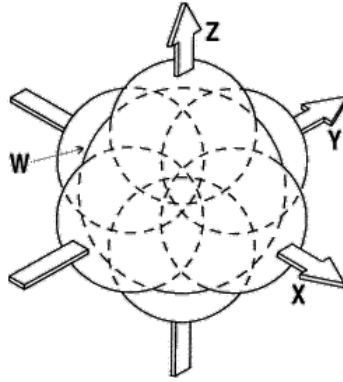
## 2.7. Le système du Grenier à Son

Enfin, le système présenté par Le Grenier à Son (P-A.Signoret et J.Warnier), est un mélange de deux techniques : la technique Ambisonic de M.Gerzon ([10]) avec le système Soundfield Mark V, et la technique binaurale avec une tête artificielle Sennheiser équipée de 2 microphones BK4003. Les deux systèmes étaient posés l'un sur l'autre à l'enregistrement (la tête étant dirigée vers l'avant).

Le microphone Mark V est constitué de 4 capsules hypocarboïdes quasi-coïncidentes disposées en tétraèdre (Figure 14) qui captent l'information de l'espace sonore environnant. Ces 4 signaux (ou Format A) donnent après matricage le Format B, correspondant aux signaux qu'auraient fourni un microphone omnidirectionnel positionné à l'endroit du Mark V (signal W) ainsi que 3 microphones bidirectionnels orientés selon 3 directions orthogonales de l'espace (signaux X, Y et Z) ; cela correspond en quelque sorte à une base acoustique (au sens mathématique) de l'espace (Figure 15). Ce Format B peut alors être traité en post-production de façon à s'adapter à toute situation de diffusion ; pour l'expérience, il a été décodé pour une restitution 5.1 à l'aide du processeur Soundfield DSP451.



Figure 14



**Figure 15**

Le Grenier à Son a voulu exploiter les qualités de localisation de ce système tout en améliorant la sensation d'enveloppement et le timbre de celui-ci ; il a alors été envisagé de lui adjoindre un système complémentaire constitué par la tête artificielle. Les signaux captés par celle-ci ont été mélangés aux canaux gauche et droit. Les concepteurs précisent qu'ils auraient souhaité traiter les signaux de la tête artificielle par un système transaural, mais n'ont pas disposé du temps nécessaire.

Nous allons maintenant nous attacher à décrire l'analyse des résultats des écoutes comparatives organisées au cours de la Convention de l'AES.

## **Troisième partie**

### **L'analyse des commentaires d'écoute de la Convention de l'AES**

### **3.1. La matière**

Les séances d'écoute de la Convention de l'AES se sont déroulées comme suit : un extrait du motet et de l'improvisation aux percussions pour chaque système était diffusé à l'auditoire. Afin de récolter le plus d'informations possibles, nous avons distribué à chaque auditeur une feuille blanche, contenant uniquement une case prévue pour chaque extrait afin qu'ils inscrivent librement leurs commentaires. C'est l'analyse de ces feuilles, très riches d'informations que nous allons présenter dans cette partie.

Il est à noter que l'ordre de diffusion (ordre dans lequel les systèmes étaient présentés, le même pour le motet ou l'improvisation) n'était communiqué aux auditeurs qu'à la fin de la séance, cela pour éviter tout préjugé - ou parti pris - qui aurait pu entacher leurs commentaires. 3 ordres différents étaient enregistrés afin de les changer au cours des séances.

Pour chaque séance, 20 auditeurs étaient disposés en 4 rangées de 5 sièges ; le siège du milieu de la 3<sup>e</sup> rangée se situait exactement au centre du cercle constitué par les 5 enceintes, de 8m de diamètre. Certaines places étaient donc relativement excentrées, c'est pourquoi il était important pour l'exploitation des commentaires que les auditeurs indiquent la place qu'ils occupaient au cours de la séance.

Suite à ces séances, près de 400 feuilles m'ont été remises. Deux exemples de feuilles de commentaires remplies (une en français et une en anglais) sont donnés Annexe 3.

### **3.2. L'analyse qualitative**

Après élimination des feuilles illisibles, inexploitables, ou n'étant ni en anglais ou en français, il restait environ 300 feuilles à analyser.

La première étape a été une analyse qualitative des commentaires afin d'élaborer une grille de lecture commune à l'ensemble des feuilles. En effet, du fait de la liberté de commentaire, la quantité de termes ou de descriptions utilisés était très importante et il fallait dégager les critères sous-jacents. Globalement, les auditeurs appartenaient au monde de l'audio et utilisaient donc un vocabulaire assez précis et explicite, mais une partie des auditeurs ne possédait pas le vocabulaire des ingénieurs du son et il fallait "décoder" leurs commentaires.

Une première lecture des feuilles a permis, après regroupements des termes synonymes ou similaires, d'élaborer une grille de lecture commune se voulant à la fois exhaustive et non redondante. Cette grille pouvait elle-même se subdiviser en 5 grandes catégories : spatialisation, localisation, timbre, définition et dynamique. La grille de lecture ainsi élaborée est donnée Figure 16.



		Motet	Percussions
<b>Spatialisation</b>			
<b>Equilibre avant / arrière</b>	Bon équilibre		
	Déséquilibré		
	Trop d'avant		
	Trop d'arrière		
<b>Côtés</b>	Présents		
	Absents		
	Trop présents		
	Eloignés		
	Peu clairs		
<b>Avant</b>	Timbre pauvre		
	Bien reproduit		
	Mal reproduit		
	Trou au centre		
	Etroit		
<b>Arrière</b>	Large		
	Bien reproduit		
	Mal reproduit		
<b>Trop de</b>	Trou au centre		
	Gauche		
	Centre		
	Droit		
	Arrière Gauche		
<b>Manque de</b>	Arrière Droit		
	Gauche		
	Centre		
	Droit		
	Arrière Gauche		
<b>Haut-parleurs</b>	Arrière Droit		
	Trop perçus		
<b>Profondeur</b>	Se font oublier		
	Bonne, avec du relief		
<b>Largeur des sources</b>	Mauvais relief		
	Trop faible		
<b>Enveloppement</b>	Trop importante		
	Ressenti		
<b>Image</b>	Non ressenti		
	Equilibrée, homogène		
	Peu équilibrée		
	Réaliste, naturelle		
	Peu naturelle		
	Manque d'air		
<b>Localisation</b>	Agréable		
	Bonne		
	Mauvaise		
	Mauvaise à l'arrière		
	Confusions avant/arrière		
<b>Stabilité</b>	Stable		
	Instable		
<b>Précision</b>	Précise		
	Diffuse		
<b>Sensation</b>	D'élévation du son		
	De son dans la tête		
<b>Problème de phase</b>			

		Motet	Percussions
<b>Timbre</b>			
<b>Qualité</b>	Beau		
	Pas très beau		
	Agressif, dur, acide		
	Doux, rond, chaud		
	Pauvre		
<b>Aigus</b>	Terne		
	Clair		
	Trop		
	Manquent		
<b>Médiums</b>	Bons		
	Détimbré, manque de corps		
	Bons		
<b>Graves</b>	Trop		
	Manquent		
<b>Définition</b>			
<b>Précision</b>	Bonne		
	Confuse		
<b>Dynamique</b>			
<b>Qualité</b>	Grande		
	Faible		

Figure 16

### 3.3. L'analyse quantitative

La deuxième étape a été de faire entrer chaque feuille de commentaire dans cette grille d'analyse. Afin de ne pas mélanger les commentaires de personnes se trouvant à des places d'écoute différentes et ayant donc une perception de l'image sonore différente, une grille d'analyse a été remplie pour chaque position d'écoute.

Un des intérêts de cette expérience était la quantité d'avis recueillis, avis la plupart du temps éclairés car émanant d'un public appartenant au monde de l'audio ; il fallait donc tirer parti de cette richesse de données. L'idée a été de regrouper les grilles de lecture d'une zone centrale constituée par la place située au centre du système de diffusion ainsi que les sièges adjacents, ce qui constituait donc une zone de 9 places. Au vu des dimensions du cercle d'écoute, cette zone pouvait être considérée comme ce qu'on pourrait nommer une "sweet area"; cela éliminait les sièges trop excentrés.

La Figure 17 fournit la grille d'analyse concernant cette zone centrale, pour chacun des systèmes et chaque extrait

Spatialisation		Motet							Percussions							Total
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	
<b>Equilibre avant / arrière</b>	Bon équilibre	15	21	9	6	7	10	1	8	11	6	4	7	2	5	112
	Déséquilibré		1				1								2	
	Trop d'avant	4	12	13	5	2	12	7	6	9	14	3	7	13	2	109
	Trop d'arrière	20	7	18	24	39	33	19	10	7	9	10	30	31	24	281
		<b>RF</b>	<b>J/M</b>	<b>KFM</b>	<b>CNSM</b>	<b>LGS</b>	<b>BR</b>	<b>W/LD</b>	<b>RF</b>	<b>J/M</b>	<b>KFM</b>	<b>CNSM</b>	<b>LGS</b>	<b>BR</b>	<b>W/LD</b>	<b>504</b>
<b>Côtés</b>	Présents	11	8	6	8	11	4	6	2	3	11	5	6	2	4	87
	Absents	12	15	8	5	10	10	7	3	5	2	3	10	3	2	95
	Trop présents	4	1		3	1	4	1	2		2	1		1	1	21
	Eloignés	1										1				1
	Peu clairs											1				1
	Timbre pauvre				1											1
		<b>RF</b>	<b>J/M</b>	<b>KFM</b>	<b>CNSM</b>	<b>LGS</b>	<b>BR</b>	<b>W/LD</b>	<b>RF</b>	<b>J/M</b>	<b>KFM</b>	<b>CNSM</b>	<b>LGS</b>	<b>BR</b>	<b>W/LD</b>	<b>206</b>
<b>Avant</b>	Bien reproduit	1	7		1	1	1		1	1	1		1			15
	Mal reproduit	1	1	1	1	1	3		1	1	1		1	3		15
	Trou au centre	5	1	2	2	1	4	11	1	1		2	1	6	2	39
	Etroit								1							1
	Large											1				1
		<b>RF</b>	<b>J/M</b>	<b>KFM</b>	<b>CNSM</b>	<b>LGS</b>	<b>BR</b>	<b>W/LD</b>	<b>RF</b>	<b>J/M</b>	<b>KFM</b>	<b>CNSM</b>	<b>LGS</b>	<b>BR</b>	<b>W/LD</b>	<b>71</b>
<b>Arrière</b>	Bien reproduit	1	2	2	1	2	2	1	1			2	3	1	1	19
	Mal reproduit	1	1		1		1		1	1		2			1	9
	Trou au centre	5	5	1	3	4	3		1	1	1	3	2		1	30
		<b>RF</b>	<b>J/M</b>	<b>KFM</b>	<b>CNSM</b>	<b>LGS</b>	<b>BR</b>	<b>W/LD</b>	<b>RF</b>	<b>J/M</b>	<b>KFM</b>	<b>CNSM</b>	<b>LGS</b>	<b>BR</b>	<b>W/LD</b>	<b>58</b>
<b>Trop de</b>	Gauche	1		4	2		8		1		10	3	1	7	4	41
	Centre		1		1	1	2		1	7	1	1		1	1	17
	Droit		1	3				2		3	1		4	2	3	19
	Arrière Gauche			1		2	3	2				1	1	5	3	18
	Arrière Droit	2				1	2	1			2	1	2	3	1	15
<b>Manque de</b>	Gauche		1			3				1			1			6
	Centre						3									3
	Droit		2			1		2	1		1	2	1	1	1	12
	Arrière Gauche										2	1			1	4
	Arrière Droit		1													1

Figure 17 (début)

		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	
<b>Haut-parleurs</b>	Trop perçus	12	9	17	9	10	5	18	5	3	5	3	6	3	4	109
	Se font oublier	3	2	4	3	3			2	1			4		1	25
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	134
<b>Profondeur</b>	Bonne, avec du relief	1	3	1	1	5	2	3	6	5	3	4	2			36
	Mauvais relief	5	5	5	3	4		4	9	9	9	3	11	9	3	79
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	115
<b>Largeur sources</b>	Trop faible	3		3	1			3	1	2	1	1	1		1	17
	Trop importante	1		1					5	1	2	6	7	3	2	28
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	45
<b>Enveloppement</b>	Ressenti	17	16	16	12	9	5	9	8	10	3	9	1	2	2	119
	Non ressenti	1	1	1	2	1	2			1	1	1		1	1	13
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	132
<b>Image</b>	Équilibrée, homogène	6	13	10	7	8	3		13	11	8	7	6	2	5	99
	Peu équilibrée	3		2	1	2	5	1			1	4		1	1	21
																120
	Réaliste, naturelle	10	12	11	7	5	9	4	23	7	14	11	2		4	119
	Peu naturelle					3	1	1		1		1	3	1		11
																130
	Manque d'air	1				1										
	Agréable		1	2					5	3	2	2		1		
<b>Localisation</b>		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	
<b>Justesse</b>	Bonne	11	14	13	8	4	3	11	23	39	11	18	10	3	16	184
	Mauvaise	3		2	6	3	6		7	3	6	11	10	24	7	88
	Mauvaise à l'arrière											2				2
	Confusions avant/arrière	1					3			2	1	3	3	4	4	21
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	295
<b>Stabilité</b>	Stable	1	1	1			2	1	3	4	2	2	1		2	20
	Instable	7	11	9	5	7	1	7	5	4	10	10	10	5	10	101
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	121
<b>Précision</b>	Précise	6	7	5	6	4	7	7	11	9	10	7	6		4	89
	Diffuse	4	4	4	3	7	8	1	6	2	8	5	7	18	12	89
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	178
<b>Sensation</b>	D'élévation du son	1	4	2	1	1	3	1	2	2	3	1	3	21	5	
	De son dans la tête		1	1					2		3					
<b>Problème de phase</b>		1		1	1	2	3		1	1	5	1		2		
<b>Timbre</b>		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	
<b>Qualité</b>	Beau	11	11	9	7	4	5	4	13	6	7	7	4	3	5	96
	Pas très beau		1	2	2	5	7	1		1	5	4	3	6	1	38
	Agressif, dur, acide	1	1			1	3	2		2	1			1		12
	Doux, rond, chaud	1	3	4	2			4		1					3	18
	Pauvre			1			3	2								6
	Terne				2	4					1	2	1			10
	Clair	2			1	1	1	1	1						1	8
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	188
<b>Aigus</b>	Bons	1	1	1	1		2	1	1			2				10
	Trop			2	1										1	4
	Manquent		1		3											4
<b>Médiums</b>	Bons	1														1
	Trop			2	1					2						5
	Détimbré, manque de corps							1	1							2
<b>Graves</b>	Bons			6			1		1	1	4		5	1		19
	Trop			3	1				2		3	1				10
	Manquent	1	1	1		4	2	3				1	2	3	3	21
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	76
<b>Définition</b>		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	
<b>Précision</b>	Bonne	9	7	3	2	3	4	5	10	11	5	4	6		3	72
	Confuse	3	6	10	15	14	11	7	4	2	5	7	5	12	5	106
		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	178
<b>Dynamique</b>		RF	J/M	KFM	CNSM	LGS	BR	W/LD	RF	J/M	KFM	CNSM	LGS	BR	W/LD	
<b>Qualité</b>	Grande			1			1		1					1		4
	Faible		1					1		1						3
																7

Figure 17 (fin)

### 3.4. Conclusions sur l'analyse qualitative

Un des grands apports de ces séances d'écoute, par delà la comparaison à proprement dit des différents systèmes, tient dans l'analyse qualitative des résultats. En effet, cette première analyse nous montre à quoi les auditeurs sont sensibles lors d'une diffusion d'enregistrement multicanal et quels sont les critères principaux qu'ils utilisent pour juger une image sonore multicanale. Le fait d'avoir laissé complètement ouvert le champ d'expression, s'il rend difficile une exploitation statistique, permet par contre d'obtenir des résultats très riches quant à la perception et au jugement d'images sonores multicanales.

On a vu que l'on peut séparer les commentaires en 5 grandes catégories : spatialisation, localisation, timbre, définition et dynamique. Si les commentaires concernant les 4 dernières catégories se rapprochent de ceux utilisés par les ingénieurs du son dans la description d'images stéréophoniques à 2 canaux, les commentaires concernant la spatialisation apportent des données nouvelles.

Les chiffres en gras de la dernière colonne de la Figure 17 correspondent au nombre total de commentaires exprimés par catégorie.

On constate que les auditeurs sont très sensibles à l'équilibre entre les voies avant et arrière : c'est le critère le plus abordé, tous critères confondus.

Le 2<sup>e</sup> critère abordé dans le cadre de la spatialisation concerne les sources fantômes sur les côtés (entre les enceintes avant et arrière), et on sait que c'est justement un des points critiques des systèmes de prise de son (et de restitution) multicanale.

Les commentaires, assez nombreux, concernant le manque ou l'excès de niveau de telle ou telle voie sont liés au positionnement précis de chaque auditeur et ne constituent donc pas des données exploitables.

Par contre, les critères abordés ensuite et qui concernent la perception (ou l'oubli) des haut-parleurs en tant que sources discrètes, les sensations d'enveloppement, d'équilibre (ou homogénéité) et de réalisme (ou naturel) nous indiquent une des attentes principales des auditeurs d'une diffusion multicanale (et particulièrement pour les scènes sonores réparties sur 360° comme c'était le cas dans l'expérience) : être plongé dans le son, avoir la sensation d'être avec les musiciens (sensation de naturel).

Outre ces critères de spatialisation, les critères de localisation (tant en justesse qu'en précision et en stabilité), de timbre et de définition restent des critères fondamentaux.

Si l'on classe les critères en fonction du nombre d'auditeurs y ayant fait référence, nous obtenons l'ordre suivant :

1. Equilibre entre les voies avant et arrière
2. Justesse de la localisation
3. Présence de sources fantômes sur les côtés
4. Précision de la localisation
5. Qualité de timbre
6. Définition
7. Oubli des haut-parleurs en tant que sources discrètes
8. Sensation d'enveloppement
9. Sensation de naturel
10. Stabilité des sources
11. Homogénéité

Cela nous permet de cerner l'attente des auditeurs pour des images sonores multicanales.

### **3.5. Conclusions sur l'analyse quantitative**

Si l'interprétation qualitative est d'un apport riche, l'interprétation quantitative des commentaires est plus délicate. En effet, le principe du commentaire libre, qui était d'un grand intérêt pour la connaissance et la hiérarchisation des critères utilisés, se prête mal à une interprétation statistique.

Le nombre absolu de commentaires sur tel ou tel critère ne constitue pas nécessairement une donnée exploitable. En effet, si un auditeur ne met aucun commentaire sur un critère, est-ce parce qu'il l'a jugé correct ou parce qu'il n'a pas pensé à l'évoquer ?

D'autre part, il apparaît sur la Figure 17 que certains critères sont évoqués essentiellement s'ils sont mauvais (par exemple la perception des haut-parleurs en tant que sources discrètes) ou au contraire s'ils sont bons (par exemple la sensation d'homogénéité ou de naturel).

Enfin, tirer une statistique sur des critères pour lesquels peu d'auditeurs se sont exprimés (par exemple la largeur des sources) serait de peu de validité : si un seul auditeur s'exprime sur un critère et le juge mauvais, il serait osé de dire que 100% des auditeurs exprimés jugent ce critère mauvais...

Néanmoins, étant donné que chaque auditeur s'est exprimé sur tous les systèmes, les résultats obtenus nous fournissent un bon jugement comparatif des différents enregistrements. On pourra objecter que l'écoute en groupe fausse le jugement ; en effet, tous les auditeurs de la zone centrale prise en compte et nommée "sweet area" ne se situaient pas précisément au centre du système de restitution ("sweet spot") - d'ailleurs, même l'auditeur situé à cette place centrale subissait le masquage des personnes situées autour de lui. Mais les auditeurs, chez eux, seront-ils toujours seuls et situés exactement au centre du système de restitution (en supposant que celui-ci soit correctement installé...) ?

Ces séances, si elles ne pouvaient se substituer à des écoutes individuelles, constituaient indéniablement une très bonne base de jugement subjectif. Afin d'obtenir une échelle de comparaison chiffrée des différents systèmes, je me suis limité aux 11 critères principaux vus plus haut et ai choisi d'exploiter les résultats en adoptant le procédé suivant, présenté Figure 18 :

- pour chaque critère de la grille d'analyse, regroupement des commentaires positifs (+) et des commentaires négatifs (-) d'après la grille d'analyse. Ce classement, critère par critère, donne :
  1. Equilibre entre les voies avant et arrière :
    - + : bon équilibre
    - : déséquilibré, trop d'avant, trop d'arrière
  2. Justesse de la localisation :
    - + : bonne
    - : mauvaise, mauvaise à l'arrière, confusions avant/arrière
  3. Présence de sources fantômes sur les côtés :
    - + : présents, trop présents (ce commentaire, qui peut surprendre, fait probablement allusion à un défaut d'homogénéité entre les côtés et le reste du cercle de restitution, mais traduit néanmoins la présence de sources sur les côtés)
    - : absents, éloignés (les commentaires "peu clairs" et "timbre pauvre" se rattachent plutôt à des problèmes de définition et de rendu spectral)
  4. Précision de la localisation :
    - + : précise
    - : diffuse
  5. Qualité de timbre (et d'équilibre spectral) :
    - + : beau, doux/rond/beau, clair, bons aigus, bons mediums, bons graves
    - : pas très beau, agressif/dur/acide, pauvre, terne, trop d'aigus, aigus manquent, trop de mediums, détimbré/manque de corps, trop de graves, graves manquent
  6. Définition :
    - + : bonne
    - : confuse
  7. Oubli des haut-parleurs en tant que sources discrètes :
    - + : se font oublier
    - : trop perçus
  8. Sensation d'enveloppement :
    - + : ressenti
    - : non ressenti
  9. Sensation de naturel :
    - + : réaliste/naturelle
    - : peu naturelle
  10. Stabilité des sources :
    - + : stable
    - : instable
  11. Homogénéité :
    - + : équilibrée/homogène
    - : peu équilibrée

- pour chaque système de prise de son et chacune des œuvres, calcul de la somme et de la différence des commentaires positifs ( $n_1$ ) et négatifs ( $n_2$ ) :

$$s = n_1 + n_2$$

$$d = n_1 - n_2$$

puis du coefficient :

$$c = d/s.$$

Ce coefficient, compris entre -1 et +1, traduit le jugement des auditeurs sur le critère (-1 signifie qu'il a été très mal jugé, +1 qu'il a été très bien jugé, 0 que les avis ont été partagés).

