

CONSERVATOIRE NATIONAL DES ARTS ET METIERS

PARIS

MEMOIRE

présenté en vue d'obtenir

le DIPLÔME D'INGENIEUR C.N.A.M.

en

ACOUSTIQUE

Par

Laurent GIVERNAUD

Etude de la perception de la localisation pour
une prise de son acoustique
en différence d'intensité en 5.1.

Soutenu le 12 décembre 2000

JURY :

Président : Jacques JOUHANEAU

Membres : Alexandre GARCIA
Christian HUGONNET
Jean-Marc LYZWA
Manuel MELON
Sébastien NOLLY
Edwige RONCIERE

Etude de la perception de la localisation pour une prise de son acoustique en différence d'intensité en 5.1.

Mémoire d'ingénieur C.N.A.M., Paris 2000

Cette étude concerne la prise de son acoustique en multicanal dont le secteur est réactualisé par les perspectives de nouveaux supports d'enregistrement et de communication. La perception de la localisation de ces systèmes est le paramètre d'investigation.

La première section est un tour d'horizon des techniques existantes en 5.1. et montre l'intérêt d'une étude consacrée à la prise de son en différence d'intensité.

Les résultats des tests subjectifs serviront à l'élaboration d'un système de captation et d'un enregistrement avec ce système. Les résultats obtenus sont comparés à ceux existant dans les publications.

Les données recueillies par le test conduisent à l'élaboration d'un microphone à directivité du second ordre. Celui-ci se rapproche des contraintes de la prise de son professionnelle.

Une prise de son est réalisée avec des microphones prototypes. Les résultats sont discutés pour mettre en avant les qualités et défauts de ces capteurs dans leur environnement.

Les résultats des écoutes finales montrent l'intérêt de la prise de son multicanale en différence d'intensité. Des suggestions sont faites pour en améliorer les performances.

Mots clés :

Séréophonie – Microphone – Multicanal – 5.1. – Directivité – Second ordre – Différence d'intensité – Perception - Localisation – Audition perceptive – Enregistrement acoustique.

Keywords :

Stereophony – Microphone – Multichannel – 5.1. – Directivity – Second-order – Intensity differences – Perception – Localization – Perceptual Hearing - Acoustic recording.

CONSERVATOIRE NATIONAL DES ARTS ET METIERS

PARIS

MEMOIRE

présenté en vue d'obtenir

le DIPLÔME D'INGENIEUR C.N.A.M.

en

ACOUSTIQUE

Par

Laurent GIVERNAUD

Etude de la perception de la localisation pour
une prise de son acoustique
en différence d'intensité en 5.1.

Les travaux relatifs au présent mémoire ont été effectués au sein du Conservatoire National Supérieur de Musique de Paris dans la section audiovisuelle sous la responsabilité de Monsieur Jean-Marc Lyzwa.

Remerciements

Je tiens à remercier particulièrement, pour leur implication et l'aide qu'elles m'ont apportée, les personnes suivantes :

Catherine de Boishéraud et Daniel Zalay pour avoir permis ce stage dans les locaux du CNSM de Paris.

Jean-Marc Lyzwa pour ses conseils avisés, sa gentillesse et son implication dans le projet.

Toute l'équipe chaleureuse de l'audiovisuelle du CNSM ainsi que les nombreux étudiants qui se sont prêtés aux tests d'écoute.

Radio France pour le congé individuel de formation de 2 mois.

Les collègues de Radio France pour les discussions passionnées qui ont fait naître l'idée du sujet de ce mémoire.

Ma femme Sandrine et mes trois enfants pour le soutien et la patience dont ils ont dû faire preuve.

Je voudrais également remercier Jacques Jouhaneau en tant que directeur de ce mémoire et pour la qualité de l'enseignement qu'il m'a dispensé tout au long du cursus d'acoustique.

Nous n'avons pas besoin d'écouter pour entendre...

Qui a conçu l'oreille ?

La faim du tigre René Barjavel

Sommaire

1	Choix du sujet du mémoire	1
1.1	<i>Introduction</i>	1
1.2	<i>Terminologie</i>	3
1.3	<i>Le choix de l'étude</i>	4
1.4	<i>Plan de recherche</i>	7
2	Les microphones pour le multicanal en 5.1.	9
2.1	<i>Systèmes basés sur le Decca-tree</i>	9
2.2	<i>La tétraphonie</i>	10
2.3	<i>Les combinaisons de microphones relevant de la biphonie</i>	10
2.4	<i>Prise de son avec écran</i>	11
2.5	<i>L'ambisonie</i>	12
2.6	<i>L'holophonie</i>	13
3	Tests de localisation	15
3.1	<i>Choix du test</i>	15
3.2	<i>Réalisation pratique du test</i>	17
3.3	<i>Déroulement du test</i>	18
3.4	<i>Résultats des tests</i>	19
3.4.1	Résultats d'ordre général	19
3.4.2	Résultats des tests avec le spectre tronqué	20
3.4.3	Résultats complets	20
3.4.4	Comparaison avec d'autres résultats	31
3.4.5	Résultats concernant la robustesse	34
3.5	<i>Conclusion du test</i>	34
4	Les microphones directifs	37
4.1	<i>Les technologies non utilisables</i>	37
4.1.1	Les microphones canons	37
4.1.2	Les microphones à constantes réparties	38
4.1.3	Les microphones utilisant une parabole	38
4.1.4	Une évolution du microphone ambisonic	39
4.2	<i>Les technologies possibles</i>	39
4.2.1	Les microphones à constantes réparties modifiées	39
4.2.2	Sommation de microphones	41
4.2.3	Les réseaux de microphones	41
4.2.4	Traitement de la pression à la puissance deux	42

4.2.5	Les microphones du second ordre	42
4.3	<i>La technologie retenue</i>	46
5	Mise au point pratique de la technologie retenue	49
5.1	<i>Sélection du microphone cardioïde</i>	49
5.2	<i>Influence de la pondération</i>	52
5.3	<i>Sommation théorique de deux systèmes</i>	54
5.4	<i>Sommation réelle des deux systèmes</i>	57
5.5	<i>Influence de la position désaxée du montage</i>	62
5.6	<i>Diffraction générée par une prise de son coïncidente</i>	64
5.7	<i>Synthèse de la technologie retenue</i>	65
6	Test du prototype	67
6.1	<i>Choix du test</i>	68
6.2	<i>Enregistrement du test</i>	69
6.3	<i>Déroulement du test</i>	71
6.4	<i>Résultats du test</i>	72
6.5	<i>Conclusion du test</i>	81
7	Conclusion	83
8	Références	85
9	Table des illustrations	89
10	Annexes	91
10.1	<i>Annexe 1. Texte de présentation du test</i>	91
10.2	<i>Annexe 2. Fichier Matlab de visualisation en diagrammes polaires</i>	91
10.3	<i>Annexe 3. Fichier Matlab des simulations de réponse en fréquence du prototype</i>	92
10.4	<i>Annexe 4. Fichier Matlab de visualisation en diagrammes polaires réels et calcul du DI</i>	93
10.5	<i>Annexe 5. Fichier Matlab de Visualisation des courbes exportées ou simulées du prototype</i>	94
10.6	<i>Annexe 6. Fichier Matlab de visualisation 3D de sommation de microphone</i>	96
10.7	<i>Annexe 7. Egalisation du prototype avec la console Yamaha</i>	97

1 Choix du sujet du mémoire

1.1 Introduction

La prise de son affectée à un système de reproduction multicanale est un sujet d'actualité. Chaque année, de nouveaux systèmes de captation sont proposés pour la prise de son multicanal. L'intérêt porté à ce sujet a été réactualisé grâce aux possibilités nouvelles et attendues offertes par les moyens de transmission et de stockage multipistes. Le fait de percevoir l'information de l'incidence sonore autorise entre autres, la création de fictions radiophoniques ou théâtrales totalement nouvelles, ainsi qu'une écoute de la Musique avec des dimensions perceptives supplémentaires. Certains de ces systèmes de prise de son nécessitent un traitement ou des pondérations matricielles, d'autres systèmes se suffisent à eux-mêmes et peuvent être directement affectés aux enceintes.

Le sujet de ce mémoire concerne la prise de son en différence d'intensité (ΔI) pour une localisation réaliste dans une configuration multicanale. Il y a très peu de résultats ou travaux concernant la prise de son coïncidente en multicanal hormis les études sur le système ambisonique qui reste un cas particulier.

Les systèmes de reproduction sont variés mais une norme internationale (ITU, UER...)[17] recommande une configuration physique, spécifiant notamment le nombre et le pointage des enceintes de diffusion pour une standardisation de la création sonore en multicanal. Cette recommandation est discutable mais entraîne une évolution dans le sens où elle évite l'anarchie assurée des moyens de reproduction et de diffusion.

La prise de son actuelle tend à se développer vers la norme ITU [17], il sera donc question ici d'une configuration classique en 5.1. malgré la tentation compréhensible de remédier à certaines contraintes :

Cette norme spécifie principalement le positionnement des enceintes et c'est cette précision qui pose le plus de problèmes. La configuration comprend une enceinte au centre ainsi que deux paires placées symétriquement aux angles de 30° et de 110° à $\pm 10^\circ$ (cf. figure 1). Le cinéma et la vidéo sont paradoxalement plus en avance que l'industrie de production audio en multicanal. On pourra y déceler des intérêts commerciaux et financiers mais, devant la démocratisation du «home theater» les milieux de l'audio n'ont pu faire autrement que de s'aligner sur les principes du son associé à l'image. En effet, l'enceinte centrale est principalement employée pour les voix «off» au cinéma afin de recentrer les dialogues monophoniques en fonction des différentes positions de l'auditeur dans la salle et afin d'offrir une perception robuste de la biphonie. L'enceinte centrale est une particularité de la diffusion avec image et n'a jamais été utilisée précédemment en audio. Le choix de l'angle de diffusion frontale de 60° est maintenant plus un souci de compatibilité biphonique que de respect du cône d'attention visuel. On nuancera tout de même cette affirmation sujette à discussion en face de l'interaction de la vue et de l'ouïe

et des choix artistiques sur le rendu sonore. Les preneurs de son sont divisés quant aux possibilités offertes de créer un espace auditif plus grand que le champ visuel. Malgré l'importance de ce débat, nous n'entrerons pas plus dans le détail.

Les enceintes surround, quant à elles sont employées uniquement pour envelopper le spectateur par le son et fournir quelques effets sonores spectaculaires sans souci de la provenance précise du bruitage. La position angulaire recommandée n'a semble-t-il pas d'autre raison que de couvrir le mieux possible la zone latérale et arrière.

On peut donc dire qu'un grand nombre de configurations différentes auraient été plus souhaitables (enceintes équidistantes, enceintes plus nombreuses, enceintes comblant les zones latérales...), mais en même temps on s'aperçoit que la norme ITU autorise à tous la reproduction à domicile des œuvres multicanales audio à partir d'un codage unique. On peut penser aussi que l'implantation existante du home theater facilitera l'arrivée du DVD audio et du SACD en raison de leur configuration de reproduction identique.

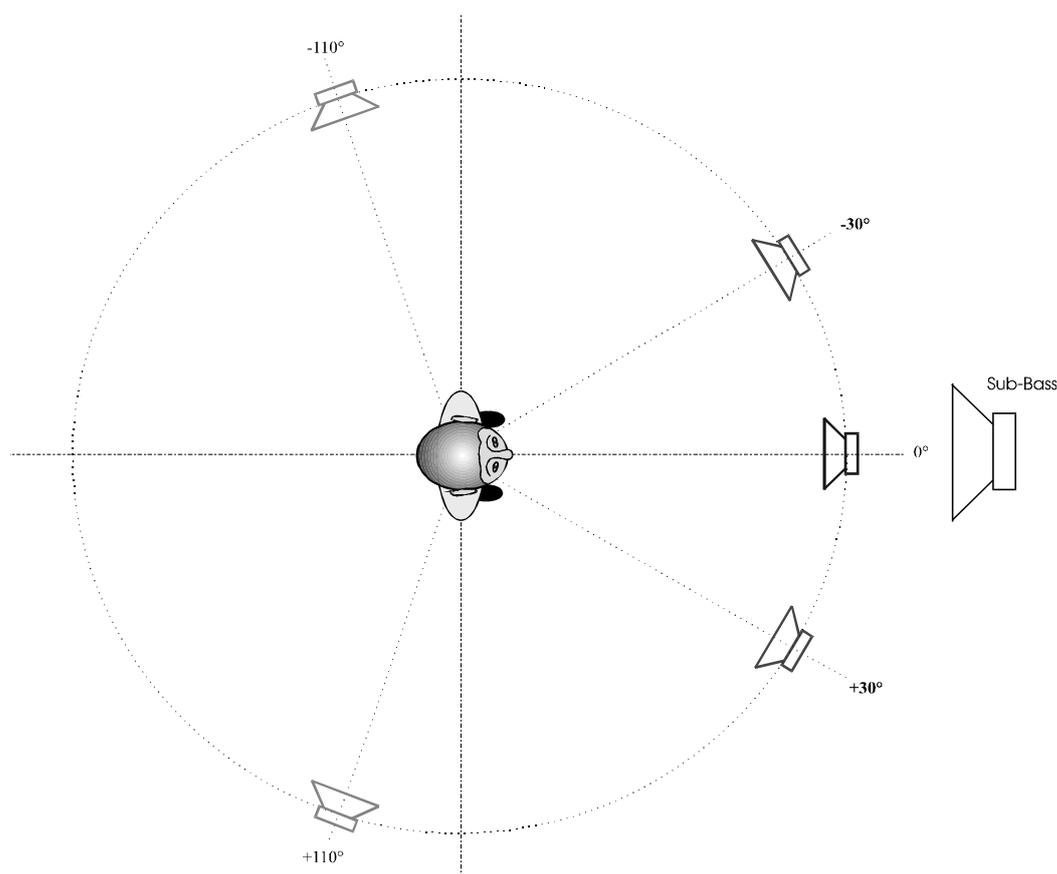


figure 1. Recommandation ITU

La configuration retenue pour ce mémoire sera donc celle du 5.1. ou plutôt la 2/2 : deux enceintes avants et deux arrières (cf. figure 1). L'utilisation d'enceinte de qualité et les réglages difficiles nécessaires au canal très basse fréquence font que le caisson de basse ne sera pas inclus à cette étude.

1.2 Terminologie

Afin d'éviter des confusions sur le sens des mots, précisons le vocabulaire utilisé dans ce mémoire :

- La reproduction multicanale est un terme qui englobe toutes les possibilités d'agencement dont le nombre de canaux de reproduction est supérieur ou égal à deux. L'écoute est restituée par un casque ou par des enceintes [14].
- La stéréophonie décrit exactement la même chose dans le sens où elle décrit la possibilité de donner une sensation de relief acoustique. Le terme stéréophonie ne devrait donc pas se limiter à une reproduction sur deux canaux. Le terme biphonie sera employé pour les systèmes composés de deux haut-parleurs et on parlera de systèmes binauraux pour les écoutes au casque.
- Le terme 5.1 (5 canaux discrets et un caisson de basse) sera préféré à 3/2 (trois enceintes avant et deux arrière pour l'enveloppement) pour son utilisation plus courante. Le terme 5.1 renvoi à la normalisation ITU. Le terme 5.1 sera donc employé même lorsque la reproduction se fera sur 4 canaux et sans canal basse fréquence, la configuration restant physiquement conforme à recommandation ITU [17].
- Dans ce mémoire, le terme de robustesse désigne la surface de réception pour laquelle une caractéristique précise de l'écoute multicanale est correcte. Cette caractéristique peut être la sensation d'enveloppement ou de précision de la localisation ou encore de la discrimination angulaire.
- La localisation quant à elle qualifie la précision moyenne de la perception angulaire d'une source sonore pour un sujet placé au centre du dispositif de reproduction. La latéralisation est le mot qui convient le mieux pour la configuration 5.1 où la perception en élévation n'est pas recherchée. Les termes de localisation et de latéralisation seront tous deux utilisés.
- Le centre de la zone frontale est le repère axial angulaire correspondant à 0° et la tête de l'auditeur est le centre du cercle. Un axe de symétrie ou médiane sépare la zone droite de 0° à 180° de la zone gauche de 180° à 360° .

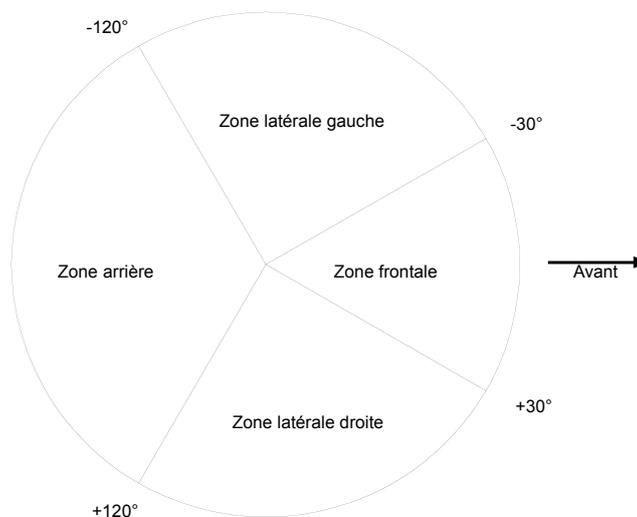


figure 2. Zones des tests délimitées par l'angle physique des enceintes

La figure 2 nous rappelle visuellement les zones angulaires (quadrants) formées par les couples d'enceintes, elle nous montre l'étroitesse de la couverture frontale par rapport aux trois autres zones.

1.3 Le choix de l'étude

La configuration angulaire des enceintes, la physique du rayonnement acoustique, la grande surface de reproduction et le nombre d'enceintes font qu'avec le 5.1, aucun système de prise de son multicanal ne sera jamais parfait et cela pour plusieurs raisons :

- La directivité des enceintes influe sur la perception du champ réverbéré et sur la réponse en fréquence lorsque l'auditeur n'est pas dans l'axe.
- Le nombre d'enceintes trop faible et la distance séparant les sources trop grande font que les ondes acoustiques produites par chaque enceinte restent distinctes. Les composantes de hautes fréquences ne peuvent pas se combiner avant d'arriver aux oreilles de l'auditeur.
- La distance des sources modifie la robustesse ainsi que le niveau de champ direct.
- L'acoustique de la salle peut modifier la réverbération, la localisation et le timbre d'une œuvre. Les réflexions des ondes sonores créées par le local peuvent être la source d'une polarisation de la salle contribuant à une localisation apparente différente de celle voulue. Ces réflexions altèrent la perception par une réverbération qui peut s'avérer excessive et donner la sensation d'un volume apparent différent de celui existant dans la salle d'enregistrement.

Si le système de prise de son est conçu dans l'optique d'une localisation réaliste, il doit se corréler à une référence physique. Il est impossible d'obtenir un rendu réaliste par une combinaison aléatoire. A partir du moment où l'on a fait le choix d'un principe, on doit s'y tenir et faire abstraction des autres possibilités. L'ajout de nouvelles théories nous conduit à des contraintes insolubles car le nombre de paramètres devient impossible à gérer. Les mariages douteux seront à utiliser à des fins artistiques ou esthétiques plutôt qu'en vue de rafistolages.

Une pléthore de théories et de procédés existe pour expliquer la perception de la localisation. Ces théories sont issues d'observations, de mesures et de tests mais aucune n'est fidèle car les configurations de la reproduction ne le permettront peut-être jamais et aucune ne peut décrire complètement la perception de la localisation auditive. Les différentes théories et techniques de la localisation comprennent entre autres les critères suivants :

- L'effet d'antériorité, phénomène bien connu des sonorisateurs qui ajoutent des retards temporels pour compenser la perception en localisation de l'enceinte la plus proche. Cet effet est la cause directe d'une faible robustesse en 5.1.
- L'effet de fusion pour deux sources décalées dans le temps $< 1\text{ms}$: c'est un des codages servant à la prise de son en biphonie pour créer une image fantôme (différence de temps ΔT).
- Les filtrages en peigne issus des réflexions acoustiques induites par la tête et les épaules ou encore par les fonctions de transfert de l'oreille (HRTFs). Ces filtres temporels mesurés ou simulés peuvent générer un codage multicanal. C'est une version améliorée des codages ΔI et ΔT ; par l'ajout de composantes fréquentielles, elle contient aussi et plus précisément ces différences pour la création d'images fantômes mais malheureusement différemment pour chaque individu. Cette technique fonctionne bien en écoute binaurale, l'écoute sur enceintes est soumise à un double codage, celui du système et celui généré par les oreilles de l'auditeur.
- Les effets cognitifs (interaction avec la vision, connaissance préalable des stimuli, l'effet « cocktail party » ...). Vaste domaine de la psychologie où de nombreuses études tentent de dévoiler les mécanismes de l'audition.
- La synthèse de front d'onde (WFS) se base sur la physique fondamentale. L'idée est de discrétiser la périphérie de la salle d'écoute par un nombre suffisant de haut-parleurs afin que les ondes sonores émises soient similaires aux sources réelles. Cette solution est tributaire du

repliement de spectre spatial [28] (l'espacement entre les sources est dépendant de la fréquence maximale que l'on désire restituer et provient de la longueur d'onde associée à cette fréquence). La captation acoustique en W.F.S reste un problème complexe car elle nécessite en théorie autant de microphones que d'enceintes. Cette technique est la seule à fournir une bonne robustesse mais elle n'utilise pas 5 enceintes...

- La métathéorie de la localisation auditive [11][24] est composée de trois approches, le degré de non-linéarité du système, l'ordre de directivité et la prise en compte des mouvements de la tête. C'est cette approche qui est à l'origine de l'ambisonie. Elle se décrit par un formalisme trop mathématique dont les bases perceptives sont vagues, mal expliquées et anciennes.
- Les différences d'intensité sont aussi un moyen de codage de la biphonie. Ce codage s'explique par l'atténuation de l'onde sur l'oreille opposée à la source (controlatéral) due à l'ombre acoustique de la tête.

La mise en application de certaines de ces théories entraîne des traitements avant restitution et la prise de son est souvent une sorte de monophonie dirigée. En règle générale, on peut dire qu'une captation acoustique est synonyme d'une simplicité de mise en œuvre. Si la prise de son acoustique en ΔT en 5.1 a été largement étudiée, il s'avère qu'il n'en est pas de même en ΔI .

Pour la théorie qui nous intéresse, celle du ΔI , nous nous dirigerons vers le principe d'une sorte de recombinaison du front d'onde associée à des lois de panning entre paires d'enceintes pour la génération de sources fantômes. Cette théorie entraîne la reconstruction du champ sonore à l'aide du même nombre d'enceintes que de microphones. Cette technique se retrouve dans les publications sous le terme anglais de « pair wise mixing » et fait appel aux ΔI . Cette simplicité apparente cache de nombreux problèmes techniques :

- Il est très difficile de réaliser des microphones coïncidents à n capsules en raison de l'encombrement dans l'espace imparti. De plus, on génère de la diffraction par la présence des corps des autres microphones.
- Chaque type d'enceinte rayonne différemment, il est très probable que les caractéristiques technologiques de certaines enceintes soient plus appropriées que d'autres pour une écoute multicanale.
- Les enceintes ne sont pas assez nombreuses et le son ne se combine pas du tout lorsque les enceintes sont trop écartées (spatial aliasing [28]). La recombinaison du front d'onde si elle est ressentie ne peut donc pas être du domaine de la physique mais plutôt de la psychologie.
- Une captation acoustique, contrairement au pan-pot enregistre le son direct ainsi que la signature de la salle. Comment le champ précoce et diffus est-il reproduit avec 5 enceintes ?
- Les lois en différence d'intensité de l'audition varient en fonction de la fréquence et des individus. Comparable à une sorte de fonction de transfert relative à la tête, la perception des images fantômes dépend de nombreux critères propres à chacun.
- Le phénomène d'image fantôme est encore mal connu. Perception auditive, l'image fantôme n'est qu'un leurre que l'on soumet au cerveau. La biphonie utilise ce phénomène sans rapport avec les mécanismes psychoacoustiques et cognitifs de la perception auditive humaine. L'interprétation est complexe car du point de vue physique par exemple, les ondes acoustiques parvenant à l'oreille sont tributaires de l'angle azimutal des enceintes et des fonctions de transfert individuelles de l'oreille [38]. La sensation de la localisation auditive en ΔI est la conséquence de la perception d'une fusion de deux ou de plusieurs sons dont on ne connaît pas le principe.

Le caractère individuel de la perception de la localisation nous conduit à dire que les résultats ne sont que des tendances. Quant à la reconstruction de l'image fantôme, il semble que ce sont de multiples facteurs qui provoquent un choix d'interprétation par le cerveau. On peut physiquement expliquer la raison pour laquelle une image fantôme est créée au centre lorsque deux stimuli de même nature sont affectés aux enceintes avant. Dans ce cas, l'oreille droite perçoit le même message que l'oreille gauche et de cette similitude, le système auditif déduit une perception de localisation centrée, mais c'est une des seules perceptions angulaires que l'on puisse expliquer simplement. D'ailleurs, une perception au centre ne veut pas dire devant, problème constaté par tous avec l'écoute au casque. Dès qu'une image n'est plus perçue au centre, un mélange pondéré très complexe de facteurs physiques, biologiques et cognitifs est à mettre en jeu.

Les avantages d'une prise de son coïncidente, dans l'hypothèse où le dispositif de captation enregistre en un point unique et où l'auditeur est placé à équidistance des enceintes pour la diffusion sont les suivants :

- Si les conditions ci-dessus sont respectées, aucun déphasage entre les enceintes ne risque d'introduire de la réverbération artificielle ou des décalages dans la perception de la localisation. Les systèmes de prise de son utilisant le delta T sont faussés par les interactions non désirées des paires de microphones (par exemple le microphone avant gauche avec le microphone arrière droit) et ceci malgré l'utilisation de microphones directifs. Un second codage vient s'ajouter à celui désiré entraînant généralement une réverbération artificielle. Pour certains systèmes, comme le Decca-tree, on génère plusieurs images fantômes pour la captation avant (voir chapitre 2.1), c'est la localisation qui devient faussée.
- Contrairement à la prise de son biphonique coïncidente, la zone de directivité non utile pour le codage du quadrant avant est utilisée en 5.1 pour le codage des quadrants latéraux. En biphonie le son enregistré parvenant d'un azimut autre que le quadrant avant de 60° , est reproduit aux angles correspondants à l'angle physique des enceintes acoustiques (figure 3).
- Affecter un microphone à une enceinte ne nécessite pas de matriçage ni de calcul DSP en temps réel.
- Le fait de faire coïncider les angles de la prise de son à ceux de la reproduction autorise une captation homogène dans la mesure où la somme de la directivité des capsules s'approche d'une directivité omnidirectionnelle. On peut toutefois s'interroger sur le rapport de capture qui sera modifiée selon les zones en fonction des directivités différentes des microphones utilisés.
- Si l'on associe la captation en delta I à une reproduction issue d'une synthèse de front d'onde, on peut supposer que tous les indices de repérage de la localisation sont respectés (Approche théorique du WFS ou de l'holophonie par ambisonie sur un nombre élevé de haut-parleurs [29]) à la condition d'être situé dans la zone d'écoute centrale.
- Un décalage de la tête de 20 cm du point d'écoute central idéal provoque dès la fréquence de 1kHz en ΔT , une confusion de la localisation [19]. En ΔI , l'atténuation due à la décroissance du son en fonction de la distance est pour ces mêmes 20cm négligeable. Le codage par intensité est donc par nature plus robuste. Il n'empêche que ce décalage pour le système en ΔI crée aussi malheureusement des différences de temps codant elle aussi la localisation. On est alors dans une situation d'écoute de codage mixte ΔI et ΔT ou ΔI et effet d'antériorité.
- En ΔI la captation des premières réflexions sera reproduite suivant son angle d'incidence réel (l'incidence du champ précoce suit les mêmes lois de pan-pot que le son direct). En ΔT les

capsules omnidirectionnelles enregistrent le champ réverbéré par tous les azimuts pour chaque microphone (les premières réflexions ne peuvent être fidèles).

- Les décalages de phases induits par la prise de son en ΔT n'autorisent pas la compatibilité de réduction en biphonie et monophonie.

Ce bref aperçu montre que contrairement à la biphonie, un système coïncident est sûrement plus intéressant que le ΔT pour une prise de son multicanale. On lui reprochera tout de même l'utilisation de microphones directifs dont la réponse en basse fréquence est médiocre.

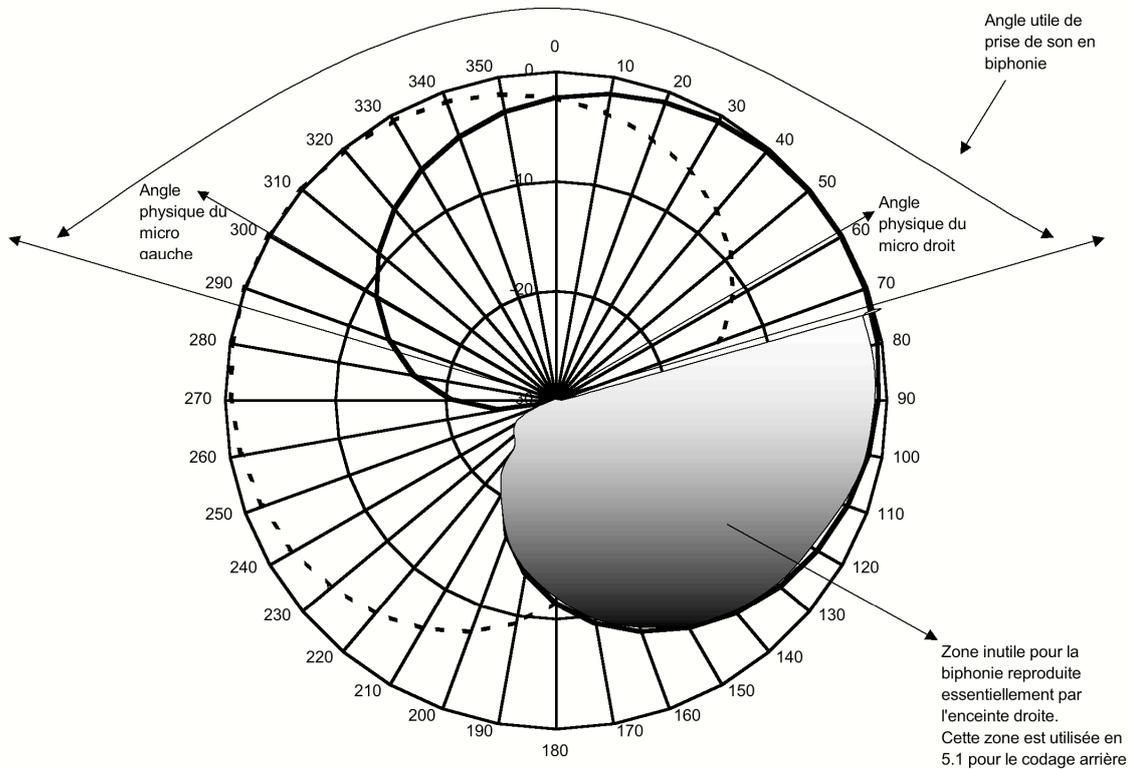


figure 3. La prise de son en ΔI pour la biphonie

1.4 Plan de recherche

La section 2 de ce mémoire est une présentation technologique non exhaustive des systèmes de prise de son existants.

La section 3 traite du test et des résultats subjectifs de la localisation en ΔI par pan-pot.

Des directivités de microphone souhaitées sont tirées de ces résultats et la section 4 relate les technologies des microphones directifs pour une utilisation envisageable en 5.1.

A partir de cet éventail, on retiendra une technologie et la section 5 est la mise au point d'un prototype proposant des caractéristiques s'approchant des directivités requises.

Finalement, en section 6, on testera ce microphone pour vérifier son respect de la perception de la localisation.

2 Les microphones pour le multicanal en 5.1.

Un bref survol des systèmes de prise de son en 5.1 s'impose car les expériences sont riches en informations et certains résultats donnent des orientations de recherche à suivre ou à ne pas suivre sur l'aptitude du système d'enregistrement à respecter la localisation des sources.

2.1 Systèmes basés sur le Decca-tree

Pour la zone frontale de captation, ce système est composé d'un triplet microphonique espacé entre eux de 1.5m à 2m. Les microphones sont omnidirectionnels et en raison de leurs écartements, ils ne sont pas adaptés à une bonne localisation : l'utilisation de trois capsules pour la reproduction frontale engendre trois images fantômes qui tendent à rendre plus confuse la précision [39](figure 4). Deux microphones placés à l'arrière servent à nourrir les enceintes surround afin d'assurer une compatibilité en 5.1. La distance séparant les microphones surround des principaux, permet d'ajuster l'image subjective de la prise de son. Ce système donne de bons résultats quant aux critères de spatialisation et du renforcement de l'image centrale.

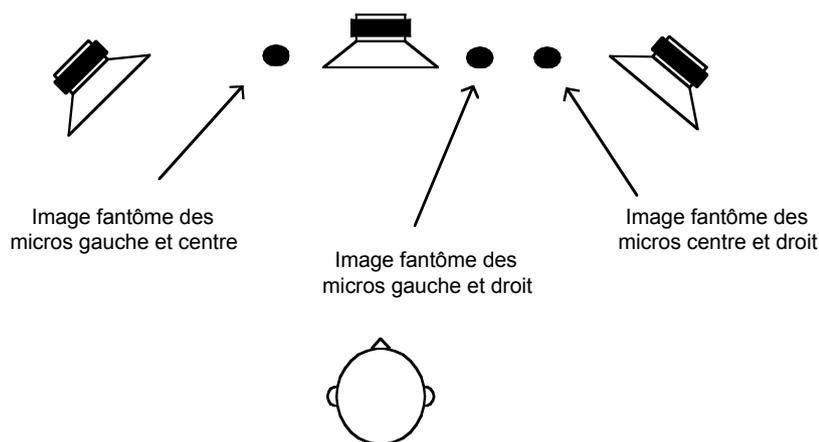


figure 4. Formation de 3 images fantômes en ΔT

Le Fukada-tree est une variante de ce système qui reprend l'agencement des microphones en remplaçant les capsules omnidirectionnelles par des cardioïdes. La localisation est ainsi améliorée au détriment de la réponse aux basses fréquences inhérentes aux capsules directives.

Dans le mémoire d'Arnaud Mora et de Nicolas Jacques [27], le sujet est une prise de son en différence de temps pure basée sur un Decca-tree. Devant la constatation de résultats de localisation médiocres, ils améliorent progressivement leur modèle en ajoutant du codage en différence d'intensité. Ils utiliseront des microphones omnidirectionnels plus directifs (grandes capsules), qu'ils pointeront ensuite vers la source et, ils sous mixeront les voies surround afin de

minimiser les inversions entre l'avant et l'arrière. Les écoutes finales leurs semblent satisfaisantes.

2.2 La tétraphonie

En tétraphonie l'enregistrement avec deux couples ORTF donne des résultats intéressants car il génère un codage frontal et arrière de la localisation. Cette captation est assez souple car la distance entre les deux couples a peu d'influence. L'interaction entre les deux couples contribue à recréer une spatialisation agréable. C'est une méthode de captation qui ne prend nullement en compte la localisation latérale mais qui devant la difficulté de perception de la direction des images fantômes dans cette zone [38], est certainement le système d'enregistrement multicanal donnant les meilleurs résultats en regard de sa facilité de mise en œuvre. C'est dans les situations où deux masses sonores seront réparties l'une à l'avant, l'autre à l'arrière que ce système donnera ces meilleures capacités (enregistrement d'un orchestre et d'un orgue dans une église par exemple). La restitution peut se conformer à l'agencement du 5.1 en éliminant l'enceinte centrale.

Une variante composée de deux couples MS [16], construit à base de la même capsule bidirective et de deux capsules directives dirigées de part et d'autre de l'axe de symétrie est une des solutions de prise de son coïncidente. Par rapport au couple ORTF, on regrettera l'impossibilité d'adapter le rapport de capture entre le champ réverbéré et le champ direct.

Les études liées à l'hexaphonie et la tétraphonie avec les enceintes placées en tétraèdre sont des études qui sortent du contexte 5.1.

2.3 Les combinaisons de microphones relevant de la biphonie

Le travail de Guillaume Le Du [21] est intéressant, il utilise les résultats de la biphonie pour obtenir une bonne localisation avant. Il se base sur les résultats mixtes de la biphonie [45][46] (prise de son en ΔT et ΔI), pour obtenir des angles de captation compatibles avec la diffusion. Le respect théorique des angles de pointage et des distances entre les capsules n'étant pas réalisable, il introduit un réglage supplémentaire. Cet ajustement se matérialisant par un décalage en distance qui introduit un offset temporel. Cela provoque une modification des références axiales. Ainsi le repère de l'axe central en ΔT est différent de celui du ΔI . Le nombre de combinaisons de codage possible augmente. Les valeurs à affecter aux différences d'intensité et de temps sont tributaires d'un nouveau paramètre et autorisent par ce biais à se rapprocher d'une configuration théorique idéale.

Les enceintes arrières servent à restituer l'enveloppement. Il constate que les capsules omnidirectives en 5.1 génèrent trop de diaphonie et ne permettent pas la perception de la latéralisation en multicanal. Il observe aussi que la distorsion angulaire entre 0° et 30° est réduite par l'utilisation de l'enceinte centrale.

Avec le système ASM5 commercialisé, Brauner propose un produit tout en un, composé de 5 microphones appairés. Chaque microphone possède une directivité ajustable du 1^{er} ordre allant de l'omnidirectionnelle à la bidirectionnelle. Les microphones avants sont placés aux sommets

d'un triangle rectangle et espacés de 17.5 cm du centre d'enregistrement. Chaque microphone « surround » est positionné à 60 cm vers l'arrière avec un angle de 60° par rapport à l'axe médian (voir figure 5).

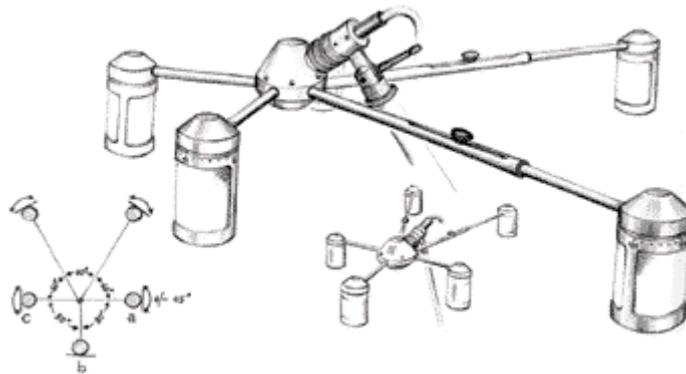


figure 5. La solution Brauner

Son intérêt principal est sa possibilité d'ajuster les directivités pour modifier la captation de l'ambiance de la salle où même de diriger une source lors de l'enregistrement.