Afin d'avoir une validité statistique, j'ai choisi de considérer ce coefficient uniquement si au moins 10 auditeurs se sont exprimés sur le critère (c'est-à-dire si  $s$  est au moins égal à 10), dans le cas contraire il n'a pas été évalué (NN)

- pour chaque système, moyenne du coefficient  $c$  sur les deux œuvres et classement des systèmes sur chaque critère d'après cette moyenne
- moyenne générale pour chaque système et classement sur l'ensemble des critères

Cela est présenté dans le tableau de la Figure 18, page suivante.





### **3.6. Conclusion**

Cette première expérience nous a conduit vers une connaissance et une hiérarchisation des principaux critères d'évaluation d'une image sonore multicanale par le public. De plus, elle a fourni un classement des enregistrements réalisés par les 7 systèmes selon chaque critère ainsi qu'un classement général, d'après les séances d'écoute communes réalisées à la Convention de l'AES.

Afin d'aller plus loin dans l'évaluation des différents systèmes, j'ai mis en place une deuxième série de tests au CNSMDP, sur 4 jours en novembre 2000. Il s'agissait d'écoutes individuelles, avec un système disposé de façon plus resserrée que pour les test de la Convention de l'AES (cercle constitué par les enceintes de 3m de diamètre). Ces conditions plus "expérimentales" allaient permettre de s'affranchir des limites de l'écoute en groupe, évoquées plus haut. De plus, la méthode d'évaluation choisie allait permettre une interprétation statistique solide des résultats. Ce test individuel d'analyse d'écoute est présenté ci-dessous.

## **Quatrième partie**

### **Les tests individuels d'analyse d'écoute**

#### 4.1. L'élaboration et le déroulement

L'idée était d'approfondir les résultats des séances d'écoute de la Convention de l'AES, dans des conditions d'écoute individuelles et une interprétation statistique rigoureuse. Ces séances d'écoutes étaient basées sur la comparaison de l'extrait du chœur diffusé au cours de la Convention.

Contrairement aux séances d'écoute de la Convention, les commentaires des auditeurs n'allaient plus être libres mais guidés par une grille d'analyse avec une échelle de notation. L'élaboration de cette grille d'analyse s'est faite à partir des résultats des séances d'écoutes de la Convention : les critères choisis ont été les 10 critères principaux qui s'étaient dégagés de l'analyse qualitative des commentaires de la Convention (seul le critère "justesse de localisation" n'était pas soumis, vu que la localisation allait faire l'objet de tests complémentaires). La description des critères fournie aux auditeurs ainsi que la grille de notation dont ils disposaient sont données en Annexe 4 ; cette notation était basée sur une échelle de 1 à 5 (1 : très mauvais ; 5 : très bon).

Deux constatations qui s'étaient dégagées du dépouillement des feuilles de commentaires de la Convention de l'AES ont été prises en compte dans l'élaboration de ce test. La première est que les auditeurs, lors des séances d'écoute, avaient tendance à se montrer très sévère dans le jugement du premier extrait, quel qu'il soit ; cela venant certainement du fait que les auditeurs ne sont pas habitués à écouter en multicanal, manquent de références et ont peut-être des exigences importantes (qu'ils pondèrent au fur et à mesure des écoutes...). Par conséquent, il a été décidé dans l'organisation des séquences d'écoute de cette deuxième série de tests de ne pas tenir compte du jugement concernant le premier enregistrement diffusé, qui était répété à la fin de la séquence (l'auditeur n'étant évidemment pas au courant).

Une deuxième constatation était apparue lors du dépouillement des feuilles de commentaires : il semblait qu'un certain nombre d'auditeurs avaient tendance à juger les systèmes non de manière absolue, mais en fonction du système précédent. Or, même si 3 séquences étaient alternées lors des séances d'écoute de la Convention, il s'agissait du même ordre de déroulement : si on repère les systèmes par des lettres, on avait comme séquences ABCDEFG, DEFGABC, FGABCDE. Ainsi, chaque système était toujours précédé du même système (sauf s'il était au début), ce qui pouvait biaiser le jugement. Dans les 3 séquences créées pour la série de tests individuels, j'ai pris soin de changer systématiquement l'enchaînement. L'idéal aurait bien sûr été d'avoir un ordre aléatoire, mais cela aurait été lourd de mise en œuvre et n'aurait pas nécessairement modifié les résultats sur une petite population telle que celle ayant participé à cette deuxième série de tests.

La diffusion s'effectuait par des enceintes JBL 4412 et le monitoring par une console Yamaha O3D. 15 personnes ont pris part à ces écoutes individuelles ; il s'agissait d'étudiants avancés de la FSMS et d'ingénieurs du son, donc de personnes à même de juger les critères proposés.

## 4.2. Les résultats et la comparaison avec les écoutes collectives

Les tableaux présentant les résultats complets pour chacun des systèmes sont donnés en Annexe 5. La Figure 19 récapitule l'ensemble des moyennes et établit le classement des différents systèmes à partir des moyennes générales.

	Radio France	KFM360	CNSMDP	Brauner	Jacques/Mora	Williams/Le Dû	Le Grenier à Son
<b>Equilibre av/ar</b>	3,73	3,53	3,2	4	3,73	2,2	2,73
<b>Sources côtés</b>	3,47	3,67	3,6	3	3,33	3,4	3
<b>Présence HP</b>	3,27	2,87	3,27	3,6	3,2	3,2	2,87
<b>Enveloppement</b>	4,07	3,67	3,47	3,93	3,33	2,8	3,2
<b>Homogénéité</b>	3,2	2,87	3,47	3,6	3,07	2,27	2,67
<b>Naturel</b>	3,4	3,4	3,4	3,33	3,13	2,47	2,53
<b>Précision loca.</b>	3,53	3,8	3,4	2,47	3,2	3,07	2,87
<b>Stabilité loca.</b>	3,53	3,87	3,47	3,27	3,67	3,27	3,27
<b>Timbre</b>	3,4	3,53	3,33	3,33	3,2	3,33	2,47
<b>Définition</b>	3,53	3,4	2,93	2,47	2,8	2,87	2,27
<b>Moy. générale</b>	3,51	3,46	3,35	3,3	3,27	2,89	2,79
<b>Classement</b>	1	2	3	4	5	6	7

Figure 19

Il était alors intéressant de comparer ces résultats à ceux des séances d'écoute de la Convention de l'AES ; la Figure 20 présente la comparaison du classement pour chaque système et chaque critère entre les deux tests (le classement pour la Convention avait été établi Figure 18). On peut globalement noter une grande similitude entre ces classements - ce qui confirme la validité des séances d'écoute de la Convention de l'AES - cependant, certains résultats sont différents (on a grisé les différences notables Figure 20). Pour essayer de comprendre ces divergences, il faut répertorier les différences de condition d'écoute entre les deux tests ; celles-ci sont de plusieurs ordres.

		Radio France	KFM360	CNSMDP	Brauner	Jacques/Mora	Williams/Le Dû	Le Grenier à Son
<b>Equilibre av/ar</b>	Test	2ex	4	5	1	2ex	7	6
	AES	2	3	4	6	1	7	5
<b>Sources côtés</b>	Test	3	1	2	6ex	5	4	6ex
	AES	3	2	1	6	7	4	5
<b>Présence HP</b>	Test	2ex	6ex	2ex	1	4ex	4ex	6ex
	AES	3	4	2	NN	5	6	1
<b>Enveloppement</b>	Test	1	3	4	2	5	7	6
	AES	1	2	5	NN	3	NN	4
<b>Homogénéité</b>	Test	3	5	2	1	4	7	6
	AES	1ex	3	5	NN	1ex	NN	4
<b>Naturel</b>	Test	1ex	1ex	1ex	4	5	7	6
	AES	1ex	1ex	4	5	1ex	NN	NN
<b>Précision localis.</b>	Test	2	1	3	7	4	5	6
	AES	2	4	3	7	1	6	5
<b>Stabilité localis.</b>	Test	3	1	4	5ex	2	5ex	5ex
	AES	NN	3	1ex	NN	5	1ex	4
<b>Timbre</b>	Test	2	1	3ex	3ex	6	3ex	7
	AES	1	3ex	5	7	2	3ex	6
<b>Définition</b>	Test	1	2	3	6	5	4	7
	AES	1	4	6	7	2	3	5
<b>Moyenne générale</b>	Test	1	2	3	4	5	6	7
	AES	1	3	4	7	2	6	5

Figure 20

Tout d'abord, les enceintes de diffusion étaient différentes : Cabasse Sirocco lors de la Convention de l'AES, JBL 4412 lors des séances au CNSMDP. Or l'équilibre spectral des enceintes, leur "couleur" propre, intervient sur le jugement de timbre. Ainsi, le timbre de l'enregistrement par le système Brauner est apparu meilleur sur les JBL que sur les Cabasses qui accentuaient certainement trop la brillance naturelle des microphones utilisés (VM1 en directivité hypercardioides), qui était par contre compensée par la coloration plus médium des JBL. Par contre, la coloration propre de l'enregistrement par le système de Nicolas Jacques et Arnaud Mora semble avoir été favorisée par les enceintes Cabasses, qui justement rajoutaient un peu de brillance.

Ensuite, il est à noter une différence fondamentale dans les expériences : les écoutes lors de la Convention de l'AES étaient collectives et les résultats concernent une zone centrale relativement étendue, alors que pour les tests individuels, les auditeurs étaient précisément au "sweet spot". Or on constate deux phénomènes intéressants dans les résultats : l'enregistrement réalisé avec le système du CNSMDP a été beaucoup plus apprécié à la fois en termes d'homogénéité, de naturel que de définition lors des tests d'écoute individuels que lors de la Convention alors que pour ces mêmes critères, ainsi qu'en termes de précision de localisation, l'enregistrement réalisé avec le système de Nicolas Jacques et Arnaud Mora a été beaucoup plus apprécié lors des écoutes collectives que lors des écoutes individuelles. Cela porte à croire que la zone d'écoute est assez restreinte pour le système du CNSMDP alors que le système de Nicolas Jacques et Arnaud Mora permet une zone d'écoute plus étendue. Cette zone d'écoute est aussi tributaire des enceintes : des enceintes très directives vont restreindre celle-ci, tout en améliorant le rendu pour le "sweet spot" lui-même ; d'ailleurs, si les écoutes collectives avaient été réalisées avec les JBL, on aurait certainement obtenu des résultats différents quant aux critères d'homogénéité, d'enveloppement, de définition et de précision de localisation dans la zone centrale.

Penchons nous aussi sur le critère d'équilibre avant/arrière concernant le système Brauner, jugé très différemment pour les deux tests. Si on se reporte à la Figure 20, on constate que, lors des écoutes de la Convention de l'AES, un grand nombre de personnes ont jugé l'enregistrement réalisé par ce système déséquilibré vers l'arrière mais un nombre également important de personnes l'ont jugé déséquilibré vers l'avant. Or, lors des écoutes individuelles, il a été jugé comme étant le mieux équilibré. On peut donc supposer que l'équilibre avant/arrière pour ce système est très critique selon la position d'écoute et nécessite un positionnement précis de l'auditeur entre les haut-parleurs avant et arrière.

Enfin, le local d'écoute et la dimension du système de restitution étaient aussi totalement différents. Lors de la Convention de l'AES, les écoutes se déroulaient dans une grande salle de conférence, relativement mate mais sans traitement acoustique particulier sur les murs, et le système de diffusion était disposé sur un cercle de grand diamètre (8m), très éloigné des murs. Par contre, les tests individuels se situaient dans un local beaucoup plus petit, avec des enceintes plus rapprochées (diamètre de 3m) et assez proches des murs. Ces conditions d'écoute différentes modifiaient certainement la perception de l'image sonore et l'appréciation esthétique de celle-ci. En particulier, je pense que les enregistrements assez réverbérés (en particulier celui réalisé par le système Brauner, situé très en hauteur lors des séances d'enregistrement) pouvaient apparaître assez gênants avec une écoute elle-même très écartée (cas des écoutes collectives) alors qu'ils pouvaient être plus appréciés dans un petit local avec une écoute rapprochée (cas des écoutes individuelles).

D'autre part, on a vu plus haut que l'équilibre entre les voies avant et arrière constitue un critère fondamental d'appréciation des enregistrements. Or, cet équilibre a été effectué à l'enregistrement ou au mixage dans des régies différentes suivant les systèmes. A la restitution, cet équilibre est dépendant de l'acoustique propre du lieu d'écoute, en particulier à cause de ses symétries ou dissymétries acoustiques entre l'avant et l'arrière et la présence ou non de parois réfléchissantes à l'arrière des enceintes renforçant leur niveau par la présence de premières réflexions. Les lieux de diffusion, différents pour la Convention de l'AES et les tests individuels, restituaient donc l'équilibre entre les voies avant et arrière de façon différente et favorisaient certains mixages au détriment d'autres. Cela peut aussi permettre d'expliquer les jugements différents concernant l'équilibre avant/arrière entre les deux tests.

### 4.3. Influence de l'ordre d'écoute

J'ai noté plus haut une tendance qui semblait se dégager du dépouillement des feuilles de réponse des écoutes collectives : les auditeurs jugeaient souvent les systèmes non de manière absolue mais en référence au précédent.

Afin d'observer cela au niveau des écoutes individuelles, j'ai analysé plus précisément la grille de réponse concernant le système Brauner, classé quatrième en moyenne générale sur les écoutes individuelles. Il se trouve que pour l'un des ordres de diffusion, il se situait juste après le système de Radio France, premier classé et pour un autre ordre, il succédait au système du Grenier à Son, dernier classé. Cela allait permettre de voir si le système précédent influençait le jugement.

La Figure 21, page suivante, reprend le tableau des résultats concernant le système Brauner, fourni en Annexe 5, en précisant pour chaque auditeur le système précédent (les lignes grisées correspondent au troisième ordre de diffusion, non pris en compte). On s'aperçoit dans le calcul des moyennes générales que celle-ci est de 2,64 lorsque le système précédent était celui de Radio France alors qu'elle est de 4,08 lorsque le système précédent était celui du Grenier à Son. Cet écart très conséquent met en évidence ce qui avait été pressenti pour les écoutes collectives : le jugement est d'autant plus favorable qu'il suit un système qui a été peu apprécié.

	Brauner										Ecart type	
	Equilibre av/ar	Sources côtés	Présence HP	Enveloppement	Homogénéité	Naturel	Précis. localis.	Stabilité localis.	Timbre	Définition		Moyenne
	5	4	5	5	4	3	2	3	2	3		
Après RF	5	1	4	5	3	2	3	3	5	3	3,4	
Après LGS	5	4	5	5	5	5	4	5	5	4	4,7	
	4	4	3	3	3	4	2	3	2	2		
Après RF	3	2	2	4	4	4	2	3	2	1	2,7	
Après LGS	4	5	5	5	5	5	1	4	5	2	4,1	
	4	3	3	4	4	3	2	1	4	2		
Après RF	2	1	2	4	2	2	1	1	3	2	2	
Après LGS	4	2	3	4	4	4	4	4	4	3	3,6	
	5	4	3	3	4	4	4	4	4	3		
Après RF	3	3	3	4	4	3	2	3	2	2	2,9	
Après LGS	5	5	5	5	3	4	2	5	5	4	4,3	
	5	1	4	1	3	2	2	3	2	2		
Après RF	1	2	2	2	2	2	3	4	2	2	2,2	
Après LGS	5	4	5	5	4	3	3	3	3	2	3,7	
Moyenne après RF	2,8	1,8	2,6	3,8	3	2,6	2,2	2,8	2,8	2	2,64	0,56
Moyenne après LGS	4,6	4	4,6	4,8	4,2	4,2	2,8	4,2	4,4	3	4,08	0,67

Figure 21

#### 4.4. Synthèse des résultats

Afin d'avoir une vision plus synthétique des résultats, il était intéressant d'essayer de regrouper les critères très corrélés afin d'en réduire le nombre. Les corrélations entre les critères pris deux par deux, sur l'ensemble des résultats des tableaux de l'Annexe 5 sont données Figure 22. Les corrélations élevées (supérieures à 0,5) y sont grisées. Au vu de ce tableau, on peut dégager deux groupes de critères à forte corrélation.

	Eq. av/ar	Sources côtés	Prés. HP	Enveloppement	Homogénéité	Naturel	Précis. loca.	Stabil. loca.	Timbre	Définit.
Eq. av/ar	1,00	0,38	0,42	0,54	0,58	0,57	0,21	0,24	0,49	0,34
Sources côtés		1,00	0,49	0,50	0,34	0,55	0,14	0,25	0,37	0,26
Prés. HP			1,00	0,58	0,51	0,49	0,17	0,33	0,45	0,30
Enveloppement				1,00	0,56	0,63	0,17	0,22	0,48	0,31
Homogénéité					1,00	0,55	0,21	0,16	0,30	0,26
Naturel						1,00	0,26	0,33	0,59	0,41
Précis. loca.							1,00	0,64	0,30	0,65
Stabil. loca.								1,00	0,36	0,55
Timbre									1,00	0,49
Définit.										1,00

Figure 22

Le premier regroupe l'équilibre avant/arrière, les sensations d'enveloppement, d'homogénéité et de naturel ainsi que la présence de sources sur les côtés et l'oubli des haut-parleurs : ces 6 critères décrivent l'espace recréé dans la restitution multicanale.

Le deuxième groupe est constitué des critères de précision et de stabilité de localisation ainsi que de définition : cet ensemble de 3 critères décrit plutôt la précision dans le rendu des sources elles-mêmes.

Le critère de timbre peut être classé à part (il concerne d'ailleurs à la fois le rendu spatial au niveau de la réverbération et le rendu des sources).

La Figure 23 donne la moyenne pour chacun des systèmes dans ces trois catégories et leur classement. Les Figures 24 et 25 présentent ces résultats sous forme graphique. Il est intéressant de noter, au niveau du classement des écoutes individuelles, que le système Brauner semble privilégier le rendu de l'espace au dépend du rendu des sources et, au contraire, que le KFM360 et, à moindre échelle, le Multichannel Microphone Array de M. Williams et G. Le Dû, semblent privilégier le rendu des sources par rapport à l'espace ; les autres systèmes ont plus d'homogénéité dans le rendu de ces deux critères.

Ces différents résultats seront exploités dans l'analyse approfondie de chaque système, dans la dernière partie.

	Radio France	KFM360	CNSMDP	Brauner	Jacques/Mora	Williams/Le Dû	Le Grenier à Son
<b>Espace</b>	3,52	3,33	3,4	3,58	3,3	2,72	2,83
<b>Classement</b>	2	4	3	1	5	7	6
<b>Sources</b>	3,53	3,69	3,27	2,73	3,22	3,07	2,8
<b>Classement</b>	2	1	3	7	4	5	6
<b>Timbre</b>	3,4	3,53	3,33	3,33	3,2	3,33	2,47
<b>Classement</b>	2	1	3ex	3ex	6	3ex	7

Figure 23

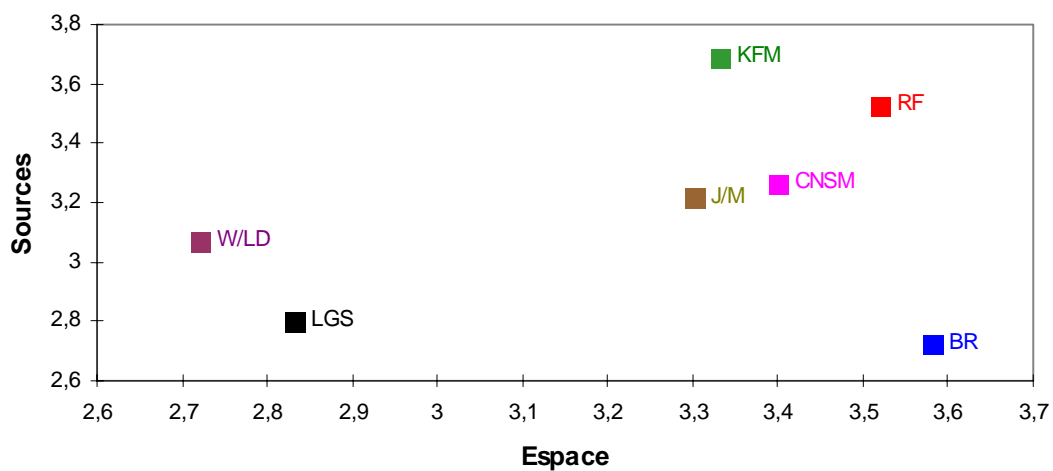


Figure 24

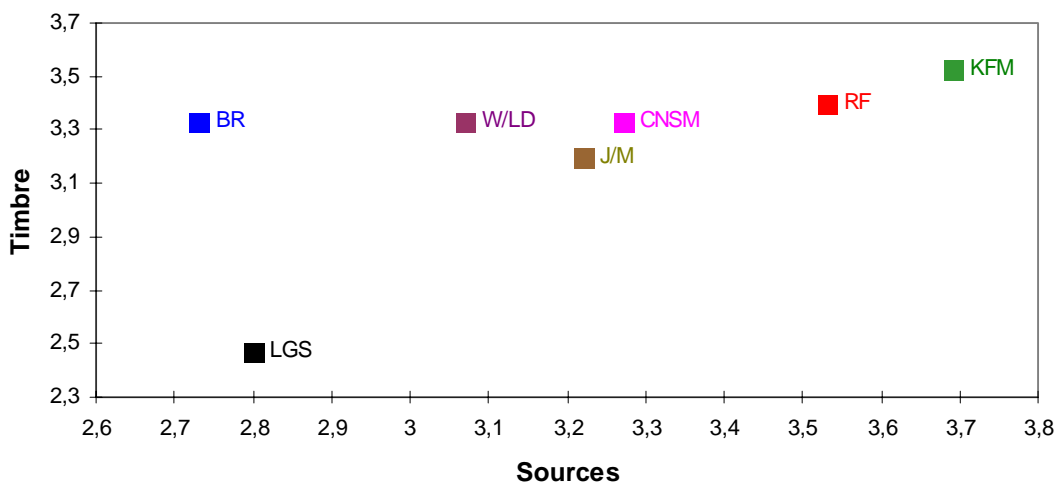


Figure 25

Suite à l'analyse des images sonores, je me suis penché sur les capacités des systèmes en localisation. Les deux parties suivantes sont consacrées à cela.



## **Cinquième partie**

### **Méthodologie des tests de localisation**

## 5.1. Méthodes de réponse pour des tests de localisation

Comme je l'ai précisé plus haut, la localisation allait faire partie d'un test indépendant, basé sur les enregistrements des signaux tests dans différentes directions. Au cours du test de localisation, les auditeurs allaient avoir à communiquer la direction perçue pour certains signaux tests enregistrés. Au moment de l'élaboration de ce test de localisation, je me suis questionné quant à la méthode à utiliser pour obtenir les réponses des auditeurs sur la direction perçue.

Pour les tests de localisation sur la base stéréophonique à 2 canaux, la méthode généralement adoptée consiste à placer des marqueurs visuels (par exemple des papiers suspendus sur une corde tendue entre les deux enceintes) régulièrement disposés (par exemple tous les 5 degrés) entre  $-30^\circ$  et  $+30^\circ$  ; l'auditeur doit alors indiquer celui correspondant à la direction perçue à l'écoute. Pour cela, l'auditeur n'a pas à bouger la tête étant donné que les marqueurs se situent dans son champ de vision. Pour de la localisation en multicanal, sur  $360^\circ$ , cela se complique et les méthodes utilisables sont plus diverses. M.J.Evans ([11]) a répertorié et discuté la validité théorique de plusieurs méthodes de réponse possibles pour les auditeurs.

La méthode la plus "sophistiquée" et la plus lourde à mettre en œuvre, mais peut-être la plus valide, consiste à utiliser une source sonore (haut-parleur sur un bras pivotant par exemple) que l'auditeur peut déplacer à partir de sa position d'écoute par une commande qui met en action le bras, afin de positionner cette source sonore dans la direction du signal perçu (diffusé par le système multicanal). La source sonore peut être choisie identique au signal diffusé (ce qui permet une meilleure comparaison, mais empêche une écoute simultanée avec le signal diffusé) ou totalement différente, tel qu'un clic (ce qui permet une écoute simultanée des deux sources).

Cette méthode, qui nécessite un appareillage évolué, peut être simplifiée et adaptée comme suit : après une première audition du signal, l'auditeur annonce à l'expérimentateur la direction approximative de provenance, en utilisant par exemple le système des heures (espace découpé en 12 directions, "midi" correspondant à l'avant, "1 heure" à  $+30^\circ$ , ...); l'expérimentateur se positionne dans la direction indiquée et déclenche la source sonore (par exemple un clic avec un métronome qu'il a en main) puis, au cours des diffusions suivantes, l'auditeur lui fait ajuster la position du clic en communiquant par des indications telles que "plus en avant" ou "plus en arrière", qui auront été décidées préalablement. Bien entendu, l'auditeur garde durant toute l'expérience la tête dirigée vers l'avant. Cette méthode sera ici dénommée M1.

Une autre méthode possible, beaucoup plus simple à mettre en place, consiste à faire noter à l'auditeur par une croix sur un cercle la direction perçue. Ce cercle symbolise l'ensemble des directions autour de l'auditeur. Pour cette méthode (M2), l'auditeur garde la tête dirigée vers l'avant (dans la mesure du possible, on place les feuilles de notation sur un pupitre suffisamment haut afin que l'auditeur n'ait pas à trop baisser la tête pour cocher).

Une autre méthode utilisée est l'adaptation de celle décrite plus haut pour les tests de localisation en stéréophonie à 2 canaux, qui utilise des marqueurs visuels. Elle consiste à disposer régulièrement sur une corde tendue le long du cercle formé par les 5 enceintes des papiers numérotés (par exemple tous les 5 degrés). L'auditeur, immobile au cours de la diffusion du signal, se tourne ensuite vers la direction perçue afin de repérer le marqueur correspondant à la direction perçue. Cela peut être validé par une deuxième diffusion. Nous nommerons cette méthode M3.

M.J.Evans ([11]) discute les avantages et inconvénients des différentes méthodes. L'avantage de M1 est que tous le test se passe dans le domaine auditif, mais la répétition du signal diffusé pour ajuster le positionnement du clic provoque de la fatigue auditive et cette méthode ne peut être utilisée sur un nombre important de signaux. L'avantage de la méthode M2 est que l'auditeur a accès à l'ensemble des directions possibles sur un continuum constitué par le cercle ; par contre, il peut être difficile de transcrire précisément une direction perçue (en particulier sur l'arrière) sur le cercle, sans références visuelles particulières. La méthode M3 permet d'avoir des références visuelles en se retournant, mais pour certains, l'information visuelle pourrait constituer un trouble de l'information auditive ; de plus, il n'y a pas simultanément entre l'audition du signal et le repérage visuel, ce qui nécessite une mémorisation de la direction perçue.

Afin d'avoir une idée plus précise de la validité de ces méthodes, j'ai décidé d'élaborer un test de comparaison de celles-ci, afin d'adopter pour les tests de localisation à venir la méthode de réponse se révélant la plus fiable.

## **5.2. Test de comparaison des méthodes**

L'idée était donc de comparer les méthodes M1, M2 et M3 décrites ci-dessus. Plus précisément, je souhaitais obtenir des résultats quant à l'écart des réponses fournies par les méthodes M2 et M3 par rapport à la méthode M1, que l'on peut considérer dans un premier temps comme la plus fiable, étant donné son processus qui ne fait intervenir que l'audition. De plus, je souhaitais obtenir un "degré de confiance" de chaque méthode, c'est-à-dire évaluer pour chacune d'elles l'écart de réponses si on présentait deux fois le même signal à un même auditeur (à des instants différents).

Le test allait donc se dérouler en trois phases, comportant les mêmes signaux, pour chaque auditeur, chaque phase utilisant une méthode de réponse différente (M1, M2 puis M3). Afin de ne pas fatiguer les auditeurs, j'ai choisi de travailler sur 4 directions données ; le signal était une voix d'homme diffusée sur 2 haut-parleurs chaque fois et positionnée par différence d'intensité. Le positionnement des signaux tests par différence d'intensité s'est fait en utilisant les résultats de L.Givernaud ([12]) concernant la localisation par différence d'intensité sur 360°. Le choix absolu des niveaux s'est effectué de sorte que le niveau perçu soit le même pour chaque signal (référence à 0 dB). Les 4 signaux ont donc été créés comme suit à partir du signal mono enregistré :

- signal 1 vers  $+60^\circ$  : D = -1,8 dB ; SD = -4,8 dB ;
- signal 2 vers  $-130^\circ$  : SG = -0,3 dB ; SD = -12,3 dB ;
- signal 3 vers  $-90^\circ$  : G = -4,1 dB ; SG = -2,1 dB ;
- signal 4 vers  $+160^\circ$  : SG = -5,5 dB ; SD = -1,5 dB.

(G, D, SG, SD désignent respectivement les haut-parleurs gauche, droit, arrière gauche et arrière droit).

Pour la méthode M1, les 4 signaux étaient diffusés 4 fois chacun afin de positionner correctement le clic. Cette phase était répétée 2 fois (pour juger du degré de confiance). Pour les méthodes M2 et M3, les 4 signaux apparaissaient également à 2 reprises et étaient mélangés à d'autres signaux, qui n'étaient pas pris en compte.

### 5.3. Les résultats

La Figure 26, page suivante, présente les résultats de ce test dans un tableau. Les 3 premières colonnes récapitulent, pour chacun des 4 signaux et chacun des 11 auditeurs, les valeurs (en degrés) fournies par chaque méthode de réponse, pour les 2 itérations du signal. Pour quelques auditeurs, et pour des contraintes d'organisation, certains signaux n'ont pu être présentés et NN apparaît alors dans la grille.

Lorsque, pour le signal positionné autour de  $+60^\circ$  (resp. les signaux positionnés autour de  $-130^\circ$  et  $+160^\circ$ ), les auditeurs ont fourni une réponse sur le demi-espace arrière (resp. avant), on a considéré qu'il s'agissait d'une erreur de confusion avant/arrière et ces résultats, grisés sur le tableau, n'ont pas été pris en compte. Ces indétermination, expliquées par Blauert ([13]), sont principalement dues au fait que 2 signaux symétriques par rapport au plan vertical passant par les oreilles provoquent une différence de temps et une différence d'intensité du même ordre et sont discriminés par les petits mouvements de tête qui permettent de faire varier les fonctions de transfert binaurales ; or, en conditions expérimentales, il arrive que cette discrimination ne se fasse pas (en particulier si les auditeurs se forcent à rester parfaitement immobiles, ce qui ne leur était pas imposé).

Les 6 colonnes suivantes contiennent l'angle moyen perçu et les écarts entre les 2 présentations du signal. Ces écarts permettent d'évaluer la fiabilité des méthodes.

Les 2 dernières colonnes contiennent l'écart absolu de réponse entre les méthodes M2 (resp. M3) et M1, ce qui permet d'évaluer la justesse des méthodes M2 et M3 par rapport à M1. Les moyennes et écart-types sont calculés pour chacun des 4 signaux ainsi que sur l'ensemble des résultats.

Ech.	Méthode 1 (M1)		Méthode 2 (M2)		Méthode 3 (M3)		Moy. M1	Moy. M2	Moy. M3	Ecart M1	Ecart M2	Ecart M3	M2-M1	M3-M1
1	55	NN	125	55	55	65	55	55	60	NN	NN	10	0	5
1	70	NN	105	120	75	60	70	NN	67,5	NN	NN	15	NN	2,5
1	45	60	65	100	70	65	52,5	65	67,5	15	NN	5	12,5	15
1	40	50	50	45	45	35	45	47,5	40	10	5	10	2,5	5
1	45	45	45	50	55	55	45	47,5	55	0	5	0	2,5	10
1	45	50	50	85	65	55	47,5	67,5	60	5	35	10	20	12,5
1	65	65	65	60	60	60	65	62,5	60	0	5	0	2,5	5
1	65	70	40	35	45	40	67,5	37,5	42,5	5	5	5	30	25
1	105	120	50	65	100	110	NN	57,5	NN	NN	15	NN	NN	NN
1	75	60	105	45	55	70	67,5	45	62,5	15	NN	15	22,5	5
1	70	NN	45	40	105	50	70	42,5	50	NN	5	NN	27,5	20
<b>Moy.</b>							<b>58,5</b>	<b>52,75</b>	<b>56,5</b>	<b>7,14</b>	<b>10,71</b>	<b>7,78</b>	<b>13,3333</b>	<b>10,5</b>
<b>Ec. type</b>							<b>10,55</b>	<b>10,24</b>	<b>9,59</b>	<b>6,36</b>	<b>11,34</b>	<b>5,65</b>	<b>11,92</b>	<b>7,53</b>
2	-110	NN	-140	-115	-120	-125	-110	-127,5	-122,5	NN	25	5	17,5	12,5
2	-145	NN	-115	-105	-110	-95	-145	-110	-102,5	NN	10	15	35	42,5
2	-125	-120	-125	-125	-110	-110	-122,5	-125	-110	5	0	0	2,5	12,5
2	-90	-75	-90	-85	-90	-115	-90	-90	-102,5	NN	NN	25	0	12,5
2	-135	-140	-120	-140	-130	-150	-137,5	-130	-140	5	20	20	7,5	2,5
2	-110	-105	-90	-120	-105	-105	-107,5	-105	-105	5	30	0	2,5	2,5
2	-135	-120	-145	-120	-145	-150	-127,5	-132,5	-147,5	15	25	5	5	20
2	-130	-65	-60	-90	-120	-130	-130	-90	-125	NN	NN	10	40	5
2	-115	-120	-135	-125	-130	-125	-117,5	-130	-127,5	5	10	5	12,5	10
2	-125	-140	-140	-130	-140	-135	-132,5	-135	-137,5	15	10	5	2,5	5
2	-125	NN	-155	-100	-130	-125	-125	-127,5	-127,5	NN	55	5	2,5	2,5
<b>Moy.</b>							<b>-122,27</b>	<b>-118,41</b>	<b>-122,5</b>	<b>8,33</b>	<b>20,56</b>	<b>8,64</b>	<b>11,59</b>	<b>11,59</b>
<b>Ec. type</b>							<b>15,43</b>	<b>16,78</b>	<b>15,73</b>	<b>5,16</b>	<b>16,09</b>	<b>8,09</b>	<b>13,84</b>	<b>11,69</b>
3	-75	NN	-120	-115	-110	-120	-75	-117,5	-115	NN	5	10	42,5	40
3	-110	NN	-60	-75	-100	-60	-110	-67,5	-80	NN	15	40	42,5	30
3	-110	-90	-115	-110	-115	-100	-100	-112,5	-107,5	20	5	15	12,5	7,5
3	-70	-80	-55	-60	-35	-50	-75	-57,5	-42,5	10	5	15	17,5	32,5
3	-90	-100	-115	-105	-75	-60	-95	-110	-67,5	10	10	15	15	27,5
3	-100	-105	-80	-105	-80	-75	-102,5	-92,5	-77,5	5	25	5	10	25
3	-110	-100	-110	-90	-120	-115	-105	-100	-117,5	10	20	5	5	12,5
3	-70	-70	-60	-70	-60	-70	-70	-65	-65	0	10	10	5	5
3	-95	-115	-135	-120	-95	-125	-105	-127,5	-110	20	15	30	22,5	5
3	-120	-80	-145	-90	-80	-75	-100	-117,5	-77,5	40	55	5	17,5	22,5
3	-125	NN	-140	-150	-105	-100	-125	-145	-102,5	NN	10	5	20	22,5
<b>Moy.</b>							<b>-96,59</b>	<b>-101,14</b>	<b>-87,5</b>	<b>14,38</b>	<b>15,91</b>	<b>14,09</b>	<b>19,09</b>	<b>20,91</b>
<b>Ec. type</b>							<b>16,82</b>	<b>27,89</b>	<b>24,44</b>	<b>12,37</b>	<b>14,46</b>	<b>11,36</b>	<b>12,86</b>	<b>11,85</b>
4	155	NN	150	145	155	150	155	147,5	152,5	NN	5	5	7,5	2,5
4	150	NN	135	135	155	150	150	135	152,5	NN	0	5	15	2,5
4	180	170	155	170	170	175	175	162,5	172,5	10	15	5	12,5	2,5
4	150	165	150	140	160	150	157,5	145	155	15	10	10	12,5	2,5
4	170	160	170	175	170	165	165	172,5	167,5	10	5	5	7,5	2,5
4	150	125	145	150	150	145	137,5	147,5	147,5	25	5	5	10	10
4	150	160	170	180	160	165	155	175	162,5	10	10	5	20	7,5
4	150	160	165	180	170	165	155	172,5	167,5	10	15	5	17,5	12,5
4	165	170	165	160	170	165	167,5	162,5	167,5	5	5	5	5	0
4	155	165	170	175	170	170	160	172,5	170	10	5	0	12,5	10
4	NN	NN	150	30	165	35	NN	150	165	NN	NN	NN	NN	NN
<b>Moy.</b>							<b>157,75</b>	<b>158,41</b>	<b>161,82</b>	<b>11,88</b>	<b>7,5</b>	<b>5</b>	<b>12</b>	<b>5,25</b>
<b>Ec. type</b>							<b>10,24</b>	<b>13,93</b>	<b>8,45</b>	<b>5,94</b>	<b>4,86</b>	<b>2,36</b>	<b>4,68</b>	<b>4,32</b>
<b>Moyenne globale</b>										<b>10,69</b>	<b>13,78</b>	<b>9,02</b>	<b>14,09</b>	<b>12,26</b>
<b>Ecart type global</b>										<b>8,32</b>	<b>12,99</b>	<b>8,23</b>	<b>11,51</b>	<b>10,82</b>

Figure 26

## 5.4. Conclusion

En observant les moyennes générales et les écart-types des 2 dernières colonnes, on peut constater que la méthode M3 conduit à des résultats légèrement plus proches de M1 que la méthode M2.

Mais surtout, en observant les colonnes des écarts, on constate que la méthode M3 semble plus fiable que M2 : écart moyen de 9 contre 13,8 avec, de plus, un écart-type plus faible, de 8,2 contre 13. C'est cela qui m'a conduit à adopter cette méthode de réponse pour les tests de localisation à venir.

Une remarque enfin concernant la méthode M1. Au vu des résultats, on peut douter de la précision de cette méthode qui conduit à un écart moyen de 10,7 (écart-type de 8,3). Mais il faut signaler que certains auditeurs se sont plaints de ne pas avoir suffisamment de répétitions du signal afin d'affiner le positionnement du clic, ce qui peut en partie expliquer ces écarts. D'autre part, la différence de matériau sonore entre le signal parlé diffusé et le métronome utilisé pour le clic pouvait rendre difficile leur superposition. Enfin - et cela est certainement une limitation de cette méthode - le signal diffusé pouvait parfois apparaître étalé en largeur à certains auditeurs, ce qui rendait difficile la superposition d'une source réelle très localisée. On ne peut donc pas affirmer qu'en l'état, cette méthode constitue une référence totalement fiable.

La partie suivante s'attache à la description des tests de localisation à proprement parler.

## **Sixième partie**

### **Les tests individuels de localisation**

## 6.1. Principe

Lors des séances d'enregistrement du chœur et des percussions organisées au CNSMDP, nous avons enregistré un certain nombre de signaux tests. Les signaux étaient de deux sortes : impulsifs (coups de wood-block et de maracas répétés) et entretenus (voix chantée a capella de Suzanne Vega, diffusée par une enceinte mobile). Le wood-block et les maracas ont été enregistrés tous les  $10^\circ$  sur le cercle occupé par les chanteurs (rayon de 4,77m) ; la voix a été diffusée tous les  $10^\circ$  sur ce même cercle, ainsi que tous les  $36^\circ$  sur des cercles concentriques de différents rayons (1,12m, 2,71m et 6,10m). Cela est représenté Figure 27.

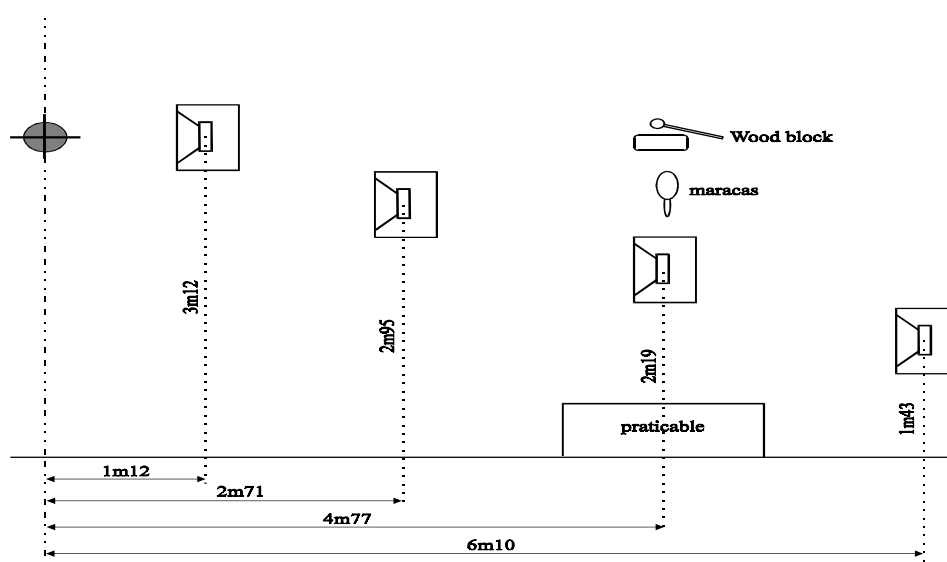


Figure 27

Cette matière a servi à monter des bandes pour des tests visant à évaluer les capacités de localisation des différents systèmes. En ce qui concerne la voix, et étant donné que l'on était dans une acoustique assez réverbérante, j'ai choisi d'utiliser les signaux diffusés sur le cercle de 2,71m de rayon, afin d'avoir un rapport son direct/son réverbéré plus important et donc une meilleure possibilité de localisation. Pour le système de prise de son du CNSMDP, installé sur un cercle de rayon 2,04m, j'ai par contre choisi de travailler avec la voix diffusée à 4,77m afin d'avoir une distance apparente à peu près équivalente aux autres systèmes et éviter l'effet de zoom sur les signaux diffusés à 2,71m.

Les voix était diffusée tous les  $36^\circ$  pour le cercle de 2,71m de rayon, ce qui donne 6 positions en tenant compte de la symétrie :  $0^\circ$ ,  $+/-36^\circ$ ,  $+/-72^\circ$ ,  $+/-108^\circ$ ,  $+/-144^\circ$ ,  $180^\circ$ . Pour le système du CNSMDP, pour avoir une correspondance avec les autres systèmes, j'ai sélectionné les 6 signaux diffusés à  $0^\circ$ ,  $+/-40^\circ$ ,  $+/-70^\circ$ ,  $+/-110^\circ$ ,  $+/-140^\circ$ ,  $180^\circ$ .

En ce qui concerne les coups de wood-block, enregistrés tous les  $10^\circ$ , on disposait par symétrie de 19 positions ( $0^\circ$ ,  $+/-10^\circ$ , ...,  $180^\circ$ ).



Le montage des bandes s'est effectué sur système AKAI DD1500. Un montage différent a été effectué pour chaque système. Il s'agissait de mettre bout à bout, dans un ordre arbitraire, les différents signaux. Pour chaque direction, j'ai choisi l'un ou l'autre des demi-espaces (angle positif ou négatif) afin d'équilibrer les sources diffusées sur la droite et sur la gauche au cours du test. Pour la voix, étant donné qu'on avait moins de positions, j'ai présenté chaque position à 2 reprises au cours du test, ce qui donnait 12 signaux. Pour le wood-block, étant donné le nombre de positions, elle n'ont été proposées qu'une fois chacune ; par contre, chaque signal était présenté 2 fois de suite, ce qui permettait d'avoir une meilleure validité de la position perçue. En effet, la localisation de signaux brefs s'avère plus difficile que celle des signaux entretenus, en partie par le fait que les petits mouvements de tête ne peuvent pas vraiment avoir lieu durant la diffusion du signal.

Au vu des résultats du test sur les méthodes de réponse, présentés au paragraphe précédent, j'ai opté pour la méthode des marqueurs de direction suspendus à une corde, tous les 5°. Les tests se sont déroulés avec la même installation que pour les tests individuels d'analyse.

## 6.2. Les résultats

Les Figures 28 à 34 présentent les résultats des tests, pour chacun des systèmes et chacun des auditeurs. La moyenne et l'écart-type pour chaque position a été calculée sur l'ensemble des auditeurs, ainsi qu'une moyenne des écarts-type.