Ce système offre ainsi à l'ingénieur du son, un outil pour l'expérimentation sonore et des possibilités artistiques étendues.

Ce microphone n'a pas pour vocation d'être fidèle au niveau de la perception de la localisation.

2.4 Prise de son avec écran

Dans le même esprit que Brauner, c'est à dire bénéficier d'une directivité ajustable au niveau de la console, Bruck a commercialisé avec Schoeps, une sphère de 18 cm de diamètre plus petite que celle utilisée pour les enregistrements biphoniques (20 cm) afin d'y loger de part et d'autre deux capsules bidirectionnelles. Ces deux capsules supplémentaires mixées avec les capsules omnidirectionnelles effleurant la sphère (principe de la prise de son M-S) permettent de créer 4 canaux audio individuels (voir figure 6).

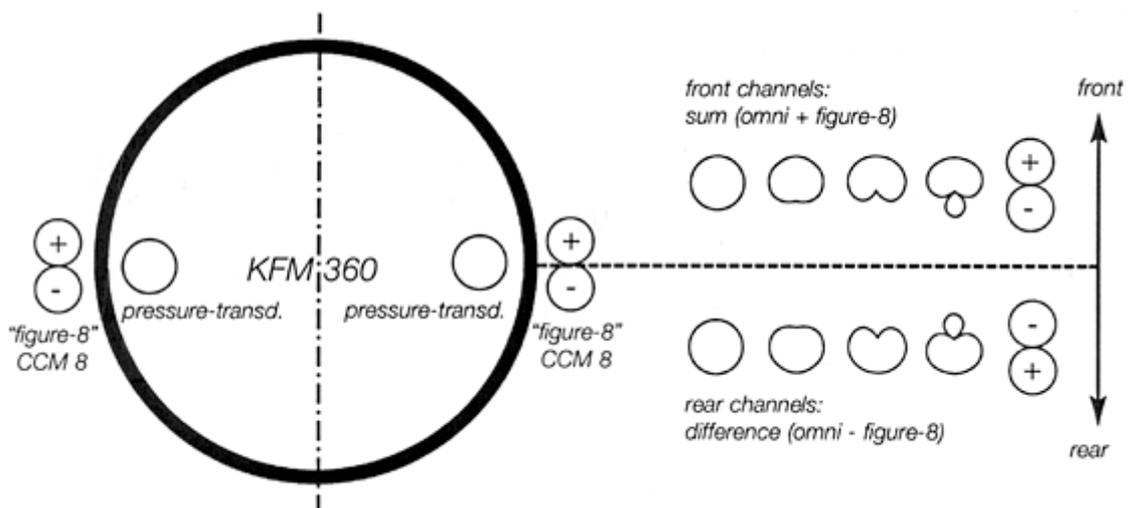


figure 6. La solution Schoeps

Le canal central est fabriqué à partir des signaux gauche et droite à l'aide d'un matriçage emprunté à l'ambisonie. Ce microphone autorise le preneur de son à modifier les directivités pour s'adapter aux contraintes de la prise de son ou de régler la qualité sonore suivant ses goûts. La vocation de ce procédé est avant tout, un microphone de captation d'ambiance.

« The Holophone » est un produit commercialisé. Ce système est construit à partir d'une tête elliptique de 19 cm par 15 sur laquelle viennent se greffer 7 microphones omnidirectionnels (figure 7).



figure 7. The Holophone

Ce système ne se base sur aucun principe ni aucune théorie, cela ressemble à une extension de la tête artificielle. Ce microphone est voué à une captation fortement génératrice de réverbération artificielle en raison de ces nombreux capteurs. Les courbes de directivité et les distances entre les capsules données par le constructeur ont des valeurs trop faibles pour une utilisation respectant la localisation.

2.5 L'ambisonie

C'est vers le début des années 70 qu'émerge le principe de la prise de son ambisonique. Issue de la métathéorie de la localisation auditive (cf. chapitre 1.3), elle propose un système de prise de son comportant un matriçage pour une diffusion sur un nombre quelconque d'enceintes à condition qu'elles soient placées à équidistance l'une de l'autre. La matrice de reproduction sera différente en fonction de la quantité d'enceintes et des indices de la métathéorie pris en compte.

La réalisation pratique du système de prise de son impose une restriction et une simplification très importante de la théorie (utilisation de la directivité du 1^{er} ordre uniquement par exemple, la captation d'ordre 2 nécessite des microphones de directivité du second ordre non disponibles). La prise en compte de chaque nouveau paramètre conduit à des matrices non solvables et des simplifications sont nécessaires pour une application pratique.

Ces systèmes, pourtant précurseurs, tardent à s'imposer pour ces raisons ainsi qu'à cause du flou entretenu quant à la théorie et à la complexité de mise en œuvre. Des matrices offrant la compatibilité 5.1 sont d'ores et déjà disponibles pour les futurs DVD Audios. L'effervescence

engendrée par ce procédé n'a pas permis d'obtenir à ce jour, de résultats subjectifs concernant la localisation.



figure 8. Capsules en montages tétraédrique

La figure 8, montre l'assemblage coïncident du microphone Soundfield, principal capteur acoustique de la prise de son ambiphonique. La distance entre les capsules autorise la restitution des fréquences jusqu'à 12 kHz sans aliasing spatial. L'ambisonie est tributaire d'un codeur et d'un décodeur.

2.6 L'holophonie

L'holophonie est un des moyens de reproduction nécessitant autant de microphones que d'enceintes. Pour une recombinaison parfaite du front d'onde cette technologie demande une quantité de l'ordre de 250 haut-parleurs. Une captation acoustique qui sans le renfort d'une simplification drastique est actuellement irréalisable. De nombreuses études en cours vont dans ce sens. Notons que c'est actuellement la seule et unique méthode pour faire une reproduction tridimensionnelle sonore réelle car elle se base sur l'équation fondamentale de la propagation des ondes en acoustique.

La modélisation des phénomènes cognitifs et physiologique est certainement le seul moyen de rendre ce système viable en réduisant la quantité de matériel nécessaire.

3 Tests de localisation

Une prise de son coïncidente utilise les caractéristiques directives des microphones. La prise de son en biphonie se base sur des résultats psychoacoustiques. Les différences d'intensité entre deux canaux sont représentatives d'un angle de perception. A ce sujet, beaucoup d'études ont été menées, et l'on a accès facilement aux nombreux résultats les concernant dans des livres ou des publications [16][45].

Il existe aussi beaucoup de résultats concernant les courbes de fonctions de transfert de la tête suivant l'incidence du son et la fréquence [4]. Malheureusement ces résultats sont inexploitable pour le cas qui nous intéresse, c'est à dire les lois de panning à affecter à chaque paire d'enceintes dans une configuration 5.1. Les lois de panning mentionnées par les auteurs font plutôt référence à des lois mathématiques, afin de les implanter directement sur les consoles de production [43].

Il est donc nécessaire de connaître le plus précisément possible les différences d'intensité à affecter par paire d'enceinte pour une localisation précise. Pour cela des tests perceptifs seront mis en place. Notons que des résultats ont été publiés dernièrement par Woszczyk [51] après la décision de mettre le test en place. D'autres résultats issus d'une étude de Bruhnes [6] qui a aussi été effectuée au CNSM, apportent des informations intéressantes sur la localisation. Les valeurs de ces deux résultats seront comparées au chapitre 3.4.4.

3.1 *Choix du test*

Un test psychoacoustique nécessite des conditions optimales pour sa fiabilité. Ici, on cherchera à déterminer l'angle azimutal ressenti par l'auditeur en fonction d'un stimulus sonore distribué avec des niveaux différents pour chaque enceinte. L'objectif pourtant simple, pose de multiples problèmes de fond pour sa validation. On citera tout d'abord les différents problèmes inhérents à la réalisation du test puis on donnera les solutions pratiques choisies avec les conséquences qu'elles entraînent :

- Quelle matière sonore doit être utilisée pour correspondre au mieux à des conditions générales ?
- Quelles sont les enceintes les plus représentatives pour ce test ?
- A quelle distance doit-on placer les enceintes de l'auditeur ?
- Doit-on utiliser un local type pour effectuer cette expérience ?
- Quelles sont les directives à donner à l'auditeur pour que la localisation ne soit pas faussée ?
- Quelle durée doit-on prescrire à l'ensemble du test pour ne pas engendrer de fatigue auditive au sujet effectuant les écoutes ?
- Comment le sujet doit-il juger l'angle d'incidence de sa perception auditive ?
- Combien d'auditeurs sont nécessaires pour des résultats pertinents ?

Les réponses aux questions concernant le déroulement du test sont dictées par les résultats recherchés et par les conditions matérielles disponibles. On s'aidera des expériences relatives à des tests déjà effectués dans des conditions similaires. [6] [27] [51].

Le sujet du mémoire concerne une prise de son comparable à celle effectuée en biphonie, cela veut dire que les microphones auront une directivité constante sur tout le spectre audible. En conséquence, il ne sera pas nécessaire de faire des essais par bandes de fréquence. De même, la captation multicanale concerne la localisation du son en général. Pourquoi recréer ici artificiellement des salves de bruit ou autres stimuli représentant des caractéristiques particulières du son ? Le choix se finalisera donc sur un stimulus d'origine musicale qui est représentatif des sonorités écoutées généralement. On reprochera tout de même à cette méthode, le fait de privilégier certaine zone formantique propre à l'échantillon musical.

Le morceau choisi est un passage chanté a cappella par Suzanne Vega, échantillon sonore ayant été utilisé avec succès pour d'autres travaux. [27][6]. Ce morceau à l'avantage de proposer un spectre plus étendu qu'une voix parlée tout en privilégiant les zones spectrales de la voix nécessaires à la clarté du message. Il contient des attaques intéressantes pour la richesse des transitoires. Devant la répétition successive du stimulus sonore pendant la durée du test, la musique offre un côté plus séduisant que des salves de bruit.

La restitution se fait sur les enceintes disponibles (modèle 4412 de JBL) présentant des réponses en fréquence et en directivité de qualité correcte. Peu de travaux concernent le choix des enceintes en restitution multicanale. Le choix est donc dicté par un modèle ayant des caractéristiques qui ne s'écartent pas trop de la moyenne (nombre de voies, directivité, encombrement, appairage entre elles, qualité générale). Il me semble que des études approfondies concernant la reproduction par enceintes du son multicanal en 5.1 donneraient des informations précises quant au matériel à retenir. Néanmoins, on se doute qu'une large directivité des transducteurs risque de perturber l'écoute en excitant davantage le local de reproduction augmentant ainsi la quantité d'énergie diffuse. Un pré test identique utilisant des Genelec 1032A donne des résultats similaires avec les mêmes sujets.

La distance entre les enceintes et l'auditeur dépend d'un compromis entre les dimensions des enceintes, le temps de réverbération du local, le volume du local et la sensation pour l'auditeur de ne pas être gêné par la présence des enceintes. Les tests se sont déroulés à la salle Schaeffer du Conservatoire National Supérieur de Musique de Paris. C'est une salle un peu réverbérante avec un temps de réverbération de 0.5s car de grand volume (500m³). Elle présente indéniablement des modes propres car la géométrie du local est de forme légèrement parallépipédique. Les enceintes sont des trois voies avec des transducteurs qui ne sont pas placés sur le même axe de symétrie. Ce sont donc des enceintes de moyenne portée qui nécessitent d'être placées à distance raisonnable. L'angle de diffusion total est de l'ordre de 120° à 2 kHz. Les conditions d'écoute optimales ne sont donc pas réunies, néanmoins, malgré une sensation sonore très différente de celle effectuée au studio de production, les résultats de localisation convergent dans le même sens. Un local avec un temps de réverbération inférieur à 0.5s aurait été plus adapté mais s'apparente plus aux conditions d'écoute chez le particulier. Si la reproduction biphonique nécessite une certaine réverbération pour s'exprimer au plus près de la sensation voulue par le preneur de son, on doit réévaluer la question pour le multicanal si la volonté est de reproduire une écoute réaliste et localisable. Le rayon entre la tête de l'auditeur et les enceintes est de 3.2m ce qui est très proche de la distance critique dans cette salle. En augmentant le

nombre de haut-parleurs, on atténue la clarté du local ce qui fait qu'une écoute plus proche des transducteurs ou un local plus mat auraient été préférables.

Un texte de préambule que l'on trouvera retranscrit en annexe 10.1 permet de spécifier à l'auditeur les conditions du déroulement du test. Il lui conseille particulièrement de garder la tête vers l'avant durant la phase d'écoute. La perception de la localisation est totalement modifiée si l'on dirige sa tête vers la source fantôme. Il est recommandé aussi à l'auditeur de fixer une lumière clignotante matérialisant le point de référence frontale basée sur la médiane. Le repère visuel contribue à fixer une référence spatiale évitant ainsi des biais de perception déjà observés dans différents tests. La perception de la localisation auditive et visuelle est étroitement liée tant de manière cognitive que subjective. Des résultats expérimentaux prouvent que des tests effectués les yeux clos sont moins précis que si le sujet possède un repère visuel qui l'ancre dans l'espace [26].

La partie choisie pour le test dure 12s, elle est répétée deux fois consécutivement pour confirmer la première impression, et créer une sorte de préapprentissage. Une pause entre chaque test autorise une certaine relaxation et laisse un temps important pour juger de la localisation et inscrire sa réponse. Une pause trop longue pourrait déconcentrer le sujet dans la phase du test.

La bande audio comprend 40 tests avec des localisations aléatoires pour une durée totale de 30 minutes. Ce qui fait que l'auditeur écoute 80 fois le même passage sonore. Le changement aléatoire de la direction du son émis requiert une concentration du sujet pour la perception relative de l'image fantôme ce qui coupe la monotonie évidente du test.

Les tests de localisation décrits dans la littérature font références à plusieurs méthodes pour le jugement angulaire de la perception. Une solution consiste à demander au sujet de se tourner vers la localisation présumée du son et de noter le numéro d'étiquette correspondant le mieux à sa sensation. Ces étiquettes fixées sur une corde à linge à la hauteur des enceintes segmentent la perception azimutale. Une autre solution est de faire correspondre une source réelle (petit haut-parleur ou petit instrument de musique) à la source fantôme perçue maniée par un manipulateur qui reportera lui-même les incidences angulaires ressenties par l'auditeur. Cette méthode nécessite beaucoup d'interactions, elle est très contraignante pour le manipulateur qui subit l'écoute obligatoire du test pour chaque individu. Une autre possibilité de jugement est possible en proposant un diagramme polaire comme feuille de réponse à la personne testée. Dans ce cas, il n'est pas nécessaire pour le sujet de pivoter afin de valider son choix. On peut craindre ici une imprécision supplémentaire due à l'abstention visuelle. Il aurait pu être intéressant de choisir une solution mixte comprenant la méthode des étiquettes et celle du diagramme polaire.

Le choix du test s'est porté sur la première solution : la division spatiale par pas de 5° à l'aide de 72 étiquettes numérotées. Il est aussi demandé à l'auditeur de spécifier l'étiquette représentant la zone centrale lors d'une perception apparente plus large que la résolution de l'étiquette. Dans le cas où l'auditeur serait dans l'incapacité de donner une réponse, on lui demande de noter un point d'interrogation.

3.2 Réalisation pratique du test

Chaque étape du test correspond à une direction azimutale différente et la séquence du test est aléatoire. Une différence de niveau est appliquée aux paires d'enceintes relatives à la zone

d'incidence. Une compensation de niveau est affectée aux deux canaux pour garder une sensation d'intensité sonore équivalente par la formule suivante :

$$L_{pT} = 10 \text{Log}_{10} \left(10^{\frac{L_{p1}}{10}} + 10^{\frac{L_{p2}}{10}} \right)$$

Le support d'enregistrement est une bande magnétique HI8 au format D.A. de Tascam. Le montage et les ajustements de niveaux ont été réalisés à l'aide d'un DD1500 AKAI et diffusés par des enceintes Genelec 1037 via une console O2R Yamaha. Les annonces qui comprennent le texte de préambule et les numéros de chaque test ont été montés sur le DD1500.

Les différences de niveau utilisées (ΔI en dB) sont les suivantes et s'inspirent des résultats obtenus par Woszczyk [51] et Bruhnes [6] :

Pour l'avant entre -30° et $+30^\circ$: 15 , 10 , 5 , 0 , 3 , 6 , 12.

Pour l'arrière entre 120° et 240° : 17 , 14 , 11 , 8 , 5 , 2 , 0 , 3 , 6 , 9 , 12 , 15 , 18.

Pour le côté gauche entre 240° et 330° : 21 , 18 , 15 , 12 , 9 , 6 , 3 , 0 , 3 , 6.

Pour le côté droit entre 30° et 120° : 5 , 3 , 1 , 1 , 3 , 5 , 7 , 9 , 11 , 13.

La configuration en 5.1 étant symétrique, une redondance d'ordre deux apparaît quant aux résultats. Les différences de niveau utilisées pour la zone située à droite ne sont pas les mêmes que pour celles de gauche, afin d'obtenir une meilleure définition.

La reproduction de la bande HI8 s'effectue par une console 03D. Cette table de mixage sert à la distribution des canaux aux enceintes 4412 JBL et à la calibration des niveaux. Les amplificateurs sont de marque Crest. Une calibration est faite à chaque changement de configuration à l'aide d'un Audioprécision System One et du bruit rose à l'aide d'un micro 4006 de B&K placé à 1.5m.

Même s'il a été convenu de ne pas prendre en compte les variations de localisation en intensité dues à la fréquence, des tests complémentaires à vocation indicative concernant la partie spectrale ont été créés pour séparer la reproduction du test en deux zones :

- Une partie basses fréquences avec un filtre passe bas du 2^{ème} ordre coupant à 2kHz.
- Une partie hautes fréquences avec un filtre passe haut placé à 1kHz.

Seule la partie test est filtrée, les annonces conservent leurs spectres d'origine. Le rendu sonore est encore intelligible avec les spectres tronqués.

De la même façon, un test complémentaire nous fournira des indications sur la robustesse du pan-pot dans une configuration 5.1. Les personnes testées seront décalées de 50 cm du centre idéal pour quatre directions : avant arrière gauche et droite par rapport à l'axe médian et transaural.

3.3 Déroulement du test

Les résultats ont été saisis au fur et à mesure des tests permettant de connaître le nombre d'auditeur nécessaire à l'obtention de résultats convergents. Sur les cinq jours impartis aux tests d'écoutes, cette méthode m'a autorisé à l'élaboration de tests complémentaires. Ils concernent d'une part la perception de la localisation aux basses fréquences et aux hautes fréquences par filtrage du stimulus original et d'autre part, la robustesse des systèmes 5.1 par panning.

Les intervenants sont issus des étudiants et du personnel du Conservatoire de Musique de Paris, ils sont majoritairement très sensibilisés par le son et la musique. 40 personnes ont effectué le test dont 75% de sexe masculin. Ils proviennent des milieux suivants : preneurs de son professionnels, étudiants aux métiers du son, étudiants en musique, accordeurs de piano et des personnes dont la profession n'est pas liée à l'audio.

La phase de test comprend l'écoute proprement dite et un dialogue d'échange sur la perception individuelle de chacun.

Devant la constatation d'un décalage angulaire général de la perception latérale vers la gauche dans les résultats, la configuration des enceintes a été remaniée (rayon plus faible de 2.8m, recentrage et retournement du dispositif par rapport à la salle, nouvelle calibration). La déviation à gauche perçue a été légèrement réduite mais elle n'a pas disparue des valeurs résultantes. Cet écart est peut être issu des modes propres de la salle ou de la réponse en fréquence des enceintes.

3.4 Résultats des tests

3.4.1 Résultats d'ordre général

Hormis les résultats recherchés par l'expérience, les commentaires donnés sous forme écrite et par l'échange oral après le test sont des plus intéressants. La durée du test semble correcte, elle n'est jugée ni trop courte ni trop longue sauf pour un cas (durée trop longue) où le sujet se focalisait sur la première attaque du chant et négligeait le reste du message (notons que ses résultats concordent avec les autres). La répétition du même message sonore a gêné certaines personnes, principalement celles qui n'appréciaient pas particulièrement Suzanne Vega et ceux qui ont fait le test avec le spectre tronqué. Pour un auditeur, le test s'est avéré devenir un supplice.

Pour quelques auditeurs, la concentration imposée par le test est devenue éprouvante, ils avaient l'impression de ne plus être objectif à partir de la moitié du test. D'autres ont ressenti avec certaines séquences un son qui émanait de deux enceintes distinctes. En fait, ils percevaient grossièrement une image fantôme mais avec des zones spectrales de fréquence qui provenaient d'une enceinte précise. Phénomène ressenti dans l'aigu pour les attaques et les composantes sibilantes de la voix.

Aucune inversion de localisation avant arrière n'a été répertoriée durant le test. Cette constatation positive provient sans doute du contenu musical du stimulus comparé aux salves de bruits souvent utilisées. De plus, l'exemple placé à l'avant, permet de référencer qualitativement de manière cognitive l'auditeur sur la couleur du son. Le son latéral et arrière est détimbré naturellement par les HRTFs, il n'y a pas de confusion possible avec la référence préalablement entendue.

La perception de la localisation des sons émis par les deux enceintes arrières est différente suivant les auditeurs (en proportion identique). Soit la perception est ressentie en élévation, soit à la hauteur des enceintes. Il est possible d'avoir deux perceptions différentes en se concentrant mentalement. Cela amène à faire une distinction entre les auditeurs sensibilisés aux écoutes

électroacoustiques et ceux qui ont une écoute plus naturelle . L'ingénieur du son a appris à placer des sons dans un espace sonore qui n'est pas une représentation respectueuse de la réalité. Néanmoins, par apprentissage, il saura poser exactement des sources sonores à la position angulaire qu'il désire. On constate ainsi que les auditeurs travaillant en stéréophonie sont plus aptes à discriminer précisément en azimut une image fantôme. Les autres perçoivent une sensation dans la même direction mais qui pour eux, reste floue ou est perçue en hauteur.

3.4.2 Résultats des tests avec le spectre tronqué

Les réponses données sont totalement similaires à celles du test global que ce soit pour le grave ou l'aigu. Ce résultat n'indique pas si la localisation en intensité est indépendante de la fréquence émise, mais elle nous informe de la similitude globale de sensation de la localisation en haute fréquence et en basse fréquence en bande large. Pour un confort d'écoute, les deux bandes passantes des filtres utilisés ne sont pas complémentaires car une zone de recouvrement de 1kHz existe. Il se peut que l'essentiel de l'information de localisation en fonction du morceau musical choisi se situe à cette frontière mais il est plus judicieux de croire que la localisation de la partie basse du spectre est de même ordre de grandeur qu'en haute fréquence ou que la largeur des bandes passantes est trop grande pour influencer le jugement.

3.4.3 Résultats complets

Les résultats complets concernent les réponses à la perception de la localisation ressentie par les différents sujets. Tous les résultats ont été pris en compte même ceux qui semblent dévier anormalement de la tendance générale. Dans la mesure où le test a été effectué avec la même rigueur par tous les participants, il ne semble pas pertinent d'écarter certains mauvais résultats. La démarche inverse pourrait d'ailleurs être jugée non scientifique. N'oublions pas que la perception de la localisation d'une image fantôme n'est pas forcément similaire entre les individus.

On trouvera dans les pages suivantes, les résultats sous forme graphique. Tout d'abord des graphiques avec les moyennes, les écarts types ainsi que les minima et maxima des réponses angulaires dans les quatre quadrants formés par les paires d'enceintes. Ces mêmes moyennes seront représentées sous une forme repliée sur la médiane en raison de la redondance par deux du test. Les résultats seront ensuite mis sous forme graphique polaire pour une meilleure lisibilité et corrélation physique. On trouvera à la fin du chapitre une comparaison des résultats à ceux que l'on peut trouver dans la littérature. Les résultats des tests à spectre tronqué sont incorporés dans les tests complets.

Toutes les courbes présentées sont obtenues avec le panel complet des auditeurs. Pour les premiers graphiques, les points en forme de losange représentent les valeurs extrêmes, les marques en forme de trait symbolisent l'écart type centré sur la moyenne qui est caractérisée par les courbes continues. Les courbes de moyenne sont lissées.

Tests de localisation

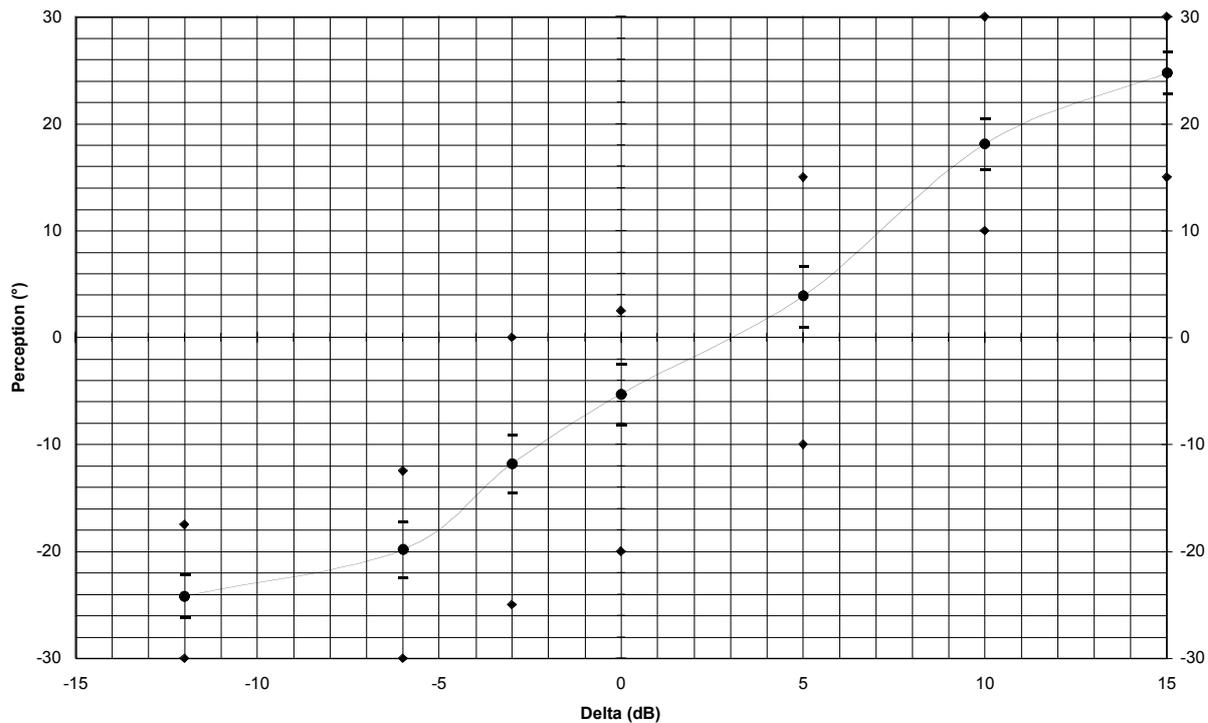


figure 9. Résultats du quadrant avant -30 à +30°

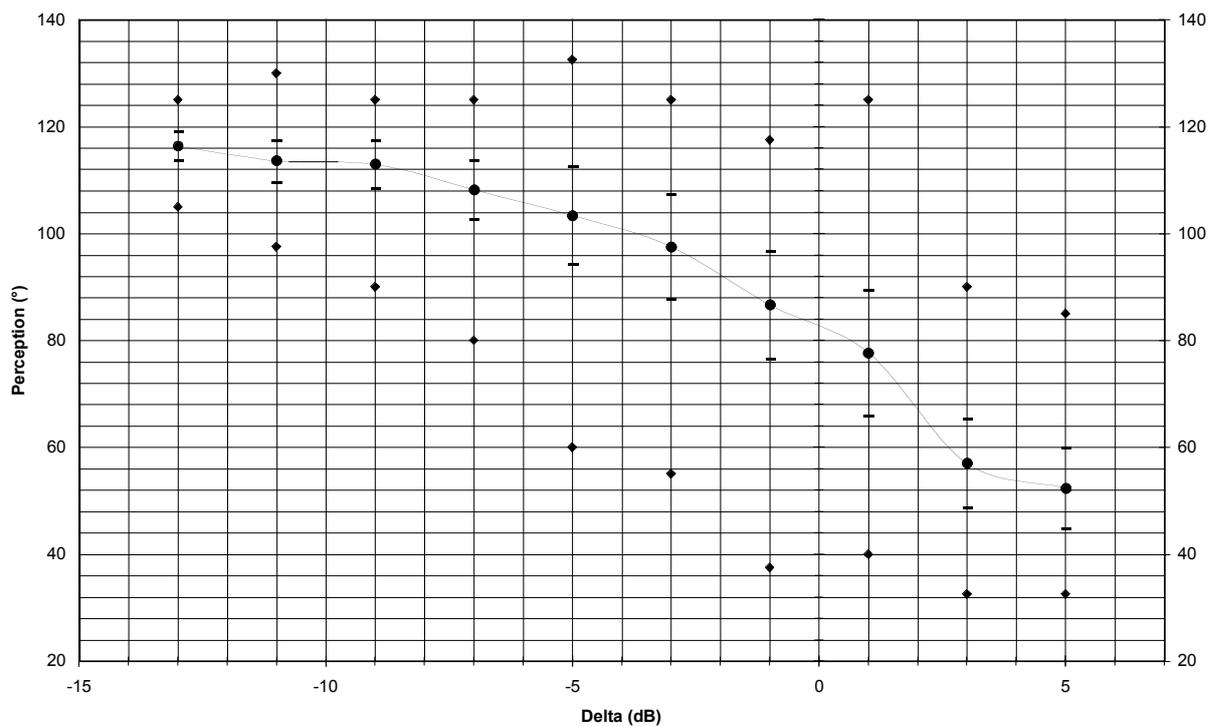


figure 10. Résultats du quadrant latéral droit +30 à +120°

On remarque instantanément pour le quadrant frontal de la figure 9, une dissymétrie de perception entre la partie droite et gauche. Ce problème d'offset n'a pas pu être élucidé malgré un changement de placement d'enceintes et une nouvelle calibration. Ce décalage vers la gauche est de l'ordre de 5°. La courbe présente une allure assez linéaire avec des pentes variant légèrement et symétriquement en fonction des zones angulaires. L'écart type est très faible.

Le quadrant latéral droit figure 10 montre que les valeurs de différence d'intensité attribuées pour le test n'ont pas été bien distribuées, il aurait été préférable de travailler sur une gamme plus large s'étendant au moins jusqu'à 12 dB. Les écarts types élevés de l'ordre de 20° indiquent une faiblesse de localisation latérale. Les valeurs des minima et des maxima montrent la forte variance inter individuelle. La courbe moyenne ne présente pas d'accident notable hormis la valeur pour une différence d'intensité de 3 dB qui s'écarte de la tendance

La figure 11 représentant le quadrant arrière met en évidence le même problème, le mauvais choix des valeurs initiales affectées aux pan-pots pour la création des tests. On remarque un accident à 6 dB. Ces accidents émanent peut être de réflexions murales. La figure 10 et la figure 11, représentant les quadrants latéraux ont des écarts types beaucoup plus grand pour les angles situés au centre des deux enceintes. On retrouve ici les constatations de Theile [38].

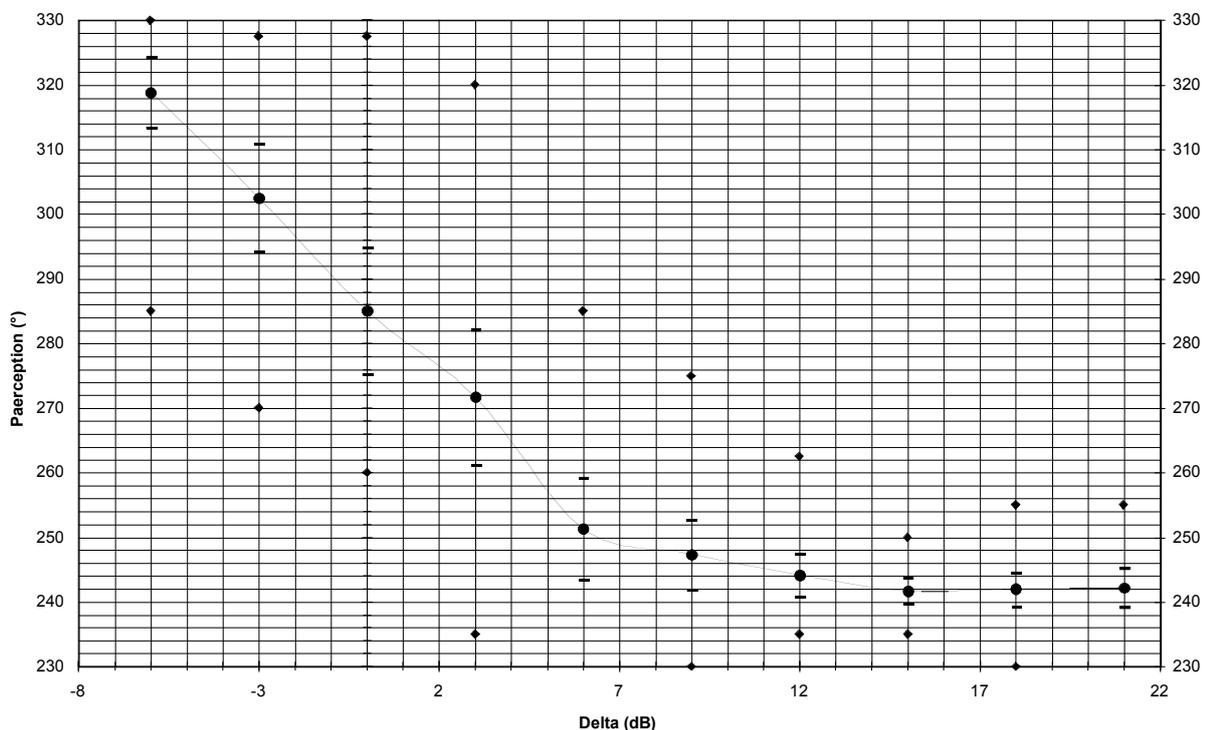


figure 11. Résultats du quadrant latéral gauche +240 à +330°

La courbe de la figure 12 est très régulière, il est presque surprenant de trouver une moyenne si régulière pour des sons qui émanent de l'arrière. La localisation est relativement stable entre individus avec la disposition des enceintes arrières du 5.1. On remarquera ici aussi un offset de l'ordre de 7° pour la position extrême de 180°. Les écarts types et les valeurs extrêmes sont beaucoup plus important que pour le quadrant avant et s'approchent d'avantage des données latérales. Ces quatre graphiques prouvent que la précision du jugement est plus

fiable à l'avant, résultats se corrélant aux résultats de la discrimination angulaire en azimut [4]. On constate aussi que la taille du quadrant est proportionnelle à cette précision.

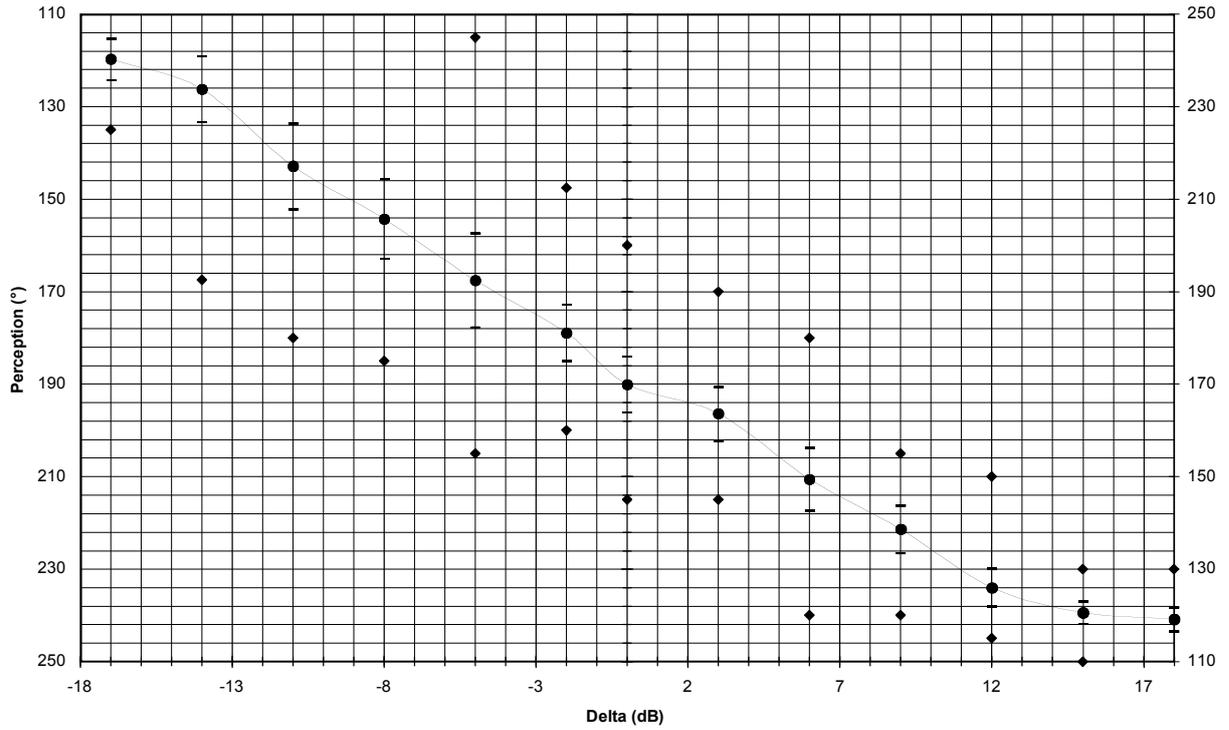


figure 12. Résultats du quadrant arrière +120 à +240°

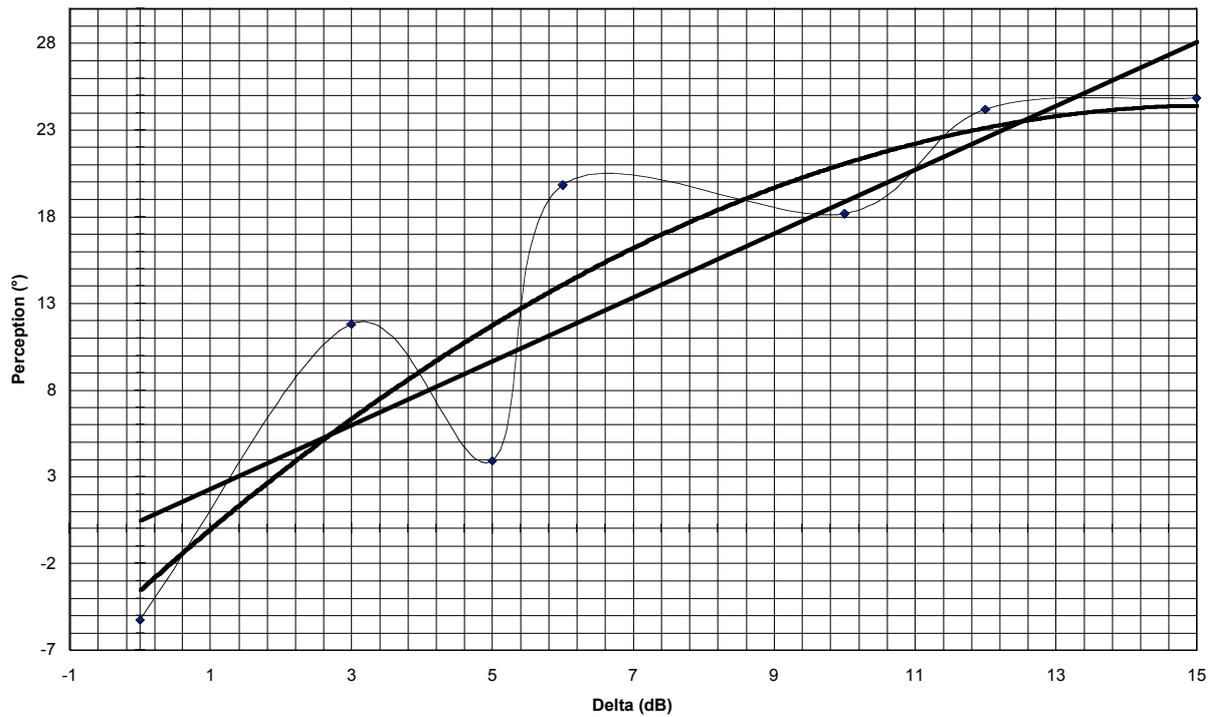


figure 13. Résultats du quadrant replié avant 0 à +30°

Les courbes de la figure 13 à la figure 15 montrent ces mêmes résultats repliés sur l'axe médian central. On visualise aussi la tendance linéaire qui est calculée en fonction des résultats repliés pour s'affranchir du décalage de perception de l'avant et de l'arrière. La tendance polynomiale d'ordre deux quant à elle, est une manière rapide d'obtenir des valeurs concernant la pente de la courbe.

Pour le quadrant frontal figure 13, la représentation n'est pas idéale mais les courbes de tendance donnent grossièrement les résultats désirés : une pente moyenne de 10° par une différence d'intensité de 5dB. On observe également une pente plus raide pour la zone angulaire proche du centre de l'ordre de 15° par 5dB.

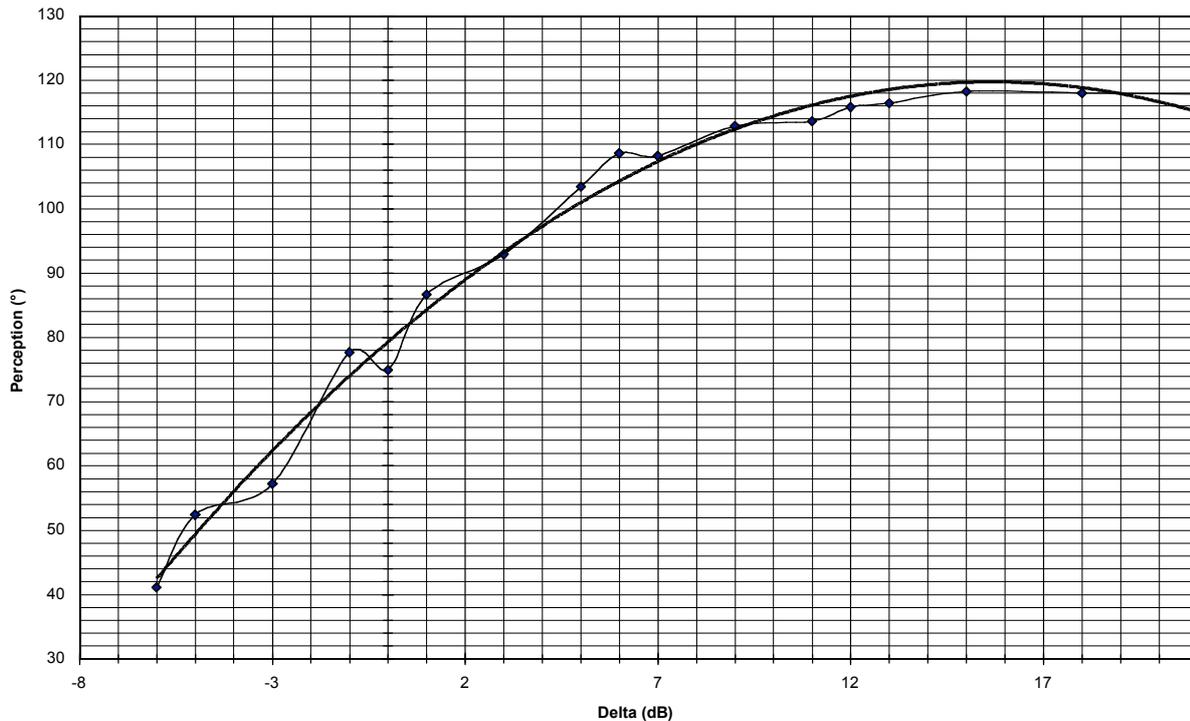


figure 14. Résultats des quadrants repliés latéraux de $+30$ à $+120^\circ$

La tendance polynomiale se confond avec la moyenne pour les quadrants latéraux de la figure 14. Entre les résultats latéraux droits et gauches, aucun biais ne vient perturber les données et l'allure constante de la courbe est remarquable. On ne peut pas extrapoler les valeurs de différence d'intensité manquantes pour l'avant entre 30 et 40° car on se situe au niveau du coude où la courbe va rejoindre doucement sans jamais l'atteindre la position angulaire de l'enceinte avant. On a une pente de $30^\circ/5\text{dB}$ pour la zone frontale et de $20^\circ/5\text{dB}$ pour la partie arrière. Une différence d'intensité de 0dB correspond à un angle de 79° environ, valeur très proche du centre géométrique des deux enceintes qui est de 75° .

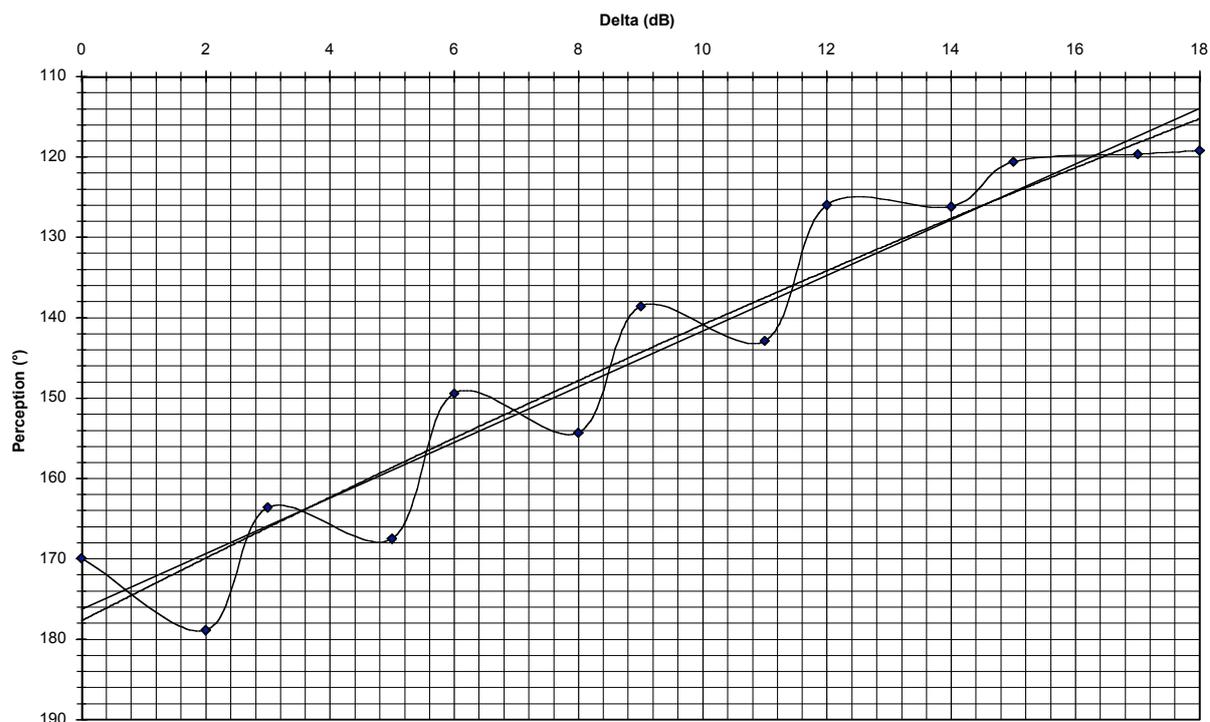


figure 15. Résultats du quadrant replié arrière de +120 à +180°

Sur le graphique replié de l'arrière figure 15, les tendances du 1^{er} et du 2nd ordre sont très semblables, cette pente continue nous informe que la progression est purement linéaire pour l'arrière. On lit ici, une pente de l'ordre de 20°/5dB.

Le Tableau 1 suivant récapitule les différentes asymptotes obtenues. Les quadrants latéraux et arrière ont des pentes plus raides mais elles ne sont pas totalement proportionnelles à la taille du quadrant. La perception latérale de la localisation nécessite des ΔI plus faibles.

Zone angulaire approximative en degré	Pente asymptotique	Enceintes concernées
0 à 15	15°/5dB	L et R
15 à 30	10°/5dB	L et R
30 à 90	30°/5dB	L et LS ou R et RS
90 à 120	20°/5dB	L et LS ou R et RS
120 à 180	20°/5dB	LS et RS

Tableau 1. Résultat sous forme de pente angulaire en fonction des différences d'intensité

Ces résultats repliés sont maintenant interprétés dans leur ensemble à l'aide d'une visualisation polaire. On y trouve des représentations avec des axes linéaires permettant de détailler les résultats ainsi que des représentations logarithmiques sur une échelle de 30 décibels correspondant aux diagrammes polaires employés pour l'affichage de la directivité des microphones de qualité professionnelle. La visualisation en linéaire est plus claire mais n'est pas un axe représentatif de la perception auditive.

Sur les diagrammes polaires de la figure 16 et de la figure 17, les résultats apparaissent pour les quatre quadrants qui nous intéressent. Les symboles qui différencient les quatre courbes sont les valeurs de ΔI choisies préalablement pour le test. On voit que certaines zones angulaires sont plus détaillées que d'autres. On retrouve dans cette représentation le biais de perception à l'avant et à l'arrière. Hormis ce défaut, la symétrie de directivité est respectée. Comme on le pressentait, les diagrammes de directivités exigés pour la zone frontale sont plus directifs que ceux de la zone arrière.

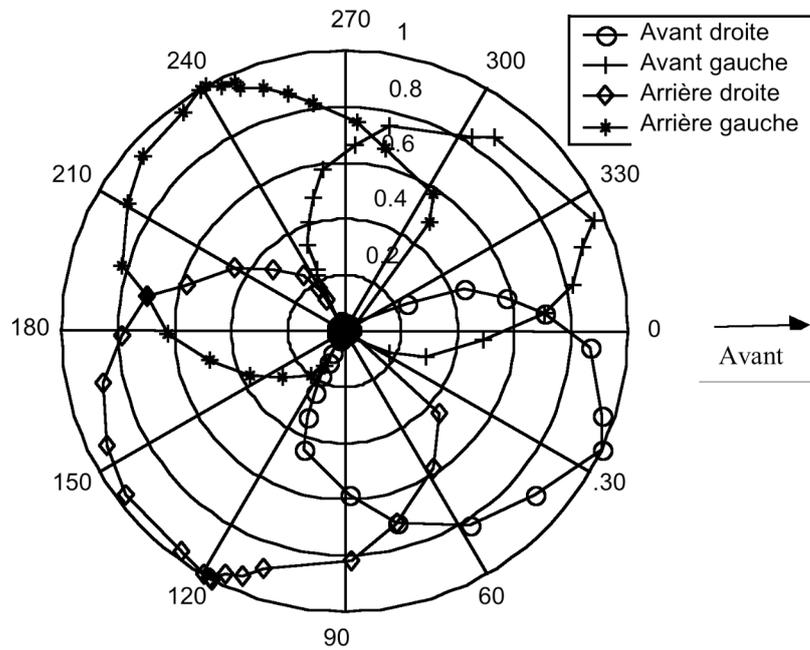


figure 16. Diagrammes polaires du test en valeurs linéaires

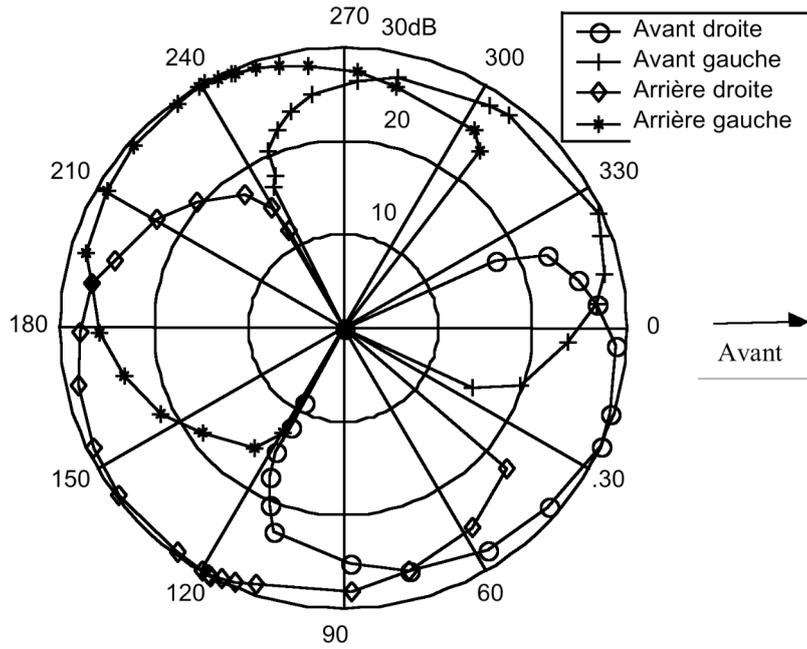


figure 17. Diagrammes polaires du test en valeurs logarithmiques

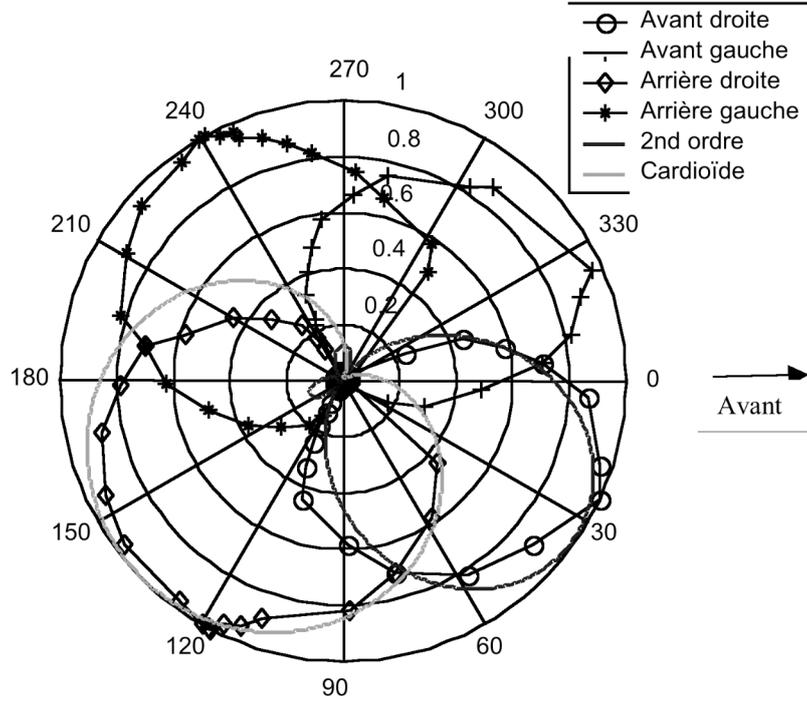


figure 18. Diagrammes polaires du test en valeurs linéaires comparés à des courbes théoriques

Nous pouvons maintenant comparer ces diagrammes aux directivités théoriques des différentes technologies de microphones. La clarté de vision obtenue par la représentation linéaire de la figure 16 nous autorise à superposer les représentations polaires souhaitées aux directivités possibles (figure 18). Les courbes théoriques retenues pour cette comparaison sont celles du 1^{er} et du 2nd ordre (voir chapitre 4.2.5. Les microphones du second ordre). L'œil sert ainsi d'outil de comparaison entre les diagrammes. On choisira les directivités des microphones qui s'approchent le plus des résultats : cardioïde, hypercardioïde et second ordre simple. La visualisation des diagrammes utilise une routine Matlab (voir annexe 10.2). Pour une meilleure précision, les courbes suivantes sont représentées avec les résultats finaux repliés.

La courbe des données frontales superposée avec la courbe théorique du 2nd ordre de la figure 19 est d'une grande similitude et plus particulièrement pour la zone frontale comprise entre 0 et 30°. La directivité du 2nd ordre correspond à notre attente et elle sera retenue pour la suite de l'étude. L'axe du diagramme polaire théorique a été calé pour un angle de 45° pour épouser au mieux les contours de la courbe subjective, ce qui fait qu'il se confond avec le centre géométrique de captation égale à $(90+120)/2-30 = 45^\circ$. Il faudra prendre en compte cette valeur lors de la prise de son pour se situer au plus près des résultats subjectifs. Cet angle correspond à l'angle physique de prise de son pour les microphones captant la zone frontale.

Les résultats du test concernant la captation arrière sont comparés avec les différentes directivités de la figure 20 à la figure 22. Pour la simulation cardioïde et du 2nd ordre, l'allure des courbes comparées diffère sensiblement, des différences supérieures à 5dB sont constatées dans la zone utile avec le cardioïde et supérieures à 10 dB avec le 2nd ordre. La figure 21 montre qu'une directivité hypercardioïde se confond bien plus avec les résultats des tests. Le problème évident concerne le lobe arrière beaucoup trop important (6dB d'atténuation par rapport à l'axe) qui vient capter le son à -50° dans la configuration. Le 2nd ordre ne sera pas sélectionné pour la captation arrière, pas plus que l'hypercardioïde qui n'est pas conforme aux exigences du multi-canal en raison de son lobe important, on gardera dans un premier temps la solution cardioïde. L'angle choisi pour la représentation des courbes théoriques est de 130° alors que le centre géométrique de captation se situe à $210/2+30=135^\circ$.

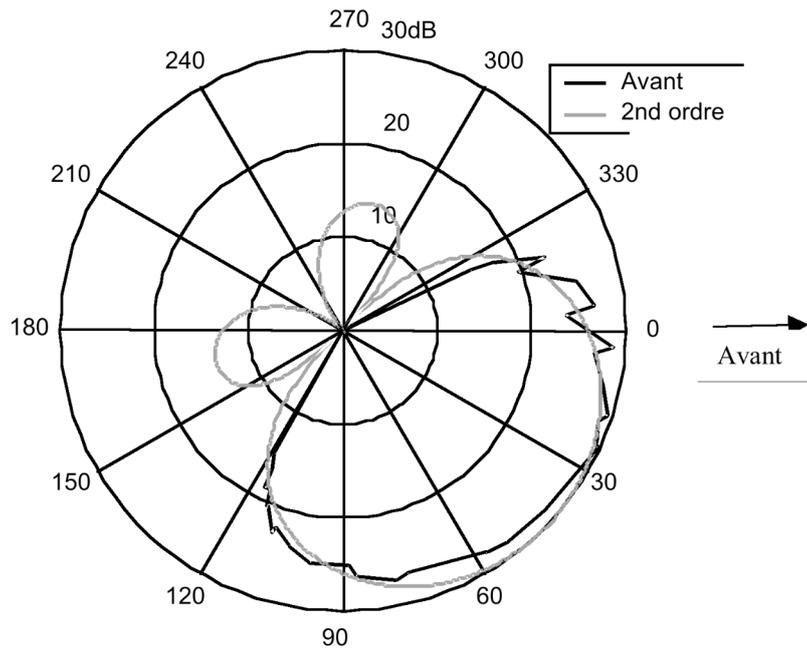


figure 19. Comparaison des valeurs repliées du test et une directivité du 2nd ordre pour l'avant

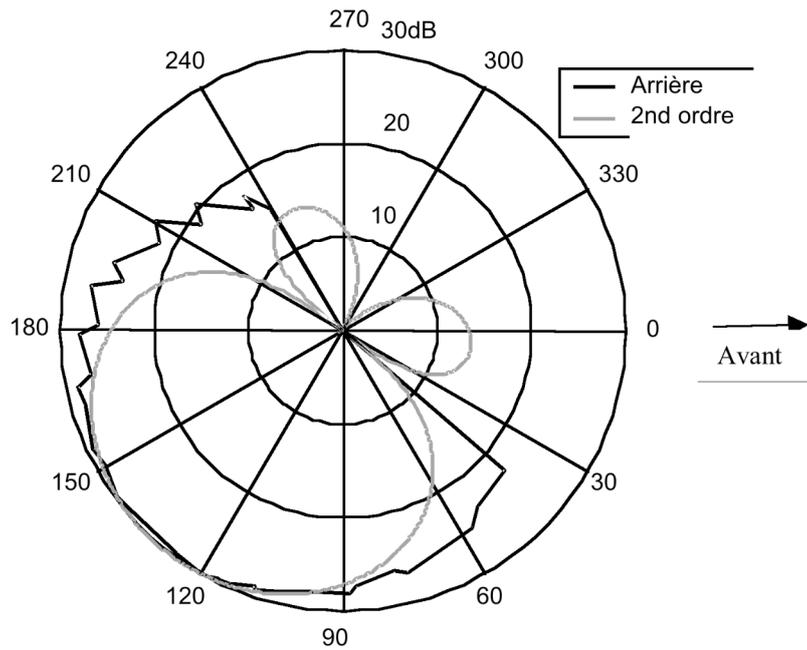


figure 20. Comparaison des valeurs repliées du test et une directivité du 2nd ordre pour l'arrière

Tests de localisation

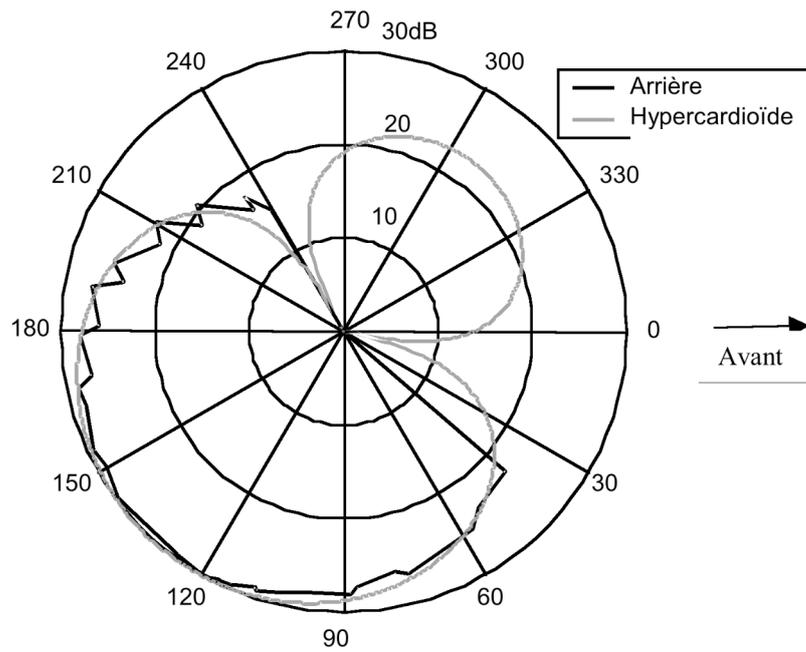


figure 21. Comparaison des valeurs repliées du test et une directivité hypercardioïde pour l'arrière

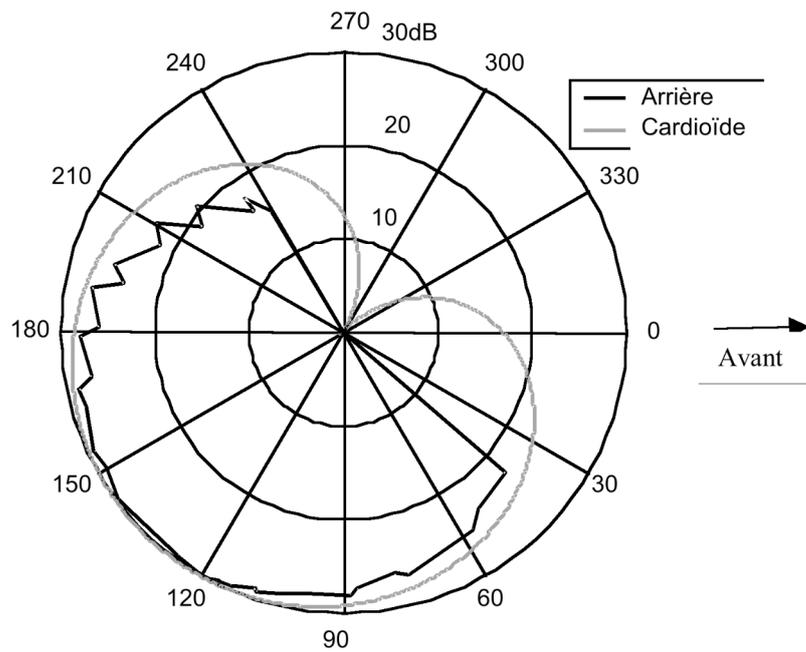


figure 22. Comparaison des valeurs repliées du test et une directivité cardioïde pour l'arrière

3.4.4 Comparaison avec d'autres résultats

Les résultats publiés concernant ce type de tests ne sont pas nombreux. En fait seul les récents résultats de Woszczyk [51] et de Bruhnes [6] existent à ma connaissance. Dans leurs études, ils se fient à la recommandation ITU [17].

Pour son test, Woszczyk utilise cinq canaux en incorporant l'enceinte centrale. Ce qui fait que ses résultats pour la zone frontale sont divisés en deux zones -30 à 0° et 0° à 30° et ne sont donc pas compatibles avec ceux obtenus dans ce mémoire. Les autres quadrants quant à eux correspondent tout à fait (choix de l'angle de diffusion des enceintes arrières de 120°). Le stimulus utilisé est une voix de femme enregistrée en milieu anéchoïque et cinq personnes ont effectué le test. La notation des résultats s'est faite à l'aide d'un tableur représentant un diagramme polaire.

Bruhnes ne fait pas l'étude du quadrant avant, il estime que ces résultats sont connus. Il utilise les stimuli suivant : impulsions de bruit rose, impulsion de grosse caisse et vibraphone. Son test est effectué par 10 sujets. Il recense en moyenne 15% d'erreur (des inversions notamment qui ne seront pas prises en compte dans les résultats). Les personnes testées n'avaient pas le droit d'utiliser les petits mouvements de tête et avaient les yeux bandés. Les éléments de comparaison sont les valeurs moyennes des résultats des trois tests. La méthode de détermination de la perception angulaire se fait par la méthode de la comparaison d'un son réel par rapport à la source fantôme. Voir chapitre 3.1.

Une étude réalisée par West [43] recense quelles sont les meilleures lois de panning pour un environnement multicanal en 5.1. Des lois mathématiques affectées aux boutons de pan-pot des consoles numériques peuvent se rapprocher de la perception de la localisation. On comparera donc deux lois avec les résultats du test : une loi en sinus et une loi linéaire.

La loi en sinus est décrite ainsi :

$$\frac{\sin \varphi}{\sin \varphi_0} = \frac{A - B}{A + B}$$

φ est l'angle de la source fantôme et φ_0 , l'angle physique entre les deux haut-parleurs.

A et B sont les coefficients linéaires à affecter aux deux canaux sachant que leur somme est égale à 1.

La loi linéaire est :

$$\frac{\varphi}{\varphi_0} = \frac{A - B}{A + B} \quad \text{sachant que} \quad A + B = 1$$

Le quadrant arrière de 120° n'est pas utilisable dans la formule en sinus car elle correspond aussi à un angle φ_0 de 60° .

On comparera aussi les résultats du test avec ceux largement diffusés dans la littérature concernant la zone frontale [4] [45]. On comparera les données de Mertens, de Simonsen et de Wendt.

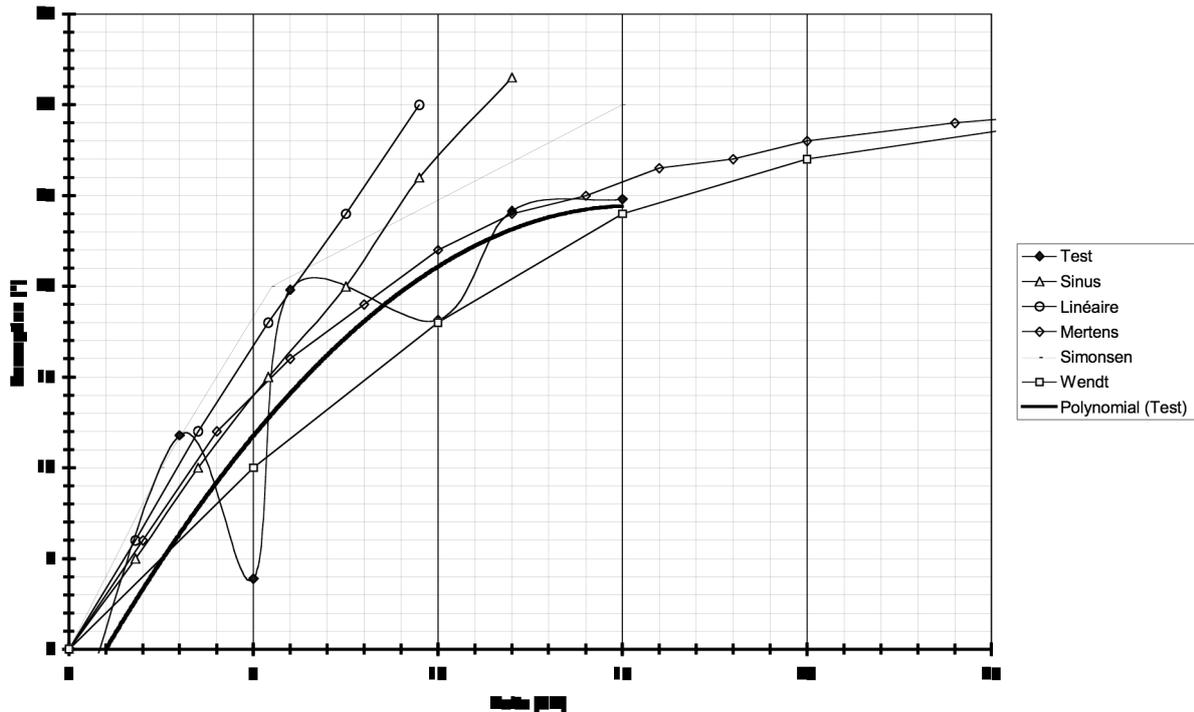


figure 23. Comparaison de divers résultats pour le quadrant frontal

Pour la comparaison des résultats pour le quadrant avant (figure 23), la pente est sensiblement la même pour la zone proche de l'axe. Les lois de pan-pot en linéaire et en sinus codent les différences d'intensité trop faiblement comparées aux données des résultats subjectifs. Les résultats de Mertens suivent de très près, les résultats obtenus ici. Ces résultats concernent ceux de la biphonie et il est étonnant qu'après des décennies d'utilisation, de telles variations de valeurs existent encore. On aurait pu s'attendre à des résultats plus proches les uns des autres car les prises de son en ΔI sont issues directement de ces données.

La figure 24 représentant les quadrants latéraux nous indique que la loi linéaire correspond mieux aux valeurs du test dans la zone centrale du quadrant que la loi en sinus. Les résultats se corrélaient assez bien avec ceux du test alors que ceux de Woszczyk semblent plus variables. Cette courbe capricieuse est certainement due au faible nombre de sujets. Paradoxalement, c'est pour ce quadrant que les différents résultats sont les plus proches alors que les divergences entre les individus sont les plus grandes.

Sur la figure 25, les résultats du test s'écartent très sensiblement de tous les autres. Est-ce que ce sont les conditions des tests (méthode et mode de notation) qui sont trop différents pour donner des résultats comparables ? Seul le test présent utilise l'interaction visuelle pour juger la perception de la localisation. Mais des effets de salle ou le choix des stimuli contribuent peut-être aussi à ces variations. Les résultats du test seront conservés pour la suite du mémoire pour deux raisons : le nombre d'auditeurs est supérieur à celui des autres tests et les résultats ne possèdent pas d'inversion avant arrière. Remarquons tout de même la similitude entre les résultats de Bruhnes et la loi linéaire.

Ces comparaisons reflètent un triste constat : les variations des différentes courbes entre elles sont la preuve d'une non reproductibilité des tests subjectifs de la perception de la localisation en multicanal.

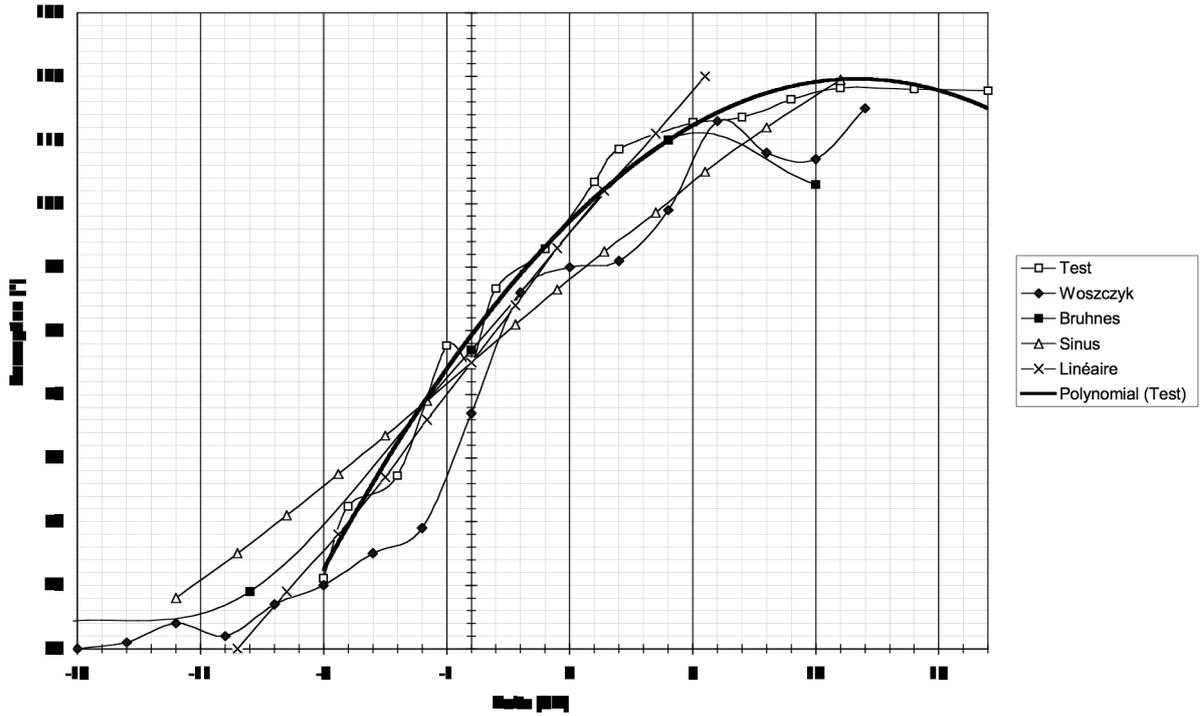


figure 24. Comparaison de divers résultats pour le quadrant latéral

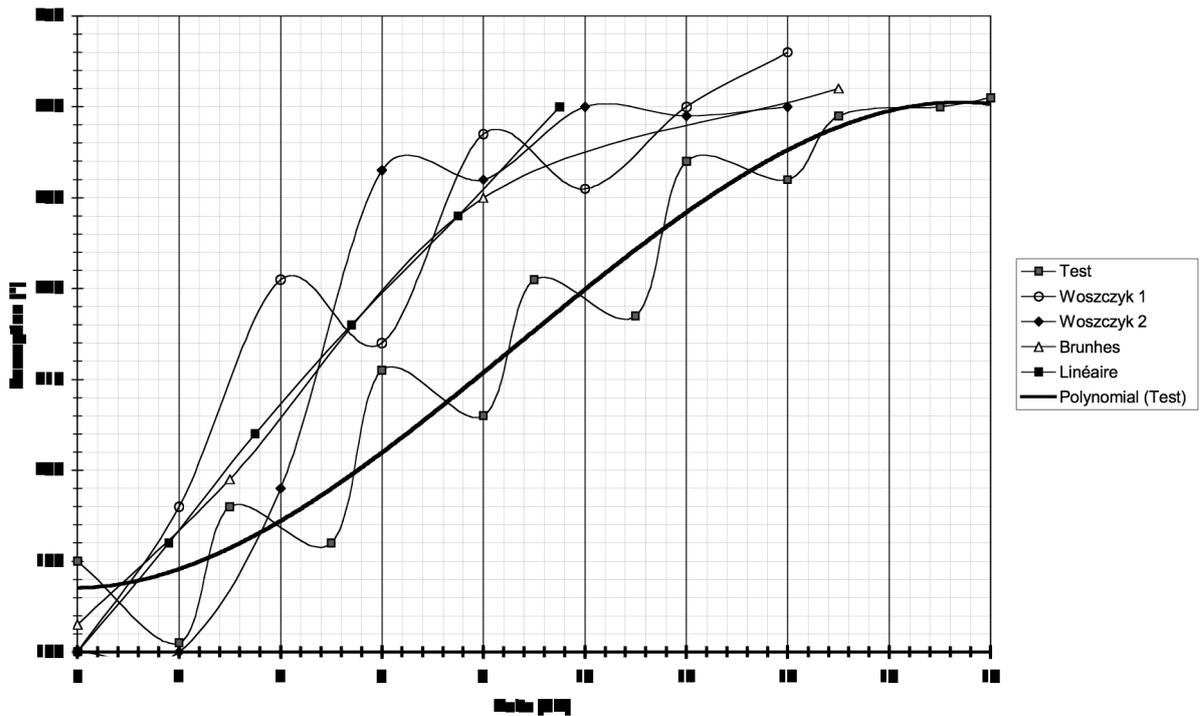


figure 25. Comparaison de divers résultats pour le quadrant arrière

3.4.5 Résultats concernant la robustesse

6 personnes ont effectué le test en étant décalées de 50 cm par rapport au centre ce qui correspond à 1/5^{ème} du rayon environ entre l'auditeur et les enceintes. Aucune inversion de localisation n'a été recensée dans les feuilles de résultats.