De même que pour les tests sur les méthodes de réponse (voir partie précédente), se pose le problème des confusions avant/arrière ; mais ici, il résulte de deux phénomènes. Premièrement, la confusion due à l'écoute avec le problème du cône de confusion de l'écoute binaurale, déjà évoqué lors des tests sur les méthodes de réponse et non imputable aux systèmes de prise de son. Mais aussi - et cela n'apparaissait pas pour des signaux synthétisés - les inversions avant/arrière créées par les systèmes de prise de son eux-même, et à considérer comme un défaut des systèmes. Il est à priori impossible de savoir si une confusion avant/arrière est à attribuer au cône de confusion ou au système de prise de son. Par contre, on peut avoir un bon indice en regardant les réponses de l'ensemble des auditeurs pour une position donnée : si une majorité des réponses dénote une confusion avant/arrière, on peut supposer que celle-ci est le fait du système lui-même, sinon on peut considérer que les confusions sont à rattacher à l'écoute individuelle et au problème du cône de confusion, qui ne se manifeste pas de la même façon chez tout le monde.

C'est cette méthode d'analyse que j'ai adoptée dans le dépouillement des résultats présentés sur les Figures 28 à 34. Les confusions avant/arrière sont grisées dans les tableaux et leur nombre est reporté dans la dernière colonne (ainsi que les absences de réponse notées NN dans le tableau). Si les inversions sont minoritaires, elles ne sont pas prises en compte dans le calcul de la moyenne et de l'écart-type ; par contre, si elles sont majoritaires, on considère que le système de prise de son est à l'origine de ces inversions et ce sont elles qui sont prises en compte pour les calculs de moyenne et d'écart-type (la case "inversion" de la dernière colonne est alors grisée). Lorsque aucune majorité ne se dégageait, j'ai fait figurer les calculs de moyenne et d'écart-type d'une part pour les réponses sans inversion et d'autre part, dans des colonnes supplémentaires, en grisé, ces calculs pour les réponses avec inversion ; ces cas seront nommés "indéterminations".

**Radio France**

Chant	Auditeur 1		Auditeur 2		Auditeur 3		Auditeur 4		Auditeur 5		Auditeur 6		Moyenne	Ecart type	Inversions
0	0	5	0	0	0	0	0	5	0	0	0	0	1	1,9	0
-36	5	-5	0	0	10	0	0	-5	0	0	5	5	1	4,3	0
-72	-125	-145	-55	-40	-145	-160	-150	-150	-140	NN	-160	-155	-148	10,9	9+1
108	65	75	65	95	110	135	40	45	90	90	120	120	109	17,5	5
-144	-120	-110	-110	-100	-140	-120	-125	-115	-120	-120	-125	-125	-119	10,0	0
180	-170	-165	-155	-170	-145	-160	-155	-160	180	180	-155	-160	-163	10,5	0
													Moyenne	9,2	

Woodblock	Aud.1	Aud.2	Aud.3	Aud.4	Aud.5	Aud.6	Moy.	Ec.type	Inversions
0	NN	0	0	0	0	0	0	0	0+1
10	NN	0	10	0	0	0	2	4,5	0+1
-20	-175	-5	180	-5	0	-5	-4	2,5	2
30	-175	0	10	175	0	5	4	4,8	2
40	165	10	20	5	30	25	18	10,4	1
-50	-165	-20	-30	-10	0	-15	-15	11,2	1
60	45	45	160	30	35	45	40	7,1	1
-70	-50	-30	-25	-155	-25	-35	-33	10,4	1
-80	-150	-40	-40	-30	-25	-40	-35	7,1	1
90	55	40	55	40	35	125	45	9,4	1
100	80	65	40	40	35	135	52	19,6	5
-110	-135	-140	-150	-145	-125	-110	-134	14,6	0
120	115	90	105	115	120	110	109	10,7	0
-130	-115	-75	-120	-135	-120	-135	-125	9,4	1
-140	-125	-105	-120	-125	-115	-120	-118	7,5	0
150	115	120	105	100	110	130	113	10,8	0
-160	-140	-90	-120	-125	-120	-115	-118	16,3	0
170	140	145	155	50	125	150	143	11,5	1
180	175	180	-170	180	160	-170	179	11,1	1
							Moy.	9,4	

Figure 28

**CNSMDP**

Chant	Auditeur 1		Auditeur 2		Auditeur 3		Auditeur 4		Auditeur 5		Auditeur 6		Moyenne	Ecart type	Inversions
0	0	0	0	0	-10	0	10	5	0	0	0	0	0	4,5	0
40	90	100	35	35	45	45	50	75	35	40	40	45	49	17,8	1
-70	-105	-105	-90	-75	-130	-130	-100	-125	-120	-100	-115	-120	-115	11,8	10
-110	-105	-115	-95	-90	-110	-125	-95	-90	-120	-120	-110	-115	-108	12,3	0
140	110	110	105	110	130	110	135	140	115	120	105	125	118	12,0	0
180	180	180	180	175	-165	-170	-165	-175	180	170	180	180	-178	7,5	0
													Moyenne	11,0	

Woodblock	Aud.1	Aud.2	Aud.3	Aud.4	Aud.5	Aud.6	Moy.	Ec.type	Inversions	Inversions	
										Moy.	Ec.type
0	0	5	-170	-5	0	0	0	3,5	1		
-10	-5	5	0	5	0	0	1	3,8	0		
-20	-25	-30	-170	-40	0	-25	-24	14,7	1		
30	100	30	50	65	40	45	46	12,9	1		
-40	-45	-60	-140	-45	-30	-65	-49	13,9	1		
50	95	35	70	40	45	50	48	13,5	1		
-60	-105	-120	-135	-110	-30	-70	-118	13,2	4		
-70	-115	-90	-140	-105	-40	-65	-65	25	3	-120	18,0
80	105	60	110	55	40	105	52	10,4	3	107	2,9
90	90	40	130	130	90	125	113	21,1	1		
-100	-105	-115	-130	-90	-110	-115	-111	13,2	0		
110	110	110	120	125	100	125	115	10	0		
-120	-105	-115	-140	-95	-120	-115	-115	15,2	0		
130	110	120	120	140	110	115	119	11,1	0		
140	110	120	135	150	125	135	129	13,9	0		
-150	-110	-120	-135	-80	-120	-120	-121	8,9	1		
-160	NN	-165	-165	-140	NN	-140	-153	14,4	0+2		
170	-155	180	-170	-170	165	180	-175	13,4	0		
180	180	180	180	-165	180	180	-178	6,1	0		
							Moy.	12,4			

Figure 29

Jacques/Mora

Chant	Auditeur 1		Auditeur 2		Auditeur 3		Auditeur 4		Auditeur 5		Auditeur 6		Moyenne	Ecart type	Inversions
0	10	5	5	15	5	5	0	0	0	10	0	0	5	5,0	0
36	35	30	30	30	30	40	30	35	40	35	35	35	34	3,8	0
-72	-110	-130	-50	-145	-60	-60	-40	-40	-65	-60	-35	-90	-56	16,9	3
-108	-120	-100	-130	-140	-120	-110	-95	-65	-85	-80	-120	-120	-117	13,9	3
144	110	90	135	120	160	125	45	125	165	170	120	120	131	24,7	1
180	160	160	180	150	180	180	-175	-170	-170	-160	170	170	176	14,6	0
													Moyenne	13,1	

Woodblock	Aud.1	Aud.2	Aud.3	Aud.4	Aud.5	Aud.6	Moy.	Ec.type	Inversions	Inversions	
										Moy.	Ec.type
0	0	5	0	5	0	0	2	2,6	0		
10	15	30	15	15	15	10	17	6,8	0		
20	35	40	30	35	25	30	33	5,2	0		
-30	-25	-25	-25	-115	-30	-20	-25	3,5	1		
40	30	30	30	40	20	30	30	6,3	0		
-50	-35	-70	-45	-25	-40	-25	-40	16,7	0		
-60	-25	-40	-45	-45	-55	-25	-39	12,0	0		
70	50	40	45	50	55	40	47	6,1	0		
80	40	55	60	45	65	55	53	9,3	0		
-90	-35	-110	-140	-115	-75	-30	-122	16,1	3	-47	24,7
100	45	90	100	135	85	NN	108	23,6	2+1		
-110	-30	-30	-130	-130	-95	-25	-118	20,2	3	-35	2,9
120	85	105	110	150	80	120	114	23,8	2		
130	75	95	110	155	100	90	110	26,2	1		
-140	-140	-150	-165	-135	-120	0	-142	16,8	1		
-150	-170	-160	-165	-150	-145	0	-158	10,4	1		
160	135	120	150	170	175	100	142	29,1	0		
-170	175	165	-165	-160	-140	NN	-155	13,2	2+1		
180	170	155	-175	-170	-165	5	179	16,4	1		
							Moy.	13,9			

Figure 30

Brauner

Chant	Auditeur 1		Auditeur 2		Auditeur 3		Auditeur 4		Auditeur 5		Auditeur 6		Moyenne	Ecart type	Inversions
0	165	160	0	180	5	5	0	0	0	0	0	-5	1	3,0	3
36	110	85	35	35	160	165	180	150	125	NN	40	25	148	26,2	6+1
72	80	75	90	95	155	155	95	30	110	30	30	35	53	27,4	5
-108	-150	-90	-20	-125	-130	-125	-130	-110	-30	-100	-125	-140	-123	18,0	2
-144	-95	-130	-115	-125	-150	-155	-120	-110	-125	-100	-130	-135	-124	17,9	0
180	150	150	165	180	180	-170	180	180	180	175	165	175	173	12,5	0
													Moyenne	17,5	

Woodblock	Aud.1	Aud.2	Aud.3	Aud.4	Aud.5	Aud.6	Moy.	Ec.type	Inversions	Inversions	
										Moy.	Ec.type
0	0	0	0	0	0	0	0	0	0		
-10	-5	180	-170	180	-10	0	-5	5	3	-177	5,8
20	170	20	165	180	10	170	171	6,3	4		
-30	180	-15	-175	-145	-170	-10	-168	15,5	4		
40	25	150	160	20	40	20	26	9,5	2		
-50	-140	-160	-165	-20	-145	-15	-153	11,9	4		
60	50	40	155	150	NN	40	43	5,8	2+1		
70	50	110	145	105	160	35	130	26,8	4		
-80	-80	-160	-140	-95	-120	-110	-125	25,5	5		
90	65	95	130	130	140	45	124	19,7	2		
-100	-50	-125	-110	-105	-105	-85	-111	9,5	2		
110	100	115	140	120	115	90	113	17,2	0		
-120	-105	-110	-120	-125	-130	-130	-120	10,5	0		
-130	-110	-130	-125	-115	-130	-95	-122	9,1	1		
140	100	120	150	135	130	100	123	19,9	0		
-150	-110	-105	-125	-115	-130	-120	-118	9,4	0		
160	85	90	130	115	125	85	115	17,8	2		
170	60	135	155	130	125	90	127	23,6	1		
180	165	180	180	-155	180	-5	-178	14,4	1		
							Moy.	13,2			

Figure 31

KFM360

Chant	Auditeur 1		Auditeur 2		Auditeur 3		Auditeur 4		Auditeur 5		Auditeur 6		Moyenne	Ecart type	Inversions
0	-20	-170	15	10	180	5	NN	NN	180	180	NN	-15	-1	15,6	4+3
-36	-175	170	-30	20	-170	-10	-20	0	-140	-135	-10	-15	-9	15,9	5
72	35	110	45	30	NN	25	50	45	45	40	25	35	38	8,9	1+1
108	70	70	50	40	115	125	110	40	55	30	40	25	47	16,0	9
-144	-150	-155	-150	-150	-110	-150	-160	-135	-30	-140	-25	-20	-144	14,9	3
180	175	180	-175	180	10	-155	180	180	180	-175	0	0	-177	8,7	3
													Moyenne	13,3	

Woodblock	Aud.1	Aud.2	Aud.3	Aud.4	Aud.5	Aud.6	Moy.	Ec.type	Inversions	Inversions	
										Moy.	Ec.type
0	NN	15	170	5	35	10	16	13,1	1+1		
10	35	20	170	30	25	10	24	9,6	1		
20	95	30	170	50	25	45	38	11,9	2		
-30	-165	120?	-160	-75	-10	-10	-32	37,5	2+?		
40	80	40	155	40	40	55	51	17,5	1		
-50	-145	-35	-130	-60	-30	-20	-36	17,0	2		
-60	-150	-20	-140	-60	-25	-20	-31	19,3	2		
70	120	30	130	40	40	80	48	22,2	2		
-80	-170	-20	-140	-85	-150	-15	-40	39,1	3	-153	15,3
90	100	50	130	110	35	60	48	12,6	3	113	15,3
-100	175?	-25	-155	-80	-30	-20	-39	27,8	4+?		
110	115	85	145	45	35	90	117	27,5	3	55	26,5
120	135	95	140	60	35	50	123	24,7	3	48	12,6
-130	-135	-140	-135	-150	-145	-25	-141	6,5	1		
140	110	30	150	35	35	50	38	8,7	4		
-150	-150	-40	-150	-55	-30	-15	-35	16,8	4		
-160	-140	-20	-155	-130	-170	-5	-149	17,5	2		
-170	180	0	-170	-165	-170	5	-171	6,3	2		
180	165	180	-175	180	180	10	178	7,6	1		
							Moy.	17,9			

Figure 32

Williams/Le Dù

Chant	Auditeur 1		Auditeur 2		Auditeur 3		Auditeur 4		Auditeur 5		Auditeur 6		Moyenne	Ecart type	Inversions
0	180	165	NN	165	-175	-170	0	0	-15	-5	-165	-170	-179	12,1	7+1
-36	-130	-135	-145	-150	-140	-130	-100	-100	-100	-90	-130	-145	-128	19,0	11
72	90	85	110	110	125	115	105	105	105	95	130	125	113	11,1	10
-108	-95	-100	-105	-110	-120	-100	-110	-90	-115	-105	-120	-125	-108	10,8	0
144	100	100	110	110	130	150	150	145	140	135	125	140	128	18,6	0
180	170	160	165	150	-160	-155	-165	-165	-140	-160	-160	-170	-173	21,2	0
													Moyenne	15,5	

Woodblock	Aud.1	Aud.2	Aud.3	Aud.4	Aud.5	Aud.6	Moy.	Ec.type	Inversions	Inversions	
										Moy.	Ec.type
0	180	180	180	0	-5	170	178	5	4		
-10	180	180	-175	-155	-175	-140	-168	16,4	6		
-20	-170	NN	-165	-145	-175	-140	-159	15,6	5+1		
30	60	95	165	35	155	50	48	12,6	3	136	37,9
40	60	100	140	50	90	135	67	20,8	3	125	21,8
-50	-100	-125	-95	-85	-115	-120	-111	12,9	5		
60	85	100	130	75	95	125	113	17,6	4		
70	80	95	105	105	125	140	114	18,2	5		
-80	-85	-100	-85	-90	-90	-120	-88	2,9	2		
-90	-95	-100	-90	-90	-95	-110	-97	7,5	0		
100	85	105	110	110	125	135	117	12,5	1		
-110	-90	-105	-100	-100	-95	-115	-101	8,6	0		
120	95	100	130	140	120	130	119	18,0	0		
130	100	105	140	125	115	125	118	14,7	0		
-140	-120	-90	-115	-105	-110	-125	-111	12,4	0		
-150	-135	-150	-130	-120	-125	-120	-130	11,4	0		
160	75	100	170	140	120	40	133	29,9	2		
-170	-175	180	-160	-160	-140	-140	-159	16,9	0		
180	165	155	-150	-160	-105	190	-164	35,7	0		
							Moy.	16,6			

Figure 33

**Le Grenier à Son**

Chant	Auditeur 1		Auditeur 2		Auditeur 3		Auditeur 4		Auditeur 5		Auditeur 6		Moyenne	Ecart type	Inversions
0	5	NN	0	0	-5	0	5	5	5	5	0	0	2	3,4	0+1
-36	130	140	30	25	30	160	30	25	25	25	20	30	27	3,5	3
-72	-160	-145	-125	-125	-155	-150	-130	-85	-135	-125	-30	-25	-139	13,9	9
-108	-140	-165	-120	-120	-140	-150	-120	-105	-135	-120	-70	-60	-132	17,8	2
144	145	145	120	115	180	180	145	145	50	115	25	15	143	24,6	3
180	180	175	180	180	-170	-170	180	175	180	175	0	5	-180	5,5	2
													Moyenne	11,5	

Woodblock	Aud.1	Aud.2	Aud.3	Aud.4	Aud.5	Aud.6	Moy.	Ec.type	Inversions	Inversions	
										Moy.	Ec.type
0	175	180	0	180	15	15	10	8,7	3	178	2,9
-10	-165	0	180	-10	-175	0	-3	5,8	3	-173	7,6
-20	-165	-170	180	-170	-165	-15	-170	6,1	5		
30	165	20	150	40	150	25	28	10,4	3	155	8,7
40	140	25	160	60	55	50	48	15,5	2		
-50	-145	-160	-165	-140	-155	-35	-153	10,4	5		
60	75	35	140	140	45	35	48	18,9	2		
70	145	55	140	60	50	65	58	6,5	2		
-80	-155	-135	-160	-120	-145	-45	-143	16,0	5		
-90	-150	-115	-150	-130	-140	-125	-135	14,1	0		
100	135	120	150	135	65	40	135	12,2	2		
-110	-160	-125	-150	-120	-150	-50	-141	17,5	1		
120	135	110	150	120	115	55	126	16,4	1		
130	150	145	150	120	125	50	138	14,4	1		
-140	-160	-155	-165	-130	-145	-15	-151	13,9	1		
150	160	165	150	145	145	20	153	9,1	1		
-160	-175	170	-170	-155	-175	165	-177	14,4	0		
-170	180	180	180	-150	-175	175	-175	12,6	0		
180	170	180	-175	180	175	5	178	5,7	1		
							Moy.	11,3			

**Figure 34**

Signalons ici que les critères "précision en localisation" et "stabilité en localisation" évalués lors du test d'écoute individuel ne correspondent pas au critère "justesse de localisation" abordé lors des écoutes collectives, et ne traduisent pas les performances en localisation des systèmes. En effet, pour le test individuel, les auditeurs se basaient uniquement sur un enregistrement du chœur et n'avaient aucune référence pour vérifier de la justesse de localisation de tel ou tel chanteur. Les auditeurs pouvaient par contre juger de la précision de localisation (absence de flou de localisation) ainsi que la stabilité (absence de déplacement des sources).

Pour les écoutes collectives, les auditeurs pouvaient par contre se baser sur l'enregistrement des percussions et du plan qui leur était fourni (Figure 3), qui leur permettait de juger de la justesse de localisation. Afin de pouvoir tirer parti de ces résultats dans la synthèse que nous effectuons dans la 7<sup>e</sup> partie du mémoire, nous fournissons Figure 35 le classement des systèmes à partir du critère "justesse de localisation" établi à partir des résultats de la Figure 18 (contrairement au classement de cette Figure, nous ne faisons pas la moyenne entre l'écoute du chœur et des percussions, mais nous nous basons uniquement sur l'écoute des percussions).

	Radio France	KFM 360	CNSMDP	Brauner	Jacques/Mora	Williams/Le Dü	Le Grenier à Son
Justesse de loc	2	3	5	7	1	4	6

**Figure 35**

### 6.3. Commentaires

Les Figures 36 à 42 représentent les résultats pour le wood-block sous forme de graphique. On a considéré le rendu des systèmes symétrique et on a tout rapporté à des angles positifs (de 0° à 180°) ; lorsque on était en présence d'une indétermination, les deux valeurs ont été représentées (d'où l'apparition de deux courbes dans certaines zones).

Ces courbes nous serviront dans la dernière partie pour une interprétation des résultats récoltés au cours du mémoire. Afin de faciliter l'interprétation de ce test, on peut déjà examiner les résultats en portant notre attention sur un certain nombre de points, tels que la portion d'espace perçue entre les enceintes avant gauche et droite (équivalent de l'angle utile de prise de son en stéréophonie à 2 canaux), ainsi qu'entre les enceintes arrière gauche et droite, la présence de sources perçues sur les côtés (entre 30° et 110°), la linéarité de la localisation, la quantité d'inversions avant/arrière ainsi que d'indéterminations (zones où les courbes se dédoublent), enfin la portion d'espace perçue dans le demi-espace frontal (de -90° à +90°) par rapport à celle perçue dans le demi-espace arrière.

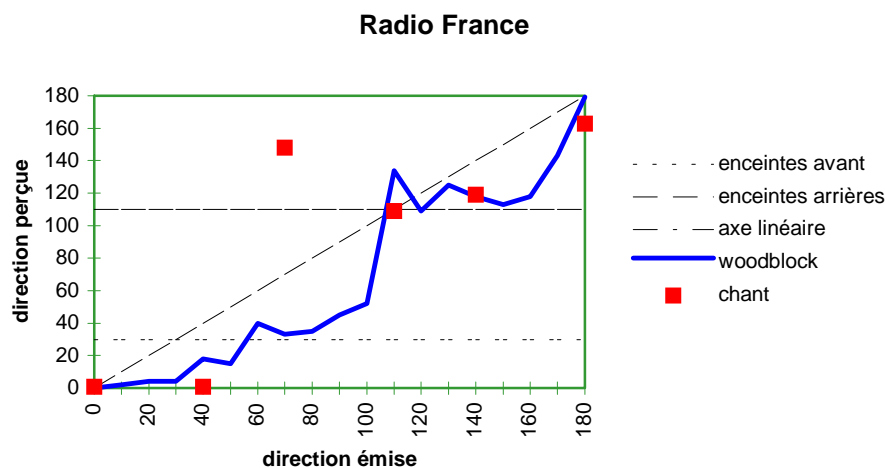
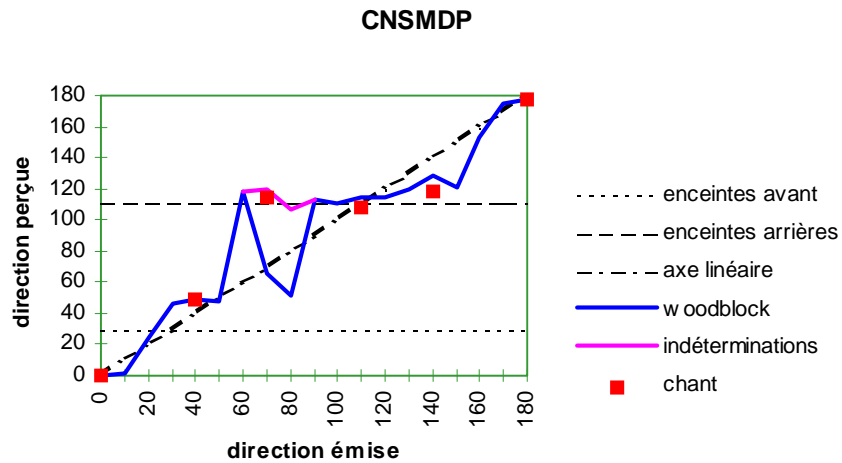


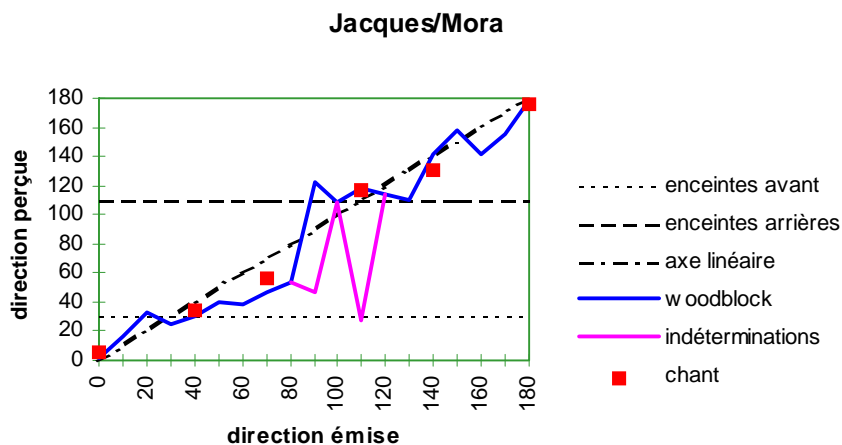
Figure 36

Le système de Radio France fait apparaître deux "blocs" dans la perception : globalement le secteur 0°/100° perçu dans la portion 0°/50°, avec un rendu très linéaire, puis le secteur 110°/180° perçu dans la portion 110°/180° avec une accumulation autour de 120° ; cela fait apparaître un "trou" dans la localisation entre 50° et 130°. Il est important de noter l'absence d'inversion avant/arrière ainsi que d'indétermination pour ce système. Globalement, l'angle de prise de son avant (angle physique perçu entre -30° et +30°) est de l'ordre de 120° (car, d'après les courbes, les signaux de 0° à 60° sont perçus entre 0° et 30°) et l'angle de prise de son arrière (portion d'espace perçue entre les enceintes arrière) de l'ordre de 120° également. La portion d'espace perçue dans le demi espace frontal est de l'ordre de 200°, celle perçue dans le demi-espace arrière de l'ordre de 160°.