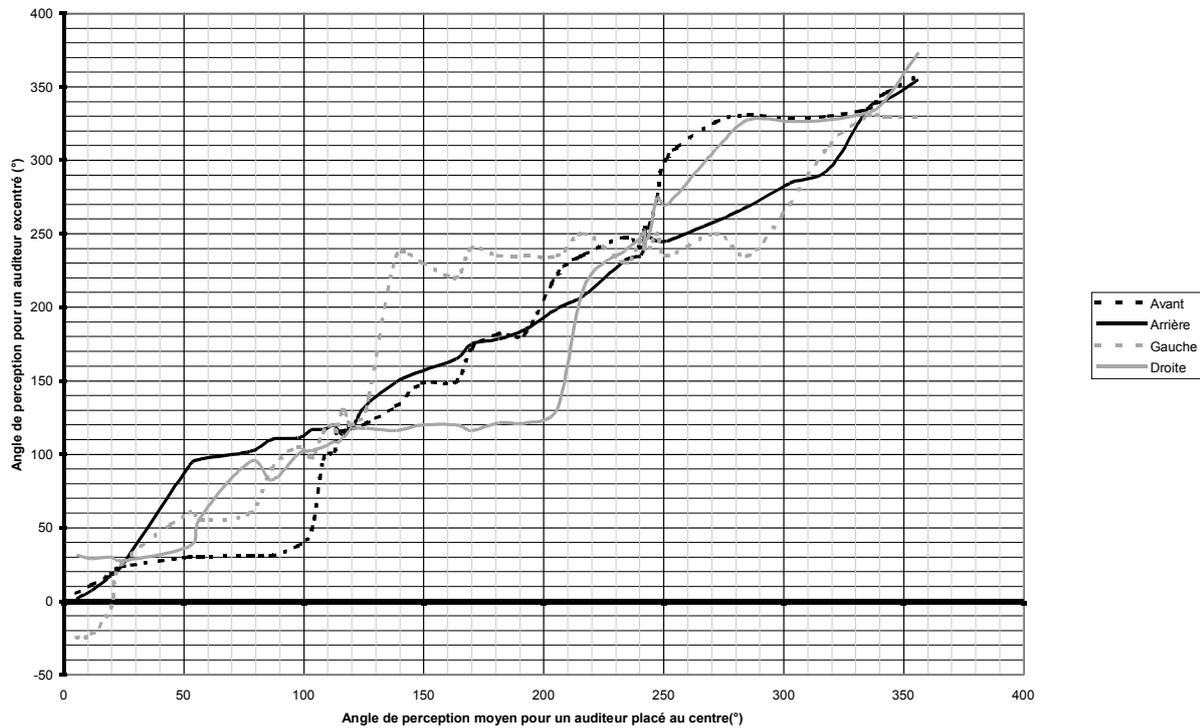


figure 26. Essai de la robustesse des lois de pan-pot

La figure 26 montre la déviation générée pour ces positions excentrées. Chaque courbe ne correspond qu'à un ou deux individus, les valeurs ne doivent donc pas être considérées comme des résultats globaux. On dira donc avec précaution que la position décalée de l'auditeur génère des écarts angulaires de perception qui sont de l'ordre de 50° en moyenne et qui peuvent aller jusqu'à 90° maximum. On peut aussi dire à la vue de ces résultats que les positions décalées sur le côté sont moins robustes. On peut l'expliquer par le fait que ces placements ne respectent pas la symétrie axiale de la médiane.

On remarque que la perception stagne sur des paliers à des angles précis (ceux des enceintes évidemment) mais que la perception de la localisation des sujets est continue. L'auditeur possède, compte tenu sa position décentrée une perception faussée certes, mais non corrompue par des changements de direction. On peut penser qu'avec des stimuli en mouvement, l'effet dynamique voulu sera encore ressenti.

3.5 Conclusion du test

Refaire le test une seconde fois améliore les résultats. Cela met en jeu un apprentissage. La difficulté d'organiser un planning et le fait de ne pas contraindre les auditeurs volontaires à une deuxième écoute m'ont dissuadé d'imposer une seconde écoute.

Les résultats des tests sont globalement très satisfaisants, ils répondent à la question concernant la directivité des microphones à employer dans une prise de son multicanal en ΔI et nous renseignent sur l'angle physique de prise de son à utiliser.

La comparaison des résultats des différents tests montre une grande divergence. Les différents échantillons musicaux et sonores ainsi que les conditions d'écoutes variées en sont certainement les facteurs d'origine. Des moyennes décalées entre elles de plus de 20° sont inquiétantes en ce qui concerne la validité des courbes, une normalisation des conditions de test s'avère nécessaire pour une meilleure fiabilité.

4 Les microphones directifs

Même sans faire de mesures subjectives, on se doutait que les angles physiques de reproduction en 5.1 allaient imposer un facteur de directivité Q élevé pour la prise de son. Une valeur de Q supérieure à 7 sera nécessaire pour la captation avant et de 6 environ pour la captation arrière. Le test précédent nous conduit à utiliser une captation avant du 2nd ordre et une captation hypercardioïde sans lobe parasite pour l'arrière.

La directivité des microphones souhaitée n'existe pas dans les catalogues électroacoustiques des différents fournisseurs du marché et on devra se contenter de caractéristiques approchantes. En raison des caractéristiques intrinsèques aux différentes technologies employées pour la réalisation de microphones directifs, certains types de microphones ne correspondent pas à l'objectif que nous nous sommes fixé et ne sont donc pas utilisables pour cette application. Les microphones utilisables sont ceux qui par nature ne présentent pas de défauts prononcés ou qui pourrait dans une certaine mesure s'adapter pour tendre vers les caractéristiques souhaitées.

4.1 Les technologies non utilisables

4.1.1 Les microphones canons

Un microphone canon ou interférentiel est conçu de telle sorte que l'onde sonore arrive sur la membrane avec des durées de parcours différentes en fonction de l'incidence de la source. On réalise cette technologie par la disposition de multiples fentes tout le long du corps de la capsule.

Ces microphones génèrent des lobes de directivité non désirés et très prononcés conduisant à une qualité sonore médiocre. Comme on peut le constater sur le diagramme polaire de la figure 27 pour deux fréquences (8kHz et 12kHz), le microphone composé ici de 10 fentes espacées d'une distance d de 3mm possède une fonction de directivité qui n'est pas constante en fréquence.

La fonction de directivité d'un microphone constitué par n capsules cardioïdes est la suivante :

$$h(\theta) = \frac{\sin(nkd \sin(\theta))}{2n \sin(kd \sin(\theta))} (1 + \cos(\theta))$$

Ce type de microphone n'est pas utilisable pour une prise de son musical. Une amélioration très significative de la réponse polaire de ce principe est possible si l'on se sert de lignes de microphones au lieu de fentes. Voir 4.2.3 Les réseaux de microphones.

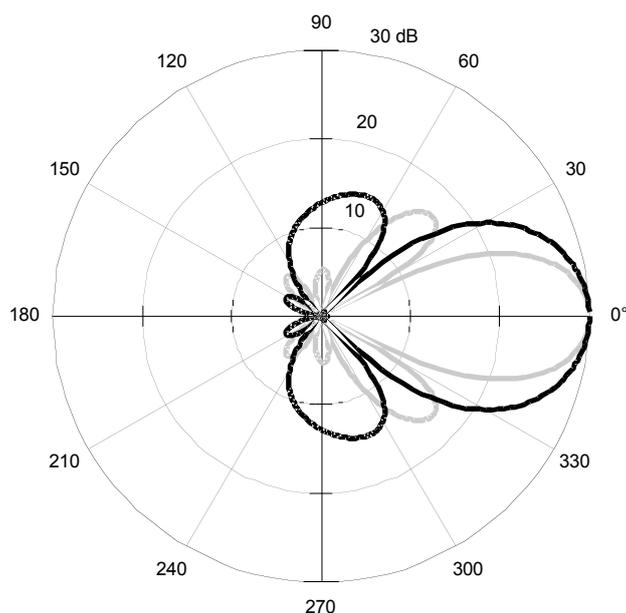


figure 27. Courbe polaire d'un microphone interférentiel pour deux fréquences

4.1.2 Les microphones à constantes réparties

Les microphones à constantes réparties sont les microphones directifs les plus couramment employés de nos jours [20]. Le principe est de ramener sur la membrane par l'arrière, une onde retardée par un filtre RC composé d'une cavité et d'un matériau résistif. Les directivités créées sont de la forme :

$$h(\theta) = A + B \cos(\theta) \text{ avec } A + B = 1$$

Elles comprennent les directivités omnidirectionnelles à bidirectionnelles en passant par toutes les variantes cardioïdes. Ce sont les directivités du 1^{er} ordre.

Le facteur de directivité des microphones donné pour hypercardioïde est plus grand que celui d'un cardioïde, malheureusement la technologie employée entraîne l'apparition d'un lobe arrière prononcé. Plus on augmente le facteur de directivité, plus la captation arrière augmente. Le lobe arrière non désirable est trop important pour une prise de son stéréophonique.

4.1.3 Les microphones utilisant une parabole

Configuration élégante, la parabole permet de focaliser le rayonnement sonore en un foyer de la même manière qu'en électromagnétisme. En raison de son ombre, la parabole empêche toute prise de son multiple coïncidente. La bande passante limitée par les dimensions de la

parabole et la directivité dépendante de la fréquence sont la cause d'une improbable utilité de cet accessoire pour le cas qui nous concerne.

4.1.4 Une évolution du microphone ambisonic

Gerzon déclare possible la réalisation d'un microphone d'un second ordre à l'aide de douze capsules coïncidentes dans une disposition en dodécaèdre [12]. L'assemblage géométrique complexe d'un tel ensemble nécessite l'utilisation de cardioïdes de très petite taille et de bonne qualité.

Gerzon n'explique pas comment il combine les capsules entre elles afin d'obtenir l'effet désiré, ni le traitement éventuel qu'il leur applique. Si l'on raisonne d'un point de vue mathématique, il faut chercher un résultat composé d'une matrice de pondération positive ou négative liée aux positions géométriques des directivités cardioïdes dans l'espace donnant une directivité du second ordre dans un espace tridimensionnel quelconque.

L'expérience s'est limitée à simuler graphiquement à l'aide de Matlab (cf. annexe 10.6) la sommation de capsules. Que ce soit en 2D ou en 3D, la sommation ou la soustraction de cellules de 1^{er} ordre n'engendre que des directivités du 1^{er} ordre. Il faudrait pouvoir démontrer mathématiquement ce résultat.

De surcroît, la complexité de réalisation de matrices complexes et d'un microphone de prise de son dodécaèdre coïncident fait, que cette possibilité de réalisation ne sera pas plus approfondie.

Le champ sonore peut être décomposé en harmoniques sphériques [29]. Actuellement le microphone ambisonic ne capte que celles d'ordre 1. Une évolution de l'ambisonic consisterait à prendre en compte les harmoniques d'ordre deux, ce qui nécessite des microphones à directivité du 2nd ordre.

4.2 Les technologies possibles

4.2.1 Les microphones à constantes réparties modifiées

Pour les microphones à constante répartie, le réseau retardateur se situe à l'intérieur du boîtier de la capsule et agit de la même manière pour toutes les incidences des ondes sonores. La taille de l'ouverture de l'ouïe au niveau du boîtier de la capsule permet d'ajuster le réseau déphaseur RC pour toutes les fréquences, en compensant la différence de marche par l'onde de contournement.

Une solution consiste à faire une ablation du lobe arrière par l'ajout externe d'une mousse agissant comme un filtre acoustique retardateur. Si l'onde acoustique ne rencontre pas ce filtre, seul le retard interne agira. Si l'incidence de la source traverse la mousse externe, l'onde est alors retardée et ce retard s'ajoute à celui du réseau interne. La mousse poreuse judicieusement placée génère un retard supplémentaire par ses caractéristiques de résistivité et de tortuosité (cf. figure 28). Malheureusement, les mousses acoustiques ne sont pas linéaires en fréquence et les caracté-

ristiques acoustiques non spécifiées. Les mousses sont plus généralement utilisées pour du traitement acoustique des salles et sont caractérisées la plupart du temps dans le domaine des fréquences ultrasonores.

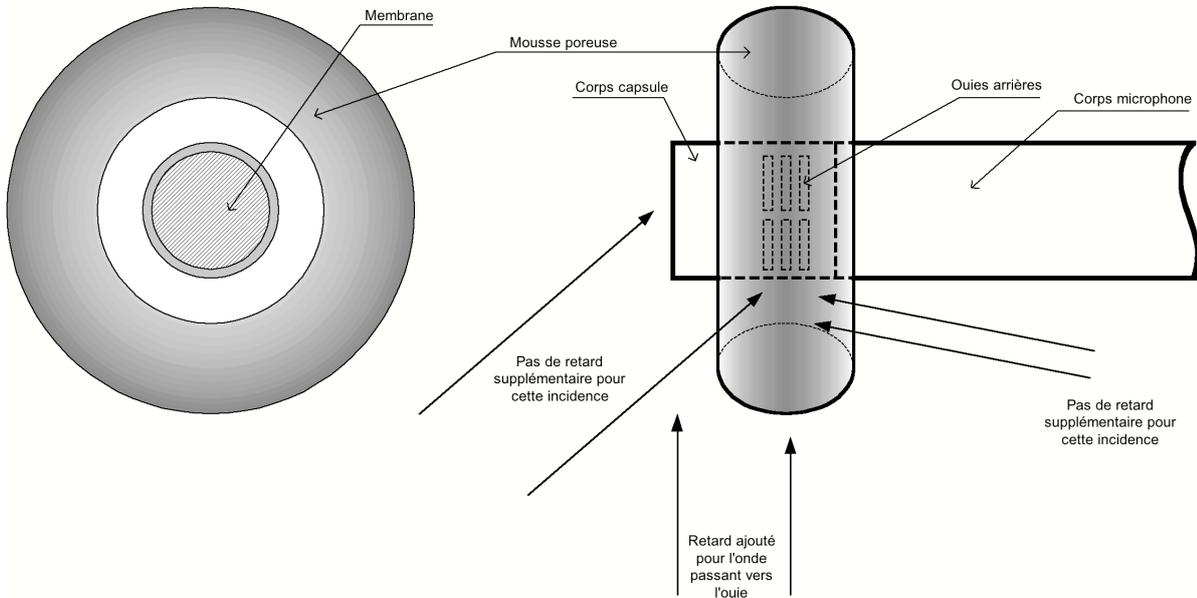


figure 28. Ajout d'un retard supplémentaire en fonction de l'incidence sonore

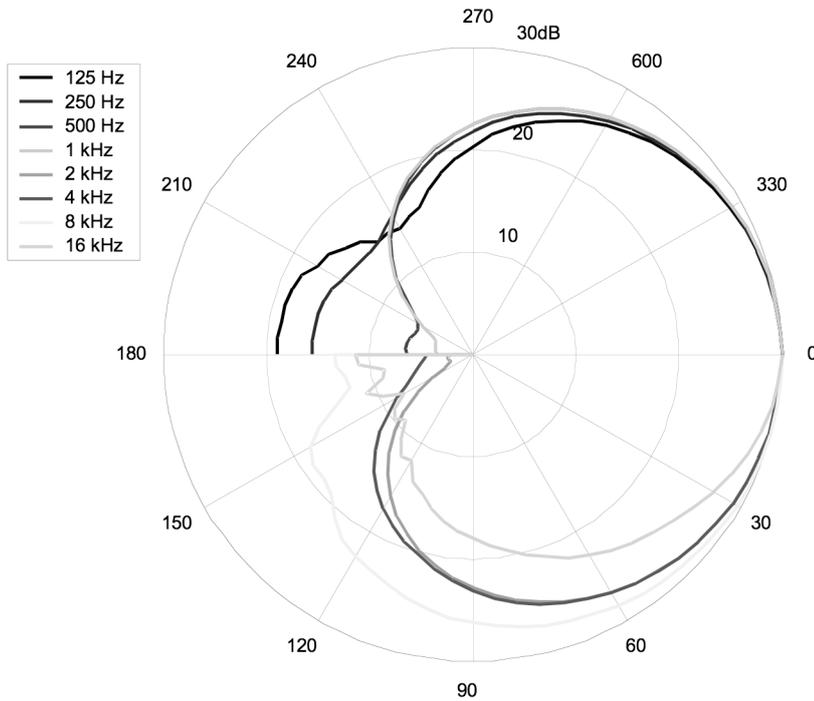


figure 29. Diagramme polaire obtenu avec une mousse externe

C'est la forme et le positionnement de la mousse qui déterminent si l'onde sonore subit un retard supplémentaire. Pour parvenir à une directivité précise, il faudrait modéliser ce système pour obtenir la forme et l'épaisseur adéquates de la mousse à positionner en fonction de la

distance entre le matériau poreux et le corps, des caractéristiques du matériau et de l'incidence de la source.

Les essais effectués montrent que l'atténuation latérale est faible (figure 29). L'encombrement dû à ce dispositif ne va pas dans le sens d'une prise de son coïncidente où la mousse risque d'interférer avec la captation des autres microphones.

4.2.2 Sommation de microphones

Une solution simple de réalisation de microphone directif pourrait être envisagée en fonction des défauts de directivité des microphones omnidirectionnels. La sommation de deux microphones cardioïdes génère un microphone omnidirectionnel. Or, la réalisation pratique d'un microphone omnidirectionnel offre rarement des caractéristiques de directivité complètement omnidirectionnelles. Un microphone omnidirectionnel auquel on retranche la directivité d'un micro cardioïde nous donne un microphone dont la directivité est plus prononcée qu'un cardioïde.

Les simulations réalisées ont démontré l'ambiguïté de cette solution. Les résultats recherchés sont générés par des défauts qui ne sont pas forcément linéaires en fréquence et qui sont spécifiques à chaque modèle de microphone issu de la même technologie.

4.2.3 Les réseaux de microphones

Ce type de microphone numérique est une solution composée d'un mélange de technologie issue des microphones interférentiels ainsi que du principe de base de la réalisation des microphones du second ordre (voir chapitre 4.2.5). Cette technologie est similaire aux antennes électromagnétiques. Les capteurs sont placés sur une ligne et sont pondérés individuellement par filtrage. La directivité de l'ensemble microphonique dépend de la fréquence, de la distance entre les capsules et de la pondération [15]. Pour compenser l'effet de propagation, les filtres numériques utilisés pour la pondération peuvent inclure des retards. Des capsules espacées en progression logarithmique donne de meilleurs résultats. La réalisation d'un tel microphone est un compromis dans le choix des variables qui sont parfois contradictoires : obtention d'un faisceau étroit et constant en fréquence, distance entre capsules faible pour éviter l'aliasing spatial, nombre important de capteurs et dimensions du système raisonnables.

Yves Grenier [15] a réalisé deux prototypes contenant 12 cellules pour l'un et 36 pour l'autre. Les mesures du prototype montrent que les lobes arrière sont de très faible amplitude (atténuation de 20 dB pour une fréquence inférieure à 250 Hz et de 30 dB pour les fréquences supérieures). La bande utile de fréquence se situe entre 70 et 14 kHz. L'atténuation à -6 dB se positionne entre 40 et 45° entre 500Hz et 8 kHz, ce qui est remarquable.

Ce microphone est trop directif pour notre application, mais son concepteur nous informe que n'importe quelle directivité peut être réalisée pour s'approcher de celle désirée. C'est la seule solution à court terme pour réaliser des prises de son coïncidentes pour un système multi-canal composé d'un nombre élevé d'enceintes. Ce pourrait d'ailleurs être une solution pour l'ambisonie ou la visioconférence. Ces prototypes sont à ce jour uniques et encombrants ce qui fait qu'ils ne seront pas retenus pour cette étude.

Une variante des réseaux classiques de pondération est proposée par la société Audio-Technica [54]. Ce microphone à réseau adaptatif est composé de cinq capsules (cf. figure 30). Il possède trois modes de fonctionnement : micro canon associé soit avec les 4 capsules cardioïdes, soit avec deux seulement, soit aucunes. Les modes adaptatifs se servent des capsules cardioïdes et de traitement analogique et numérique pour rendre le système plus directif.

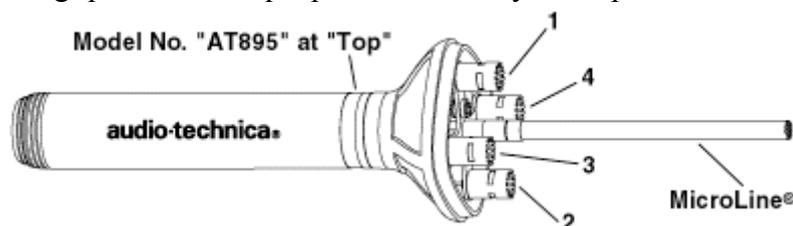


figure 30. Microphone directif Audio-Technica

Ce produit a la particularité d'offrir une directivité inférieure à celle des microphones du type canon. Ces caractéristiques attestent d'une atténuation maximum arrière de 80 dB. Par contre sa directivité pointue n'est pas constante en fréquence et son angle d'ouverture à -3 dB varie du simple au triple entre 2 kHz et 4 kHz. Sa fréquence maximale n'est que de 12 kHz ce qui pourrait être préjudiciable à une application musicale.

Sa directivité trop efficace ne nous est pas utile mais elle pourrait l'être dans le cas d'une reproduction multicanale composée d'une dizaine d'enceintes.

4.2.4 Traitement de la pression à la puissance deux

Une élévation au carré des tensions issues des microphones conduit à une directivité importante mais elle ne respecte plus la dynamique du signal. Ce traitement est similaire à un extenseur de dynamique. On remarquera que ce traitement génère une directivité en cosinus carré. Des directivités ajustables ou plus modérées pourraient être réalisées en fonction du facteur d'expansion.

Si la compression de dynamique omniprésente sur tous les systèmes de diffusion ne semble pas perturber outre mesure l'auditeur, une expansion de la dynamique est peut être gênante sur le plan auditif mais elle est surtout un risque d'avoir des passages avec des niveaux d'écoute trop importants, qui obligeraient l'auditeur à de fréquents réajustements de volume. On pourrait éventuellement corriger cet effet par une compensation automatique de gain comportant une constante de temps lente appropriée. L'expansion de la dynamique résultante risque de noyer une partie du signal utile dans la zone de bruit des enregistreurs.

4.2.5 Les microphones du second ordre

On réalise un microphone du 2nd ordre avec deux microphones cardioïdes décalés sur un même axe. Les capsules pointent dans la même direction et la polarité de la capsule arrière est retranchée à celle de l'avant (figure 31). On réalise en fait un microphone à gradient de pression construit avec des cardioïdes. La courbe de directivité créée par ce traitement est un diagramme polaire en cosinus carré.

Commercialisé par RCA pour la télévision depuis 1956 [30][31][32][33], le microphone BK-10A unidirectionnel de second-ordre fut réalisé à l'aide de deux microphones à ruban. Cette technologie n'a pas évolué en raison de son spectre limité et de sa réponse en fréquence médio-cré.

Cette solution a été reprise par Electro-Voice ou par Sanken [55], JVC, Sony [36] ou encore réadaptée par Woszczyk [49].

Un procédé similaire introduit l'apport de baffle (non utilisable ici en raison de l'encombrement) pour la réalisation de la directivité d'ordre 2 [44].

Avec P1, la pression recueillie au microphone avant et P2 celle du microphone retranché, A et B les coefficients de pondération associés à chaque capsule et θ l'angle d'incidence du son référencée par rapport à l'axe des deux microphones, on obtient :

$$P1 = A \frac{(1 + \cos \theta)}{2} e^{j\omega t}$$

$$P2 = B \frac{(1 + \cos \theta)}{2} e^{j(\omega t - kD \cos \theta)}$$

$$P = P1 - P2 = A \frac{(1 + \cos \theta)}{2} e^{j\omega t} - B \frac{(1 + \cos \theta)}{2} e^{j(\omega t - kD \cos \theta)}$$

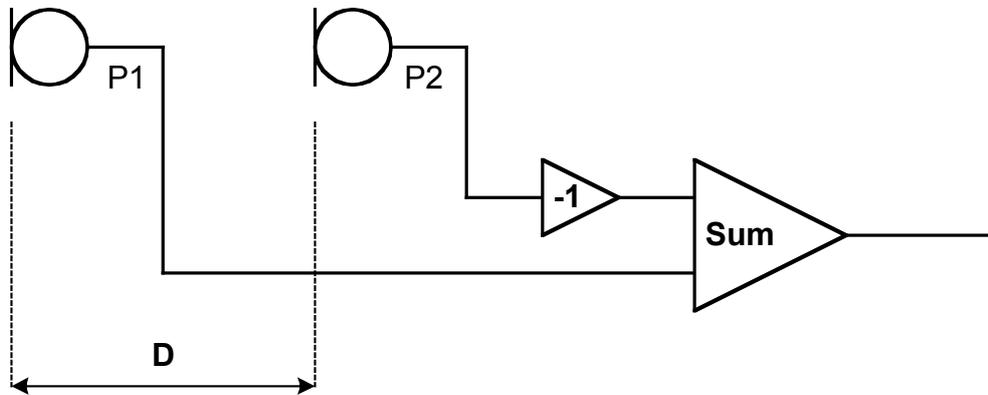


figure 31. Réalisation d'un microphone du 2nd ordre avec deux cardioïdes

Si $A=B$ on a :

$$P = A \frac{(1 + \cos \theta)}{2} e^{-jkD \cos \theta}$$

pour une distance D inférieure à 5cm et aux basses fréquences ($\lambda \gg D$) on a :

$$e^{-jkD \cos \theta} \cong kD \cos \theta$$

on a donc une directivité de la forme :

$$h(\theta) = M(1 + \cos \theta) \cos \theta \quad \text{avec} \quad M = \frac{AkD}{2}$$

Le calcul du facteur de directivité Q et de l'indice de directivité du microphone DI du second ordre donne les résultats suivants [20]:

$$Q = \frac{1}{2} \int_0^\pi h^2(\theta) \sin(\theta) d\theta$$

$$h(\theta) = \cos(\theta) + \cos^2(\theta)$$

$h(\theta)$ étant la fonction de directivité en pression on a après calcul :

$$Q = 7.5$$

$$DI = 10 \text{Log}(Q) = 8.75$$

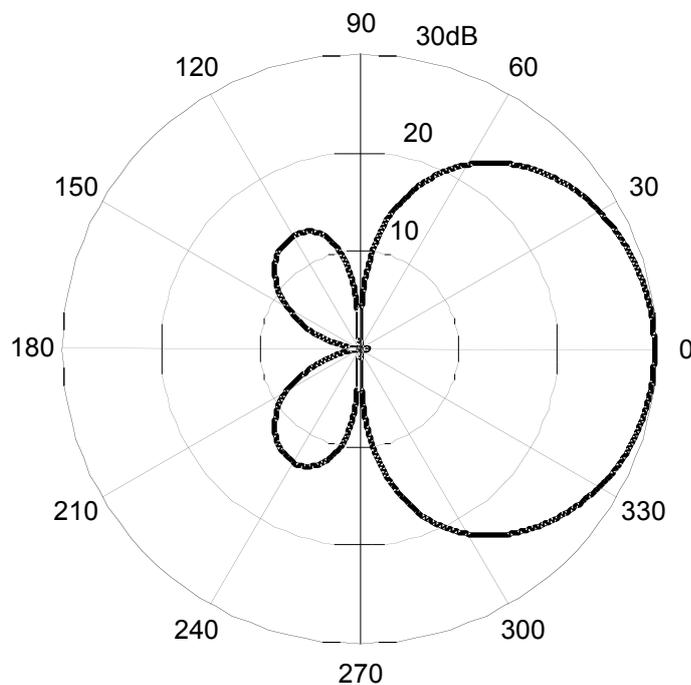


figure 32. Diagramme polaire du microphone du second ordre pour $f=100\text{Hz}$ et $D=5\text{cm}$

On observe sur le diagramme représentant la fonction de directivité de la figure 32 une atténuation maximale pour les incidences de $\pm 90^\circ$ et 180° . Deux lobes arrières sont générés dont le niveau reste inférieur à 15 dB par rapport à l'axe.

Si la directivité reste constante en fréquence pour $\lambda \gg D$, la réponse en fréquence n'est pas linéaire et varie comme un filtrage en peigne dont la première bosse dépend de la distance inter microphones et qui réagit pour les fréquences basses comme un filtre passe haut du 1^{er} ordre (figure 33). Il est évident qu'un microphone ayant une telle courbe est inutilisable.

La fréquence de la première bosse est relative à la distance inter capsule $D=\lambda/2$. On a donc pour une distance D de 5cm, un maximum d'amplitude à la fréquence f de:

$$f = c/\lambda$$

$$f = c/(2*D)$$

$$f = 340/(2*0.05)=3400\text{Hz}$$

La partie utile de la courbe se situe avant cette fréquence. Un espacement inter capsule de 8.5 mm est nécessaire pour obtenir une réponse en fréquence s'étalant jusqu'à 20kHz. On pourra augmenter cette distance si l'on s'autorise à travailler sur la partie décroissante de la courbe juste après le premier maximum d'amplitude. Dans cette zone, la directivité reste proche du 2nd ordre et on compensera la perte d'amplitude par une égalisation. Ce choix permettra l'utilisation d'une distance D de l'ordre de 1.5cm.

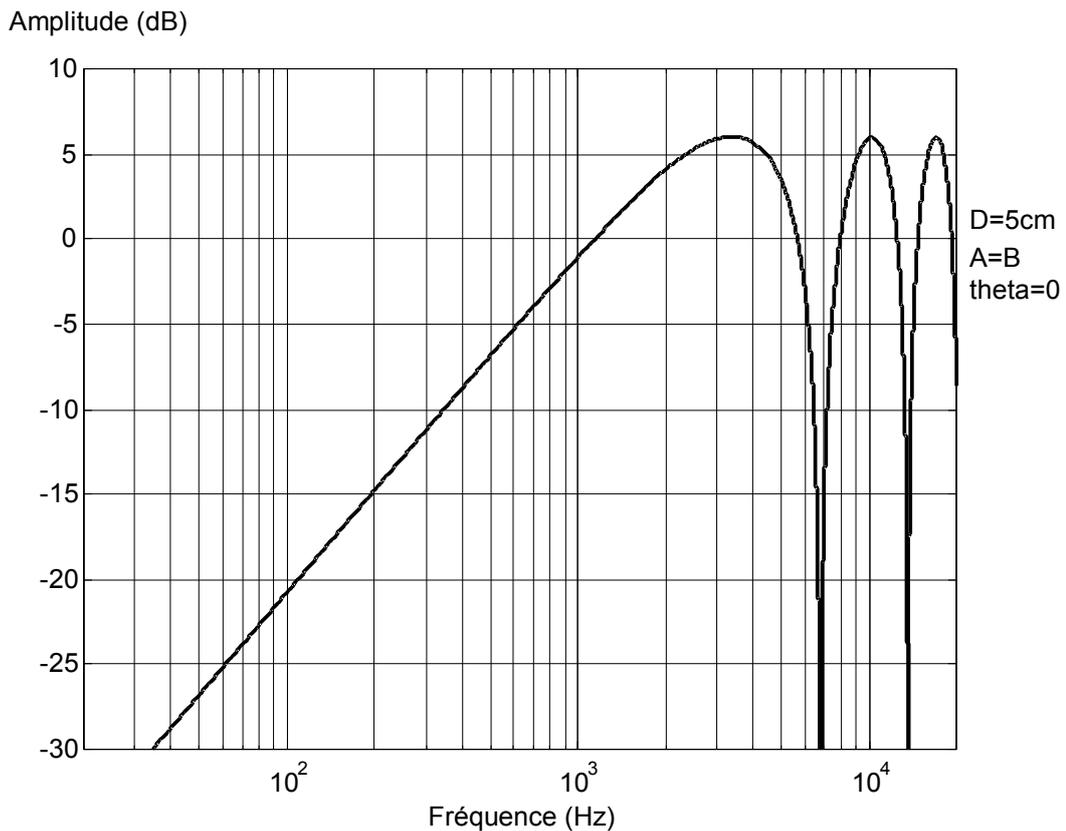


figure 33. Réponse en fréquence dans l'axe avec un espacement D de 5cm

Les fréquences graves nécessitent une égalisation importante pour linéariser la courbe. Pour une utilisation en champ proche, l'effet de proximité offre une compensation naturelle corrigeant le manque de grave [49]. Pour notre application où la distance d'enregistrement est en règle générale supérieur à 1m, nous ne pourrons pas utiliser cette correction. Une égalisation classique est donc requise mais elle est trop importante et aura pour conséquence de ramener à un niveau audible le bruit intrinsèque du microphone.

On s'aperçoit figure 34, qu'en pondérant d'un facteur deux le microphone retranché au principal, la courbe de réponse ne décroît plus continuellement avec une pente de 6 dB par octave mais qu'elle se stabilise sur un palier. Une zone plane est obtenue en basse fréquence car le microphone avant devient prépondérant. La courbe polaire dans cette zone se rapproche de la forme cardioïde. La restitution des basses fréquences à un niveau correct est réalisable si l'on s'accommode de cette directivité plus grande pour cette zone de fréquence. Son influence est réduite en raison de la mauvaise localisation perceptive des très basses fréquences et de la transition progressive du mode cardioïde au mode du 2nd ordre.

Pour la simulation de la figure 34, une pondération de 0.5 (-6dB) est faite sur la capsule arrière. Le niveau maximum sera donc de : $20\text{Log}(1+0.5) = 3.5\text{dB}$ et le minimum de $20\text{Log}(1-0.5) = -6\text{dB}$. Avec une pondération de 0.707 (-3dB) on aura respectivement $20\text{Log}(1+0.7) = 4.6\text{dB}$ pour le niveau max. et $20\text{Log}(1-0.7) = -10.5\text{dB}$ pour le niveau min.

Le problème en hautes fréquences décrit pour la configuration comportant des pondérations équivalentes est resté identique hormis le fait que les atténuations de la courbe sont moins prononcées. Si on désire couvrir le spectre de 100 Hz à 16 kHz, l'utilisation d'un double réseau s'impose.

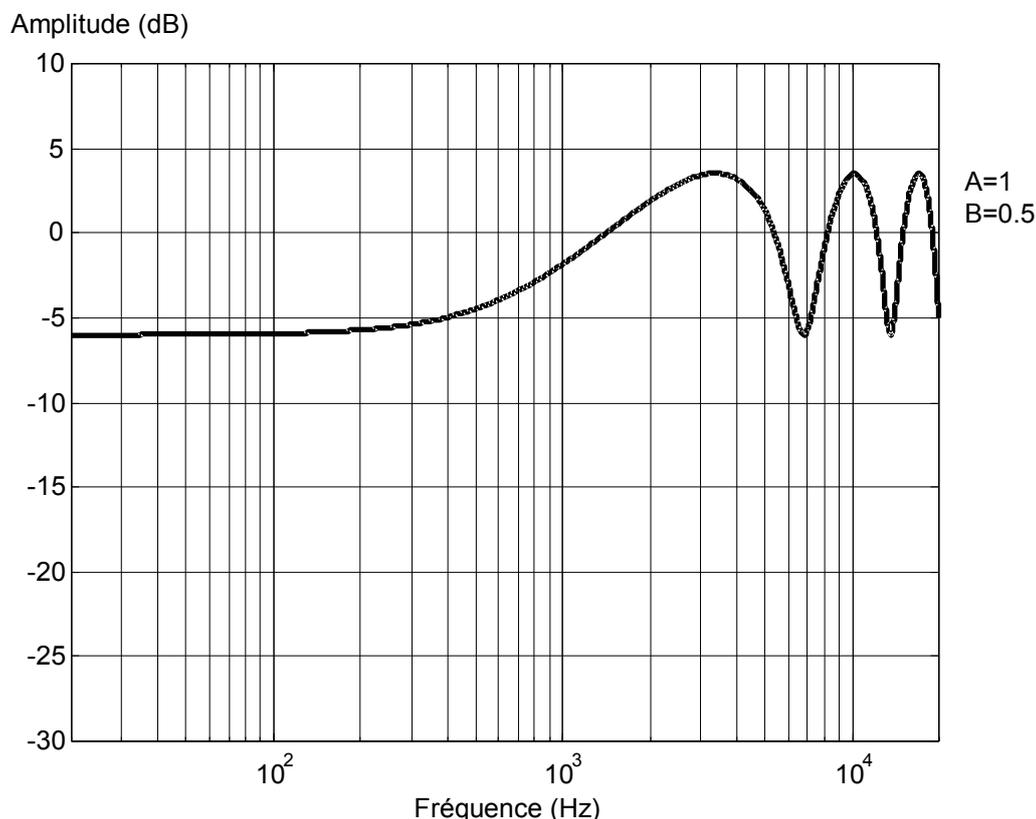


figure 34. Réponse en fréquence avec -6dB sur le microphone arrière

4.3 La technologie retenue

Aucune des solutions n'est réellement satisfaisante : soit la technologie utilisée impose une directivité qui n'est pas constante sur le spectre utile, soit elle impose des ajustements difficilement reproductibles, soit elle est trop complexe à mettre en œuvre.

La technologie des microphones du second ordre reste la plus intéressante car son lobe arrière est faible, sa directivité semble assez régulière en fonction de la fréquence, de plus, la directivité recherchée pour l'avant est similaire à cette captation. Malheureusement, la bande passante est très faible et le niveau de correction des basses fréquences important. Ces deux dernières caractéristiques devraient être atténuées en sommant deux systèmes du 2nd ordre où chacun des systèmes servira à la captation d'une zone fréquentielle particulière.

La section suivante traite de la mise au point de ce système.

5 Mise au point pratique de la technologie retenue

Cette étape concerne la mesure et les réglages pour parvenir à un microphone du second ordre large bande.

Des microphones miniatures auraient été plus souhaitables mais aucun ne convenait pour l'expérience. Les directivités des électrets miniatures ne sont pas assez constantes en fréquence et on peut craindre des caractéristiques divergentes en fonction de la série de fabrication. Pour obtenir un microphone du second ordre fonctionnant jusqu'à 20kHz, l'écartement nécessaire entre les deux capsules est de l'ordre de 8 à 15mm (cf. 4.2.5). Il est donc impossible de placer les différents microphones sur le même axe. Les capsules équipées de câble actif que l'on trouve chez les différents constructeurs sont encore trop longues.

Les microphones devront donc être décalés de leur axe dans le plan vertical afin de ne pas influencer le diagramme polaire en azimut. Nous observerons l'influence de ce montage vertical pour la directivité en élévation.

Les mesures sont réalisées à l'aide d'un logiciel de mesure électroacoustique MLSSA équipé d'une table tournante Outline. L'acquisition du signal se fait à l'aide d'un microphone de mesure B&K 4133. La représentation graphique est traitée par le logiciel Matlab (cf. Annexe 10.5) car la représentation en 'Waterfall' de MLSSA est difficilement exploitable. Pour obtenir des informations de directivité sur tout le spectre utile, une représentation tridimensionnelle est préférée à une représentation polaire qui ne donne pas les réponses en tiers d'octave et qui privilégie certaines fréquences si on ne fait pas de moyenne. La représentation sous Matlab utilisée génère des courbes isophoniques par pallier de 3 dB d'intervalle avec la fréquence en abscisse et la position angulaire en ordonnée. Les graphiques sont normalisés par rapport à l'axe pour s'affranchir des variations de la réponse en fréquence axiale et ne visualiser que les écarts d'amplitude relatifs à l'angle d'incidence.

5.1 Sélection du microphone cardioïde

Pour que la réalisation d'un gradient de pression du 2nd ordre (voir chapitre 4.2.5) puisse correspondre à nos exigences, il est nécessaire de choisir des microphones cardioïdes dont les caractéristiques soient les plus proches des courbes théoriques. Pour déterminer quels sont les microphones les plus performants, des mesures comparatives entre les microphones Neumann KM84, Schoeps MK4 et DPA 4011 sont réalisées (figure 35 figure 36 et figure 37).

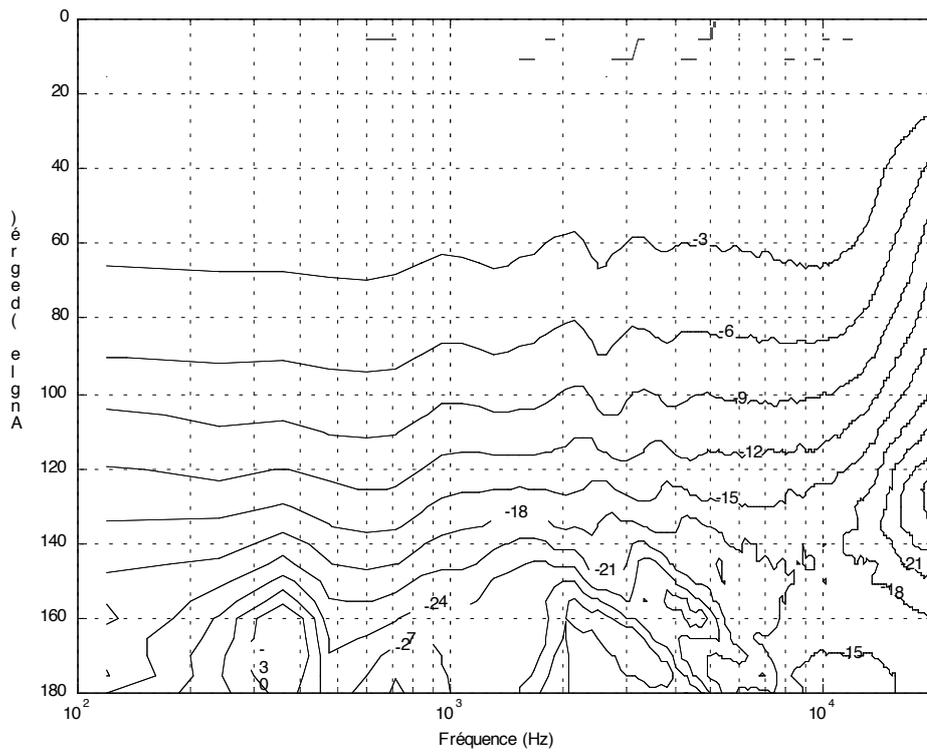


figure 35. Directivité du microphone DPA 4011

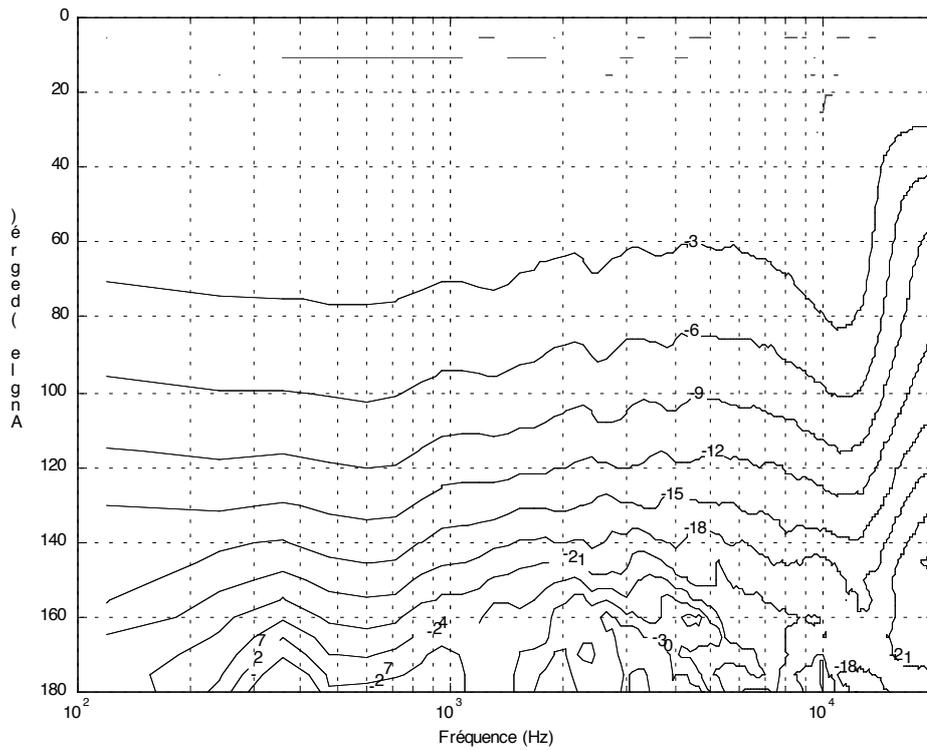


figure 36. Directivité du microphone Schoeps MK4

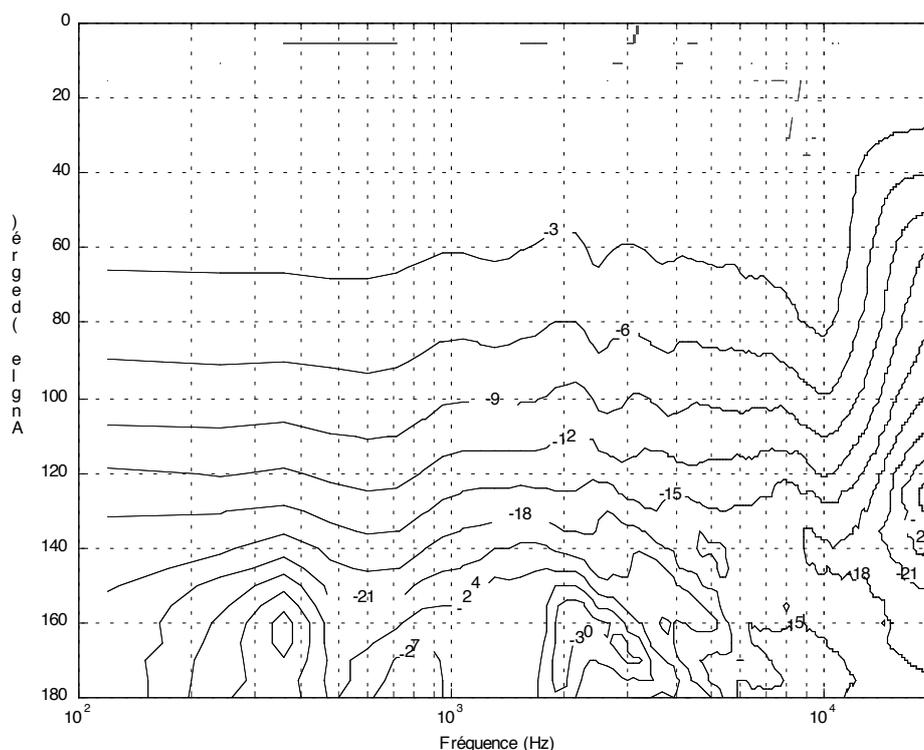


figure 37. Directivité du microphone Neumann KM84

Pour une directivité cardioïde, l'atténuation de 6db correspond en théorie à un angle de 90° et à un peu plus de 65° pour le demi angle de captation à -3dB.

Le microphone DPA 4011 (figure 35) est assez linéaire en fréquence sauf pour l'aigu. Une atténuation proche des 30 dB pour l'arrière est aussi le signe d'un bon comportement d'une capsule cardioïde. Sa directivité est constante et les angles à -3 et -6 dB sont proches de la théorie.

Le microphone Schoeps (figure 36) possède aussi une bonne atténuation arrière mais il est moins linéaire en fréquence que le DPA 4011 (un peu moins de 20° de divergence avant la fréquence de 10kHz pour une atténuation de -3dB). On observe aussi un comportement oscillant (sorte de coude) que l'on ne constatait pas avec le DPA à partir de 10kHz.

Assez linéaire en fréquence, le KM84 (figure 37), présente le même défaut que le microphone Schoeps à partir de 10 kHz et le resserrement de la directivité s'amorce plus tôt.

Ces résultats sont conformes aux données des constructeurs mais la représentation polaire fournie par les spécifications des fabricants ne permet pas la visualisation de certains comportements erratiques mentionnés ci-dessus. Ces trois microphones ont de très bonnes qualités. Les courbes du microphone DPA se révèlent légèrement meilleures. Le CNSM étant équipé de 6 microphones de ce type, la suite des essais se fera avec eux.

5.2 Influence de la pondération

Observons maintenant l'influence de la pondération à affecter à chaque microphone. Nous avons sur les figure 38 à la figure 40, le signal théorique obtenu par la différence de deux capsules espacées de 8cm avec des pondérations similaires, puis avec 1.5dB retranché à la capsule arrière et enfin avec -3 dB toujours sur la capsule arrière.

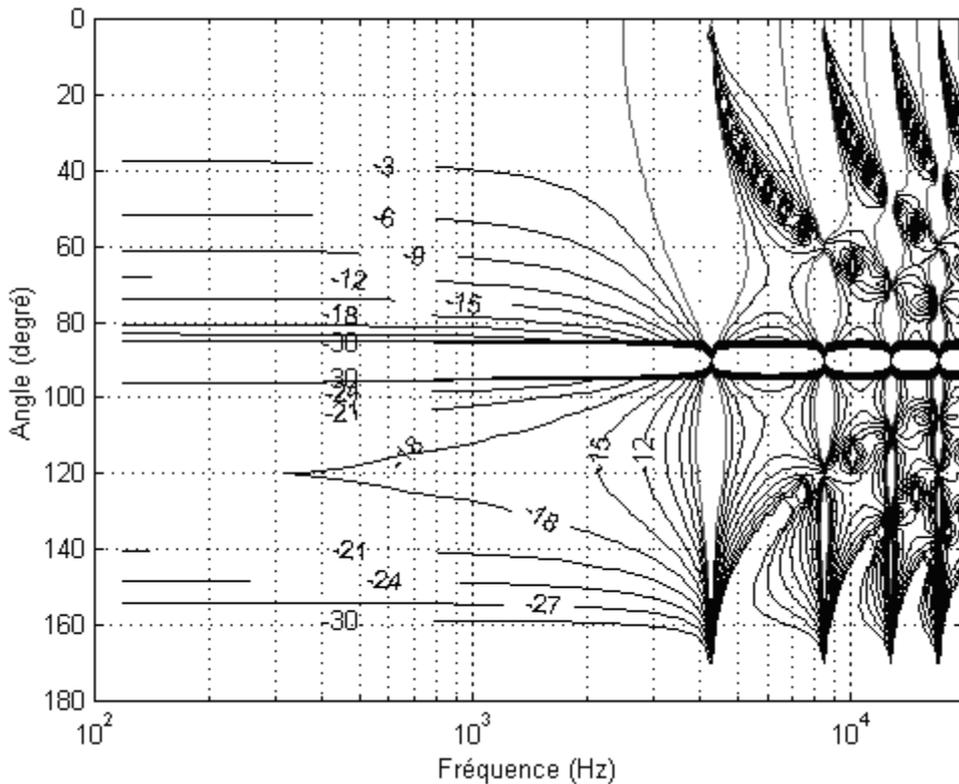


figure 38. Directivité de deux capsules espacées de 8 cm sans pondération

Avec une pondération identique sur les deux capsules, la directivité du second ordre (figure 38) est bien représentée pour les fréquences inférieures à 2000Hz, au-dessus on retrouve le filtrage en peigne dû au déphasage correspondant à la longueur d'onde (cf. figure 33). On trouve un demi-angle d'ouverture à -3 dB du capteur qui est inférieur à 40° . On remarque l'émergence du lobe arrière à 120° .

Si l'on observe la courbe (figure 39) dont la capsule arrière est atténuée de 3dB on voit que l'atténuation affectée à la seconde capsule tend à lisser la courbe résultante. La directivité de la zone bas médium et médium sur la décade 200 à 2000Hz reste grossièrement du 2nd ordre.

Dans la configuration avec -6 dB de pondération sur la capsule arrière (figure 40), le microphone simulé passe progressivement de cardioïde vers le second ordre (on tend vers une atténuation de 6dB à 90° dans les graves). Cette constatation s'explique par le fait que pour cette zone spectrale le microphone avant contribue essentiellement au niveau résultant. Les deux dernières configurations ont tendance à atténuer le lobe arrière du second ordre initial. Cette caractéristique peut être utile pour s'approcher de la directivité du test subjectif qui est exempt de lobes arrières. Le demi-angle d'ouverture à -3 dB du montage est supérieur à 40° .

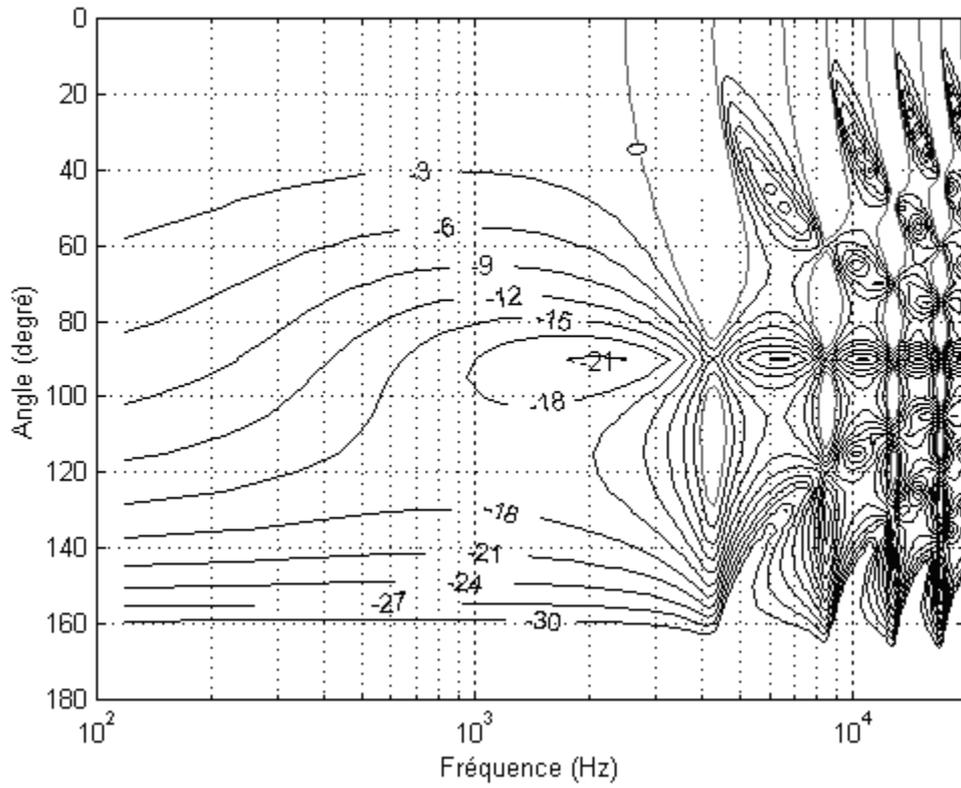


figure 39. Directivité de deux capsules espacées de 8 cm avec 3 dB d'atténuation sur la seconde capsule

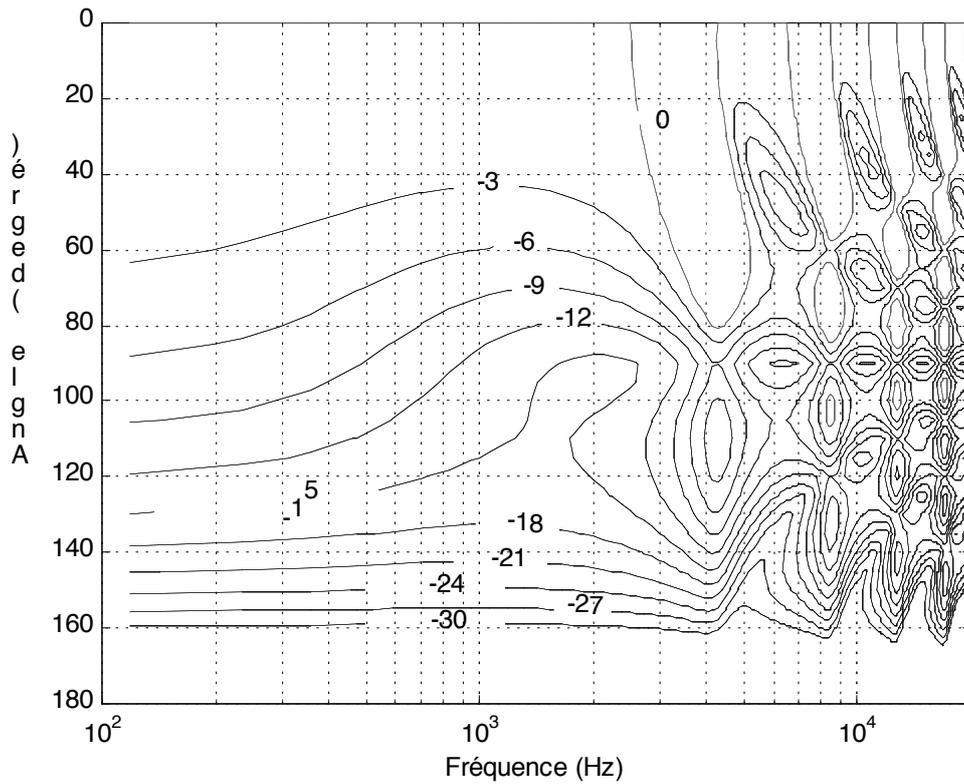


figure 40. Directivité de deux capsules espacées de 8 cm avec 6dB d'atténuation sur la seconde capsule

La conclusion de cette manipulation, c'est que nous avons à gérer un compromis entre une directivité du second ordre étendue sur le spectre et un niveau dans les graves respectable. Il serait superflu d'exiger une directivité marquée dans les basses fréquences sachant que celles-ci sont peu directives par nature. Comme on le constatera par la suite l'égalisation nécessaire au réajustement du grave ramène un souffle très gênant ! Dans un premier temps on s'orientera donc vers un montage comportant 3 dB de pondération sur la capsule arrière.

5.3 Sommation théorique de deux systèmes

L'addition de deux systèmes se chargeant respectivement d'une partie du spectre nous contraint à l'optimisation des phases respectives de chacun des systèmes. Si l'oreille ne semble que peu perturbée par des déphasages inférieurs à 45° [11], il en va tout autrement lorsque l'assemblage conduit à des oppositions de phases. De plus pour ne pas interférer dans les zones fréquentielles utiles, il est important de pouvoir filtrer chaque zone avec un filtre comportant une pente raide. Ce filtrage induit des déphasages par la même occasion car ce sont des filtres numériques non récurrents. On se retrouve en face d'une adéquation entre la fréquence de coupure des filtres, l'ordre des filtres, la distance inter capsules et la pondération à associer à chaque capsule.

La figure 41 montre le synoptique pratique servant à la génération des directivités du second ordre pour les deux zones fréquentielles.

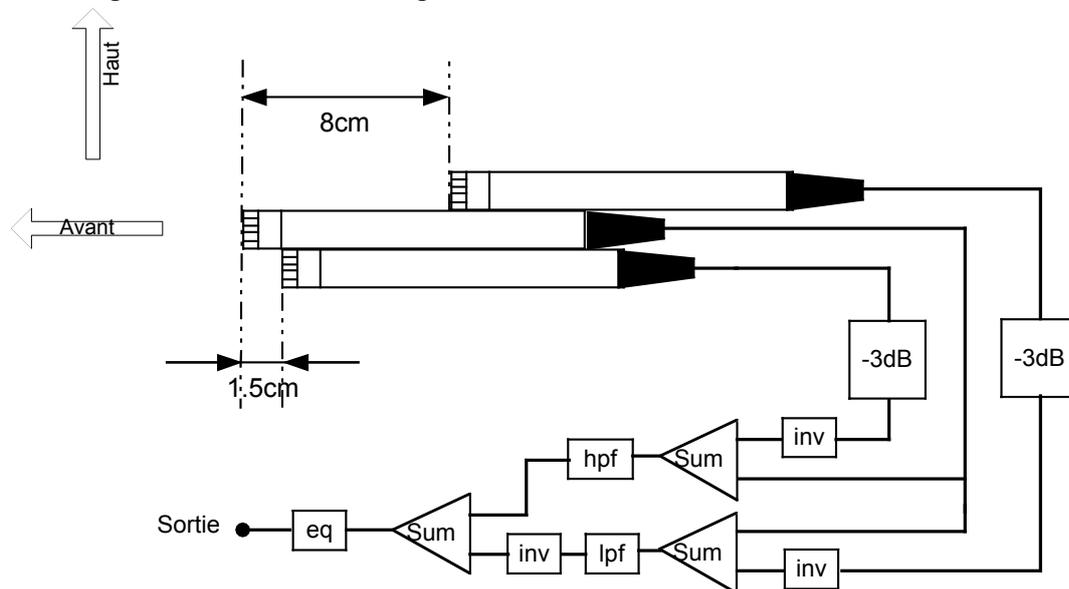


figure 41. Synoptique de la sommation de deux microphones du 2nd ordre.

Les courbes des figure 42 à la figure 44 sont les réponses en fréquence des deux systèmes et celle de la sommation. Le module est représenté en trait plein, la partie réelle avec des marques + et la partie imaginaire avec des marques o. Ces tracés ont été simulés sous Matlab (voir annexe 10.3). Ces simulations sont présentées pour une atténuation de -3 dB pour la capsule arrière, avec des filtres du deuxième ordre pour une fréquence de 1kHz et pour les distances intercapsules retenues de 8cm et 1.5cm.

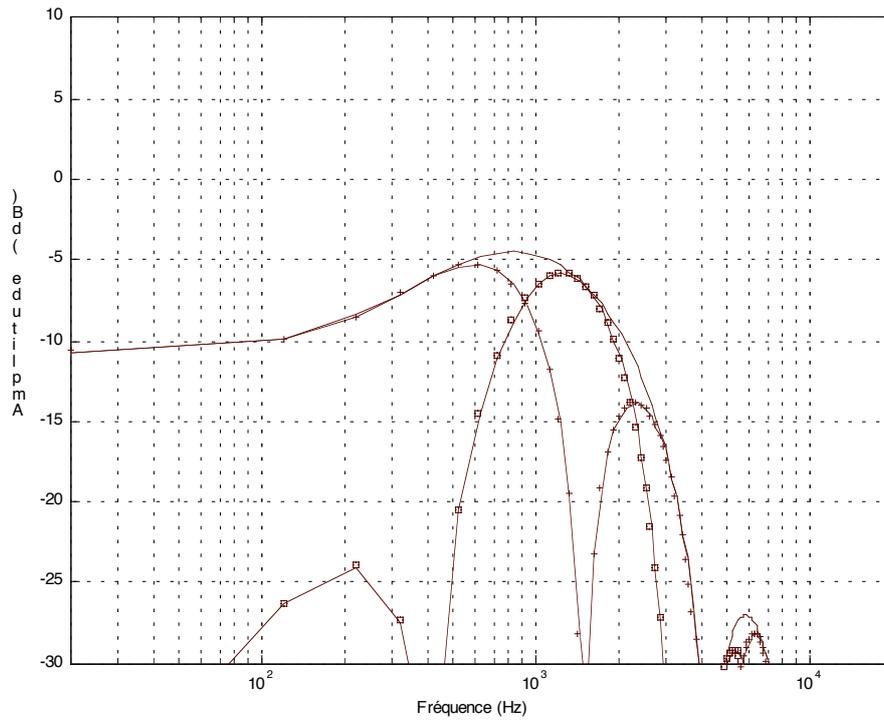


figure 42. Réponse en fréquence du couple de capsules espacées de 8cm.

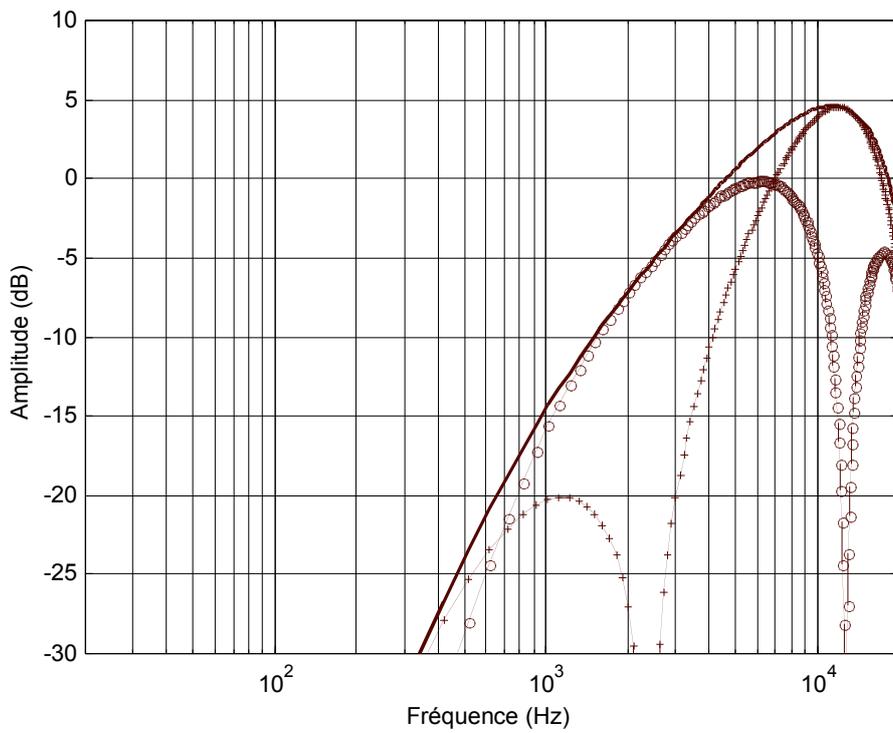


figure 43. Réponse en fréquence du couple de capsules espacées de 1.5cm.

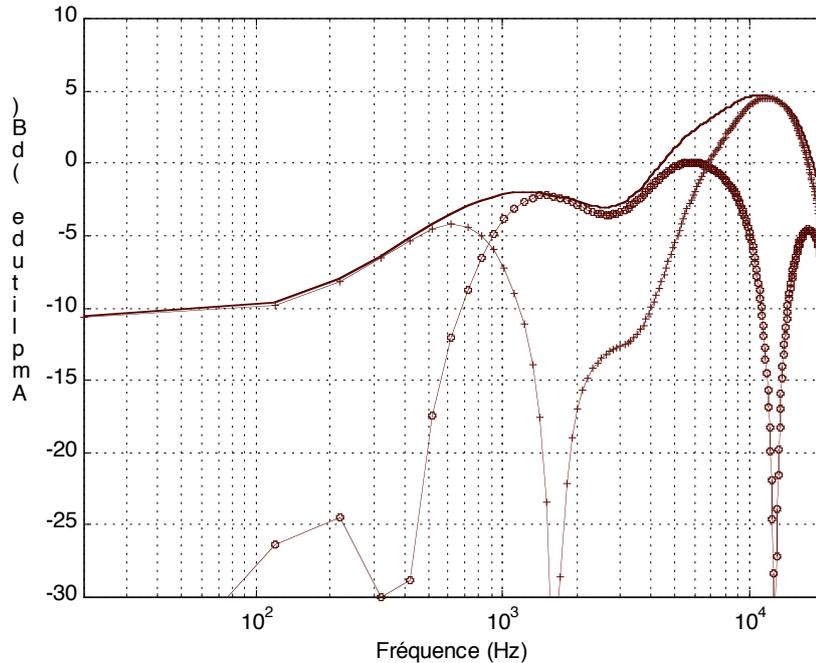


figure 44. Réponse en fréquence de la sommation des deux systèmes.

La figure 42 représente le système se chargeant des basses fréquences. Le filtre passe bas du second ordre de fréquence de coupure de 1000 Hz est activé, ce qui fait que l'on n'observe plus le filtrage en peigne comme sur la figure 34. Par contre on remarque que la phase varie énormément et que des risques d'inversion sont à craindre lors de la sommation.

La figure 43 montre le comportement du système chargé des hautes fréquences, le filtrage passe haut est aussi activé, ce qui fait que l'on observe une décroissance de 12 dB par octave.

La figure 44 montre que l'on peut arriver à une solution parmi d'autres ne présentant pas de défauts majeurs. Dans la configuration présentée, on observe que la sommation des deux montages avec un filtre du second ordre se réalise en créant une opposition de phase sur le second système (soustraction). Les filtres du premier ordre nécessitent quant à eux, une simple sommation. La courbe finale obtenue est le résultat d'une démarche empirique, les variables sont trop nombreuses pour que le modèle soit affiné par une procédure. Ces variables sont :

- les distances inter capsules,
- le gain des microphones arrières,
- le gain de chaque couple,
- les fréquences de coupure des filtres,
- l'ordre des filtres,
- les possibilités d'inversions de phase
- et le respect d'une fonction de directivité cohérente sur tout le spectre.

Chaque modification de variable nécessite une nouvelle simulation. On peut valider les valeurs à partir du moment où la courbe du module présente une allure sans accidents et que les valeurs du module sont supérieures à celles des courbes imaginaires et réelles.

La courbe du module global ne subit pas d'oppositions de phase, par contre, la phase nulle dans les graves passe à 90° dans les médiums pour revenir à zéro dans l'aigu. Il serait intéressant de connaître les nuisances qui pourraient être engendrées à l'écoute d'un son musical.

La réponse en fréquence du système complet nous informe sur le niveau d'égalisation à apporter pour obtenir une réponse en fréquence linéaire. Notons que les valeurs d'amplitude de la figure 44 correspondent à celles calculées précédemment au chapitre 4.2.5 pour une atténuation du microphone arrière de 3dB. Si on corrige par rapport au niveau relatif de 1kHz, les graves devront être augmentés au maximum de 8dB ce qui semble être raisonnable. Les aigus seront atténués de 7 dB au maximum. La sensibilité du microphone prototype restera donc quasiment inchangé (-2dB) par rapport à ceux des microphones DPA 4011.

Les figures suivantes concernent le comportement en directivité de l'assemblage des deux systèmes. Les variables sont les mêmes que celles précédemment utilisées (distance de 1.5 et 8 cm, atténuation arrière de 3dB et filtre du second ordre à 1kHz). La courbe théorique de la figure 45 est de bonne augure, la directivité s'approchant de celle désirée. Les filtres ont rejeté les zones fréquentielles indésirables de chaque doublet et la zone de transition autour de la fréquence de coupure est lisse et continue.

5.4 Sommutation réelle des deux systèmes

Tout ceci n'est qu'une vision du point de vue théorique et nécessite d'être appréhendé avec les défauts inhérents aux microphones. On trouvera les résultats simulés avec les réponses des microphones réels et ceux du prototype.

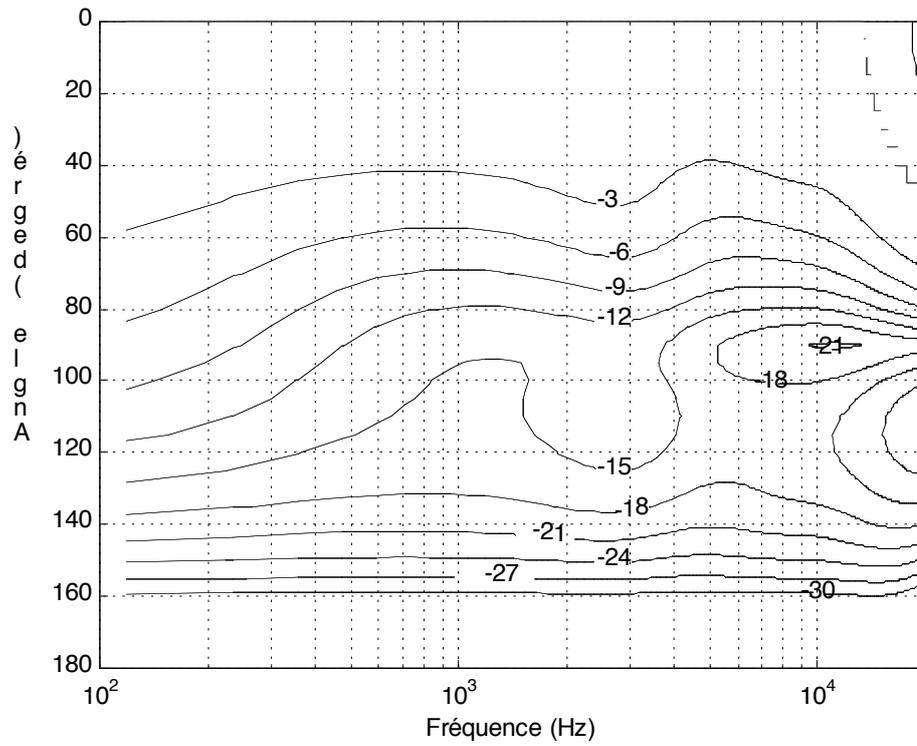


figure 45. Courbe de directivité du système complet théorique

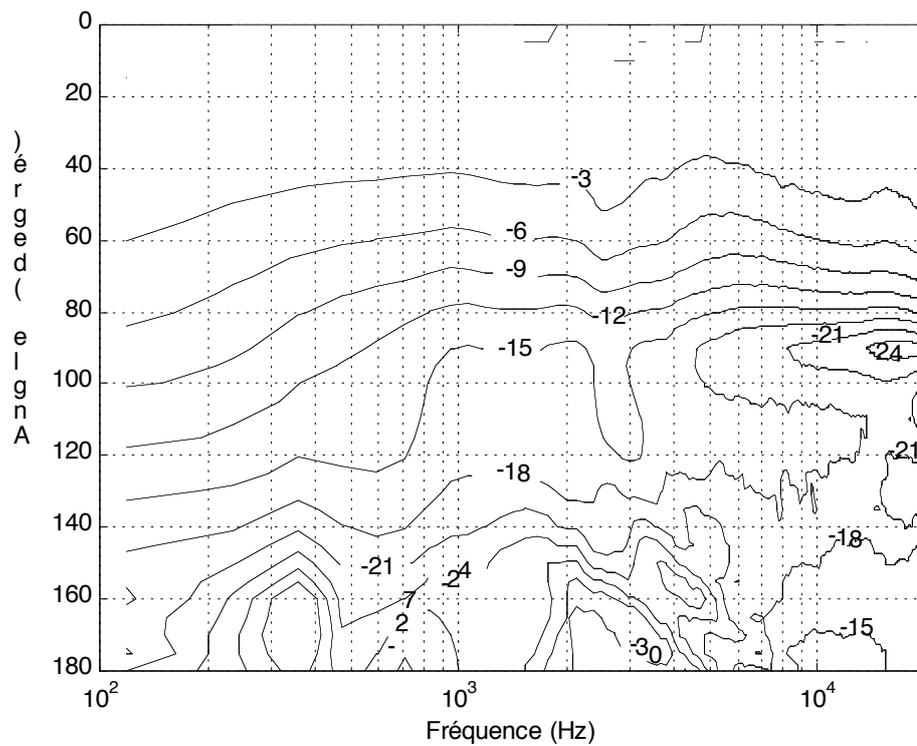


figure 46. Second Ordre simulé avec les courbes des micros DPA 4011 mesurées

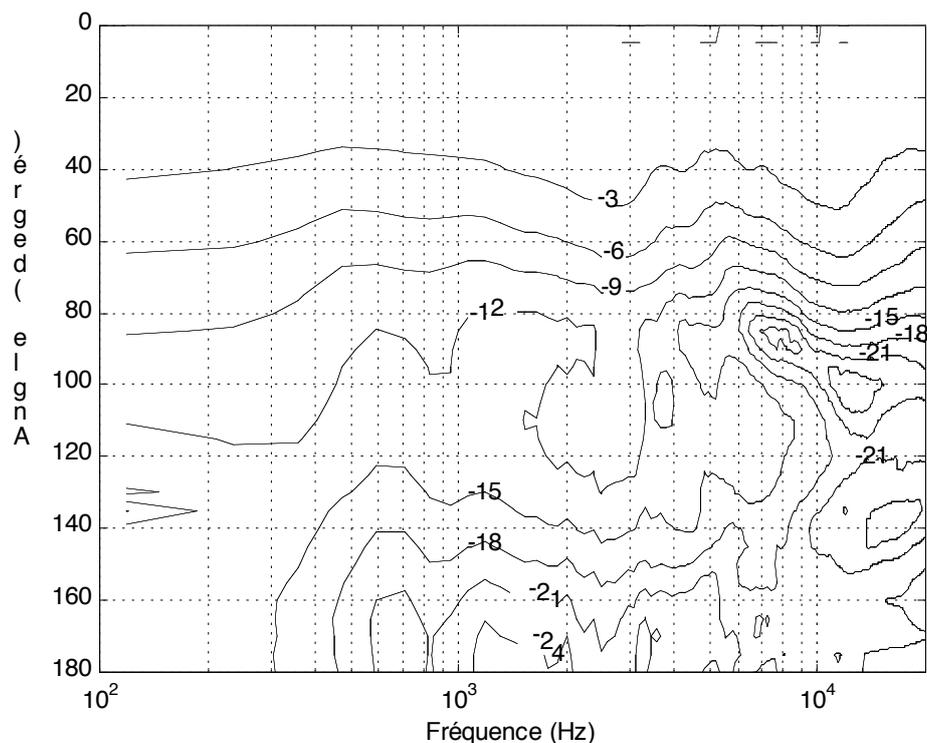


figure 47. Mesure de directivité du système réel

La figure 46 est une simulation faite avec les données réelles de la directivité des microphones cardioïdes choisis, la courbe a une allure sensiblement équivalente à la courbe purement théorique preuve de qualité des microphones DPA. On constatera avec surprise que les défauts en haute fréquences des microphones 4011 de la figure 35 sont compensés par ceux de la simulation dans cette même zone fréquentielle. La courbe se comporte beaucoup mieux pour la partie supérieure à 10 kHz ce qui justifie encore plus le choix des capsules 4011 par rapport à leurs concurrentes.