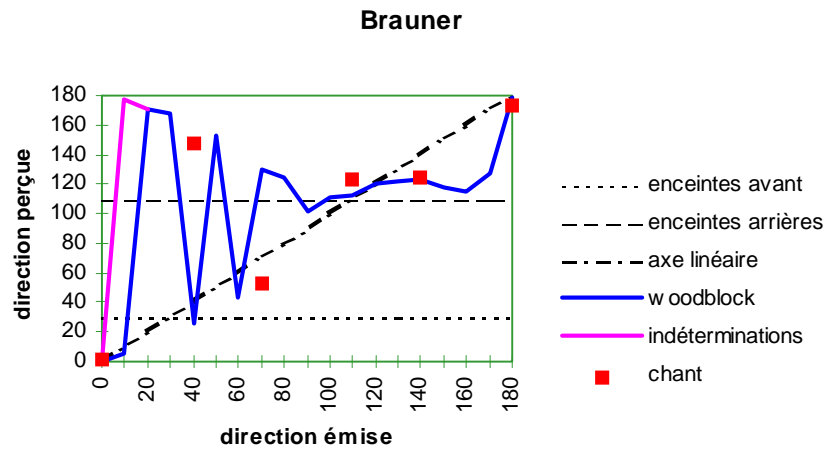
**Figure 37**

Le système du CNSMDP conduit à une localisation globalement régulière, sur l'ensemble de l'espace, avec des indéterminations pour les sources à 70° et 80°. L'angle perçu sur la base avant est de l'ordre de 40°, celui perçu sur la base arrière de 180°. Le demi-espace avant restitue environ 160° et le demi-espace arrière environ 200°.

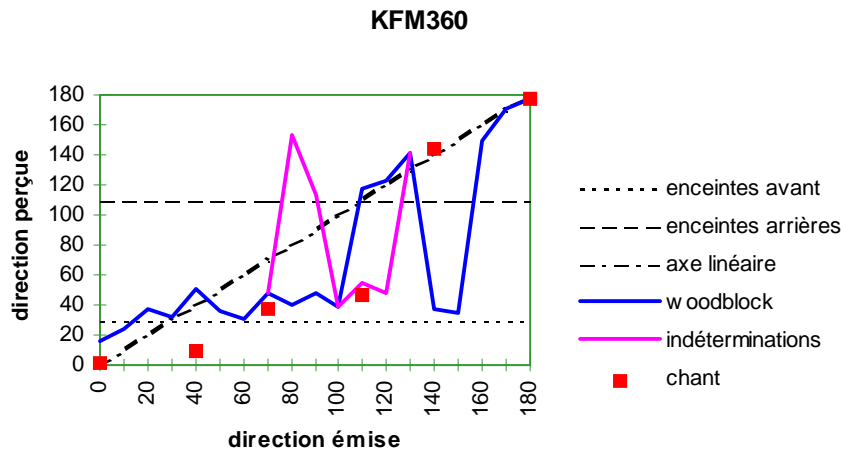


**Figure 38**

Le système de N. Jacques et A. Mora conduit lui aussi à une localisation globalement régulière avec des indéterminations pour les sources à 90° et 110°. La base avant restitue un secteur de 80° environ et la base arrière un secteur de 160°. Les demi-espaces avant et arrière restituent chacun environ 180°.

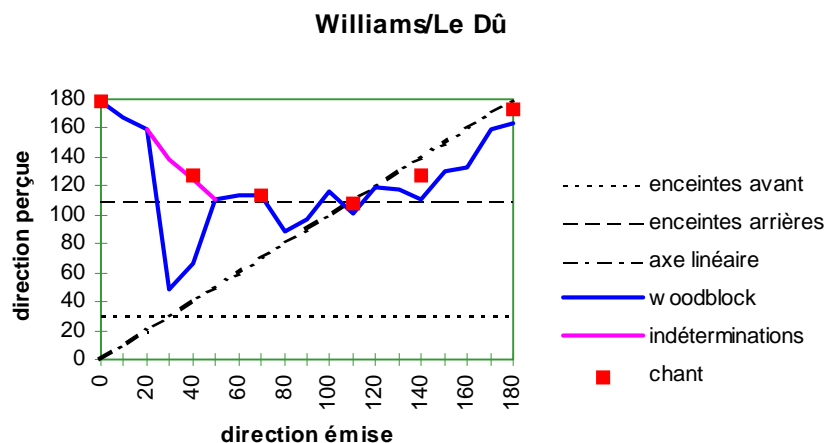


Le système Brauner conduit à de nombreuses inversions avant/arrière pour les sources de  $0^\circ$  à  $70^\circ$ , ce qui rend difficile l'établissement de l'angle perçu dans la base avant ou le demi-espace avant. On constate qu'aucune source n'est perçue entre  $40^\circ$  et  $100^\circ$  et que, par contre, beaucoup de sources sont perçues autour de  $120^\circ$ , ce qui conduit à une absence de régularité de la perception sur la base arrière. On peut évaluer le secteur perçu sur cette base arrière autour de  $220^\circ$  (si on ne tient pas compte des inversions de la base avant).



Les courbes concernant le système KFM 360 montrent de nombreuses indéterminations à partir de  $70^\circ$  ; beaucoup de sources sont perçues sur l'avant, avec en particulier des inversions de l'arrière vers l'avant pour les sources situées à  $140^\circ$  et  $150^\circ$ . Les sources sont perçues soit entre  $0^\circ$  et  $50^\circ$ , avec une forte accumulation autour de  $40^\circ$ , soit au-delà de  $110^\circ$ .





**Figure 41**

Dans le cas du système de M. Williams et G. Le Dû, mis à part pour les sources situées à 30° et 40°, tout est perçu sur l'arrière avec en particulier une accumulation de sources perçues autour de 110°.

**Erreur! Objet incorporé incorrect.**

**Figure 42**

Quant au système du Grenier à Son, si l'on examine les résultats de la Figure 34, on constate que toutes les sources utilisées dans la bande test se situant sur le côté gauche (angles négatifs), entre 0° et -110° ont été perçues soit complètement à l'avant, soit complètement à l'arrière, ce qui n'est pas le cas des sources situées sur la droite. Cela dénote un problème de décodage du Soundfield (le décodeur utilisé a d'ailleurs été renvoyé en usine peu de temps après les enregistrements). Cela rend inexploitable les résultats concernant ce système. En dehors de ce problème de décodage, il apparaît des indéterminations sur certaines sources provenant de l'avant (sources à 0° et 30°).

Maintenant que l'ensemble des résultats concernant les écoutes collectives de la Convention de l'AES, les analyses des images sonores en écoute individuelle et les tests de localisations ont été présentés, nous allons consacrer la dernière partie de ce mémoire à une synthèse de ces données.

## **Septième partie**

### **Synthèse des résultats**

Nous disposons maintenant d'une somme importante de données provenant de 3 expériences : les écoutes collectives de la Convention de l'AES, avec ses résultats qualitatifs et quantitatifs, les analyses individuelles des images sonores des différents systèmes et leur comparaison avec les résultats de la Convention, et, enfin, les résultats des tests de localisation. Nous avons d'autre part synthétisé le principe théorique et pratique de fonctionnement de chacun des systèmes.

Le but de cette partie est de regrouper et d'analyser l'ensemble des résultats pour chaque système, en les mettant en regard avec leurs principes de fonctionnement, afin d'essayer de dégager les propriétés de chacun au vu des résultats des expériences. Il n'est pas question ici d'établir un "classement" absolu des systèmes, étant donné que nous nous basons sur une expérience donnée, mais plutôt de dégager - voire d'extrapoler - les caractéristiques apparaissant au vu de cette expérience d'enregistrement.

Cette analyse générale des résultats doit d'ailleurs dépasser le cadre purement comparatif des 7 systèmes en jeu et permettre de soulever un certain nombre de problèmes ou de dégager certaines constatations permettant d'envisager le développement de nouveaux systèmes de prise de son multicanale à partir de bases concrètes. En particulier, il va être intéressant de mettre en regard les résultats par rapport aux prévisions des concepteurs, dans le cadre d'une expérience qui se veut la plus objective possible.

Nous allons tout d'abord analyser les résultats système par système puis, dans la dernière partie, essayer de dégager des constatations plus générales au vu des différentes analyses.

### **7.1. Le système de J-M.Porcher (Radio France)**

Force est de constater que ce système, alors qu'il n'était initialement pas prévu pour une captation sur 360° (voir 2.4.), s'est très bien adapté à la situation de l'expérience : il a obtenu le meilleur classement général, que se soit au cours des écoutes collectives que lors des écoutes individuelles (Figure 20).

Si l'on se reporte aux 3 catégories définies en 4.4. pour les écoutes individuelles (espace, sources, timbre), on constate que ce système réalise un très bon équilibre entre rendu spatial, rendu des sources et équilibre fréquentiel (classé second pour chacune de ces catégories sur la Figure 23).

En se basant sur les résultats des écoutes individuelles, et le classement présenté Figures 20 et 35 (écoutes collectives), on peut classer les qualités accordées à ce système ainsi :

1<sup>er</sup> classé : enveloppement, naturel, définition

2<sup>nd</sup> classé : équilibre avant/arrière, oubli des haut-parleurs, précision en localisation, timbre, justesse de localisation (écoutes collectives)

3<sup>e</sup> classé : sources sur les côtés, homogénéité, stabilité en localisation

Essayons de comprendre cela au vu de la conception du système et des résultats du test "objectif" de localisation.

Que ce système ait été plébiscité pour l'enveloppement et le naturel traduit une caractéristique connue des microphones omnidirectionnels distants, à savoir l'impression d'espace due à la décorrélation des signaux. On retrouve ici un résultat établi pour les systèmes stéréophoniques à 2 canaux ([14],[15]).

En ce qui concerne la définition, il est à noter que le système, du fait de sa disposition non ponctuelle, se situait plus proche des chanteurs que les systèmes ponctuels (c'est-à-dire tous sauf celui du CNSMDP). Néanmoins, contrairement à la plupart des autres systèmes, les microphones utilisés ici étaient omnidirectionnels et le rapprochement vers les sources compense à peine le facteur de distance par rapport aux microphones directifs. En effet, si on se réfère à la Figure 4, la base avant se trouve avancée de 74cm dans le cercle de 4m77 de rayon, soit un rapport de distance de l'ordre de  $4,77/4,03 = 1,18$  (inférieur au facteur de distance d'un hypocarديوide) ; pour l'arrière, ce rapport est de l'ordre de  $4,77/(4,77-1,86) = 1,64$  (environ le facteur de distance d'un cardioïde).

Il est évident que les qualités propres des microphones interviennent aussi sur ce critère de définition, et les microphones utilisés par J-M.Porcher sont assez appréciés pour leur définition.

Mais je pense qu'on peut trouver une bonne justification de la très bonne définition de ce système dans un article récent de G.Theile ([16]) qui s'est penché sur un des problèmes inhérents aux systèmes multicanaux : les images fantômes multiples. G.Theile part du principe qu'un système à 5 microphones peut être vu comme 5 couples de microphones fournissant chacun une image propre des sources. Si la différence de niveau ou le délai entre ces images fantômes multiples ne sont pas suffisamment importants, on risque une création de flou de localisation, de définition ainsi qu'un détimbrage par le mélange de ces différentes sources virtuelles. Or, de part ses dimensions, ce système génère des différences temporelles importantes entre les différentes images fantômes (pour une source située face au système, on calcule facilement à partir de la Figure 4 la différence de temps d'arrivée entre le microphone correspondant au sommet du triangle et un de ceux de la base : 2.8ms). L'effet de précédence (ou effet Haas) joue et diminue les "interférences acoustiques". Outre la définition, cela sert la précision en localisation et le timbre (2<sup>nd</sup> classés).

Outre la remarque précédente, le timbre est favorisé du fait de l'utilisation de microphones omnidirectionnels, et particulièrement de ceux utilisés ici.

En ce qui concerne l'équilibre avant/arrière, on a souligné en 6.3. que l'angle de prise de son avant est de l'ordre de l'angle de prise de son arrière (environ 120°) et que la portion d'espace perçue dans le demi-espace avant est de 200° (160° pour le demi-espace arrière) ; cela permet de comprendre le jugement positif quant à l'équilibre avant/arrière.

Quant à la sensation d'oubli des haut-parleurs, on a une certaine contradiction entre le jugement assez positif des auditeurs et la courbe de localisation établie Figure 36. Sur cette courbe, comme on l'a déjà signalé en 6.3., apparaît une certaine accumulation des sources sur les différents haut-parleurs (à 0°, +/-30° et +/-110°). Cette accumulation se comprend aisément par l'analyse des angles de prise de son. Elle a d'ailleurs été signalée, pour le triangle DECCA, par G.Theile dans l'article cité plus haut : la distance importante entre les microphones conduit à des angles de prise de son gauche/centre et centre/droit faibles et donc des accumulations sur les enceintes (voir Figure 43).

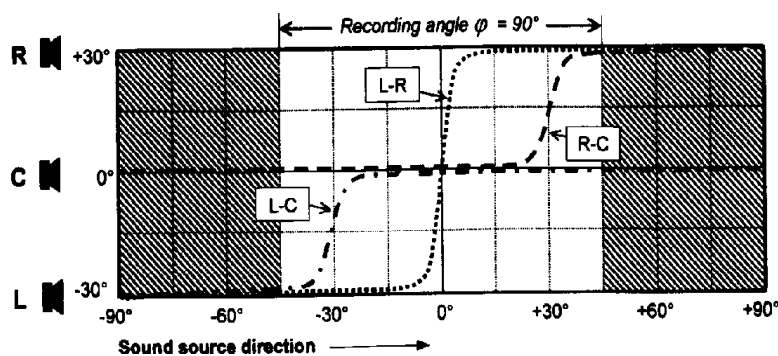


Figure 43

Pour l'arrière, les 2 microphones, bien que plus proches l'un de l'autre, sont tout de même espacés de 48cm. Le passage assez rapide sur les enceintes traduit un phénomène établi pour la prise de son à 2 canaux par différence de temps, à savoir la latéralisation assez rapide et un phénomène de trou au centre dû à la non linéarité de la perception par différence de temps.

On constate donc que la lecture des courbes de localisation ne permet pas à elle seule de prévoir la sensation d'oubli des haut-parleurs en tant que sources ponctuelles. La sensation pour un ensemble de sources (à savoir le chœur) est en effet différente de celle pour une source isolée (comme ça a été le cas pour les tests de localisation), en particulier du fait de la réverbération. On peut d'ailleurs penser que l'effet d'accumulation du son direct sur les enceintes, visible sur la courbe de localisation, a été compensé par l'espace créé entre les enceintes par la réverbération - espace d'autant plus riche que les microphones sont distants et donc décorrélés.

Il est d'ailleurs fort probable que le critère d'homogénéité, moins bien apprécié, traduise cette non linéarité de localisation des sources.

On note, toujours d'après la Figure 36, que ce système est le seul à ne créer ni inversion avant/arrière ni ce qu'on a nommé indétermination (perception à la fois sur l'avant et l'arrière). C'est la différence de temps importante entre avant et arrière qui permet d'expliquer cette absence d'artéfact, comme on l'a déjà vu en ce qui concerne le problème des images fantômes multiples.

Cette absence d'inversion avant/arrière ainsi que d'indétermination joue certainement beaucoup sur le bon jugement concernant la justesse de localisation lors des écoutes collectives alors qu'au vu de la Figure 36, la linéarité de localisation est loin d'être réalisée.

Parmi les critères les moins appréciés se trouvent les sources fantômes sur les côtés. Cela se traduit sur les courbes de localisation : comme on l'a signalé en 6.3., on n'a pas de localisation de source entre 50° et 110°. Cela s'explique par la distance importante entre les microphones avant et arrière, qui crée un angle de prise de son très faible et donc un basculement rapide de la localisation de l'avant vers l'arrière.

De même, la stabilité en localisation fait partie des critères moins appréciés. Si on compare la courbe établie pour le woodblock et les localisations pour la voix, on constate une différence importante sur les côtés (voix à 72° perçue sur la base arrière, alors que le woodblock au même endroit est perçue sur l'avant). Cela montre que la localisation des sources se situant sur les côtés est fortement dépendante de leur contenu spectral ; cela est d'ailleurs à rapprocher des constatations faites pour la stéréophonie à 2 canaux avec microphones omnidirectionnels. Des instabilités en localisations sont donc à craindre.

## 7.2. Le KFM360 de J.Bruck

Le KFM360 a été classé 2<sup>nd</sup> lors des tests individuels et 3<sup>e</sup> lors des écoutes collectives. C'est celui qui a donc été le mieux apprécié après le système de J-M.Porcher.

Il est intéressant de noter d'après la Figure 23, comme on l'a fait en 4.2., que ce système semble privilégier le rendu des sources (1<sup>er</sup> classé) au rendu spatial (4<sup>e</sup> classé). On retrouve ici un résultat établi pour la stéréophonie à 2 canaux : les systèmes coïncidents privilégient le rendu des sources (en particulier en localisation) au détriment de la sensation d'espace. Or, on a vu en 2.6. que le KFM360 est constitué pour moitié de microphones coïncidents (pour les côtés).

Si l'on regarde plus précisément le classement établi Figure 20 pour les tests individuels et Figure 35, on peut classer les critères d'appréciation de ce système ainsi :

- 1<sup>er</sup> classé : sources sur les côtés, naturel, précision en localisation, stabilité en localisation, timbre
- 2<sup>nd</sup> classé : définition
- 3<sup>e</sup> classé : enveloppement, justesse de localisation (écoutes collectives)
- 4<sup>e</sup> classé : équilibre avant/arrière
- 5<sup>e</sup> classé : homogénéité
- 6<sup>e</sup> classé : oubli des haut-parleurs

On voit donc que la précision et la stabilité en localisation ont été particulièrement appréciées. L'utilisation pour partie de capsules coïncidentes, comme on vient de le souligner, joue certainement en faveur de ces critères. Il faut ajouter à cela la différence d'intensité créée par la diffraction sur la sphère, qui permet une précision et une stabilisation de la localisation.

L'utilisation de systèmes coïncidents sur les côtés semble avoir rempli le rôle attendu par J.Bruck, à savoir la création de sources fantômes sur les côtés (voir 2.6.), critère très apprécié pour ce système.

Mais si on se penche sur la courbe de localisation Figure 40, on constate que la présence de sources sur les côtés n'est pas flagrante : les sources les plus latéralisées sont à 51°. Par contre, pour les sources à 80° et 90°, on constate des indéterminations avant/arrière. Cela peut traduire une extra-largeur de sources latérales, et irait dans le sens de plusieurs recherches concluant à l'impossibilité de création de sources fantômes latérales précises, avec la configuration de haut-parleurs telle que celle utilisée aujourd'hui en diffusion ([17]).

Ainsi, pour G.Theile, qui a récemment prolongé l'étude de l'article cité ci-dessus (datant de l'époque de la quadriphonie), sans l'ajout de haut-parleurs à +/-70°, une localisation précise de sources fantômes latérales est impossible. Dans la configuration de diffusion actuelle, répondant à la recommandation ITU-R BS 775-1 (Annexe 1), une indétermination pour les sources latérales est inévitable.

Néanmoins, les auditeurs perçoivent tout de même cette indétermination comme présence de sources fantômes latérales. Il ne faut d'ailleurs pas oublier que l'incertitude de localisation naturelle chez l'homme est assez importante sur les côtés puisqu'elle atteint 10° (jusqu'à 20° selon les fréquences), alors que le seuil de discrimination sur l'avant est inférieur à 3°. En ce sens, il y a une plus grande tolérance quant à la localisation de sources fantômes sur les côtés. Toujours est-il que le système le plus apprécié quant aux sources fantômes latérales ne crée pas de localisation entre 50° et 110°.

Dans les autres critères très appréciés, se trouvent le timbre et le naturel.

L'utilisation pour moitié de capteurs omnidirectionnels joue en faveur du timbre.

D'autre part, l'utilisation combinée de différence de temps et de différence d'intensité dépendante de la fréquence par l'utilisation de la sphère (voir 2.6.), proche du principe de la perception binaurale humaine, contribue au jugement très positif concernant le critère de naturel. Cela est d'ailleurs à rapprocher du jugement concernant le naturel de la sphère Schœps en stéréophonie à 2 canaux.