La figure 47 représente le résultat le plus important de cette section car elle correspond aux valeurs réelles mesurées du prototype. On constate que le prototype réagit globalement de la même manière que la simulation. Le demi-angle à -3 dB d'ouverture oscille autour de 45° . On en déduit que la simulation nécessaire à la validation du modèle est valable. Les angles d'ouverture sont respectés mais on a tout de même une plus grande directivité dans les graves. Les valeurs théoriques à 180° ne sont malheureusement pas atteintes. On devra se contenter d'une vingtaine de décibels d'atténuation en général. Les différences entre la théorie et la mesure proviennent essentiellement de la diffraction des corps des microphones placés en avant avec les capsules légèrement décentrées. Les différences obtenues dans le grave sont certainement issues d'erreurs de mesure car la résolution fréquentielle choisie est faible pour ne pas avoir des fichiers de valeurs trop long à traiter. De plus, les mesures faites en MLS au-dessous de 200 Hz sont tributaires des réflexions sonores du local de mesure.

La figure 48 représente les mêmes résultats sous la forme d'un diagramme en waterfall. Elle montre une décroissance régulière et similaire sur toute la gamme, elle est l'image d'une directivité cohérente. On en conclue que la réalisation par sommation d'un microphone de second ordre sur l'ensemble du spectre audio est réalisable.

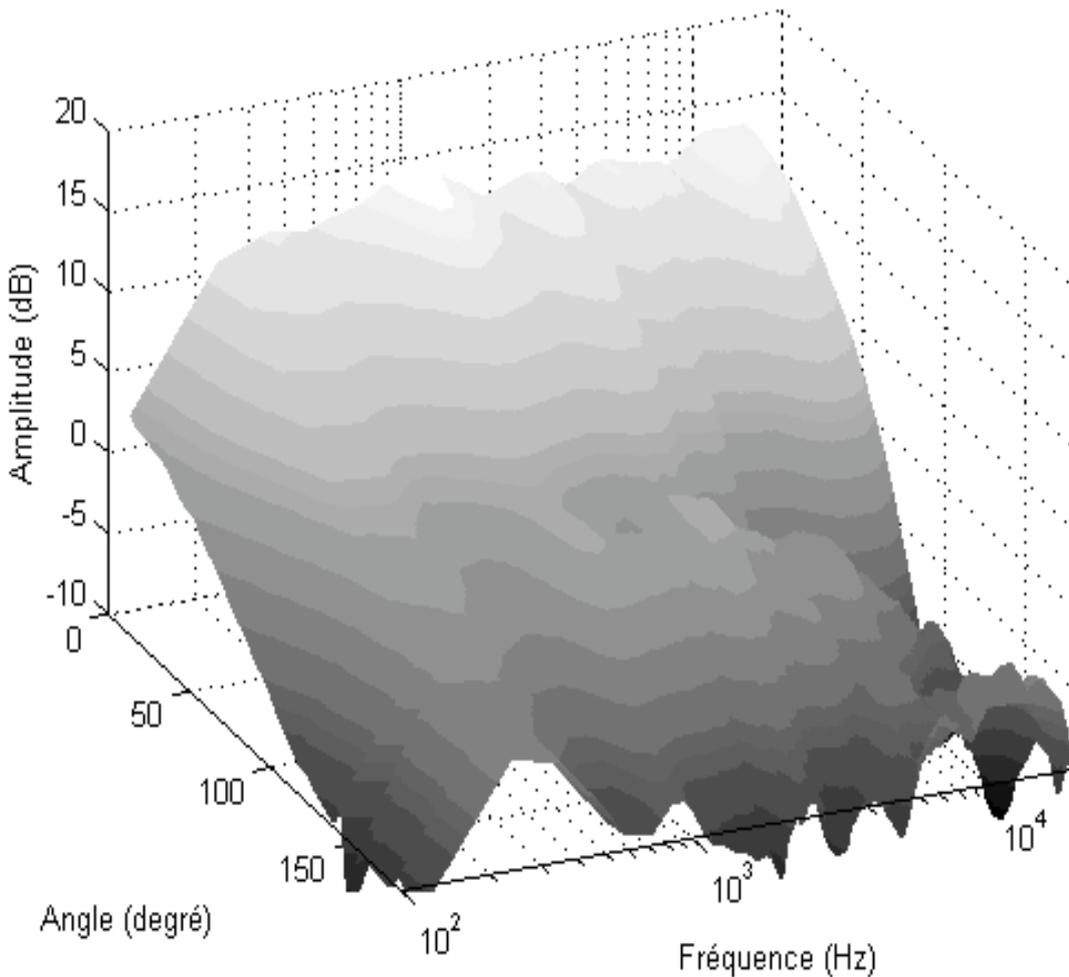


figure 48. Représentation en 'waterfall' du prototype non égalisée.

L'égalisation de la réponse en fréquence dans l'axe nécessite une console numérique pour une meilleure reproductibilité. L'utilisation d'une console analogique n'est pas envisageable car les graduations des potentiomètres sont trop approximatives et ces consoles ne possèdent pas pour la majorité d'entre elles de rappels de configuration pour la mémorisation des valeurs des niveaux et des filtres.

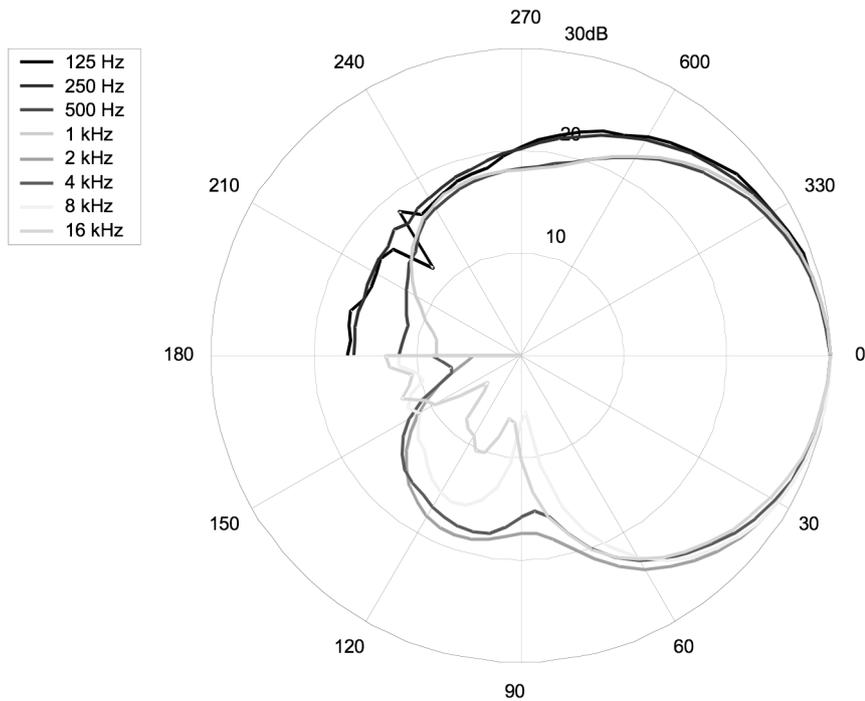
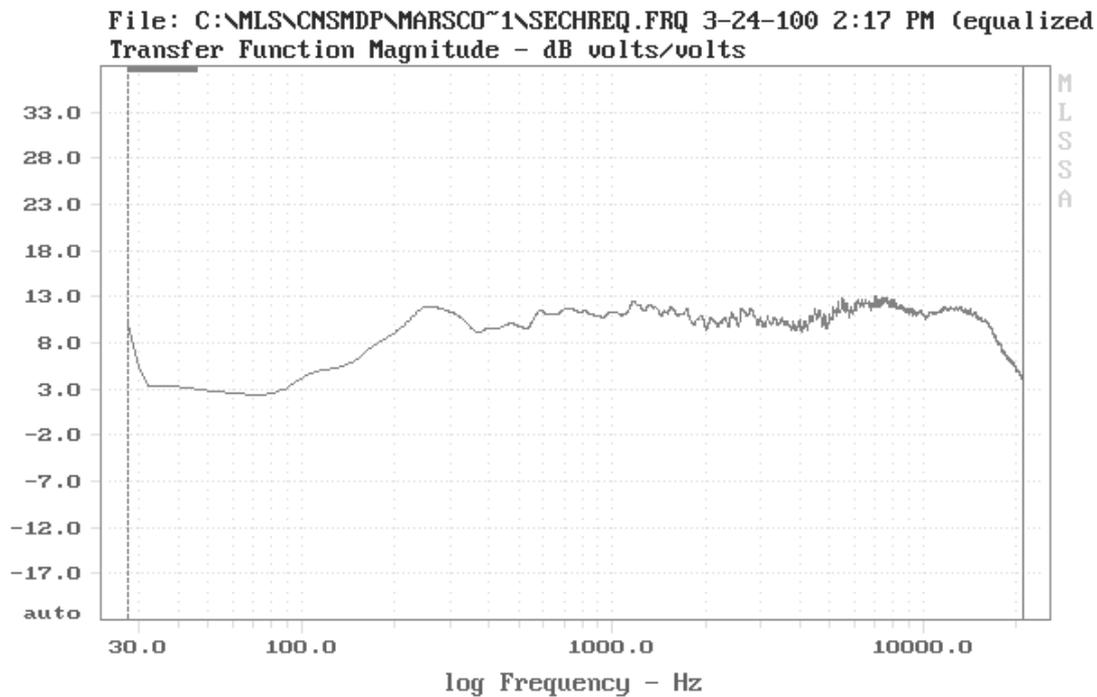


figure 49. Diagramme de directivité du prototype

La figure 49 est la représentation polaire du prototype pour chaque octave. Pour la zone supérieure représentant les courbes des fréquences basses, on constate que l'atténuation arrière n'est que de 14 dB jusqu'à 250 Hz. Pour les hautes fréquences, les lobes caractéristiques du 2nd ordre vers $\pm 120^\circ$ se sont légèrement renforcés par rapport au modèle théorique de la figure 32. Il est évident que ces défauts risquent d'interagir avec les microphones utilisés pour la captation des autres zones lors d'une prise de son multicanal et perturber la perception de la latéralisation



CURSOR: y = 4.63586 x = 21000.6993 (9118)

FREQUENCY DOMAIN MENU: Go View Reference Acquisition Setup Transfer Macro QC
Overlay Calculate Printer DOS Units Library Info Exit
F1 for Help MLSSA: Frequency Domain

figure 50. Réponse en fréquence du prototype dans l'axe après égalisation

La figure 50 est la réponse dans l'axe obtenue après égalisation, elle tient dans un gabarit de ± 2 dB entre 200Hz et 15kHz, ce qui est convenable pour un microphone de prise de son. Les basses fréquence auraient pu être rehaussées un peu plus mais le bruit de fond est déjà trop important dans cette zone fréquentielle.

Le traitement numérique et les réinsertions induisent des retards de traitement qu'il est nécessaire de compenser pour les voies arrières qui ne comportent qu'un seul microphone. Le retard mesuré par MLSSA en mode autocorrélation est de 3.9ms. On trouvera en annexe 10.7, l'égalisation utilisée pour la mise au point du prototype ainsi que le synoptique de câblage de la console Yamaha 03D.

5.5 Influence de la position désaxée du montage

Quel est le comportement directif de ce microphone dans l'axe vertical ? Si la position désaxée des microphones modifie peu la directivité sur le plan horizontal, il n'en est pas de même pour le plan vertical et cela peu modifier grandement le timbre du microphone sachant que le champ réverbéré ne sera pas capté de manière uniforme.

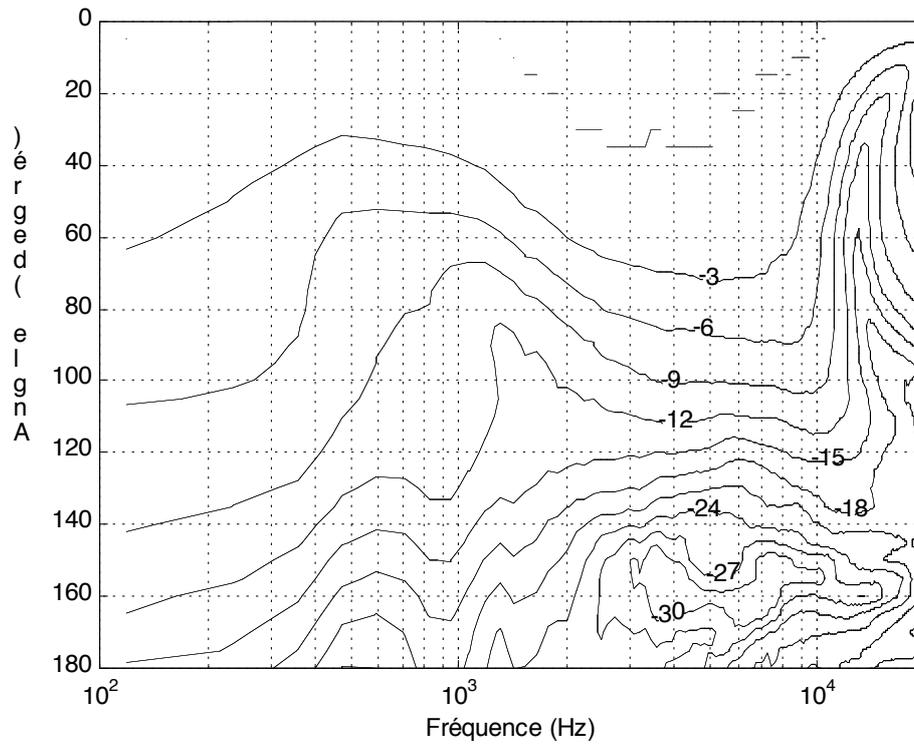


figure 51. Directivité du microphone mesurée dans le plan vertical en hauteur

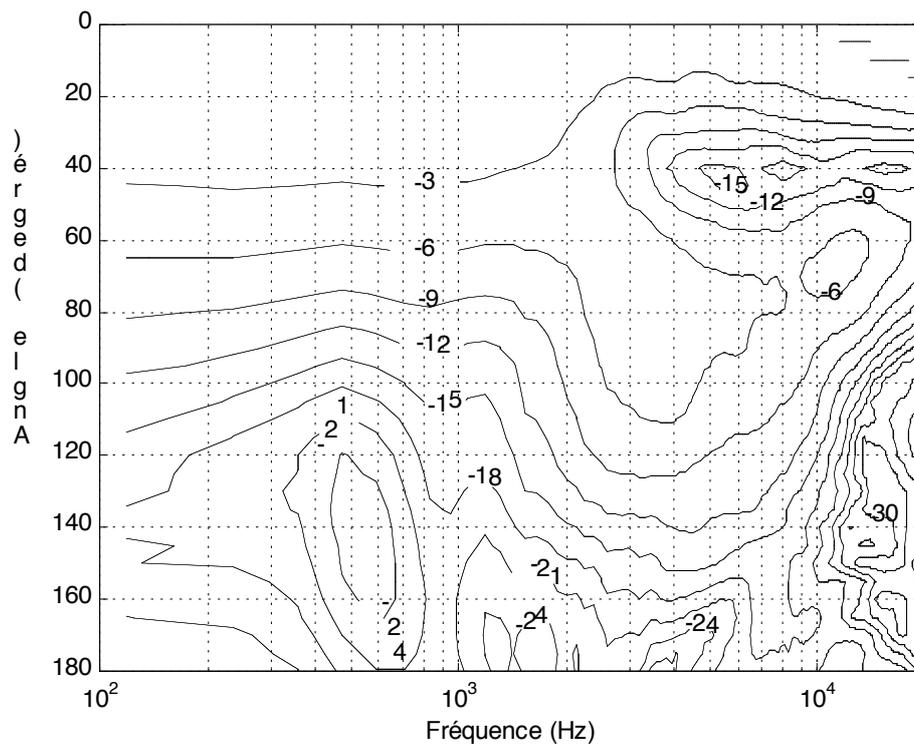


figure 52. Directivité du microphone mesurée dans le plan vertical vers le bas

La figure 51 et la figure 52 représentent le comportement de la directivité supérieure et inférieure du système. La directivité n'est pas constante en fréquence. Le demi-angle de captation à -3dB varie de plus de 40° en fonction de la fréquence. Cette captation inhomogène est synonyme d'une mauvaise qualité pour le rendu des premières réflexions. On ne pourra pas compenser ces défauts autrement qu'en positionnant les capsules sur un même axe. Il n'y a plus qu'à espérer que la latéralisation (cas où la source émet sur le même plan que les capteurs) ne sera pas perturbée par ces distorsions.

5.6 Diffraction générée par une prise de son coïncidente

Un autre dégradation du signal est générée par l'interaction réciproque des microphones pour une prise de son coïncidente (corps, capsules, pieds et genouillères de microphone). La proximité des capsules composant le système engendre des altérations de la réponse des microphones (voir la figure 53 et la photographie de la figure 57). Cette courbe est issue d'un assemblage de deux microphones du 2nd ordre et de deux du 1^{er}. Les microphones sont totalement coïncidents par rapport à l'axe vertical, c'est à dire qu'ils sont décalés en hauteur et que chaque microphone est en contact avec son voisin par le haut ou par le bas (cf. figure 57). Notons qu'un placement avec les quatre microphones en position tête bêche a donné des résultats beaucoup plus décevants.

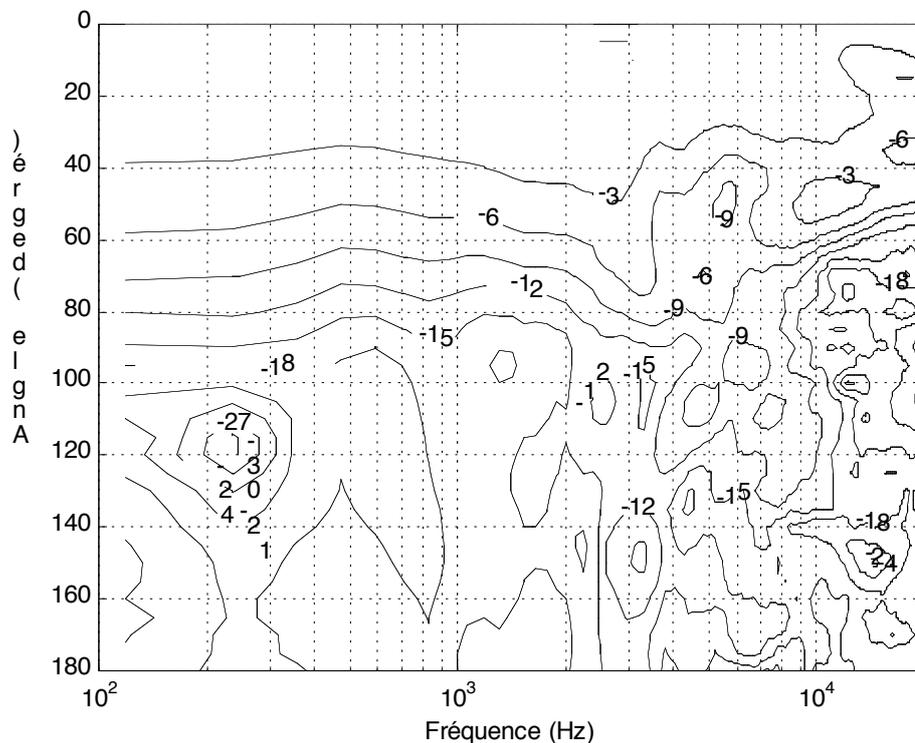


figure 53. Directivité du microphone mesurée en place dans une prise de son coïncidente

La directivité résultante de la disposition retenue tend à être légèrement supérieure à celle du prototype isolé (figure 47). Cette diffraction ne semble pas modifier exagérément la directivité résultante, elle crée de multiples irrégularités qui affecteront tout au plus la précision de la captation et la linéarité de la réponse en fréquence en fonction de l'incidence.

5.7 Synthèse de la technologie retenue

Le prototype réalisé possède un diagramme polaire correspondant à la perception avant. Cette directivité pourrait être éventuellement utilisée pour la zone arrière car elle est un peu plus large que celle de la simulation théorique.

Si la fonction de directivité est similaire à celle recherchée dans le plan horizontal, le modèle possède des défauts :

- L'écart des trois capsules sur l'axe vertical engendre une directivité verticale très médiocre.
- Des variations de phases en fonction de la fréquence issu du principe du second ordre vont probablement altérer la qualité des timbres.
- La diffraction induite par le placement des microphones pour la captation angulaire totale en différence d'intensité perturbe légèrement la directivité horizontale et probablement plus encore la directivité verticale.
- Le rapport signal/bruit est faible en basse fréquence.

De petites capsules cardioïdes d'excellente qualité permettraient de s'affranchir de deux de ses problèmes. La distance nécessaire aux deux premières capsules n'autorise pas l'usage des capsules miniatures ou à câble actif courants. Des électrets cardioïdes miniatures de haut de gamme imposeraient un système d'accroche sophistiqué qui n'existe pas actuellement. Les défauts de phase pourraient être corrigés par des filtres numériques agissant sur la phase uniquement. Des filtres passe haut et passe bas à phase linéaire pourraient être réalisés avec des filtres numériques récursifs. Le prototype accompagné d'un circuit de traitement DSP intégré améliorerait la qualité générale de la captation. Le DSP pourrait aussi se charger de la sommation et de l'ajustement des niveaux afin de fournir une modulation unique en sortie de l'assemblage microphonique.

La section suivante traite d'essais subjectifs avec le prototype retenu. Les microphones ont les mêmes caractéristiques que celles décrites ici.

6 Test du prototype

Trouver un test fiable et complet permettant de faire une critique exhaustive du système n'est pas réalisable. Une approche purement théorique n'a pas non plus de sens en raison de notre méconnaissance des artefacts mis en jeu, tant pour la captation que pour la reproduction. La comparaison subjective entre différents systèmes de prise de son est le plus sûr moyen de les qualifier, à condition de maîtriser les conséquences et distinctions physiques et psychoacoustiques de la perception ressentie. Cette démarche est longue et doit être extrêmement rigoureuse. Elle peut faire l'objet d'un mémoire à elle seule (cf. travail en cours au CNSM sur la comparaison des systèmes de prise de son en multicanal présentée au 108th congrès de l'AES de Paris). Le test du prototype qui nous intéresse devra donc s'accommoder d'une procédure plus modeste qui ne pourra en aucun cas, qualifier les performances de ce système en comparaison à un autre ni spécifier les performances brutes du système.

La démarche utilisée dans ce mémoire nous conduit à vérifier la capacité du système d'enregistrement à fournir une perception de localisation valide. Les conditions des tests initiaux peuvent être reproduites sans changements notoire. Elles permettront de valider les nouveaux résultats en raison des conditions d'écoutes identiques et des hypothèses du départ. Le prototype sera valable si les directions sonores captées par les microphones sont similaires à celles ressenties en moyenne par les auditeurs. Si elles sont semblables, on conclura que le système de prise de son est apte à restituer la localisation dans le contexte de la reproduction en 5.1.

De nombreux facteurs dissemblables entre le prototype et le modèle idéal risquent de faire dévier les résultats attendus :

- La directivité des capsules ne sont pas identiques à celles des valeurs obtenues par les écoutes initiales.
- La proximité des capsules entre elles induit des diffractions susceptibles d'altérer la précision de la localisation.
- La captation du champ réverbéré fait qu'un événement unique va être reproduit sur 4 canaux. Comment cette information supplémentaire sera gérée par notre cerveau ? Des effets de masques ne risquent-ils pas de se manifester et détruire la localisation désirée ? Quelle est l'information locative subjective ajoutée par une captation acoustique comparée à une monophonie dirigée (panning) ?
- Quelle est l'interaction de la sommation de deux réverbérations, celle captée à l'enregistrement et celle issue du local de reproduction qui, en l'occurrence sont les mêmes dans notre configuration ?
- La distance critique différente entre avant et arrière est-elle préjudiciable ?

Le test ne répondra pas à ces questions qui sont des sujets de recherche en eux-mêmes d'autant plus que ce n'est sûrement pas le paramètre perceptif de la localisation qui sera le plus altéré par ces phénomènes.

6.1 Choix du test

Le test doit valider la qualité de la localisation obtenue par la prise de son. Une prise de son utilisant la technologie retenue constitue la matière du test subjectif. Pour garder une continuité logique dans la démarche de ce travail, les principes du test antérieur seront repris en grande partie. L'enregistrement ainsi que les écoutes seront réalisés à la salle Schaeffer du C.N.S.M de Paris.

Les sons enregistrés par le prototype sont de deux sortes : des échantillons diffusés par une enceinte et des sons purement acoustiques. La chanson de Suzanne Vega est à nouveau l'échantillon retenu pour la diffusion sur haut-parleurs pour une meilleure compatibilité avec le test initial, tandis que des sons percussifs sont choisis pour la phase de l'enregistrement acoustique. Des sons d'impact ont l'avantage d'être bref et de ne pas avoir recours au service d'un musicien. Les sons générés par le woodblock et les claves ont une durée très courte et sont assimilables à des notes ne comportant que des transitoires. La sonorité du woodblock est plus grave que celles des claves. Les maracas sont le troisième instrument choisi et qui, malgré sa classification d'instrument à percussion est à cataloguer avec des sons longs en raison du type de l'excitation sonore (bruit de grenaille). Des essais ont été tentés avec un crotale, sorte de disque à son de cloche. Les sonorités sont impossibles à localiser malgré la résonance typique de ces instruments. Cette constatation peut s'expliquer par la nature inharmonique de sa structure spectrale.

L'enregistrement est simple, il suffit de décaler en azimut la source sonore autour d'un cercle dont le centre sont les microphones. Des enregistrements à différentes distances du centre permettent de choisir dans une seconde étape, les échantillons à sélectionner (figure 55).

Le dispositif utilise deux microphones du 2nd ordre pour les canaux des enceintes frontales et deux microphones cardioïdes pour les enceintes arrières car aucune solution améliorée n'a pu être trouvée. Des microphones du 2nd ordre pour l'arrière nécessitent des ressources en microphones et de console non disponibles. La configuration théorique correspondante est celle de la figure 54. Un second test comparatif, n'utilisant que des microphones cardioïdes est mis en place afin de juger la pertinence de l'emploi des microphones du second ordre pour la captation avant. Par commodité mnémotechnique, le système 1 sera l'enregistrement du 1^{er} ordre et le système 2 celui du 2nd ordre.

Les échantillons sont plus nombreux que pour le test initial, la seconde écoute de confirmation est supprimée à regret pour une durée de test approximative de 25 minutes.

Le texte de préambule en annexe 10.1 sera conservé sauf pour les passages se référant à la seconde écoutes de confirmation.

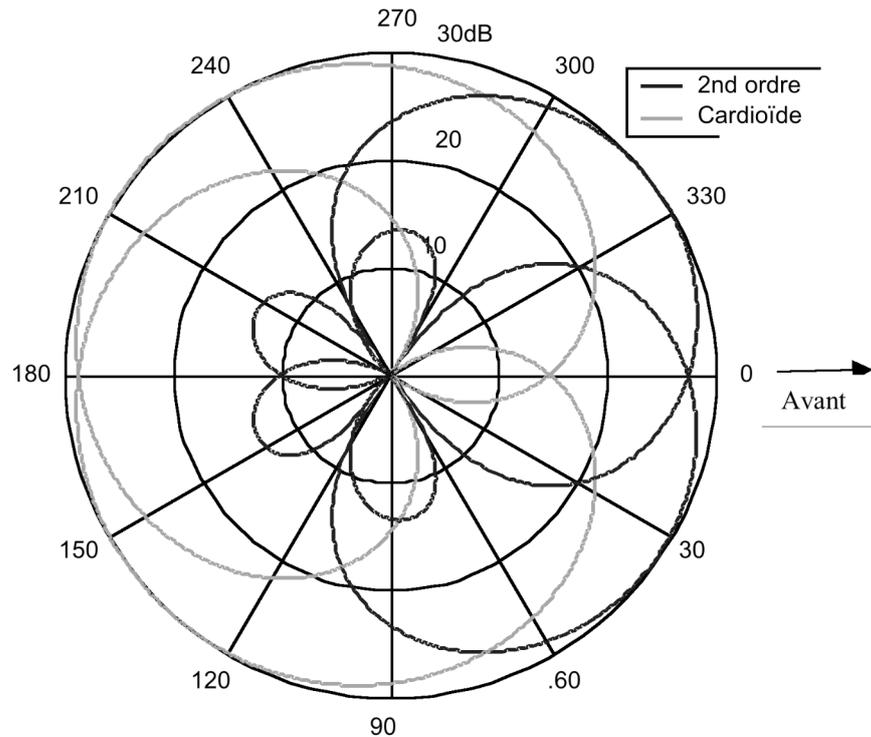


figure 54. Diagramme polaire théorique du système 1

6.2 Enregistrement du test

Des cercles concentriques matérialisés au sol de la salle Schaeffer représentent les différentes distances d'enregistrement allant de 1,5m à 3m (figure 55). Les angles physiques des microphones sont de $\pm 45^\circ$ et de $\pm 120^\circ$ pour les deux systèmes à comparer (cf. figure 56). Ces valeurs correspondent à celles des axes des courbes théoriques superposées aux courbes de test de la figure 19 à la figure 22.

Le système 1, qui est composé de 4 microphones Schoeps MK4 cardioïdes, est placé au-dessus du premier système de façon à faire correspondre le milieu émissif des enceintes au centre des deux configurations (figure 57). Une seconde table de mixage de marque Sonosax sert à la calibration et à la distribution du système 1 vers l'enregistreur Tascam 8 pistes. La calibration des niveaux est réalisée suivant la procédure décrite au chapitre 3.2.

La console de mixage n'autorise qu'un contrôle au casque de l'enregistrement. En effet l'utilisation complète de ces ressources pour la phase de la prise de son ne nous permet pas d'accéder à un monitoring général de reproduction sur les quatre canaux. Le contrôle du rendu réel n'a pu se faire qu'en différé dans un autre studio.



figure 55. Enregistrement du test

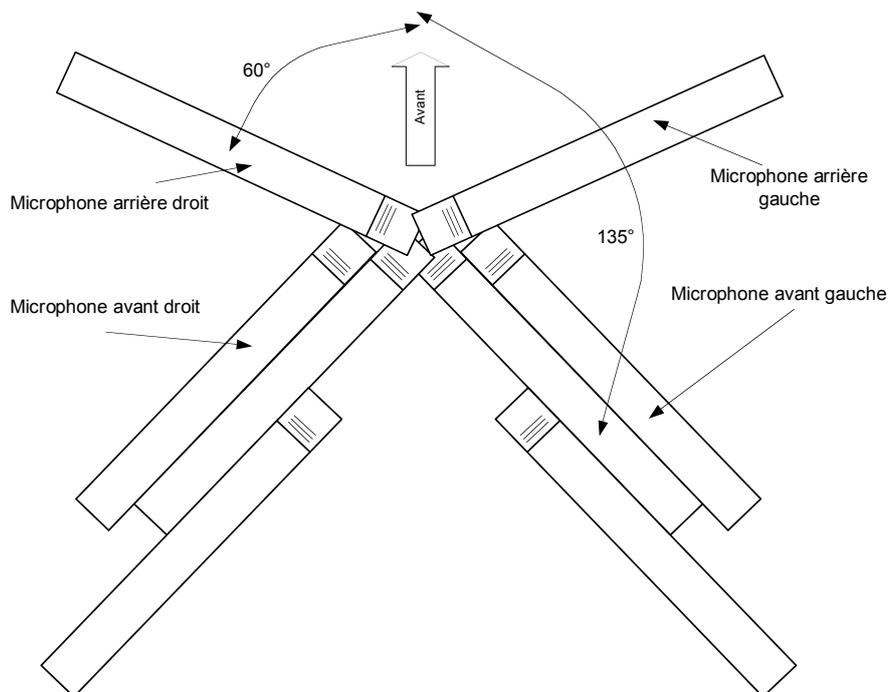


figure 56. Dispositif d'enregistrement avec le système 2

Prise de son	N° du système
4 cardioïdes	1
2 prototypes du 2 nd ordre + 2 cardioïdes arrières	2

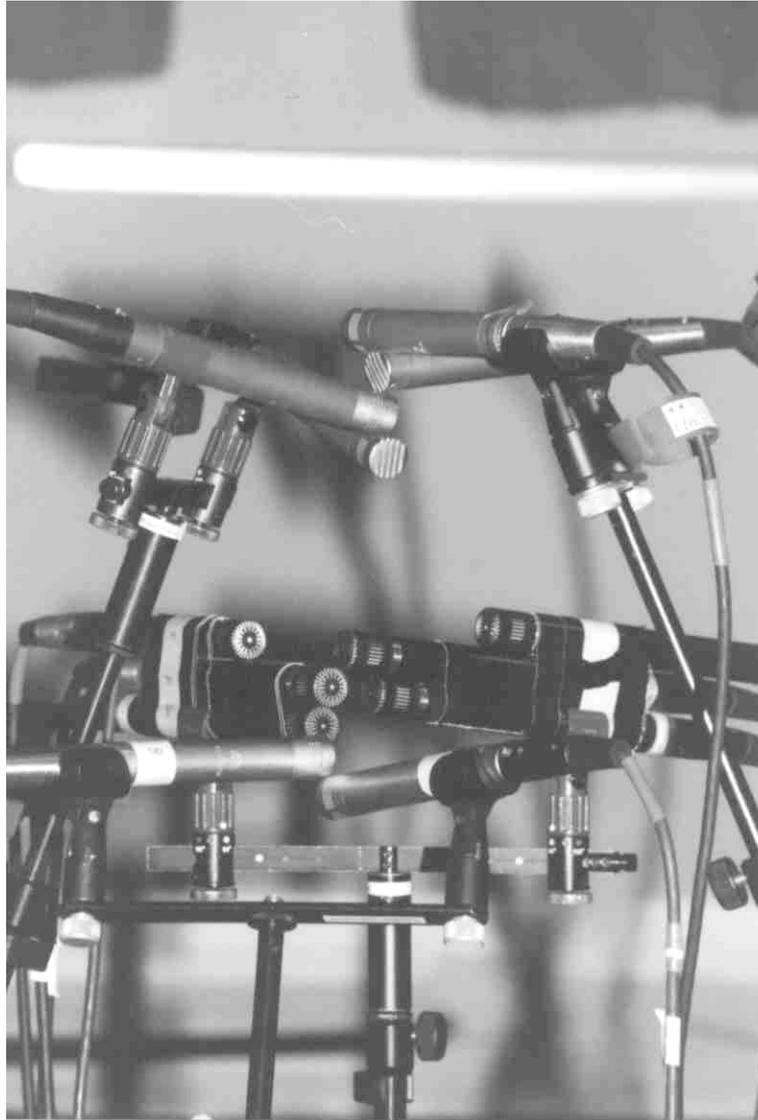


figure 57. Position des deux systèmes pour l'enregistrement

6.3 *Déroulement du test*

Contrairement aux premiers tests, les résultats n'ont pas été saisis et visualisés durant le déroulement des écoutes. Les résultats étaient plus nombreux en raison des quatre sortes de stimulus mais surtout, ils comportaient beaucoup d'inversions de localisation entre l'avant et l'arrière. En face de ces résultats déroutants et inattendus, la conduite à tenir a été plus difficile à mettre en œuvre. Le choix de faire passer plus d'individus sur le système 2 émane d'une volonté d'obtenir plus de précision sur le système étudié dans ce mémoire. Comment exploiter ces données inversées sans interférer sur le résultat final ?

Des auditeurs seront placés à 50 cm de l'axe centrale dans les 4 positions cardinales pour obtenir comme initialement des informations sur la robustesse.

35 fiches réponses composent les résultats de ce test réparties ainsi : 18 aux écoutes du système 2, 10 au système 1 et 7 pour le test de la robustesse.

Les auditeurs sont en majorité les mêmes que ceux qui ont participé au test précédent. Chaque auditeur ne fait qu'un test, il note successivement 22 localisations par stimulus. La séquence de la localisation est aléatoire. Faire passer des personnes sur les deux tests risquait d'influencer les résultats d'un des systèmes en raison de la possibilité d'apprentissage du test.

6.4 Résultats du test

Les enregistrements réalisés avec le système 2 ont un niveau de bruit de fond en basse fréquence très perceptible. L'égalisation de +8dB dans cette zone fréquentielle en est directement la cause.

La première information qui ressort instantanément des données recueillies est le nombre impressionnant d'inversions de localisation entre l'avant et l'arrière alors que le test préalable concernant la localisation par panning de la console n'en comportait aucun. Ces inversions sont ressenties sur les 4 tests mais dans une proportion alarmante pour les échantillons brefs du woodblock et des claves. En raison de leur nature spectrale et temporelle plus riche, l'échantillon chanté et les sonorités des maracas se distinguent très nettement des tests basés sur les transitoires. On retrouve ce même clivage pour la difficulté que les auditeurs ont ressentis à juger pertinemment la direction présumée des tests courts. Le jugement de la localisation des sons percussifs entraîne une concentration plus soutenue et un temps de réponse plus long.

Le tableau 2 nous donne le pourcentage des valeurs pour chaque test dont la réponse est inversée ou non indiquée. Le nombre de réponses indéterminées (point d'interrogation) contribue à environ la moitié de ces valeurs pour chaque test.

Pourcentage d'inversion total	Prise de son cardioïde système 1	Prise de son du second ordre système 2
Chant	11 %	7 %
Maracas	13 %	8 %
Claves	26 %	36 %
Woodblock	24 %	35 %

tableau 2. Pourcentage d'inversion avant arrière en fonction du test

Chaque test comportant 22 échantillons, 2 erreurs correspondent donc à moins de 10 pourcent. Ces résultats se révèlent meilleurs pour les tests maracas et chant pour le système 2 par rapport au 1 alors que c'est l'inverse avec les percussions brèves. On peut y déceler une fragilité des systèmes en face d'une calibration trop simple pour être ultra précise (la calibration est faite avec du bruit rose et un Audioprécision). Cette calibration rapide ne permet pas de s'affranchir des petits défauts de phases ou de retard. Le système 1 est de plus assemblé avec des microphone n'ayant pas pu être appairés. Le fait qu'une dizaine de fiches de résultats ne comportent pas d'inversion sur les tests avec le chant et les maracas évoque qu'un apprentissage du test permettrait de réduire sensiblement les inversions.

Curieusement ces inversions ne sont pas spécifiques à des positions angulaires précises pour tous les auditeurs. Lorsque l'on fait la moyenne, les tests comportant le moins d'erreur sont ceux de l'axe transaural à 90° et 270°, on comprend aisément que l'inversion étant de symétrie avant arrière, le nombre d'erreur est plus faible pour ces valeurs. On observe aussi une bonne moyenne pour l'azimut de 140° et -30° (angle physique des enceintes). Inversement les mauvais résultats concernent les positions arrières comprises entre 170° et 220° ainsi que les angles situés entre 50° et 70°. Comme ces derniers résultats ne sont pas symétriques et qu'ils se retrouvent à la fois sur les deux systèmes d'enregistrement, il est tentant de dire que ces erreurs sont plus liées à un mauvais appairage d'enceintes ou à un défaut acoustique de la salle plutôt qu'à un défaut de calibrage ou de phase des systèmes d'enregistrement.

Si, dans le test préalable, des perceptions étaient ressenties en élévation, elles n'influençaient pas la localisation. Ici, la sensation de hauteur est devenue un paramètre d'incertitude et de gêne. A nouveau les tests percussifs sont les plus marqués par cette sensation. On a 43% des 30 auditeurs ayant participé aux tests qui ont ressenti cette élévation.

Un auditeur quadragénaire particulièrement intéressant pour ce test en raison de son incapacité visuelle datant de l'enfance montrait du doigt systématiquement des zones plus hautes que la position des enceintes. Dans son cas, les deux tests courts sont quasiment inexploitable car il pointait des zones proches de l'aplomb.

Les notations inversées sont soit la cause d'une localisation incertaine et à ce titre, elles doivent être considérées comme des réponses indéterminées soit, elles émanent d'un effet miroir entre l'avant et l'arrière. Cette considération entraîne trois interprétations possibles de ces données :

- Affecter l'angle correspondant à l'image du miroir.
- Les classer avec les résultats d'inversion.
- Conserver leurs valeurs.

Garder les réponses telles quelles, est une solution qui va perturber les résultats car les inversions entre l'avant et l'arrière auront comme conséquence d'augmenter la zone de précision de la localisation. Ajoutons qu'un apprentissage du test, qu'un second test ou l'utilisation de stimuli spécifiques comme la voix, donneraient certainement beaucoup moins d'inversions.

Affecter un angle par symétrie axiale est douteux car certaines valeurs sont peut être dues à des incertitudes.

Les supprimer des tables reste le meilleur choix malgré le fait que certaines valeurs considérées comme inversées ne sont peut-être que des déviations. Pour prendre en compte cette hypothèse, seules les valeurs dont l'angle de symétrie est supérieur à $\pm 30^\circ$ seront omises.

Les courbes de résultats suivantes sont créées sous Excel à partir des tables de valeurs dont les résultats inversés et non localisés ont été éliminés. Elles sont présentées sur le cercle complet de latéralisation (0 à 360°) parce qu'il n'est plus nécessaire de détailler la contribution des différences d'intensité par quadrants.

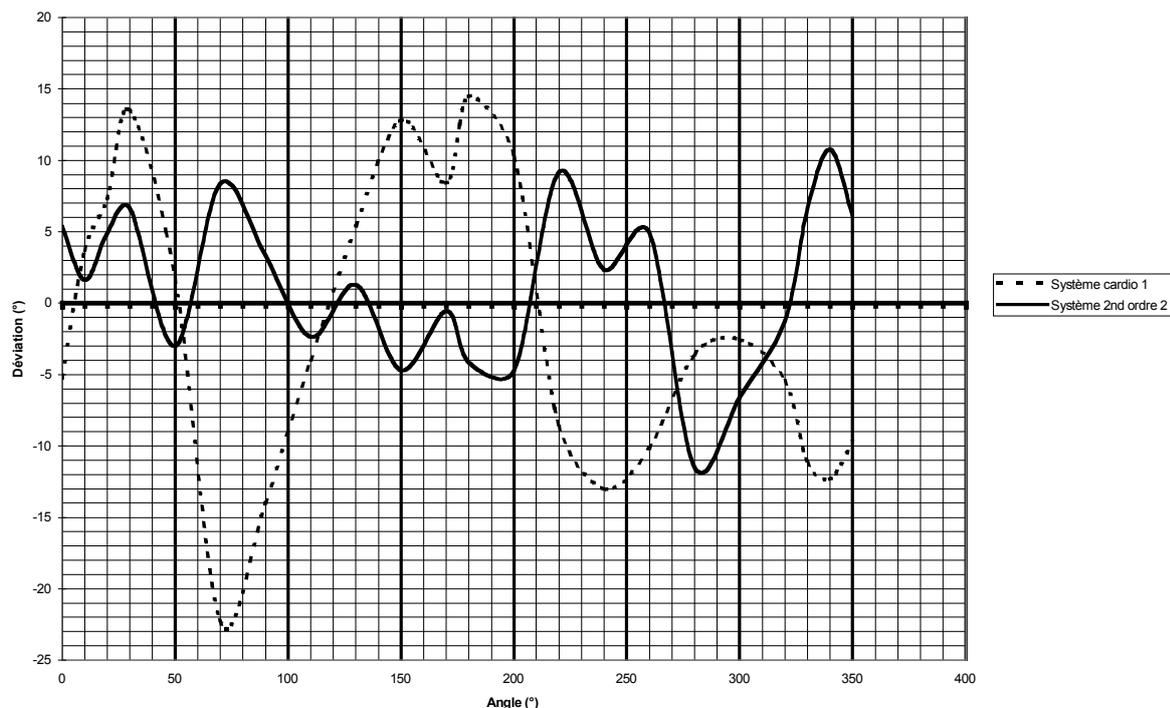


figure 58. Déviation moyenne pour les deux systèmes

Lorsque l'on observe la courbe de la figure 58, on regrettera de ne pas y voir de symétrie marquée entre les zones situées à droite et à gauche. Cette courbe représente la déviation moyenne des résultats de chaque système par rapport à la configuration physique de l'enregistrement. C'est en fait, la différence perçue entre l'enregistrement et l'écoute. Les résultats de la zone angulaire arrière auraient dû être similaires entre les deux systèmes car ils utilisent tous les deux des capsules cardioïdes. La courbe du système 2 fluctue autour de $\pm 12^\circ$ alors que le système 1 oscille entre 14° et -23° . Cette courbe prend aussi en compte l'erreur due à la directivité partiellement similaire des microphones utilisés avec les données des tests initiaux. Le système 2 est en conclusion légèrement plus fidèle que le système 1.

La figure 59 représente les déviations issues du test initial par rapport aux directivités théoriques des microphones utilisés. Cette figure est à comparer avec la précédente et aurait dû être totalement similaire. Les déviations sont plus faibles sur cette courbe, c'est à dire que la prise de son ajoute une imprécision supplémentaire, ce qui est logique. Les similitudes entre les deux figures sont plutôt de l'ordre spéculatif car certaines tendances sont inversées et asymétriques. On s'étonnera tout de même de la similitude des courbes du système 1 entre 0 et 50° et 300 et 360°

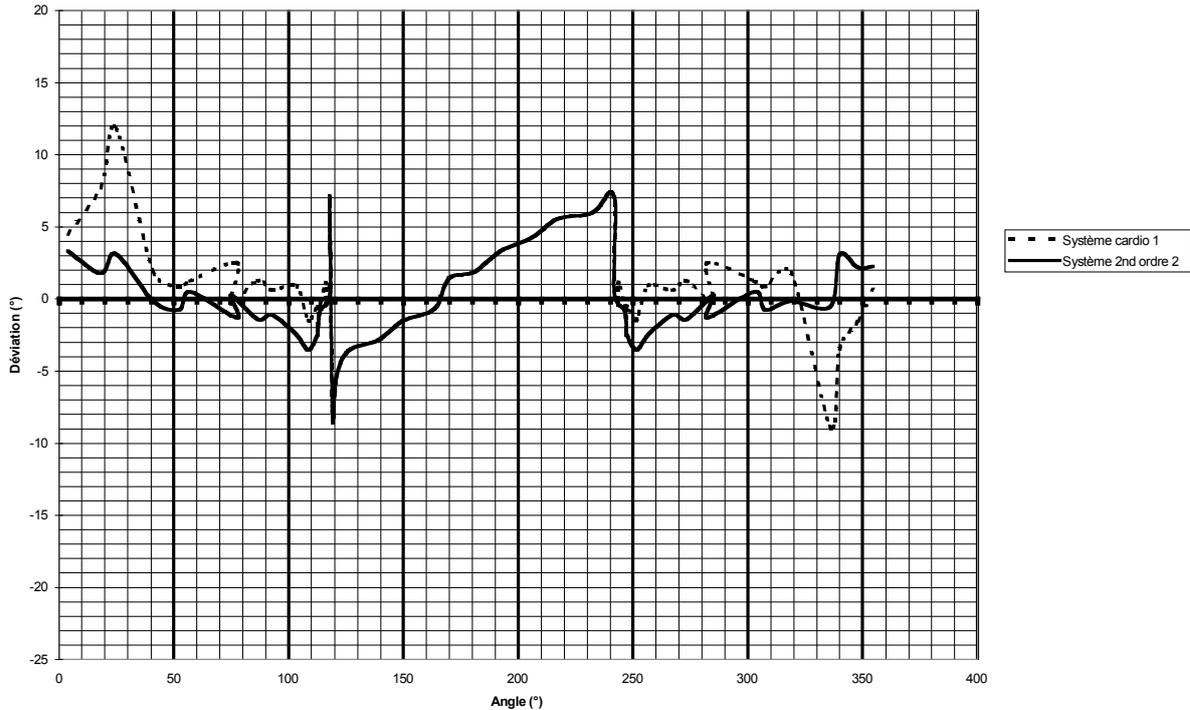


figure 59. Déviation théorique des deux systèmes

Les résultats complets sont présentés pour chaque système avec les quatre stimulus :

- figure 60 et figure 61 pour le chant.
- figure 62 et figure 63 pour les maracas.
- figure 64 et figure 65 pour les claves.
- figure 66 et figure 67 pour le woodblock.

Comme pour les premiers résultats, la ligne continue représente la moyenne obtenue, les tirets symbolisent l'écart type et les marques rondes montrent les valeurs extrêmes. La ligne continue quant à elle représente les résultats souhaités. Les courbes sont graduées en abscisse de 0 à 360°, car le nombre de stimuli ne permet pas de représenter chaque courbe pour différents quadrants, les courbes repliées par rapport à l'axe de symétrie ne donnent pas non plus d'informations supplémentaires. L'analyse directe des courbes est difficile, on donnera les résultats qui s'en dégagent dans les paragraphes suivants.

A première vue, les résultats donnent des courbes proches de la pente désirée, mais la représentation sur 360° tend à minimiser les déviations car elles sont en réalité plus importantes qu'au premier test.

Test du prototype

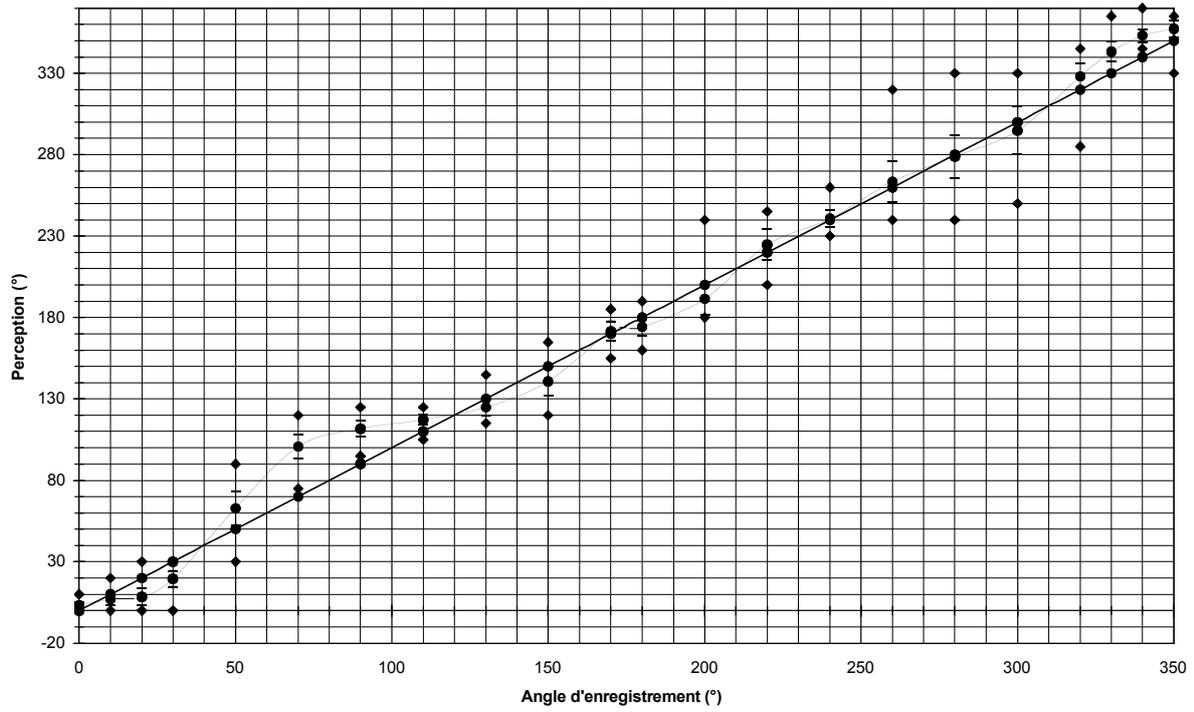


figure 60. Résultats complets pour le chant pour le système 1

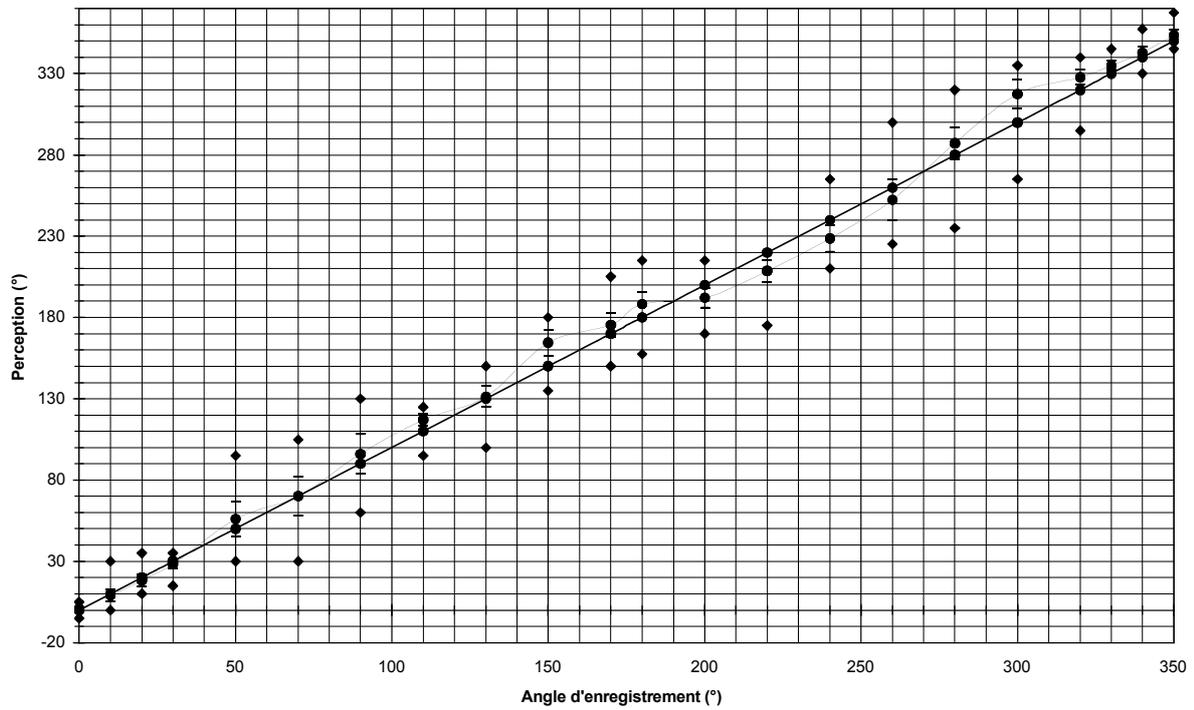


figure 61. Résultats complets avec le chant pour le système 2

Test du prototype

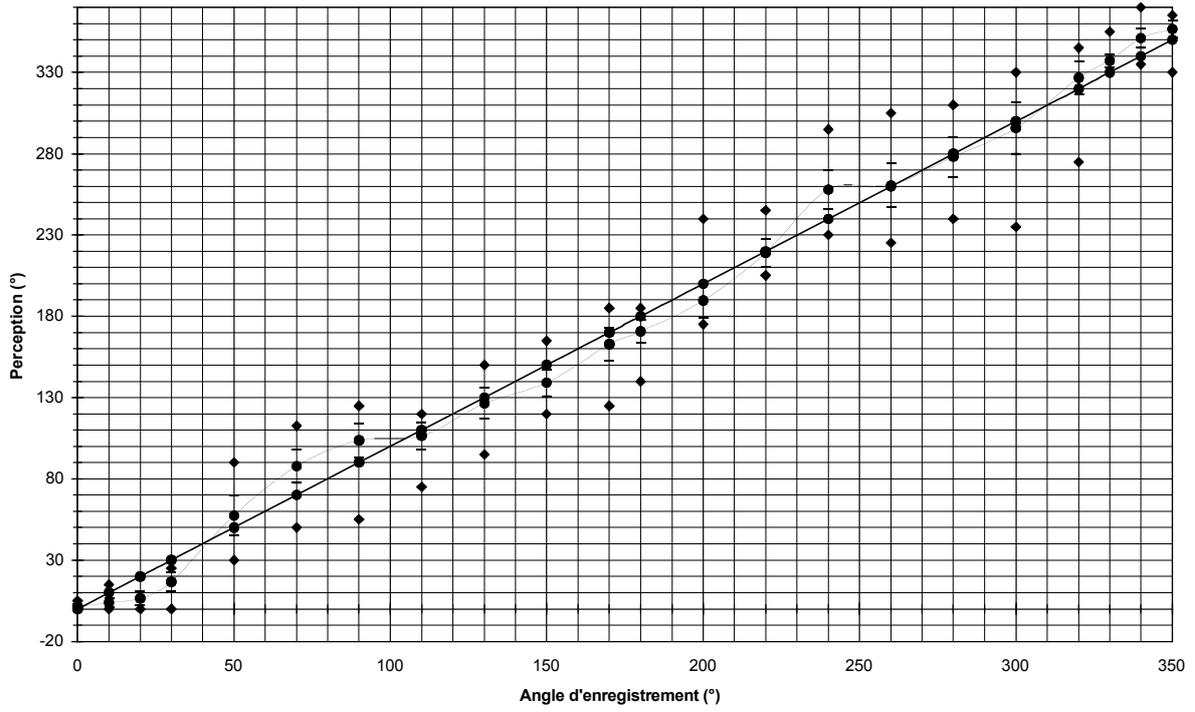


figure 62. Résultats complets avec les maracas pour le système 1

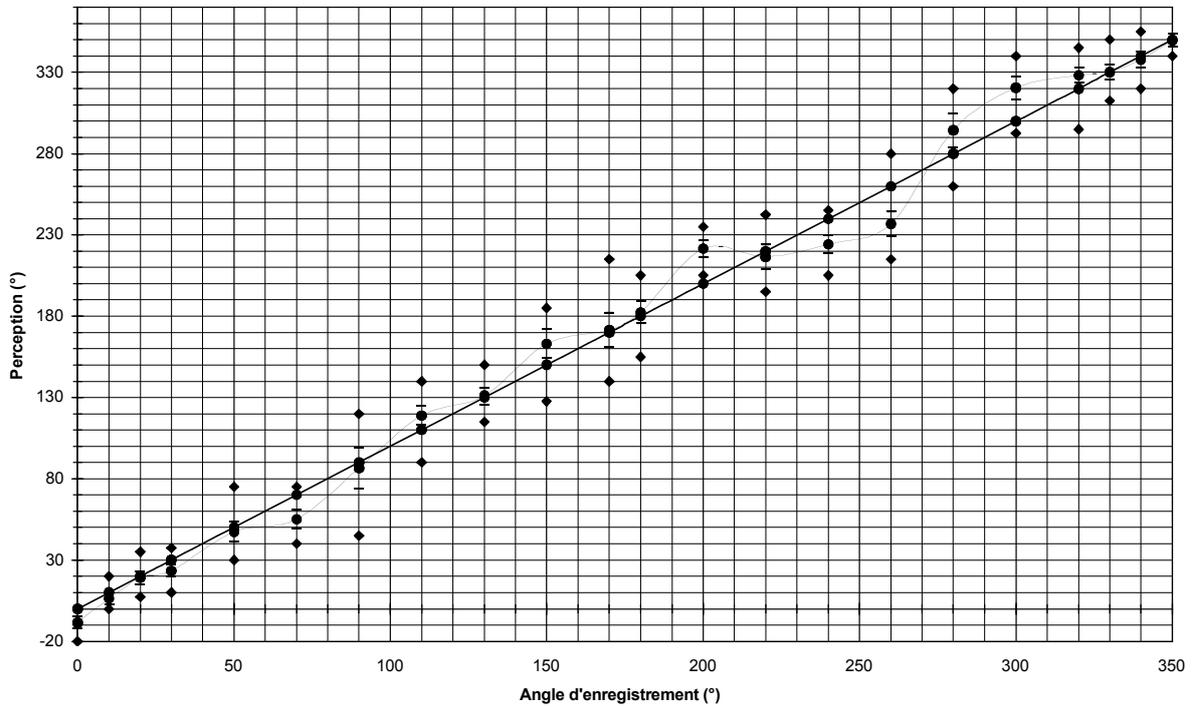


figure 63. Résultats complets avec les maracas pour le système 2

Test du prototype

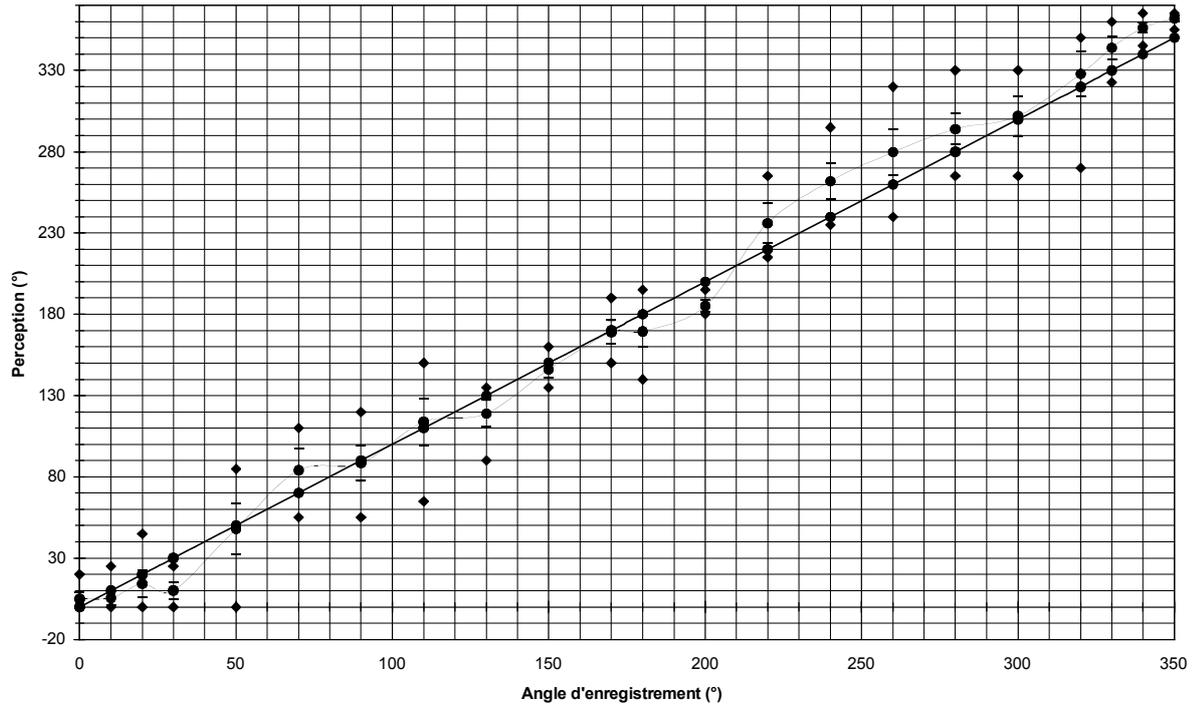


figure 64. Résultats complets avec les claves pour le système 1

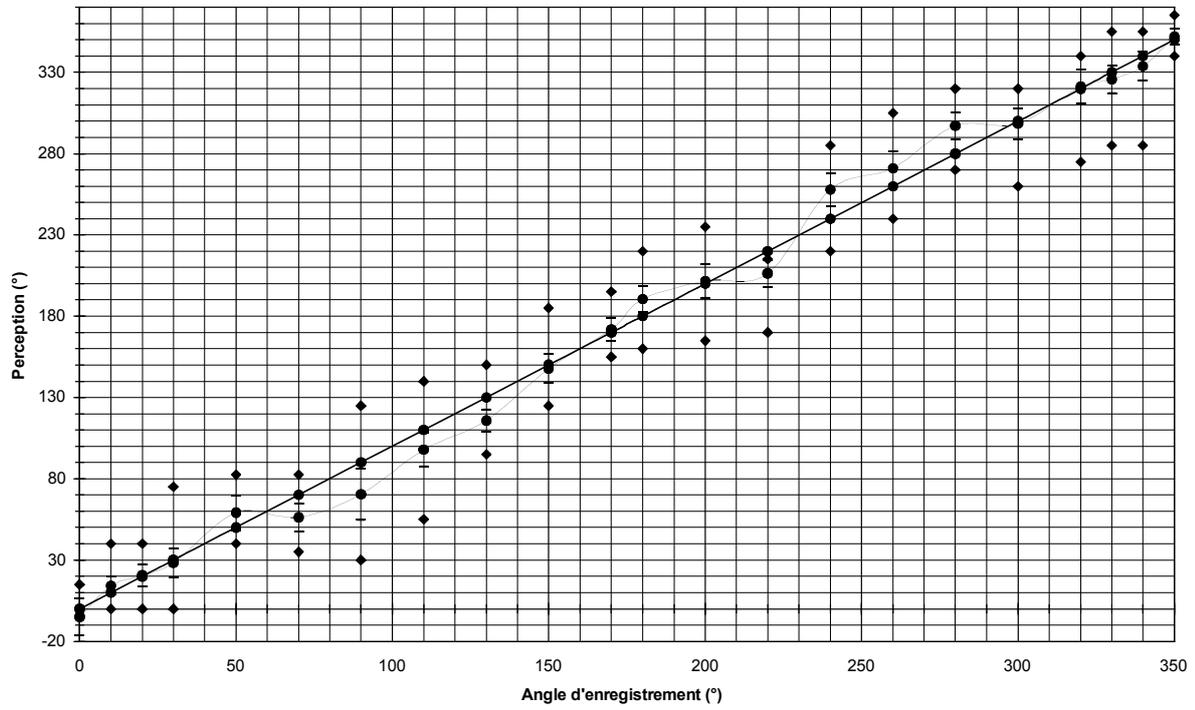


figure 65. Résultats complets pour les claves pour le système 2

Test du prototype

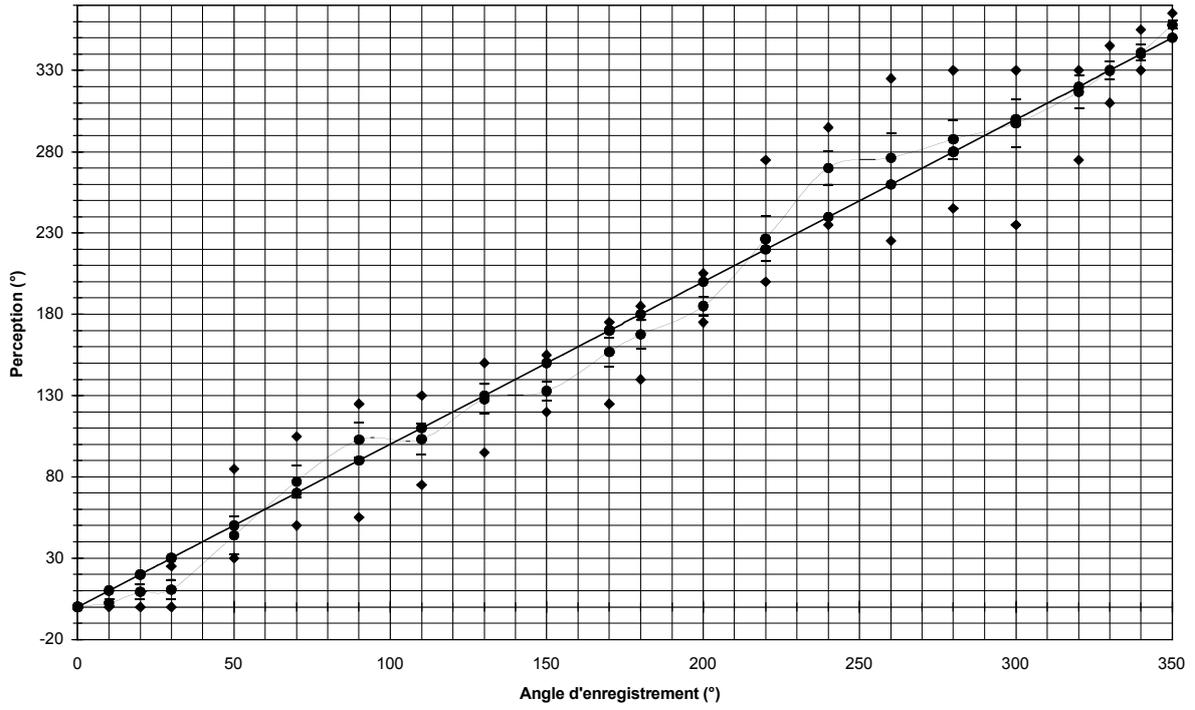


figure 66. Résultats complets avec les woodblock pour le système 1

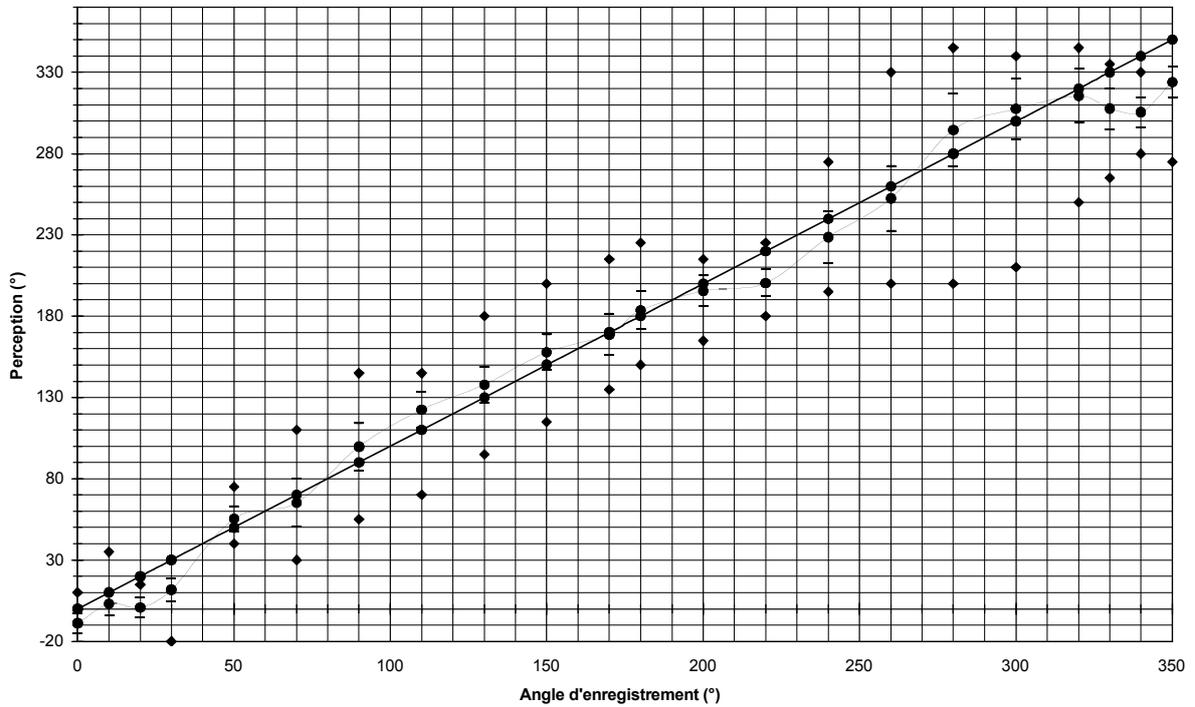


figure 67. Résultats complets avec le woodblock pour le système 2

Les résultats du système 2 sont plus précis pour la zone avant de $\pm 30^\circ$ que le système 1 ce qui donne ici, un réel avantage au 2nd ordre. Le codage fait avec les capsules cardioïdes dans cette zone est décalé d'une dizaine de degré. Seul le stimulus du woodblock ne corrobore pas cette constatation. La sonorité plus grave du woodblock modifie-t-elle le codage obtenu par le test initial ?

Les meilleurs résultats sont obtenus pour le stimulus chanté avec le système 2 (figure 61). La moyenne s'écarte rarement de plus de 10° .

Certains motifs se retrouvent pour un même système sur plusieurs courbes, ces similitudes trahissent un problème de précision de codage de la localisation provenant d'un effet de polarisation de salle ou d'un défaut intrinsèque à l'assemblage microphonique.

Lorsque l'on compare uniquement les écarts type de chaque courbe, on observe bien cette fois une symétrie axiale avec une prépondérance de 12° en moyenne proche des angles de $\pm 70^\circ$, une zone relativement faible avec un écart type de 8° en moyenne autour de 180° et enfin, une zone nettement meilleure avec un écart type moyen de 4.5° vers $\pm 30^\circ$.

Les valeurs des minima et maxima ne sont pas comparables entre les deux systèmes, car le système 1 ne comporte que 10 auditeurs comparés aux 18 du 2nd système. Les écarts entre minima et maxima devraient être plus important sur le système 2 ce qui n'est pas toujours le cas.

Echantillon	Système 1	Système 2
Chant	7.08°	6.68°
Maracas	8.54°	6.28°
Claves	8.83°	9.31°
Woodblock	8.26°	11.8°

tableau 3. Moyenne des écarts types

Le tableau 3 donne la moyenne des écarts types pour chaque échantillon sonore et pour chaque système. On retrouve les mêmes tendances que sur la moyenne d'inversion avant arrière du tableau 2, ce qui laisse à penser que le nombre d'inversions est lié au facteur de précision représenté par l'écart type. Ce tableau prouve que le système composé du prototype est plus précis que le système 1 tout au moins en ce qui concerne le chant et les maracas.

Echantillon	Système 1	Système 2
Chant	8.53°	6.2°
Maracas	10.9°	8.07°
Claves	10.1°	7.77°
Woodblock	9.31°	11.8°

tableau 4. Moyenne de l'écart type moyen

Le tableau 4 quant à lui représente la moyenne de la déviation entre les écoutes et la prise de son. C'est certainement le résultat le plus important de ce test qui montre que la directivité du second ordre améliore la justesse de la localisation. A nouveau, seule la sonorité du woodblock

se démarque de cette analyse et tend à démontrer que les basses fréquences nécessitent un codage en différence d'intensité moindre (ce qui est conforme aux expérimentations subjectives retranscrites dans la littérature [16][4][53]...).

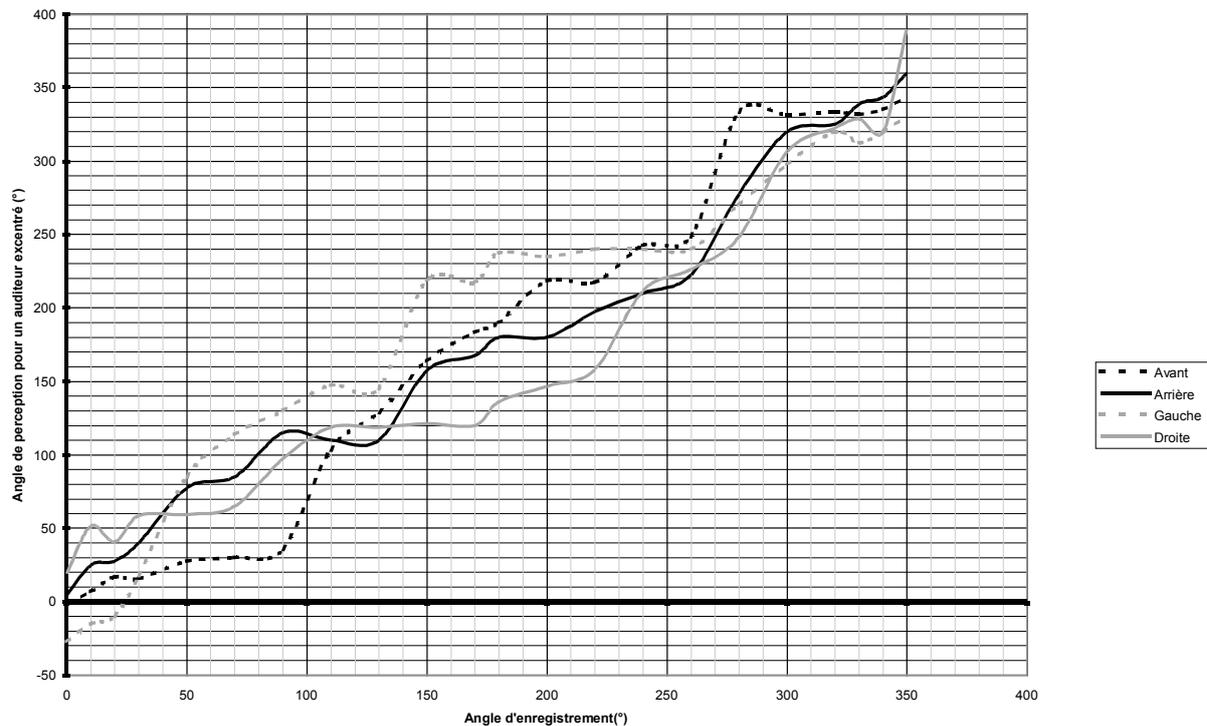


figure 68. Test de robustesse à 50 cm du centre

La figure 68 est à comparer avec la figure 26, elle représente les résultats moyens pour des auditeurs placés à 50 cm de l'axe central. Le faible nombre d'individus ayant réalisé ce test nous pousse à la prudence quant aux commentaires de l'allure de ces courbes. Le nombre d'inversions est de proportion équivalente aux résultats centrés.