La définition, appréciée, bénéficie certainement de l'utilisation des microphones bidirectionnels dont le dosage en post-production permet entre autres de gérer la présence et les problèmes d'images fantômes multiples, abordé plus haut. En effet, une augmentation du niveau de ces microphones dans le décodage des couples M/S (voir 2.6.) permet d'une part d'augmenter le rapport champ direct/champ réverbéré et, d'autre part, d'augmenter la séparation acoustique entre les couples de microphones afin de diminuer l'interférence acoustique entre images fantômes multiples.

Il est d'ailleurs intéressant de constater au vu de la courbe de localisation de la Figure 40 et de sa dissymétrie avant/arrière que le dosage en post-production a certainement été différent sur l'avant et sur l'arrière. L'avant a certainement bénéficié d'une utilisation importante des microphones bidirectionnels, ce qui crée une latéralisation assez rapide - pour le wood-block en tout cas - et une accumulation rapide sur les enceintes avant, du fait de la diminution de l'angle de prise de son. Par contre, le décodage de l'arrière a certainement moins utilisé le signal des microphones bidirectionnels, ce qui explique en partie les inversions avant/arrière (signaux à 140° et 150° perçus sur l'avant). Cela explique aussi le fait que ce système soit classé 3<sup>e</sup> en justesse de localisation : celle-ci est relative (si l'on tient compte de la voix, qui ne provoque pas d'inversion avant/arrière, l'arrière semble rendre une localisation plus juste que l'avant).

Il est intéressant de noter sur l'avant que la localisation de la voix est très différente de celle du wood-block. Or on se souvient que le canal central n'est alimenté par aucun signal réel, mais est reconstitué par matricage fréquentiellement dépendant (voir 2.6.). Il apparaît ici que ce matricage crée des différences de localisation selon la fréquence ; d'autre part, on constate qu'il ne fournit pas de centre précis pour les signaux de wood-block (signal à 0° perçu à 16°). Toutes ces remarques vont dans le sens d'un manque d'homogénéité relevé par les auditeurs (système classé 5<sup>e</sup> pour ce critère).

Le manque d'enveloppement et d'équilibre avant/arrière se lit sur la Figure 40 : beaucoup de sources sont perçues sur l'avant.

Le critère d'oubli des haut-parleurs, également mal jugé, est également visible sur la courbe avec l'accumulation déjà soulignée de sources sur les enceintes gauche et droite.

### 7.3. Le système de J-M.Lyzwa et J-C.Messonier (CNSMDP)

Ce système a été classé 3<sup>e</sup> au cours des écoutes individuelles et 4<sup>e</sup> lors de la Convention (Figure 20). Si on se reporte Figure 23, on voit qu'il réalise un bon équilibre entre rendu des sources, de l'espace et des timbres.

Comme pour les autres systèmes, classons les critères d'appréciation suivant le classement de la Figure 20 pour les écoutes individuelles et de la Figure 35 :

1<sup>er</sup> classé : naturel

2<sup>nd</sup> classé : sources sur les côtés, oubli des haut-parleurs, homogénéité

3<sup>e</sup> classé : précision en localisation, timbre, définition

4<sup>e</sup> classé : enveloppement, stabilité en localisation

5<sup>e</sup> classé : équilibre avant/arrière, justesse de localisation (écoutes collectives)

On a déjà relevé en 4.2. que le classement lors des écoutes collectives sur les critères d'homogénéité, de naturel et de définition s'est avéré moins bon que lors des écoutes individuelles. Cela semble pointer le fait, comme on l'a dit, que la zone d'écoute pour ce système est relativement réduite, ce qui s'explique en partie par l'utilisation du décodage transaural qui nécessite un positionnement précis des auditeurs (voir 2.5.).

Les critères de naturel et d'homogénéité, très appréciés, sont probablement à rattacher au fait que ce système a particulièrement été étudié pour la situation d'enregistrement : une couronne régulière de microphones à l'intérieur du cercle des chanteurs ou des percussionnistes. En soi, la captation était très homogène.

En ce qui concerne la présence de sources latérales, appréciée, on constate le même phénomène que pour le système KFM360 : pas de réelle localisation entre 50° et 110°, mais des indéterminations pour les sources à 70° et 80°. Cela nous ramène aux conclusions du KFM360. Il faut noter tout de même une perception un peu plus élargie sur les côtés pour le système du CNSMDP que pour le KFM360, vu qu'une partie de la source à 70° est perçue à 65°. A ce titre, il faut être conscient que l'effet transaural appliqué sur un signal microphonique est "brouillé" par les microphones adjacents : si on se reporte à la Figure 11, on peut voir que le signal du microphone 3 traité par le système transaural va subir la présence des signaux des microphones 2 et 4 alimentant les enceintes droites (avant et arrière) et constituant un couple microphonique créant sa propre localisation.

Quant au jugement positif concernant l'oubli des haut-parleurs, il semble être partiellement en contradiction avec la courbe de localisation de la Figure 37. En effet, on y voit une accumulation de localisation sur les enceintes arrières. Cela ne semble pas faire apparaître de gêne particulière et nous ramène à la conclusion correspondant au système de J-M.Porcher : la lecture des courbes de localisation ne permet pas à elle seule la prévision de sensation de présence des haut-parleurs.



L'avis des auditeurs a été plus mitigé sur les critères de précision en localisation, timbre et définition.

Il se trouve que ces critères sont justement ceux qui peuvent être "endommagés" par la présence d'images fantômes multiples (voir plus haut, commentaire sur le système de J-M.Porcher). Pour une source située face à un microphone, on calcule à partir de la Figure 44 que la différence de temps d'arrivée entre ce microphone et un microphone adjacent est de 1.9ms, ce qui est à la limite de l'effet Haas et peut expliquer que les images fantômes multiplient entâchent les critères de précision en localisation, timbre et définition.

D'autre part, les microphones arrières sont panoramiqués entre les enceintes. Or, il s'agit de signaux temporellement décalés que l'on mélange, ce qui peut conduire à une altération du timbre par un phénomène de filtrage en peigne. On a aussi souligné l'interaction des microphones dirigés vers les enceintes avant et arrière avec le signal transaural.

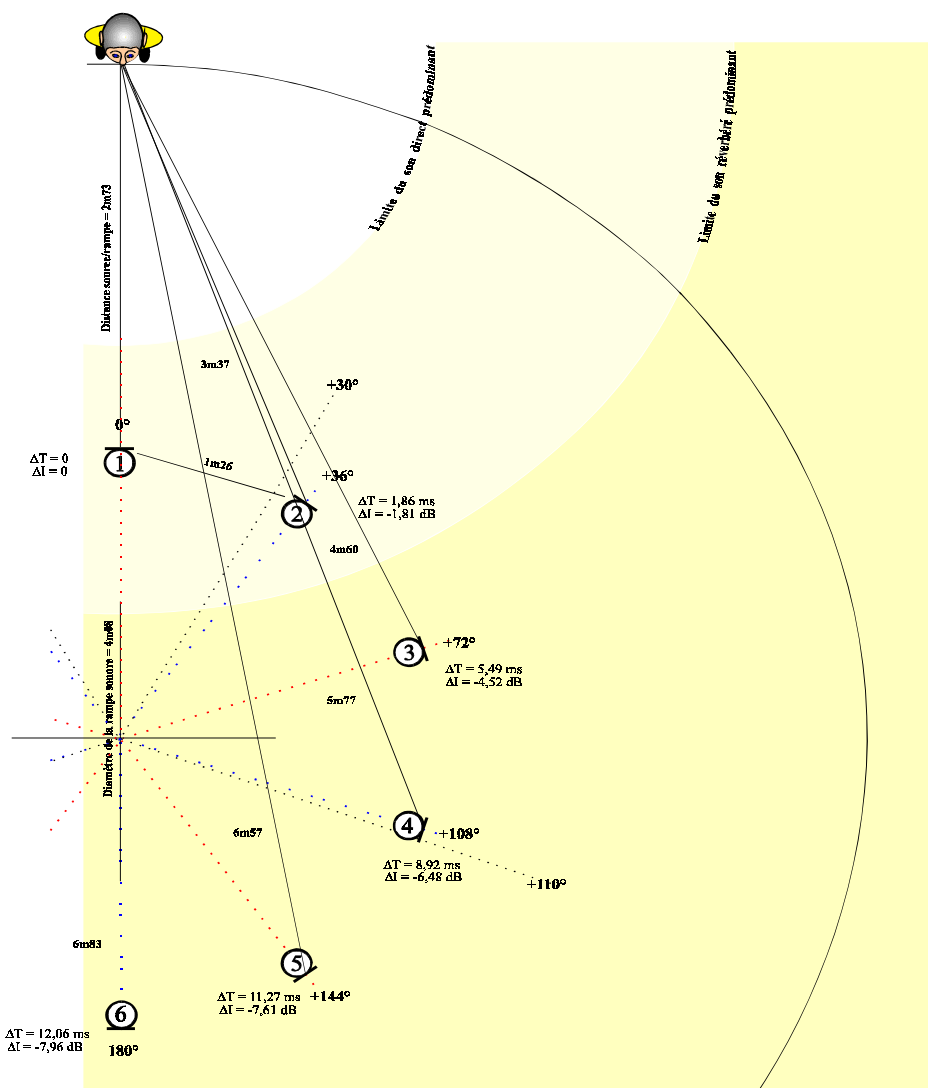


Figure 44

Le mauvais jugement concernant l'équilibre avant/arrière et l'enveloppement se lisent sur la courbe de localisation Figure 37 : l'angle de prise de son de la base avant n'est que de  $40^\circ$ , alors que celui de l'arrière est de  $180^\circ$ . De plus, d'un point de vue énergétique, on envoie les signaux de 5 microphones sur l'arrière et de 3 microphones sur l'avant. Si cela correspond à une répartition régulière vu la disposition des microphones, il n'en demeure pas moins que ça provoque une gêne.

Cela pointe une constatation faite lors des séances d'écoute collectives, qui ressort du dépouillement des résultats Figure 17 et de discussions que j'ai pu avoir avec certains auditeurs : le son à l'arrière crée parfois une gêne. On constate en effet sur la Figure 17 que pour tous les systèmes à l'exception d'un seul, les auditeurs reprochaient le fait qu'il y ait "trop d'arrière". Il y a certainement une question de manque d'habitude d'écoute multicanale qui provoque cette réaction. Mais il y a aussi une question plus psychologique qui fait qu'une source venant de l'arrière est naturellement perçue comme une "agression" (dans la vie courante, la première réaction face à une telle source est de se retourner). C'est pourquoi il faut être prudent quant à la répartition énergétique entre avant et arrière : une répartition correspondant à la réalité (comme c'est le cas pour ce système) risque de déranger. D'ailleurs, si on se reporte au système de J-M.Porcher, très bien apprécié pour l'équilibre avant/arrière, on a un angle de prise de son avant égal à celui de l'arrière ( $120^\circ$ ) alors que la base de diffusion avant correspond à  $60^\circ$  de l'espace et celle arrière à  $140^\circ$ .

Ces remarques valent pour le cas où des sources se situent sur l'arrière, comme c'était le cas pour l'expérience. Dans des situations de scènes frontales, elles ne s'appliqueraient pas nécessairement.

Les critères de justesse et de stabilité de localisation ont été mal jugés. Cela ne correspond pas vraiment aux résultats des tests de localisation présentés Figure 37. En effet, mis à part une inversion et le phénomène d'indétermination déjà évoqué plus haut, la localisation semble assez linéaire. De plus, fait remarquable, la localisation de la voix correspond quasiment à celle du wood-block.

On a pointé plus haut le fait que ce système semble avoir une zone d'écoute restreinte. Or, les commentaires concernant la justesse de localisation ressortent d'écoutes collectives alors que la courbe de localisation de la Figure 37 a été établie à partir d'écoutes individuelles. La dichotomie entre les deux semble aller dans le sens d'une dépendance importante selon la place d'écoute. A ce titre, il aurait été intéressant, mais malheureusement le temps ne l'a pas permis, d'effectuer les tests de localisation individuels avec les auditeurs excentrés afin de mesurer la distorsion de perception. Cela peut constituer une expérience complémentaire intéressante.

#### 7.4. L'ASM5 de Brauner

Comme on l'a signalé en 4.2., le système Brauner a été beaucoup moins apprécié lors des diffusions de la Convention (classé 7<sup>e</sup>) que lors des test d'écoute individuels (classé 4<sup>e</sup>). D'après la Figure 20, les différences principales concernent le timbre et l'équilibre avant/arrière. On se reportera au 4.2. pour l'analyse et une interprétation de ces différences. On rappelle seulement que les différences de jugement en timbre sont à rattacher aux enceintes, différentes lors des 2 diffusions, et que les jugements concernant l'équilibre avant/arrière tend à conclure que cet équilibre est fortement dépendant de la position de l'auditeur par rapport à l'axe  $-90^{\circ}/+90^{\circ}$ .

En se basant sur les résultats du test d'écoute individuel - le seul exploitable statistiquement de façon fiable - on obtient pour ce système la hiérarchie des critères suivantes :

1<sup>er</sup> classé : équilibre avant/arrière, oubli des haut-parleurs, homogénéité

2<sup>e</sup> classé : enveloppement

3<sup>e</sup> classé : timbre

4<sup>e</sup> classé : naturel

5<sup>e</sup> classé : stabilité en localisation

6<sup>e</sup> classé : sources sur les côtés, définition

7<sup>e</sup> classé : précision en localisation, justesse de localisation (écoutes collectives)

On voit, ce qui est résumé Figure 23, que ce système semble privilégier le rendu spatial par rapport au rendu des sources.

Le caractère critique de l'équilibre avant/arrière, évoqué ci-dessus, apparaît sur la Figure 39. En effet, en ce qui concerne le demi-espace avant, on constate une alternance d'inversions vers l'arrière (pour le woodblock : 20°, 30°, 50°, 70°, 80°) et de localisation assez bonne (0°, 40°, 60°). On ne peut donc pas dire que ce système est globalement déséquilibré vers l'arrière, mais on constate que l'équilibre est critique, avec une tendance au basculement arrière. Cela explique également le jugement négatif sur la stabilité, la précision et la justesse de localisation.

Au vu de la constatation précédente, on peut penser que la séparation acoustique entre les voies arrière et avant n'est pas suffisante, bien que la directivité adoptée ait été hypercardioïde. Si cette séparation avait pu être réalisée (par exemple par un délai plus important entre l'arrière et l'avant) et que l'on "replie" les inversions de la Figure 39 vers l'avant (par exemple, la perception du signal à 20° non plus à 171° mais à  $180^{\circ}-171^{\circ} = 9^{\circ}$ ), on constate que la localisation serait assez régulière (localisation des sources de 0° à 80° vers la zone 0°/55°).

Cette séparation aurait peut être été réalisable par l'allongement des bras arrière. Cela tend à valider les qualités théoriques du système INA5 en ce qui concerne la localisation avant (bien que la directivité prévue soit cardioïde).

Par contre, même en simulant l'absence de repliement vers l'arrière, on constate l'absence de localisation entre 55° et 110°, ce qui va dans le sens du jugement négatif concernant la perception de sources latérales.

Enfin, en ce qui concerne l'arrière, on constate une accumulation sur les haut-parleurs et peu de localisation entre. Cela s'explique par la réduction de l'angle de prise de son théorique (60° avec des cardioïdes) du fait de l'utilisation de la directivité hypercardioïde. La non-linéarité déjà prévisible sur l'arrière est ici augmentée.

On se souvient que le système Brauner était celui placé le plus en hauteur. De ce fait, ce système captait beaucoup de champ réverbéré. L'utilisation de la directivité hypercardioïde avait pour but de limiter cela, mais il faut noter que les microphones (horizontaux) pointaient au-dessus des chanteurs comme on le voit sur la Figure 4 ; de ce fait, le signal direct n'était pas totalement favorisé par le choix de la directivité.

Il est probable que cette présence importante de champ réverbéré soit intervenue dans le jugement très positif sur l'homogénéité et l'enveloppement. De plus, il faut noter que parmi les systèmes ponctuels (c'est-à-dire tous sauf ceux du CNSMDP et de Radio France), c'est ce système qui a les microphones les plus éloignés les uns des autres, ce qui crée une décorrélation favorable à la sensation d'espace. Cela permet aussi de comprendre le jugement très positif concernant l'oubli des haut-parleurs.

Par contre, il est probable que la présence importante de champ réverbéré ait nui à la définition, mal jugée. A cela s'ajoute les images fantômes multiples entre l'arrière et l'avant, comme on peut le supposer au vu des inversions avant/arrière.

## **7.5. Le Compact Surround System de N.Jacques et A.Mora**

Comme on l'a souligné en 4.2., ce système a également été jugé très différemment entre les écoutes collectives de la Convention, où il a été fort apprécié (2<sup>nd</sup> classé), et les écoutes individuelles (5<sup>e</sup> classé). L'analyse qu'on en a fait en 4.2. tend à supposer que ce système n'est peut-être pas optimal pour un auditeur correctement positionné entre les enceintes mais possède une zone d'écoute assez étendue, ce qui peut être un compromis intéressant.

A partir du résultat des écoutes individuelles, Figure 20 et du classement de la Figure 35, on peut classer les critères d'appréciation pour ce système en fonction du classement qu'il a obtenu :

1<sup>er</sup> classé : justesse de localisation (écoutes collectives)

2<sup>e</sup> classé : équilibre avant/arrière, stabilité en localisation

4<sup>e</sup> classé : oubli des haut-parleurs, homogénéité, précision en localisation

5<sup>e</sup> classé : sources sur les côtés, enveloppement, naturel, définition

6<sup>e</sup> classé : timbre

La courbe de localisation établie Figure 38 va dans le sens du très bon jugement concernant la justesse de localisation. En effet, c'est certainement celle qui apparaît la plus linéaire sur l'ensemble des systèmes (du moins sur l'avant et sur l'arrière). De plus, on n'observe aucune inversion avant/arrière. On n'a aucune localisation entre 53° et 108°, ce qui va dans le sens de ce qui a été observé sur la plupart des systèmes. On constate 2 indéterminations, à 90° et 110°. Cette courbe tend à valider la théorie utilisée concernant la localisation, avec en particulier l'attention portée à la linéarité de celle-ci (par le choix d'une égalité entre les angles de prise de son et les angles physiques entre les haut-parleurs, détaillé en 2.3.). De plus, l'utilisation de délais entre l'avant et l'arrière explique certainement l'absence d'inversion.

Il est également intéressant de constater, toujours sur la Figure 38, qu'il y a une grande similitude entre la localisation de la voix et celle du woodblock. Cela, ainsi que l'absence d'inversion avant/arrière, justifie le bon jugement concernant la stabilité en localisation.

L'équilibre avant/arrière, également plébiscité, se lit aussi sur la courbe de la Figure 38 : la même portion d'espace (180°) est perçue sur le demi-espace avant et sur le demi-espace arrière.

Par contre, la présence de sources sur les côtés ne s'est pas vraiment faite sentir. En ce sens, l'ajout de délai entre l'avant et l'arrière ne permet peut-être pas une optimisation simultanée de l'absence d'inversion avant/arrière et de réglage des angles de prise de son latéraux. En effet, les délais nécessaires pour éviter une inversion devraient être d'au moins 3ms, ce qui conduit à un basculement rapide des sources entre l'avant et l'arrière (ce qui est visible sur la Figure 38). Cela peut expliquer le mauvais jugement concernant les sources latérales, mais également l'homogénéité, l'enveloppement et le naturel pour les écoutes individuelles.

Mais c'est certainement également ces délais qui ont conduit à un bon jugement de ces critères spatiaux d'homogénéité, d'enveloppement et de naturel lors de la Convention (voir Figure 20), par l'agrandissement de la zone d'écoute et la limitation des phénomènes de basculement avant/arrière en fonction de la place de l'auditeur.

Enfin, le mauvais jugement des critères de précision en localisation, définition et timbre pourraient suggérer la présence d'images fantômes multiples, d'autant plus que les microphones sont relativement proches. Le désir d'avoir un système compact augmente ce risque d'images fantômes multiples (ce n'est pas le délai avant/arrière qui va les supprimer, tout au plus cela va les étaler).

Cependant, on a abordé en 4.2. le fait que les jugements concernant ces critères a été beaucoup plus positif lors des écoutes collectives, et on y renvoie le lecteur pour l'analyse de cela.

## **7.6. Le Multichannel Microphone Array de M.Williams et G.Le Dû**

Ce système a été classé 6<sup>e</sup>, que ce soit au cours des écoutes de la Convention au au cours des écoutes individuelles (Figure 20). On constate Figure 23 que ce système semble néanmoins privilégier le rendu des sources au rendu spatial.

En se basant sur les Figures 20 et 35, on obtient, comme pour les autres systèmes, le classement suivant :

3<sup>e</sup> classé : timbre

4<sup>e</sup> classé : sources sur les côtés, oubli des haut-parleurs, définition, justesse de localisation

5<sup>e</sup> classé : précision en localisation, stabilité en localisation

7<sup>e</sup> classé : équilibre avant/arrière, enveloppement, homogénéité, naturel

Le timbre est le critère pour lequel ce système a été le mieux classé. Il est fort probable que l'utilisation de 2 microphones hypocarديوïdes MK21 à l'avant a favorisé le timbre (par rapport à l'utilisation de microphones cardioïdes MK4).

De plus, le fait que la définition fasse également partie des critères mieux appréciés peut nous faire penser que le problème des images fantômes multiples s'est peu posé. Cela vient déjà du fait que le canal central n'a pas été utilisé (voir 2.2.), ce qui supprime un couple de microphones. D'autre part, si on compare avec le système d'A.Mora et N.Jacques vu ci-dessus et basé sur les mêmes principes, les microphones ne sont pas aussi rapprochés.

Intéressons nous maintenant à la courbe de localisation établie Figure 41. On constate sur celle-ci que quasiment tout a été localisé sur l'arrière (seule indétermination à 30° et 40°). Cela explique le très mauvais jugement concernant l'équilibre avant/arrière (bien que, comme on le verra plus loin, une modification ait été effectuée sur le mixage après coup). Le fait que le canal central n'ait pas été utilisé du tout a pu conduire à un manque d'énergie à l'avant qui a favorisé ce basculement général vers l'arrière. D'autre part, l'utilisation de microphones hypocarديوïdes sur l'avant a diminué la séparation acoustique des sources venant de l'arrière et donc également favorisé ce basculement.

Enfin, cela pose la question de la post-production et du lieu dans lequel elle s'effectue. En effet, il faut signaler qu'au début des écoutes de la Convention, dans l'auditorium prévu à cet effet, M. Williams et G. Le Dû ont constaté que le mixage qu'ils avaient effectué en studio apparaissait très déséquilibré vers l'arrière, et ont fait modifier celui-ci (les tests de localisation, qui ont conduit à la courbe Figure 41, ont par contre été basé sur le mixage d'origine). Or, le mixage initial avait été effectué dans une régie du CNSMDP, du type LEDE (Live End Dead End). Ce type de régie, très répandu pour les écoutes stéréophoniques à 2 canaux, disymétrise l'espace en créant un demi-espace absorbant sur l'avant - pour limiter l'influence des premières réflexions sur la localisation, le timbre, ... - et un demi-espace diffusant sur l'arrière - pour augmenter la zone d'écoute et homogénéiser le son. Or, lors des écoutes collectives, l'auditorium présentait un traitement acoustique symétrique. Ces différences de traitement acoustique changent la perception relative entre l'avant et l'arrière.

Cela a permis de soulever une question fondamentale, celle de l'acoustique des lieux d'écoute en multicanal. D'autant plus qu'on a constaté plus haut l'importance donnée au critère d'équilibre entre l'avant et l'arrière par les auditeurs et leur grande sensibilité à tout déséquilibre. Or, on est actuellement dans le flou en ce qui concerne les recommandations concernant l'acoustique des lieux d'écoute, et ce sont souvent d'anciennes régies LEDE qui sont transformées en régies multicanales par le simple ajout d'enceintes à l'arrière... Il semble pourtant peu approprié d'adopter des conditions d'écoute disymétriques pour un système qui se veut symétrique dans son rendu (et d'autant plus dans le cadre de l'expérience où la source elle-même, le chœur, était totalement symétrique). Au vu de l'importance que revêt le critère d'équilibre avant/arrière, il semble important de mener une réflexion sur les conditions d'écoute parallèlement à l'élaboration des systèmes de prise de son. Heureusement certains, tels que R. Walker ([18]), mènent des recherches dans ce domaine.

Comme on l'a signalé plus haut, la courbe de localisation a été établie à partir du mixage initial, alors que les tests d'analyse de l'image sonore ont été basé sur le mixage modifié, avec moins de niveau sur l'arrière - ce qui n'a pas empêché un mauvais jugement de l'équilibre avant/arrière. Il est intéressant de tracer la courbe de localisation symétrique pour les sources de 0° à 90°, en supposant que l'augmentation de niveau sur l'avant ramène la localisation de l'arrière vers l'avant en supprimant les inversions : cela a été fait Figure 45. On constate alors une très grande linéarité de localisation, avec en particulier une très bonne localisation sur les côtés. Si la modification de niveau entre l'avant et l'arrière n'a pas nécessairement conduit à cette courbe, elle a dû permettre de s'en rapprocher ; cela justifie le fait que la perception de sources sur les côtés, la justesse de localisation et l'oubli des haut-parleurs aient fait partie des critères mieux jugés.

### Williams/Le Dû "replié"

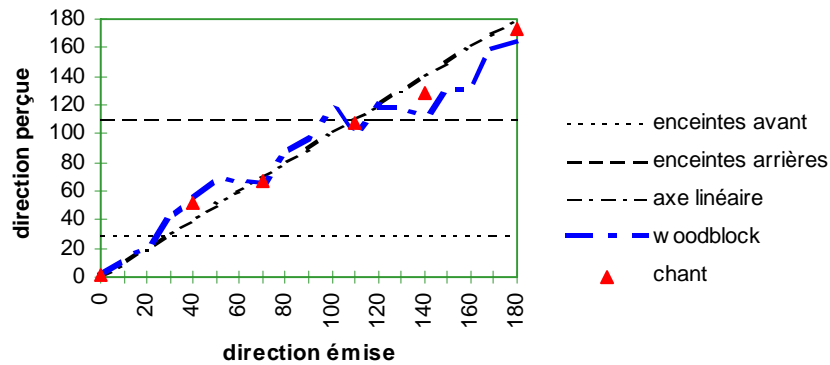


Figure 45

La précision en localisation et la stabilité ont été moins bien jugées. Il est probable que l'absence de canal central diminue la stabilité de localisation (en effet, ce canal a entre autre pour but une stabilisation de l'image frontale, et une augmentation de la zone d'écoute). Par contre, on constate sur la Figure 41 une très grande similitude de localisation entre la voix et le wood-block, ce qui dénote peu de dépendance de la localisation par rapport au contenu spectral de la source.

Les critères spatiaux d'enveloppement, d'homogénéité et de naturel ont été très mal jugés. Cela est d'ailleurs en partie relié au mauvais équilibre avant/arrière. De plus, ce système utilise des microphones assez proches, donc peu décorrélés, ce qui ne favorise pas l'impression d'espace.

Bien que ce système ait mal été jugé dans l'expérience, certaines constatations semblent nous montrer qu'il peut avoir des capacités de localisation très performantes, mais qu'elle nécessitent un dosage très fin entre le niveau des canaux avant et arrière. Cela est une contrainte importante, dans la mesure où, comme on l'a vu, la perception de l'équilibre avant/arrière est fort dépendante du traitement acoustique du lieu. On voit que, malgré la théorie, un dosage doit se faire réellement à l'écoute.

## 7.7. Le système du Grenier à Son

C'est le système le moins bien jugé lors des écoutes individuelles. Il avait été un peu mieux apprécié au cours des écoutes de la Convention. Le classement des critères, à partir des Figures 20 et 35 donne :

5<sup>e</sup> classé : stabilité en localisation

6<sup>e</sup> classé : équilibre avant/arrière, sources sur les côtés, oubli des haut-parleurs, enveloppement, homogénéité, naturel, précision en localisation, justesse de localisation

7<sup>e</sup> classé : timbre, définition

Il faut être très prudent dans l'interprétation de ces résultats, car on a constaté en 6.3., suite aux tests de localisation, qu'il y a certainement eu un problème de décodage du système Soundfield (sur le demi-espace gauche). Cela conduit inévitablement à un mauvais jugement concernant les sources sur les côtés, l'enveloppement, l'homogénéité, le naturel, la précision et la justesse de localisation.

Néanmoins, du fait de l'utilisation de capsules totalement coïncidentes dans le microphone Mark V, et donc l'absence de corrélation, il est prévisible que le rendu spatial soit limité (critères d'enveloppement, d'homogénéité et de naturel).

L'ajout des signaux de la tête artificielle (voir 2.7.) ne peut compenser cela. En effet, ces signaux alimentent uniquement les canaux avant gauche et droit et n'ont pas été traités par un système transaural. Cela ne permet pas l'enveloppement souhaité. Il serait intéressant d'écouter le résultat avec un traitement transaural des signaux de la tête artificielle. Néanmoins, l'ajout de ces signaux ne doit pas détruire l'image créée par le système Soundfield et est donc limité.

On peut en fait peut-être voir ce système hybride comme modulable en fonction des exigences : le rapport de dosage entre la tête artificielle et le système Soundfield permettant d'obtenir plus d'effet d'espace (caractéristique de la tête artificielle) ou plus de localisation (caractéristique du Soundfield).

En ce qui concerne la localisation, on a vu que les résultats du test sont entâchés d'un problème de décodage du Soundfield. Reconsidérons la Figure 42 et gardons uniquement les résultats concernant les sources présentées dans le demi-espace droit (voir Figure 34 pour le détail des sources présentées au cours du test) ; on obtient la Figure 46. On se rend compte que le rendu de ces sources est très linéaire, et que la localisation sur les côtés est présente. On peut donc entrevoir une réelle capacité en localisation du système Soundfield. On constate en particulier qu'il n'y a pas de divergence en localisation entre les signaux de woodblock et de la voix, donc une réelle stabilité en localisation (c'est d'ailleurs le critère qui a été le moins mal jugé).



### Le Grenier à Son (partiel)

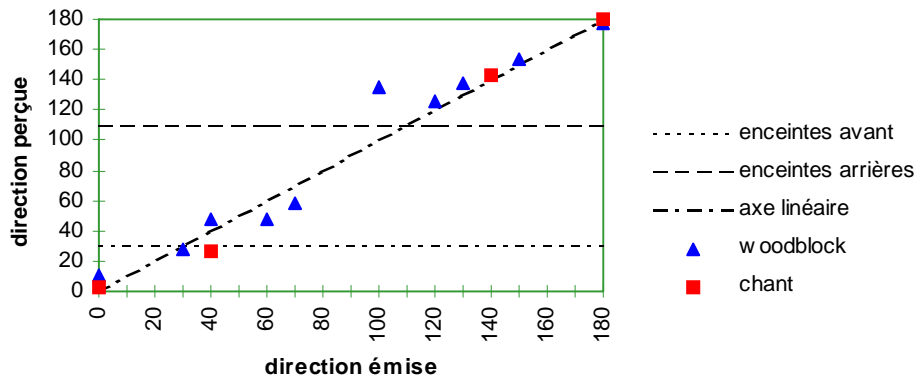


Figure 43

Les 2 critères les plus mal jugés sont le timbre et la définition. En ce qui concerne le timbre, les concepteurs soulignaient la relative pauvreté de timbre du système Soundfield ; cela est en partie dû à l'utilisation de capsules coïncidentes, ainsi qu'aux opérations sur les signaux réalisées dans le décodeur. L'utilisation des microphones omnidirectionnels de la tête artificielle avaient en partie pour but d'améliorer le rendu spectral ; néanmoins, du fait de la relative proximité de la tête et du microphone Mark V, on peut supposer que l'on a subi l'effet d'un filtrage en peigne entre leurs signaux respectifs. La définition a dû également pâtir de ce phénomène.

Ainsi, si les résultats de l'expérience sont peu exploitables, on peut néanmoins entrevoir les propriétés de ce système. Il serait intéressant de juger de l'effet transaural sur la tête artificielle.

## Conclusion

Les étapes successives de ce travail de recherche ont permis, dans un premier temps, de déterminer les critères utilisés par des auditeurs "spécialisés" et de cerner leurs principales attentes dans le rendu d'images sonores multicanales grâce aux commentaires libres récoltés lors de la Convention de l'AES.

Grâce aux tests individuels d'analyse et de localisation, nous avons pu dégager les caractéristiques de chacun des systèmes impliqués dans l'expérience. Mais comme nous l'avons signalé dès l'introduction, l'objet de ce travail n'était pas de classer de façon définitive ces différents systèmes. On a en effet vu lors de l'analyse que certains systèmes n'ont peut-être pas été exploités au mieux (cas du système de M. Williams et G. Le Dû ou de celui du Grenier à Son) ; il faudrait donc se baser sur d'autres enregistrements. De plus, comme on l'a souligné lors de la présentation de la session d'enregistrement, la disposition des systèmes n'était pas totalement libre (il fallait tous les positionner simultanément), ce qui fait que tous n'étaient pas nécessairement positionnés de façon optimale (par exemple, l'ASM5 était probablement un peu trop haut). Enfin, les œuvres enregistrées présentaient la particularité d'être spatialisées sur 360°, ce qui est une situation peu fréquente.

Plutôt que de m'arrêter au résultat de tel ou tel système dans le cadre de l'expérience, je me suis efforcé à mettre en regard ces résultats avec les principes qui sous-tendent les systèmes. Ce travail doit, d'une part, aller vers une meilleure connaissance des caractéristiques de chacun des systèmes utilisés ici et, d'autre part, donner des bases permettant le développement éventuel d'autres systèmes.

On constate, dans l'analyse des résultats, que plusieurs "acquis" de la stéréophonie à 2 canaux s'appliquent en stéréophonie à 5 canaux, en particulier le fait que des systèmes à microphones omnidirectionnels distants privilégient le rendu spatial, alors que des systèmes coïncidents privilégient la précision et la stabilité de localisation.

Mais certains problèmes nouveaux apparaissent, tels que le phénomène des sources fantômes multiples, qui peut nuire à la définition, au timbre et à la précision de localisation. De même, on a noté le caractère critique de l'équilibre entre l'arrière et l'avant et la gêne provoquée par une présence des voies arrières trop importante ou par des inversions avant/arrière.

Il est également intéressant de noter que la justesse de localisation n'est peut-être pas le critère le plus important : le système qui a été le plus apprécié a des qualités de localisation médiocres. Toujours en ce qui concerne la localisation, on constate qu'aucun système dans le cadre de l'expérience n'a réellement permis de localisation latérale, entre 50° et 110° (la sensation de source latérale traduit plutôt une extra-largeur entre l'avant et l'arrière).

On a également été confronté aux différences de taille de zone d'écoute privilégiée selon les systèmes.

Les résultats ont également soulevé des questions fondamentales telles que la réflexion à mener sur l'acoustique des lieux de mixage et d'écoute en multicanal.

Ce travail constitue une étape dans l'analyse des systèmes de prise de son multicanale, qui s'est intéressé à des systèmes principaux, captant une image sonore globale. Il serait intéressant d'étudier de quelle façon ils peuvent se combiner à des microphones d'appoint.

## **Annexe 1**

### **Recommandation ITU-R BS 775-1**

## RECOMMENDATION ITU-R BS.775-1\*

MULTICHANNEL STEREOPHONIC SOUND SYSTEM WITH  
AND WITHOUT ACCOMPANYING PICTURE

(Question ITU-R 79/10)

(1992-1994)

The ITU Radiocommunication Assembly,

*considering*

- a) that it is widely recognized that a two-channel sound system has serious limitations and improved presentation is necessary;
- b) that the requirements of cinema presentation differ from those that apply in the home, particularly with respect to room and screen size and distribution of listeners, but that the same programmes may be reproduced in either the cinema or the home;
- c) that broadcast HDTV signals, and those delivered by other media, should be capable of giving appropriate sound quality with a wide range of domestic loudspeaker configurations, including compatibility with two-channel stereophonic and monophonic listening;
- d) that for multichannel sound it is desirable to separate the requirements of production, delivery and domestic presentation, though these are mutually interacting;
- e) that investigations about multichannel sound transmission and reproduction associated and not associated with accompanying picture are being carried out with the basic requirements as laid down in Annex 2;
- f) that one universal multichannel sound system applicable to both sound and television broadcasting would be beneficial to the listener;
- g) that compromises may be necessary to ensure that the system is as universal and as practical as possible;
- h) that a hierarchy of compatible sound systems for broadcasting, cinema and recordings is useful for programme exchange and up- and down-mixing depending on the programme material;
- j) that ancillary services such as those for the visually impaired and hearing impaired are desirable;
- k) that advances in digital audio coding currently allow the delivery of multiple audio channels in an efficient manner,

*recommends*

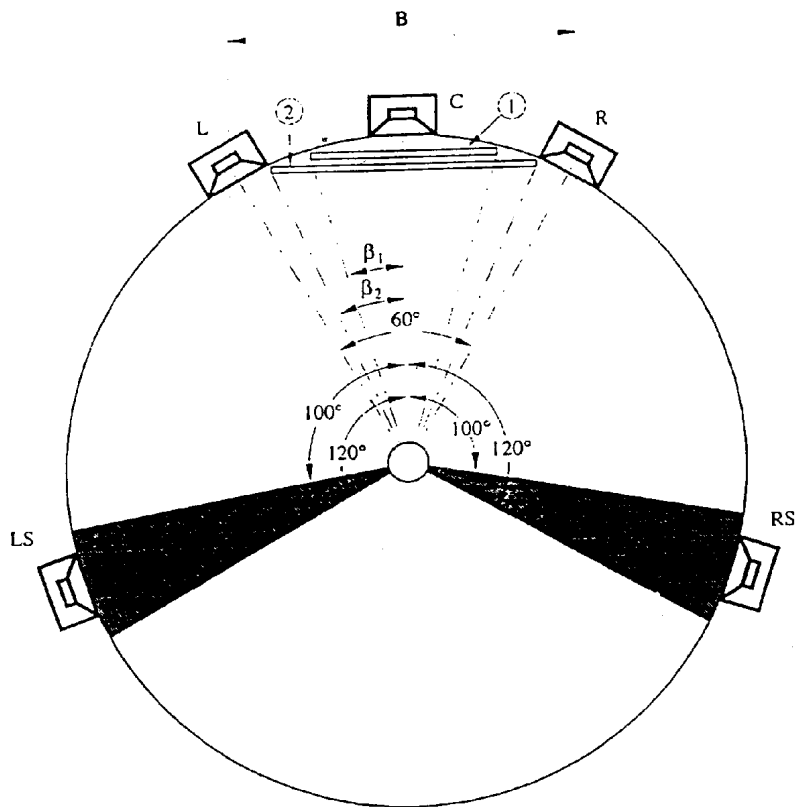
- 1. one universal multichannel stereophonic sound system, with or without accompanying picture, within a hierarchy given in Annex 1;
- 2. the following reference loudspeaker arrangement (see Fig. 1):
  - three front loudspeakers combined with two rear/side loudspeakers (Note 1);
  - the left and right frontal loudspeakers are placed at the extremities of an arc subtending 60° at the reference listening point (Notes 2 and 3);

Where for reasons of available space, it is preferred to place the frontal loudspeakers on a straight line base, then it may be necessary to introduce compensating time delays in the signal feed of the centre loudspeaker;

\* This Recommendation should be brought to the attention of the International Electrotechnical Commission (IEC) and the Society of Motion Picture and Television Engineers (SMPTE).

- both side/rear loudspeakers should be placed within the sectors from 100° to 120° from the centre front reference. Precise location is not necessary. Side/rear loudspeakers should be no closer to the listener than the frontal loudspeakers, unless compensating time delay is introduced (Note 4);
- the frontal loudspeakers should ideally be at a height approximately equal to that of the listener's ears. This implies an acoustically transparent screen. Where a non-acoustically transparent screen is used, the centre loudspeaker should be placed immediately above or below the picture. The height of side/rear loudspeakers is less critical.

FIGURE 1  
Reference loudspeaker arrangement with  
loudspeakers L/C/R and LS/RS



Screen 1 HDTV - Reference distance =  $3H$  ( $2\beta_1 = 33^\circ$ )

Screen 2 =  $2H$  ( $2\beta_2 = 48^\circ$ )

H: height of screen

B: loudspeaker base width

Loudspeaker	Horizontal angle from centre (degrees)	Height (m)	Inclination (degrees)
C	0	1.2	0
L, R	30	1.2	0
LS, RS	100 ... 120	$\geq 1.2$	0 ... 15 down

D01

3. the use of five reference recording/transmission signals for left (L), right (R), centre (C), channels for the front, and left surround (LS) and right surround (RS) channels for the side/rear. Additionally the system may include a low frequency extension signal for a low frequency extension (LFE) channel (see Annex 7);

In circumstances where transmission capacity or other constraints apply, the three front signals can be combined with one (mono surround, MS) or zero rear/side signals. In the case of mono surround, the MS signal is fed to both LS and RS loudspeakers (see Fig. 1);

4. compatibility, if required, with existing and low cost receivers by using one of the methods given in Annex 3;

5. down-mixing capability, if required, for reduction of the number of channels, either prior to transmission or at the receiver, by employing the down-mixing equations given in Table 2;

6. upward conversion where an increase in the number of channels is desired, either prior to transmission or at the receiver, by employing upwards-conversion techniques described in Annex 5;

7. overall quality to the requirements of Annex 2;

8. provision (but see also § 9 below) for the following if necessary:

- alternate multiple language principal services;
- one or more independent channels carrying descriptive information for the visually impaired;
- one or more independent channels for the purpose of supplying improved intelligibility to the hearing impaired;

9. additional data transmitted with the audio to enable the flexible use of the data capacity allocatable to audio signals (see Annex 6).

*Note 1* – Optionally, there may be an even number of more than two rear/side loudspeakers which may provide a larger optimum listening area and greater envelopment.

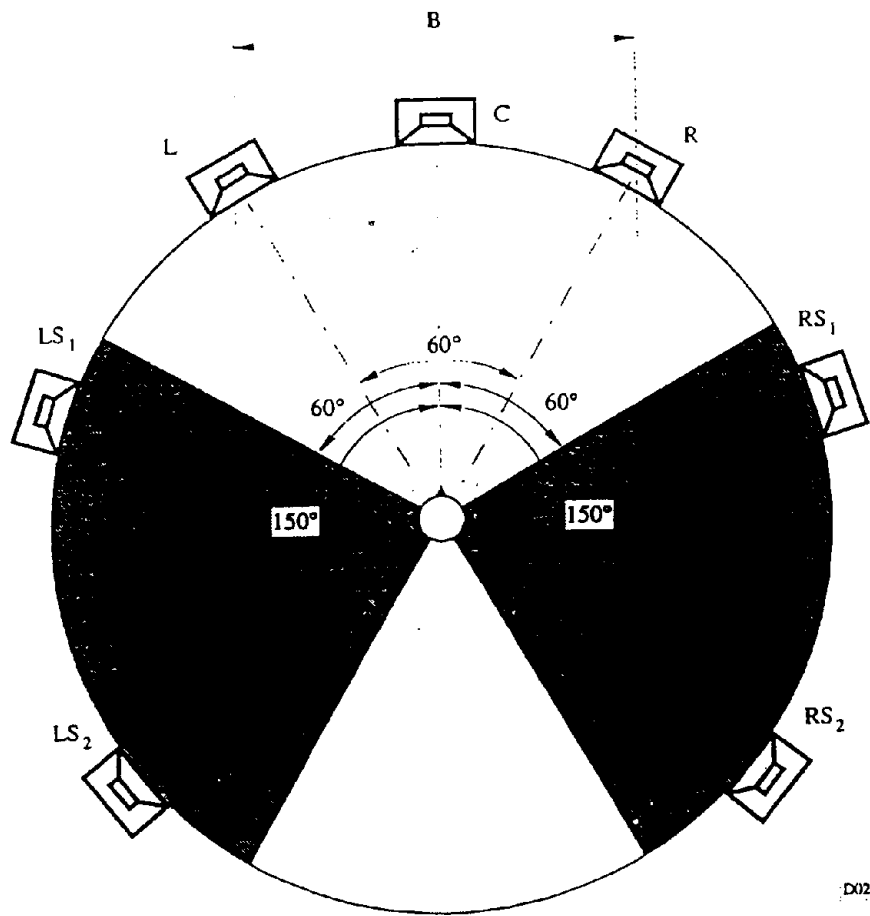
*Note 2* – Optimum sound reproduction requires use of wide angular spacing between the left and right loudspeakers of two or three front loudspeaker channel stereophonic systems (see Fig. 1). It is recognized that the television pictures accompanying stereophonic sound having such an angular width cannot, with current techniques, be displayed to the same wide angles, but are often restricted to 33° horizontal subtended angle at the reference distance, although cinema images may be displayed at such angles (see Fig. 1). The resulting mismatch between picture and sound image width leads to differences in mixing technique for cinema and television. It is expected that larger television displays will lead to better compatibility of mixes for cinema and television display.

*Note 3* – The size of the loudspeaker basewidth, B (see Fig. 1), is defined for reference listening test conditions in Recommendation ITU-R BS.1116 "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems".

*Note 4* – If more than two rear/side loudspeakers are used, then the loudspeakers should be disposed symmetrically and at equal intervals on the arc which measures from 60° to 150° from the centre front reference (see Fig. 2).

*Note 5* – If more than two rear/side loudspeakers are used, the LS signal should be fed to each of the side/rear loudspeakers on the left side of the room and the RS signal should be fed to each of the side/rear loudspeakers on the right side of the room. In doing so, it will be necessary to reduce the signal gain such that the total power emitted by the loudspeakers carrying the LS (or RS) signal is the same as if that signal had been reproduced over a single loudspeaker. For large room reproduction, it may also be necessary to delay, or otherwise decorrelate, the feeds to some or all of the side/rear loudspeakers. Further studies on such decorrelation is necessary.

FIGURE 2  
Optional 3/4 loudspeaker arrangement (3 front and 4 surround)

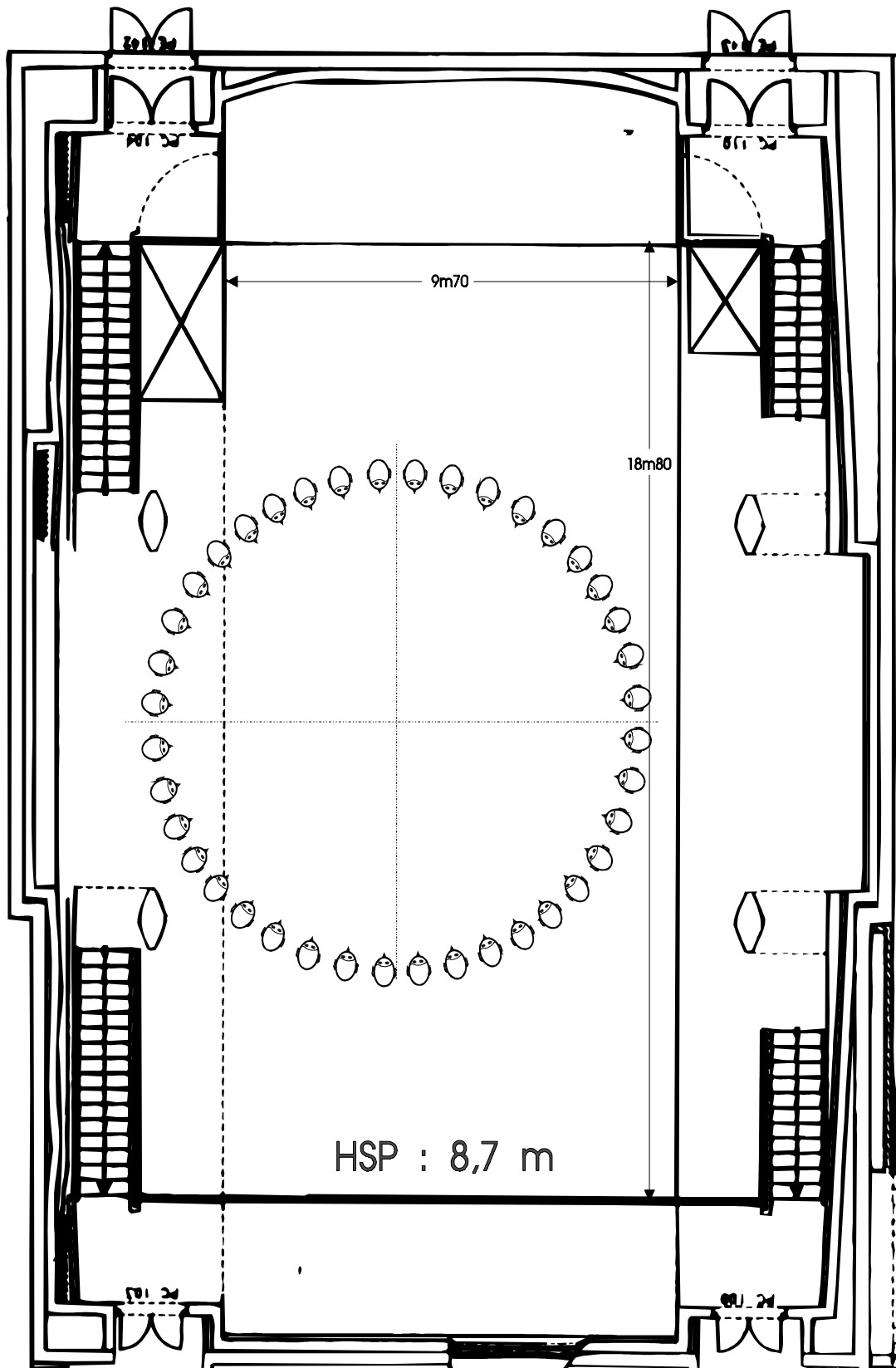


D02

## **Annexe 2**

### **Plan du grand plateau d'orchestre**





## **Annexe 3**

### **Feuilles de commentaires de la Convention**

Impressions

position n° 13

1st. sequence

Motet - Tallis	Percussion
de se'quilibre <sup>A</sup> AV - AR son pauvre.	Tout à l'AR <sup>A</sup> peu à <sup>D</sup> .
Equilibre instable <sup>B</sup> peu à l'AV peu de grave.	meilleure <sup>B</sup> localisation tjs favorable AR.
rien <sup>C</sup> AV rien sur les côtés. son + riche	Bien à l'AV <sup>C</sup>
+ homogène en spatialisation <sup>D</sup> meilleure définition sonorité + large	image v.h. <sup>D</sup> belle un peu trop sur les côtés non homogène. phase! équilibre en timbre
image trop <sup>E</sup> sur l'AV de concitèci de l'AR	+ précis sur l'AR <sup>E</sup> que sur l'AV son laid
T Bien spécialisé <sup>F</sup> en fin ça manque. le + beau timbre	le + homogène <sup>F</sup> pas de trou AV/AR + belle sonorité
légère <sup>G</sup> prédominance AR message + riche sur les voies AR, que AV pauvre en timbre sur les côtés AV	tjs pas tjs <sup>G</sup> clair sur les côtés. ou bien non homogène

Impressions

position n° 9

1 ~~X~~ sequence

Motet - Tallis	Percussion
<p><sup>A</sup> Mainly in front, and partly to the rear. Low freq &amp; High not coherent Localisation = weak</p>	<p><sup>A</sup> Front only Localisation = very weak</p>
<p><sup>B</sup> Good localisation +++ Room is clearly perceived</p>	<p><sup>B</sup> Good +++</p>
<p><sup>C</sup> Hole between front &amp; rear. Mainly in front.</p>	<p><sup>C</sup> Front &amp; Right OK left Back = absent</p>
<p><sup>D</sup> Localisation = ++ Rear sources come from loudspeakers</p>	<p><sup>D</sup> Good ++</p>
<p><sup>E</sup> Weak localisation Sources come from loudspeakers</p>	<p><sup>E</sup> Center is spread and its localisation is uncertain</p>
<p><sup>F</sup> Localisation is good. Treble sounds are sometimes not coherent spatially</p>	<p><sup>F</sup> Mean localisation Front seems wider than in reality.</p>
<p><sup>G</sup> Localisation is weak Sources come from loudspeakers</p>	<p><sup>G</sup> Localisation of Bass = weak Treble = OK</p>

## **Annexe 4**

### **Critères et feuille de notation des tests individuels d'analyse**

Le but de ce test est d'analyser l'image sonore multicanale d'un enregistrement réalisé avec différents systèmes de prise de son destinés à une diffusion 5.0 (trois enceintes frontales, deux arrières). Merci de participer à cette expérience. Vous allez entendre huit prises de son, réalisées simultanément, d'un même extrait d'un chœur de 42 chanteurs (femmes et hommes) entourant totalement les systèmes de prise de son (positionnés au centre du chœur). Vous allez donc être mis dans une situation d'écoute vous positionnant au centre du cercle formé par les choristes.

Pour chacune des prises de son, vous allez avoir à noter différents critères sur une échelle de 1 à 5 (n'hésitez pas à utiliser toute l'échelle de notation). Vous noterez en entourant la note que vous souhaitez attribuer.

Les dix critères retenus sont :

1. **Equilibre entre les voies avant et arrière** : l'équilibre entre les haut-parleurs avant et arrière vous semble-t-il bon ou déséquilibré (cela peut être au profit de l'arrière ou de l'avant) ?
2. **Sources fantômes sur les côtés** : percevez-vous des sources fantômes sur les côtés (entre les haut-parleurs avant et arrière) ou non ?
3. **Présence des haut-parleurs** : les haut-parleurs se font-ils ressentir (le son semble-t-il provenir de cinq points discrets) ou oublier ?
4. **Sensation d'enveloppement** : vous sentez-vous enveloppé par le son (sensation d'être au centre du chœur) ou non ?
5. **Sensation d'homogénéité** : l'image sonore vous semble-t-elle homogène ou déséquilibrée au profit de certaines zones angulaires ?
6. **Sensation de naturel** : l'image sonore vous paraît-elle naturelle ou non ?
7. **Précision de localisation** : localisez-vous de façon précise ou diffuse les voix ?
8. **Stabilité de localisation** : la localisation des voix vous semble-t-elle stable ou instable ?
9. **Timbre** : le timbre des voix vous semble-t-il réaliste et agréable ou non ?
10. **Définition** : la définition des voix vous paraît-elle précise ou confuse ?

Les extraits durent environ trois minutes.

Posez maintenant toutes les questions que vous désirez.

### 1. Equilibre entre les voies avant et arrière

Mal équilibré 

1	2	3	4	5
---	---	---	---	---

 Bien équilibré

### 2. Sources fantômes sur les côtés

Non perçues 

1	2	3	4	5
---	---	---	---	---

 Bien perçues

### 3. Présence des haut-parleurs

Trop ressentis 

1	2	3	4	5
---	---	---	---	---

 Oubliés

### 4. Sensation d'enveloppement

Enveloppement non ressenti 

1	2	3	4	5
---	---	---	---	---

 Enveloppement bien ressenti

### 5. Sensation d'homogénéité

Pas homogène 

1	2	3	4	5
---	---	---	---	---

 Très homogène

### 6. Sensation de naturel

Pas naturel 

1	2	3	4	5
---	---	---	---	---

 Très naturel

### 7. Précision de localisation

Localisation diffuse 

1	2	3	4	5
---	---	---	---	---

 Localisation précise

### 8. Stabilité de localisation

Localisation instable 

1	2	3	4	5
---	---	---	---	---

 Localisation stable

### 9. Timbre

Mauvais 

1	2	3	4	5
---	---	---	---	---

 Très bon

### 10. Définition

Confuse 

1	2	3	4	5
---	---	---	---	---

 Très précise

## **Annexe 5**

### **Résultats complets des tests individuels d'analyse d'écoute**



Radio France										
Equilibre av/ar	Sources côtés	Présence HP	Enveloppement	Homogénéité	Naturel	Précis. localis.	Stabilité localis.	Timbre	Définition	
3	3	2	5	2	4	2	3	4	4	
3	3	1	2	4	1	1	1	1	1	
5	4	5	5	4	5	5	4	5	5	
3	4	5	5	4	4	5	4	4	4	
5	5	5	4	3	4	4	5	5	4	
4	4	4	5	4	2	5	5	1	5	
5	4	5	5	4	4	4	3	4	4	
3	2	2	2	2	2	3	3	3	2	
4	3	2	3	3	4	4	4	3	4	
4	2	2	4	2	3	4	4	4	3	
2	3	2	4	2	4	3	3	3	2	
4	5	5	5	5	4	4	4	4	4	
5	4	2	4	4	4	3	3	4	4	
4	4	4	4	4	4	3	3	4	3	
2	2	3	4	1	2	3	4	2	4	
Moy.	3,73	3,47	3,27	4,07	3,20	3,40	3,53	3,53	3,40	3,53
Ec. type	1,03	0,99	1,49	1,03	1,15	1,12	1,13	0,99	1,24	1,13

Moy. générale	3,51
Ecart type	0,25

KFM 360										
Equilibre av/ar	Sources côtés	Présence HP	Enveloppement	Homogénéité	Naturel	Précis. localis.	Stabilité localis.	Timbre	Définition	
5	4	3	5	2	3	2	3	4	2	
2	2	1	3	2	2	2	3	2	2	
5	5	4	5	4	5	5	5	4	5	
4	4	3	3	4	3	4	4	3	4	
1	2	2	1	2	1	4	5	4	5	
4	4	1	3	3	3	5	5	4	4	
5	4	4	5	4	4	4	4	4	2	
4	4	5	4	4	4	4	4	4	2	
4	2	3	4	4	4	4	3	4	5	
2	3	2	2	2	3	4	3	3	2	
3	4	2	4	2	4	4	4	2	2	
1	5	4	4	2	4	4	5	4	5	
5	5	4	4	2	4	4	3	4	4	
4	3	3	4	4	4	3	3	4	3	
4	4	2	4	2	3	4	4	3	4	
Moy.	3,53	3,67	2,87	3,67	2,87	3,40	3,80	3,87	3,53	3,40
Ec. type	1,41	1,05	1,19	1,11	0,99	0,99	0,86	0,83	0,74	1,30

Moy. générale	3,46
Ecart type	0,34776

CNSMDP										
Equilibre av/ar	Sources côtés	Présence HP	Enveloppement	Homogénéité	Naturel	Précis. localis.	Stabilité localis.	Timbre	Définition	
2	2	1	1	1	1	2	3	2	1	
5	3	4	4	5	5	5	4	5	4	
4	4	4	4	3	4	4	3	3	3	
3	3	4	2	3	2	2	3	3	2	
4	4	4	5	5	5	3	4	4	4	
5	5	5	5	5	5	5	5	5	5	
3	4	5	4	4	5	4	4	4	4	
3	4	1	4	3	3	2	2	2	1	
2	2	1	2	4	2	5	4	2	3	
2	3	2	2	3	2	4	3	2	2	
3	5	4	5	4	3	2	2	3	2	
2	5	4	4	1	4	2	4	5	2	
4	4	4	3	4	4	4	3	4	5	
4	4	4	4	4	4	4	4	4	4	
2	2	2	3	3	2	3	4	2	2	
Moy.	3,20	3,60	3,27	3,47	3,47	3,40	3,40	3,47	3,33	2,93
Ec. type	1,08	1,06	1,44	1,25	1,25	1,35	1,18	0,83	1,18	1,33

Moy. générale	3,35
Ecart type	0,19

Brauner										
Equilibre av/ar	Sources côtés	Présence HP	Enveloppement	Homogénéité	Naturel	Précis. localis.	Stabilité localis.	Timbre	Définition	
5	4	5	5	4	3	2	3	2	3	
5	1	4	5	3	2	3	3	5	3	
5	4	5	5	5	5	4	5	5	4	
4	4	3	3	3	4	2	3	2	2	
3	2	2	4	4	4	2	3	2	1	
4	5	5	5	5	5	1	4	5	2	
4	3	3	4	4	3	2	1	4	2	
2	1	2	4	2	2	1	1	3	2	
4	2	3	4	4	4	4	4	4	3	
5	4	3	3	4	4	4	4	4	3	
3	3	3	4	4	3	2	3	2	2	
5	5	5	5	3	4	2	5	5	4	
5	1	4	1	3	2	2	3	2	2	
1	2	2	2	2	2	3	4	2	2	
5	4	5	5	4	3	3	3	3	2	
Moy.	4,00	3,00	3,60	3,93	3,60	3,33	2,47	3,27	3,33	2,47
Ec. type	1,25	1,41	1,18	1,22	0,91	1,05	0,99	1,16	1,29	0,83

Moy. générale	3,30
Ecart type	0,53

Jacques/Mora										
Equilibre av/ar	Sources côtés	Présence HP	Enveloppement	Homogénéité	Naturel	Précis. localis.	Stabilité localis.	Timbre	Définition	
5	4	3	4	3	4	4	4	4	3	
5	5	5	5	5	4	2	4	4	2	
5	3	4	4	3	4	4	4	4	5	
4	2	3	2	2	2	4	5	4	3	
2	2	3	3	5	2	1	1	2	1	
4	4	2	5	2	4	2	3	4	1	
3	3	3	3	3	4	1	3	3	2	
5	5	5	4	4	3	4	5	3	2	
2	2	3	2	3	3	3	3	2	4	
5	3	2	3	2	2	4	4	4	4	
3	4	2	2	2	2	4	4	3	3	
2	2	2	1	1	2	4	4	2	3	
4	4	2	4	3	4	4	3	3	3	
4	3	5	4	4	4	4	4	3	3	
3	4	4	4	4	3	3	4	3	3	
Moy.	3,73	3,33	3,20	3,33	3,07	3,13	3,20	3,67	3,20	2,80
Ec. type	1,16	1,05	1,15	1,18	1,16	0,92	1,15	0,98	0,77	1,08

Moy. générale	3,27
Ecart type	0,27

Williams/Le Dû										
Equilibre av/ar	Sources côtés	Présence HP	Enveloppement	Homogénéité	Naturel	Précis. localis.	Stabilité localis.	Timbre	Définition	
2	3	3	3	2	3	2	2	4	3	
1	5	4	4	2	1	2	2	2	2	
3	2	2	2	2	3	2	3	3	3	
3	4	3	4	3	3	4	4	4	4	
5	5	4	4	4	3	4	4	5	5	
2	4	1	1	1	2	2	2	4	1	
1	3	5	1	1	2	2	3	3	2	
5	4	5	4	4	4	4	5	4	4	
2	2	2	2	3	2	3	3	4	2	
1	3	4	4	2	2	4	4	4	4	
2	3	2	2	2	2	2	3	2	2	
2	4	4	2	2	2	2	2	2	1	
1	2	3	4	2	2	5	4	3	4	
2	3	4	3	2	3	4	4	3	3	
1	4	2	2	2	3	4	4	3	3	
Moy.	2,20	3,40	3,20	2,80	2,27	2,47	3,07	3,27	3,33	2,87
Ec. type	1,32	0,99	1,21	1,15	0,88	0,74	1,10	0,96	0,90	1,19

Moy. générale	2,89
Ecart type	0,4

Le Grenier à Son

	Equilibre av/ar	Sources côtés	Présence HP	Enveloppement	Homogénéité	Naturel	Précis. localis.	Stabilité localis.	Timbre	Définition
	2	2	2	1	1	1	2	4	1	1
	2	4	4	4	2	3	2	2	3	2
	4	2	2	2	2	2	1	1	2	2
	2	2	2	3	3	2	4	2	2	2
	5	4	4	5	4	4	2	4	3	3
	4	5	5	5	5	5	3	5	4	4
	3	3	5	5	4	2	4	4	3	1
	3	4	2	3	2	2	2	1	3	2
	2	2	3	3	3	2	4	3	2	2
	2	2	4	4	2	2	3	5	3	2
	2	4	2	2	2	3	2	3	1	1
	2	1	1	2	1	2	3	4	3	2
	2	4	2	2	2	3	4	4	2	5
	2	3	3	3	3	3	3	3	2	1
	4	3	2	4	4	2	4	4	3	4
Moy.	2,73	3,00	2,87	3,20	2,67	2,53	2,87	3,27	2,47	2,27
Ec. type	1,03	1,13	1,25	1,26	1,18	0,99	0,99	1,28	0,83	1,22

Moy. générale	2,79
Ecart type	0,32

## Bibliographie

- [1] WILLIAMS, M. Unified Theory of Microphone Systems for Stereophonic Sound Recording. 82<sup>e</sup> Convention de l'AES, 1987, Preprint 2466
- [2] HERRMANN, U., HENKELS, V., BRAUN, D. Vergleich von 5 Surround-Mikrofonverfahren. 20<sup>e</sup> Tonmeistertagung, 1998, p. 508-517
- [3] WILLIAMS, M., LE DU, G. Multichannel Microphone Array Design. 108<sup>e</sup> Convention de l'AES, 2000, Preprint 5157
- [4] WILLIAMS, M., LE DU, G. Microphone Array Analysis for Multichannel Sound Recording. 107<sup>e</sup> Convention de l'AES, 1999, Preprint 4997
- [5] LE DU, G. (1999) Conception de systèmes de prise de son multicanaux - Application à l'enregistrement musical (Mémoire - ENSLL)
- [6] JACQUES, N., MORA, A. (1998) Prise de son multicanale en différence de temps (Mémoire - FSMS, CNSMDP)
- [7] JOT, J-M., WARUSFEL, O. [www.ircam.fr/equipes/salles](http://www.ircam.fr/equipes/salles)
- [8] BRUCK, J. The KFM 360 Surround - A Purist Approach. 103<sup>e</sup> Convention de l'AES, 1997, Preprint 4637
- [9] GERZON, M.A. Optimal Reproduction Matrices for Multispeaker Stereo. 91<sup>e</sup> Convention de l'AES, 1991, Preprint 3180
- [10] JAGGER, D.S. Recent Developments and Improvements in Soundfield Microphone Technology. 75<sup>e</sup> Convention de l'AES, 1984, Preprint 2064
- [11] EVANS, M.J. Obtaining Accurate Responses in Directional Listening Tests. 104<sup>e</sup> Convention de l'AES, 1998, Preprint 4730
- [12] GIVERNAUD, L. (2000) Etude de la perception de la localisation pour une prise de son stéréophonique en différence d'intensité en 5.1 (Mémoire - CNAM)
- [13] BLAUERT, J. Spatial Hearing - The Psychophysics of Human Sound Localization. The MIT Press, Cambridge, Mass, 1996
- [14] HUGONNET, C. (1985) Restitution spatiale de différents systèmes stéréophoniques - caractérisation objective et subjective (Mémoire - CNAM)
- [15] WUTTKE, J. Considérations générales sur la prise de son multicanale. Conférence de la SFA, Paris, 2000
- [16] THEILE, G. Multichannel Natural Recording based on Psychoacoustic Principles. 108<sup>e</sup> Convention de l'AES, 2000, Preprint 5156
- [17] PLENGE, G., THEILE, G. Localization of Lateral Phantom Sources. Journal of AES, avril 1977, vol.25-4
- [18] WALKER, R., A Controlled-Réflexion Listening Room for Multi-Channel Sound. 104<sup>e</sup> Convention de l'AES, 1998, Preprint 4645