Les courbes paraissent en effet meilleures que celles concernant les essais de robustesse des lois de pan-pot. Peut être que le panel d'auditeur de ce test est plus entraîné ou que la réverbération présente dans le message améliore la robustesse ? L'effet d'antériorité sensible pour une position décentrée entre les 4 enceintes nous prédisposait à des résultats plus médiocres.

La déviation angulaire ne dépasse rarement 50° sur cette courbe où les valeurs des enregistrements des claves et woodblock n'a pas été prise en compte en raison du peu de données valables avec ces stimuli.

6.5 Conclusion du test

Doit-on en déduire que la durée des transitoires des tests percussifs ne laisse pas le temps au cerveau de juger avec justesse de la direction sonore ? Le cerveau a besoin d'un temps d'analyse avant de donner une conclusion fiable. Une séquence répétitive de ces stimuli aurait sans aucun doute améliorée les résultats. Le comportement des résultats du woodblock nous interroge sur la validité du test initial concernant les tests filtrés.

Le cas de la personne non voyante nous renvoie au fait que l'image fantôme n'est qu'un leurre propre à chaque individu. Cet homme bénéficie d'une oreille doublement exercée : il est accordeur de piano dans son métier et utilise chaque jour son acquis auditif de la localisation comme seul moyen de se diriger et d'appréhender les formes. Par essence, il ne peut se recalculer géométriquement que par la vue, signe que la perception auditive de la localisation est directement influencée par les autres sens de la perception .

Au premier abord, les résultats du test sont peu probants. L'avantage du système 2 par rapport au 1^{er} est faible et la précision de la localisation en moyenne reste large surtout en regard du nombre important d'inversion entre l'avant et l'arrière. Néanmoins, la localisation virtuelle attendue fonctionne bien car la perception de la localisation en général se fait dans la direction voulue. Ces résultats montrent que la localisation pour un système multicanal non matricé est possible par une simple prise de son. L'amélioration certes faible mais réelle du système 2 se ressent non seulement au niveau de la finesse mais aussi de la précision de la localisation, c'est à dire d'une part que l'azimut perçue est plus proche de l'angle de la source réelle et d'autre part que les auditeurs localisent en moyenne une zone plus restreinte. L'avantage du système 2 doit être valorisé par les défauts intrinsèques de sa conception. Le système 1, en effet est linéaire en phase et possède une directivité verticale constante.

Des essais d'enregistrement de son dynamiques (voix et instruments rotatifs autour de l'axe) avec ces systèmes rendent une très bonne perception de l'effet de mouvement attendu. Des essais à différentes distance du système de captation donnent aussi une perception semblant assez réaliste de la distance apparente des sources.

Ces résultats finaux démontrent l'intérêt d'un tel système pour la prise de son mais il devra être amélioré au niveau de la directivité et de sa miniaturisation et pourrait être éventuellement équipé de filtrages en fréquence supplémentaires pour améliorer le pourcentage des perceptions d'inversion.

Les points suivants relatent des améliorations nécessaires à une meilleure fidélité :

- Utilisation de microphones cardioïdes dont la fonction de directivité est proche des valeurs théoriques.
- Placement des microphones du prototype sur un même axe.
- Utilisation de microphones miniatures.
- Linéarisation de la réponse en phase en fonction de la fréquence. Utilisation de DSP.
- Adaptation du local et des enceintes de monitoring pour une reproduction en 5.1.
- Augmentation du rapport signal sur bruit en basse fréquence. Une captation avec un microphone omnidirectionnelle en basse fréquence pour tous les quadrants que l'on affectera par filtrage à chaque canaux ou uniquement à un caisson de basse est une solution à envisager devant la faible localisation auditive des fréquences inférieures à 100Hz.
- Intégration complète du système dans un boîtier. L'assemblage serait calibré une fois pour toutes.

Il serait maintenant intéressant de comparer des résultats du même type et dans les mêmes conditions avec d'autres systèmes de prise de son ou même des versions améliorées du prototype.

7 Conclusion

La prise de son en multiphonie ne trouvera pas un système de captation idéal avant longtemps. Plus il y a de sources et de récepteurs, plus le nombre de combinaisons possibles est important. Ceci augmentant en même temps la complexité des paramètres pour parvenir à résoudre le problème. On peut prévoir que d'innombrables solutions seront proposées ces prochaines années pour la reproduction et la captation en multicanal surtout si l'on nous propose à l'avenir une reproduction normalisée sur plus de cinq enceintes. N'oublions pas, que nous ne comprenons toujours pas encore très bien quels sont les artefacts mis en jeu dans la perception d'une image fantôme en stéréophonie sur deux canaux. Cette technique pourtant largement employée depuis plusieurs décennies renvoie les principes actuels du multicanal au tout début de leurs évolutions.

L'étude de la localisation d'un système de prise de son en stéréophonie n'a guère été étudiée qu'en biphonie. Ce travail de recherche en apportant des résultats concrets, contribue à clarifier la prise de son en multicanal. On s'étonne en effet de la tendance actuelle dans les milieux audio professionnels et acoustiques à sous estimer les résultats physiques en faveur d'un idéal artistique. Un choix culturel ou artistique ne peut pas faire l'unanimité en matière de décision car les goûts subjectifs de chacun sont libres. La seule et unique solution est donc d'établir une base de référence sur laquelle on pourra comparer les caractéristiques objectives et subjectives de chaque système. La démarche présentée dans ce mémoire permettra je l'espère, d'aider à l'élaboration de concepts similaires ou nouveaux en progressant plus rapidement sur les étapes intermédiaires et néanmoins obligatoires de la mise en œuvre.

Le prototype étudié pour cette recherche donne des résultats intéressants mais il devra être amélioré et intégré dans un module compact pour un usage courant en prise de son acoustique. Les expériences réalisées prouvent qu'une localisation azimutale en 5.1 est possible sans avoir recours à l'emploi de matrices de décodage ou de filtrages du type HRTFs. On constate qu'une prise de son coïncidente dans une configuration 5.1 autorise la reproduction de la latéralisation ce qui n'a pas été encore constatée avec les systèmes d'enregistrement utilisant les différences de temps. L'avantage de la prise de son réalisée avec le prototype du 2nd ordre comparée à une prise de son en différences d'intensité avec les capsules cardioïdes, se retrouve principalement, sur la fidélité et la précision de la perception de la localisation en zone frontale.

Des critères perceptifs de la prise de son non étudiées ici comme le respect de la distance apparente ou l'enregistrement de source dynamique semblent prometteuses avec une prise de son multicanale en différence d'intensité.

Les enregistrements musicaux et l'élaborations de fictions sont les applications directement concernées par un système d'enregistrement respectant la localisation. Les prises de sons acoustiques dans ces domaines doivent respecter le timbre, les transitoires, les formants et la clarté des

Conclusion

sons originaux mais aussi la réverbération, la spatialisation, l'enveloppement, l'équilibre spectral qui sont des nuances propres à l'art de la prise de son. Il faudra vérifier si des microphones coïncidents en prise de son multicanal sont compatibles avec les exigences d'un enregistrement sonore de qualité.

Ces divers points pourraient être le sujet d'une nouvelle étude.

8 Références

- [1]. BAUER B. BENJAMIN, ABBAGNARO A., GRAVEREAUX DANIEL W., MARSHALL TREVOR J. (1976). The Ghent Microphone System for SQ Quadraphonic Recording and Broadcasting. AES 55th Convention. New York 1976.
- [2]. BEAVER B.R. and BROWN R. (1970). Third-Order Gradient Microphone for Speech Reception. J. Audio Eng. Soc. , Vol. 18, No. 12, 1970.
- [3]. BERNFELD BENJAMIN (1975). Simple Equations For Multichannel Stereophonic Sound Localization. J. Audio Eng. Soc. , Vol. 23, No. 7, 1975.
- [4]. BLAUERT JENS (1997). Spatial Hearing : The Psychophysics of Human Sound Localization. Londres, MIT Press, Cambridge, Mass. 1997.
- [5]. BOTTE MARIE CLAIRE, CANEVET G., DEMANY L. et SORIN C. (1989). Psychoacoustique et Perception Auditive. Série Audition INSERM/SFA/ CNET. 1989.
- [6]. BRUNHES JEAN BAPTISTE (1999). Localisation d'une Source Monophonique Musical en multicanal 5.1. Ecole National Supérieur Louis Lumière 1996-1999.
- [7]. CANEVET GEORGES (1989). Audition Binaurale et Localisation Auditive. Aspects Physiques et Psychoacoustiques. Psychoacoustique et Perception Auditive. Série Audition. 1989.
- [8]. COURVILLE DANIEL (1993). Procédés et Systèmes d'Enregistrement et de Reproduction Sonore en Trois Dimensions. Montréal. 1993.
- [9]. FLANAGAN PATRICK, DICKINS GLENN and LAYTON LEONARD (1999). Real Time Virtual Acoustic for 5.1. . AES 16th Conference on Spatial Sound Reproduction.
- [10]. GERZON MICHAEL A. (1992). Microphone Technique for 3-Channel Stereo. AES 93rd Convention. San Francisco October 1992
- [11]. GERZON MICHAEL A. (1992). General Metatheory of Auditory Localisation. AES 92nd Convention. Vienna. March 1992.
- [12]. GERZON MICHAEL A. (1973). Periphony : With-Height Sound Reproduction. . J. Audio Eng. Soc. , Vol. 21, No. 1, 1973.
- [13]. GERZON MICHAEL A. (1992). Application of Blumlein Shuffling to Stereo Microphone Techniques. J. Audio Eng. Soc., Vol. 42, No. 6, 1994.
- [14]. GIVERNAUD LAURENT. (1999). Les systèmes de reproduction en multicanal. Probatoire CNAM 1999
- [15]. GRENIER YVES, PRADO JACQUES et LIEBENGUTH DANIEL (1999). A Signal Processing Digital Microphone. AES 107TH Convention. September 1999.
- [16]. HUGONNET CHRISTIAN et WALDER PIERRE (1994). Théorie et pratique de la prise de son stéréophonique. Edition Eyrolles 1995.

Références

- [17]. ITU-R BS. 775-1. (1994) Multichannel Stereophonic Sound System With and Without Accompanying Picture. Geneva 1992-1994
- [18]. JAMROZ ANTHONY (1959). The Design and Use of a Double Cardioid Stereophonic Microphone. AES 11th Convention. New York. October 1953.
- [19]. JOUHANEAU JACQUES (1997). Acoustique des Salles et Sonorisation. Collection CNAM. Acoustique Appliquée. Technique et Documentation 1997.
- [20]. JOUHANEAU JACQUES (1994). Notions élémentaires d'acoustique. Electroacoustique. Collection CNAM. Acoustique Appliquée. Technique et Documentation 1994.
- [21]. LE DU GUILLAUME (1999). Conception de Systèmes de Prise de Son Multicanaux. Application à l'Enregistrement Musical. Mémoire de L'école Nationale Supérieure Louis Lumière. Juin 99.
- [22]. LEHMANN R. (1963). Les Transducteurs Electro et Mécano-Acoustiques Haut-Parleurs et Microphones. Collection technique et scientifique du CENT. Edition CHIRON 1963.
- [23]. LYZWA JEAN MARC et MESSONNIER JEAN CRISTOPHE. CNSMDP Service Audiovisuel (2000). Comparative Experiment of Multichannel Sound Recording Systems. AES PARIS 2000.
- [24]. MAKITA Y. (1962). Localisation Directionnelle du Son dans un Champ Sonore Stéréophonique. Revue de l'UER. Cahier A. Technique N°73. Juin 1962.
- [25]. MERTENS H. (1966). L'Ecoute Stéréophonique. Conférences des Journées d'études. Editions Chiron. 1966.
- [26]. MESSONNIER JEAN CRISTOPHE (1994). Utilisation du Spatialisateur de l'Ircam pour l'étude de l'influence des Systèmes de Prise de Son et de Restitution Stéréophoniques sur la Perception Spatiale. Mémoire CNAM 1994.
- [27]. MORA ARNAUD et NICOLAS JACQUES (1998). Prise de son multicanale en différence de temps. FSMS. Conservatoire National Supérieur de Musique de Paris. Promotion 1994-1998
- [28]. NICOL ROZENN. EMERIT MARC. (1998). Reproducing 3D-Sound for Videoconferencing : A Comparaison between Holophony and Ambisonic.DAFX98.
- [29]. NICOL ROZENN. (1999). Etude de la restitution du son spatialisé dans une zone étendue : application à la téléprésence. Thèse
- [30]. OLSON HARRY F. (1976). A History of High-Quality Studio Microphones. AES 55TH Convention. New York November 1976
- [31]. OLSON HARRY F. (1977). Microphones for Recording. Anthology JAES
- [32]. OLSON HARRY F. Directional Microphones. Anthology JAES
- [33]. OLSON HARRY F. (1980). The Quest for Directional Microphones at RCA. J. Audio Eng. Soc., Vol. 28, No. 11, 1980.
- [34]. RUMSEY FRANCIS. BERG JAN (1999). Spatial Attribute Identification an Scaling by Repertory Grid Technique and other Methods. AES 16th Conference on Spatial Sound Reproduction.
- [35]. RUMSEY FRANCIS and MASON RUSSEL (1999). An Investigation of Microphone Techniques for ambient Sound in Surround Sound System. AES Convention.
- [36]. SANK JON R. (1985). Microphones. J. Audio Eng. Soc., Vol. 33, No. 7/8, 1985.

- [37]. THEODORE C. TANNER Jr. (1997). Psychoacoustic Criteria for Auditioning Virtual Imaging System. AES Convention
- [38]. THEILE G. PLENGE G. (1977). Localization of Lateral Phantom Sources. J. Audio Eng. Soc., Vol. 25, No. 4, 1977. AES 53rd Convention. Zurich. March 1976.
- [39]. THEILE GUNTHER (2000). Multichannel Natural Music Recording Based on Psycho-acoustic Principles. AES 108TH Convention of Paris. February 2000.
- [40]. VARLA A. MAKIVIRTA A. MARTIKAINEN L. PILCHNER M. SCHOUTAL R. ANET C. (1999). Design of Rooms for Multichannel Audio Monitoring. AES 16th Conference on Spatial Sound Reproduction.
- [41]. WALKER R. (1999). A Controlled-reflection Listening Room for Multichannel Sound. AES convention
- [42]. WEINGARTNER B. (1965). Two-Way Cardioid Microphone. AES 70TH Convention of New York. October 1965.
- [43]. WEST JAMES R. (1998). Five-Channel Panning Laws : an Analytical and Experimental Comparaison. Master of science in music engineering technology. Florida. 1998.
<http://www.music.miami.edu/students/personal/jwest/master.html>
- [44]. WEST JAMES R., SESSIER GERHARD M. and KURBI R. A. (1989). Unidirectional, second-order gradient microphone. J. Acoust. Soc. Am. December 1989.
- [45]. WILLIAMS MICHAEL (1985). Le 'Zoom' Stéréophonique
- [46]. WILLIAMS MICHAEL (1987). Unified Theory of Microphone Systems of Stereophonics Sound Recording. AES 82nd Convention in London. Preprint 2466.
- [47]. WILLIAMS MICHAEL et LE DU GUILLAUME (1999). Microphone Array Analysis For Multichannel Sound Recording. AES 107TH Convention. September 1999. Preprint 4997.
- [48]. WILLIAMS MICHAEL et LE DU GUILLAUME (2000). Multichannel Microphone array design. AES 108TH Convention of Paris. February 2000. Preprint 5157.
- [49]. WOSZCZYK WIESLAW R. (1984). A Microphone Technique Applying the Principle of Second-Order-Gradient Unidirectionality. J. Audio Eng. Soc. , Vol. 32, No. 7/8, 1984.
- [50]. WOSZCZYK WIESLAW R. (1992). Microphone Array Optimised for Music Recording. J. Audio Eng. Soc. , Vol. 40, No. 11, 1992.
- [51]. WOSZCZYK WIESLAW R. (1999). Sound Source Localization in a Five-Channel Surround Sound Reproduction System. AES 107TH Convention. September 1999.
- [52]. YAMAMOTO TAKEO (1972). Quadrasonic One Point Pickup Microphone. AES 43rd Convention New York. September 1972.
- [53]. ZWICKER E. et FELDTKELLER R. (1981). L'Oreille Récepteur d'Information. Stimulus et Sensation. Edition Masson 1981.
- [54]. AUDIO-TECHNICA. <http://www.audiotechnica.com/prodpro/profiles/AT895.html>
- [55]. SANKEN. <http://www.sas-mk.co.jp/CSS-5%20English.html>
- [56]. BRAUNER (SPL). http://www.spl-electronics.com/Atmos/in_detail.html
- [57]. Natural Surround Sound. <http://www.mcgill.ca/fgsr/ott/tech/91014.html>
- [58]. PANASONIC http://204.146.85/industrial_oem/elec...ponents_audio_components_microphones.htm

[59]. SCHOEPS <http://Schoeps.de/index1.html>

<http://www.open1.org>

http://www.media.mit.edu/people/araz/sss/sound_localization.html

Altieri <http://www.afit.af.mil/schools/en/enc/dreinhar/thesis>

9 Table des illustrations

figure 1. Recommandation ITU	2
figure 2. Zones des tests délimitées par l'angle physique des enceintes.....	3
figure 3. La prise de son en ΔI pour la biphonie	7
figure 4. Formation de 3 images fantômes en ΔT	9
figure 5. La solution Brauner	11
figure 6. La solution Schoeps.....	11
figure 7. The Holophone.....	12
figure 8. Capsules en montages tétraédrique.....	13
figure 9. Résultats du quadrant avant -30 à $+30^\circ$	21
figure 10. Résultats du quadrant latéral droit $+30$ à $+120^\circ$	21
figure 11. Résultats du quadrant latéral gauche $+240$ à $+330^\circ$	22
figure 12. Résultats du quadrant arrière $+120$ à $+240^\circ$	23
figure 13. Résultats du quadrant replié avant 0 à $+30^\circ$	23
figure 14. Résultats des quadrants repliés latéraux de $+30$ à $+120^\circ$	24
figure 15. Résultats du quadrant replié arrière de $+120$ à $+180^\circ$	25
figure 16. Diagrammes polaires du test en valeurs linéaires.....	26
figure 17. Diagrammes polaires du test en valeurs logarithmiques.....	27
figure 18. Diagrammes polaires du test en valeurs linéaires comparés à des courbes théoriques.....	27
figure 19. Comparaison des valeurs repliées du test et une directivité du 2nd ordre pour l'avant.....	29
figure 20. Comparaison des valeurs repliées du test et une directivité du 2nd ordre pour l'arrière.....	29
figure 21. Comparaison des valeurs repliées du test et une directivité hypercardioïde pour l'arrière.....	30
figure 22. Comparaison des valeurs repliées du test et une directivité cardioïde pour l'arrière.....	30
figure 23. Comparaison de divers résultats pour le quadrant frontal.....	32
figure 24. Comparaison de divers résultats pour le quadrant latéral.....	33
figure 25. Comparaison de divers résultats pour le quadrant arrière.....	33
figure 26. Essai de la robustesse des lois de pan-pot.....	34
figure 27. Courbe polaire d'un microphone interférentiel pour deux fréquences.....	38
figure 28. Ajout d'un retard supplémentaire en fonction de l'incidence sonore.....	40
figure 29. Diagramme polaire obtenu avec une mousse externe.....	40
figure 30. Microphone directif Audio-Technica.....	42
figure 31. Réalisation d'un microphone du 2nd ordre avec deux cardioïdes.....	43
figure 32. Diagramme polaire du microphone du second ordre pour $f=100\text{Hz}$ et $D=5\text{cm}$	44
figure 33. Réponse en fréquence dans l'axe avec un espacement D de 5cm	45
figure 34. Réponse en fréquence avec -6dB sur le microphone arrière.....	46
figure 35. Directivité du microphone DPA 4011.....	50
figure 36. Directivité du microphone Schoeps MK4.....	50
figure 37. Directivité du microphone Neumann KM84.....	51
figure 38. Directivité de deux capsules espacées de 8 cm sans pondération.....	52
figure 39. Directivité de deux capsules espacées de 8 cm avec 3 dB d'atténuation sur la seconde capsule.....	53
figure 40. Directivité de deux capsules espacées de 8 cm avec 6dB d'atténuation sur la seconde capsule.....	53
figure 41. Synoptique de la sommation de deux microphones du 2 nd ordre.....	54
figure 42. Réponse en fréquence du couple de capsules espacées de 8cm	55
figure 43. Réponse en fréquence du couple de capsules espacées de 1.5cm	55
figure 44. Réponse en fréquence de la sommation des deux systèmes.....	56
figure 45. Courbe de directivité du système complet théorique.....	58
figure 46. Second Ordre simulé avec les courbes des micros DPA 4011 mesurées.....	58

Table des illustrations

figure 47. Mesure de directivité du système réel	59
figure 48. Représentation en 'waterfall' du prototype non égalisée.	60
figure 49. Diagramme de directivité du prototype.....	61
figure 50. Réponse en fréquence du prototype dans l'axe après égalisation	62
figure 51. Directivité du microphone mesurée dans le plan vertical en hauteur.....	63
figure 52. Directivité du microphone mesurée dans le plan vertical vers le bas	63
figure 53. Directivité du microphone mesurée en place dans une prise de son coïncidente.....	64
figure 54. Diagramme polaire théorique du système 1.....	69
figure 55. Enregistrement du test.....	70
figure 56. Dispositif d'enregistrement avec le système 2	70
figure 57. Position des deux systèmes pour l'enregistrement.....	71
figure 58. Déviation moyenne pour les deux systèmes	74
figure 59. Déviation théorique des deux systèmes	75
figure 60. Résultats complets pour le chant pour le système 1.....	76
figure 61. Résultats complets avec le chant pour le système 2.....	76
figure 62. Résultats complets avec les maracas pour le système 1.....	77
figure 63. Résultats complets avec les maracas pour le système 2.....	77
figure 64. Résultats complets avec les claves pour le système 1	78
figure 65. Résultats complets pour les claves pour le système 2.....	78
figure 66. Résultats complets avec les woodblock pour le système 1	79
figure 67. Résultats complets avec le woodblock pour le système 2	79
figure 68. Test de robustesse à 50 cm du centre	81
figure 69. Synoptique du câblage de la console Yamaha 03D	98

10 Annexes

10.1 Annexe 1. Texte de présentation du test

Texte Enregistré en préambule

« Bonjour à vous qui avez la bonté de vous soumettre à un test sonore.

Ce test est réalisé pour obtenir des informations sur la localisation du son dans l'espace.

On vous demande, en vous orientant vers la lumière jaune clignotante, de noter le plus précisément possible la direction du son perçue en se référant aux numéros inscrits sur les étiquettes

Pour cela vous vous tournerez vers la direction présumée une fois le son éteint.

Vous entendrez chaque test une première fois suivi d'une pause pour vous retourner et revenir en position initiale puis vous entendrez le test une seconde fois suivi d'une pause de 7 secondes vous permettant de confirmer votre impression initiale et d'inscrire votre réponse sur la fiche test.

Vous noterez ainsi successivement toutes les positions sur la fiche devant vous.

Dans le cas où la direction vous semble imprécise, notez la référence se situant au centre de la zone d'incertitude. Si vous n'arrivez pas du tout à localiser le son, notez un point d'interrogation.

Vos commentaires sont les bienvenus, une case en bas de la feuille est prévue à cet effet.

Voici à titre d'exemple un essai.

--- essai ---

Logiquement votre sensation doit se situer entre les numéros 7 et 14.

Le test commence, à vous de jouer et merci de votre participation. »

10.2 Annexe 2. Fichier Matlab de visualisation en diagrammes polaires

```
% Visualisation des résultats des tests sous forme de diagramme polaire
clear all
hold off
% Traitement des Données
[A]=dlmread('résultarepli30.txt','\t');
Theta=A(:, [1]);Theta=Theta/180*pi;
R=A(:, [2])+10;
L=A(:, [3])+10;
LS=A(:, [4])+10;
RS=A(:, [5])+10;
%R=10.^(R/20);
%L=10.^(L/20);
%LS=10.^(LS/20);
```

```

%RS=10.^(RS/20);
LP=polar(Theta,LS);
set(LP,'LineWidth',2,'color',[0 0 0],'LineStyle','-');
hold on
%RP=polar(Theta,L)
%RSP=polar(Theta,RS)
%LSP=polar(Theta,LS)
%set(RP,'Color',[0 1 0])
%set(LP,'Color',[1 0 0])
%set(RSP,'Color',[0 0 1])
%set(LSP,'Color',[0 0 0])
%Données théoriques
t=[0:0.01:2*pi];
RHO=128/180*pi;
A=1;B=1;D=0.05;C=340;F=108;W=2*pi*F;k=W/C;
x=(A*(1+cos(t+RHO))/2.*exp(j*W*t));
y=(B*(1+cos(t+RHO))/2.*exp(j*W*t-j*k*D*cos(t+RHO)));
% Visualisation
SO=abs(x-y)/max(abs(x-y))+0.03;
SO=20*log10(SO)+30;
%SO=polar(t,SO)
%set(SO,'Color',[0.5 0.5 0])
CO=0.5+0.5*cos(t+RHO);
CO=abs(CO)/max(abs(CO))+0.03;
CO=20*log10(CO)+30;
CO=polar(t,CO);
set(CO,'LineWidth',2,'color',[0.6 0.6 0.6],'LineStyle','-');
%set(CO,'Color',[0 0.5 0.5])
%legend({num2str('Avant droite'),num2str('Avant gauche'),num2str('Arrière
droite'),num2str('Arrière gauche'),num2str('2nd or-
dre'),num2str('Cardioïde')})
legend({num2str('Arrière'),num2str('Cardioïde')})

```

10.3 Annexe 3. Fichier Matlab des simulations de réponse en fréquence du prototype

```

%Courbe de module et phase du prototype
%variables
F=[20:100:20000];
phi=pi/2;
a=1;b=0.707;c=0.707;
D1=0.08;f1=1000;
D2=0.015;f2=1000;
W=2*pi*F;
k=W/340;
t=0;

%fonctions
x=(a*(1+cos(t))/2.*exp(j*W*t));
y=(b*(1+cos(t))/2.*exp(j*W*t-j*k*D1*cos(t)));
hpf=1./(1+(2*pi*f1)./(j*W))*1./(1+(2*pi*f1)./(j*W));
z=(c*(1+cos(t))/2.*exp(j*W*t-j*k*D2*cos(t)));
lpf=1./(1+j*W/(2*pi*f2))*1./(1+j*W/(2*pi*f2));

%graphisme
figure(1)% modèle complet

```

```

semilogx(F,20*log10(abs((x-y).*lpf-(x-z).*hpf)),F,20*log10(imag((x-y).*lpf-
(x-z).*hpf)),F,20*log10(real((x-y).*lpf-(x-z).*hpf)));
xlabel ('Fréquence (Hz)')
ylabel ('Amplitude (dB)')
grid on
axis manual
axis ([20 20000 -30 10])

figure(2) %partie hf
semilogx(F,20*log10(abs((x-y).*lpf)),F,20*log10(imag((x-
y).*lpf)),F,20*log10(real((x-y).*lpf)))
xlabel ('Fréquence (Hz)')
ylabel ('Amplitude (dB)')
grid on
axis ([20 20000 -30 10])

figure(3)% partie bf
semilogx(F,20*log10(abs((x-z).*hpf)),F,20*log10(imag((x-
z).*hpf)),F,20*log10(real((x-z).*hpf)))
xlabel ('Fréquence (Hz)')
ylabel ('Amplitude (dB)')
grid on
axis ([20 20000 -30 10])

```

10.4 Annexe 4. Fichier Matlab de visualisation en diagrammes polaires réels et calcul du DI

```

% récupération fichier MLSSA et visualisation polaire
clear all
clf
hold off
pas=[0:5:36*5];
pasbis=[0:5:72*5];
grid on
z=load('K5mo3bis.txt');
set(gcf,'nextplot','add')
set(gca,'TickLength',[2 2])
n=0
%génération diagramme fréquence basse
for i=[1 2 4 8];
    val=z(:,i);val2=fliplr(val');
    c=[0 0 0 ;0 0 1;1 0 0;0 1 0];
    n=n+1;

    pl=polar([0:5:36*5]*pi/180,val2-max(val)+30);
    axis([-30 30 -30 30]);
    C=c(n,:);

    set(pl,'Color',C,'linewidth',2);
    hold on;
end
n=0;
%génération diagramme fréquence haute
for i=[17 34 67 134];
    c=[0.5 0.5 0.5;1 0 1;1 1 0;0.1 0.9 1];
    val=z(:,i)';

```

```

        n=n+1;
        a=ones(length(val)-1)*(-30)+max(val)';
        val=[a(1,:) val]';
        pl=polar(pasbis*pi/180,val-max(val)'+30);
        axis([-30 30 -30 30])
        C=c(n,:);
        set(pl,'Color',C,'linewidth',2);
        hold on;
end
legend({num2str('125 Hz'),num2str('250 Hz'),num2str('500 Hz'),num2str('1
kHz'),num2str('2 kHz'),num2str('4 kHz'),num2str('8 kHz'),num2str('16
kHz')},0)
grid on

% calcul indice de directivité à 2000Hz
val=z(:,17);
val=val-max(val);

val=10.^(val/20);
val2=fliplr(val);
pas=[0:5:36*5];
val3=val2.^2.*sin(pas/180*pi)*0.0873;
val4=sum(val3)/2
Q=1./val4
ID=10*log10(Q)

```

10.5 Annexe 5. Fichier Matlab de Visualisation des courbes exportées ou simulées du prototype

```

%Simulation des courbes de directivité du prototype
clear all
hold off
close all
% Variables
t=0;
A=1;
B=0.707;
E=0.707;
D1=0.015;f1=1000;
D2=0.08;f2=1000;
C=340;
F=100;
N=36;
S=zeros(37,169);

%z=10.^(z/20);
%z1=max(z);
%z3=z1*ones(length(pas),1)';
%z3=z./z3';
%z=20*log10(z3);
% Récupération fichier MLSSA Waterfall
z=load('4011bis.txt');
f=[118.37:118.37:20004.73];
pas=[36*5:-5:0];
x=f'*ones(length(pas),1)';
y=pas'*ones(length(f),1)';

figure(1)

```

```

set(figure(1), 'name', 'normal')
%surf(x,y',z');
zb=10.^(z/20);
zc=max(zb);
zd=zc'*ones(length(pas),1)';
zd=zb./zd';
ze=20*log10(zd);
contour(x,y',ze',[12 9 6 3 0 -3 -6 -9 -12 -15 -18 -21 -24 -27 -30])
set(gca, 'XScale', 'log', 'XDir', 'normal', 'YDir', 'reverse');
shading interp;
axis([100 20000 0 180 -30 20 ]);
rotate3d
z1=10.^(z/20);
z=z1;
N=36;
for i=[0:N];
    t=(5*(i+0.04))/180*pi;
    z2=z1((37-i),:);
    for n=[1:169];
        f=n*118.37;k=2*pi*f/C;
        z3=z2(:,n);
        %z2(:,n)=z3.*exp(-
j*k*D1*cos(t)).*1./(1+(f2./(j*f))).*1./(1+(f2./(j*f)));
        z5(:,n)=z3.*(1-B*exp(-
j*k*D1*cos(t)).*1./(1+(f2./(j*f))).*1./(1+(f2./(j*f)));

        end
        %Z1((37-i),:)=z2;
        Z5((37-i),:)=z5;
end

N=36;
for i=[0:N];
    t=(5*(i+0.04))/180*pi;
    z2=z1((37-i),:);
    for n=[1:169];
        f=n*118.37;k=2*pi*f/C;
        z3=z2(:,n);
        %z2(:,n)=z3.*exp(-
j*k*D2*cos(t)).*1./(1+(f2./(j*f))).*1./(1+(f2./(j*f)));
        %z4(:,n)=z3.*exp(-j*k*D2*cos(t));
        z6(:,n)=z3.*(1-E*exp(-
j*k*D2*cos(t)).*1./(1+(j*f)/(f1))*1./(1+(j*f)/(f1)));

        end

        %Z2((37-i),:)=z2;
        %Z4((37-i),:)=z4;
        Z6((37-i),:)=z6;
end

figure(2)
set(figure(2), 'name', 'complet')
surf(x,y',20*log10(abs(Z5'-Z6')));
%contour(x,y',20*log10(abs(Z5'-Z6')),[12 9 6 3 0 -3 -6 -9 -12 -15 -18 -21 -24
-27 -30])
set(gca, 'XScale', 'log', 'XDir', 'normal', 'YDir', 'reverse');
shading interp;rotate3d
axis([100 20000 0 180 -30 20 ]);
% Visu partie haure fréquence
figure(3)

```

```

set (figure(3), 'name', 'hf')
surf(x,y',20*log10(abs(Z5')));
set(gca,'XScale','log','XDir','normal','YDir','reverse');
shading interp;rotate3d
axis([100 20000 0 180 -30 20]);
%Visu partie basse fréquence
figure(4)
set (figure(4), 'name', 'lf')
surf(x,y',20*log10(abs(Z6')));
set(gca,'XScale','log','XDir','normal','YDir','reverse');
shading interp;rotate3d
axis([100 20000 0 180 -30 20]);
%Visu complète 2D par pas de 3dB
figure(5)
Z9=Z5-Z6;
%Z7=20*log10(abs(Z9(37,:)))*ones(length(pas),1)';Z7=Z7;
%Z8=20*log10(abs(Z9))-Z7';
%z=10.^((Z9)/20);
z1=max(Z9);
z3=z1*ones(length(pas),1)';
z3=Z9./z3';
z=20*log10(z3);

set(gca,'XScale','log','XDir','normal','YDir','reverse');
shading interp;
axis([100 20000 0 180 -30 10]);
[e,g]=contour(x,y',z',[12 9 6 3 0 -3 -6 -9 -12 -15 -18 -21 -24 -27 -30]);
axis([100 20000 0 180 -30 10]);
rotate3d
grid on
set(gca,'YDir','reverse')
shading interp
set (gca , 'xscale', 'log')
view([0,90]);
clabel(e,g,'manual');
xlabel('Fréquence (Hz)');ylabel('Angle (degré)');

colormap(gray)
caxis([-1 3])

```

10.6 Annexe 6. Fichier Matlab de visualisation 3D de sommation de microphone

```

%Visualisation 3D de sommation microphone
clear all
close all
n = 37;
% Angles des microphones
r=1;
azi1=pi/4;ele1=0;
azi2=-pi/4;ele2=0;
azi3=-2*pi/3;ele3=0;
azi4=+2*pi/3;ele4=0;

A=1;
r=0.5;
a=1;
b=1;
c=0.5;

```

```

D1=0.01;f1=1000;
D2=0.1;f2=1000;
C=340;
f=500;W=2*pi*f;
k=W/C;
theta =(-n:2:n)/n*pi;

phi =(-n:2:n)'/n*pi/2;
%x=(a*(1+cos(t))/2);
%y=(b*(1+cos(t))/2.*exp(-j*k*D1*cos(t)));
%z=(c*(1+cos(t))/2.*exp(-
j*k*D2*cos(t)))*1./(1+j*W/(2*pi*f1))*1./(1+j*W/(2*pi*f2));
%T=(abs(x-y-z));

X=(cos(phi)*cos(theta));
Y=(cos(phi)*sin(theta));
Z=(sin(phi))*ones(1,n+1);
%r=abs(A*(0.5+(0.5)*(X*cos(ele1)*cos(azi1)+Y*sin(azi1)*cos(ele1)+Z*sin(ele1))
));
r=abs(a*(0.5+(0.5)*(X*cos(ele1)*cos(azi1)+Y*sin(azi1)*cos(ele1)+Z*sin(ele1))
-...
b*(0.5+(0.5)*(X*cos(ele1)*cos(azi1)+Y*sin(azi1)*cos(ele1)+Z*sin(ele1))...
.*exp(-
j*k*D1*(X*cos(ele1)*cos(azi1)+Y*sin(azi1)*cos(ele1)+Z*sin(ele1)))+...%mic2
a*(0.5+(0.5)*(X*cos(ele2)*cos(azi2)+Y*sin(azi2)*cos(ele1)+Z*sin(ele1)))-...
b*(0.5+(0.5)*(X*cos(ele1)*cos(azi2)+Y*sin(azi2)*cos(ele1)+Z*sin(ele1))...
.*exp(-
j*k*D1*(X*cos(ele1)*cos(azi2)+Y*sin(azi2)*cos(ele1)+Z*sin(ele1)))+...%mic3
A*(0.5+(0.5)*(X*cos(ele1)*cos(azi3)+Y*sin(azi3)*cos(ele1)+Z*sin(ele1)))+...%m
ic4
A*(0.5+(0.5)*(X*cos(ele1)*cos(azi4)+Y*sin(azi4)*cos(ele1)+Z*sin(ele1))));
cosphi = cos(phi); cosphi(1) = 0; cosphi(n+1) = 0;
sintheta = sin(theta); sintheta(1) = 0; sintheta(n+1) = 0;
x = r.*(cosphi*cos(theta));
y = r.*(cosphi*sintheta);
z = r.*(sin(phi)*ones(1,n+1));
    mesh(x,y,z)
    axis equal

shading interp
rotate3D

```

10.7 Annexe 7. Egalisation du prototype avec la console Yamaha

Le tableau 5 donne les paramètres d'égalisation utilisés pour compenser les défauts des valeurs axiales visibles de la figure 48. Ces paramètres sont donnés à titre indicatif pour ceux qui souhaiteraient utiliser dans le futur cette même configuration. Ces égalisations étant ainsi reproductibles sur toutes les consoles Yamaha largement diffusées.

Egalisation	Low	Low-mid	High-mid	High
Gain	8	3	-6	0
Fréquence	250 Hz	2.82 kHz	10 kHz	13.4 kHz
Facteur Q	2	2.5	0.8	7

tableau 5. Réglage des correcteurs paramétriques de la console Yamaha 03-D

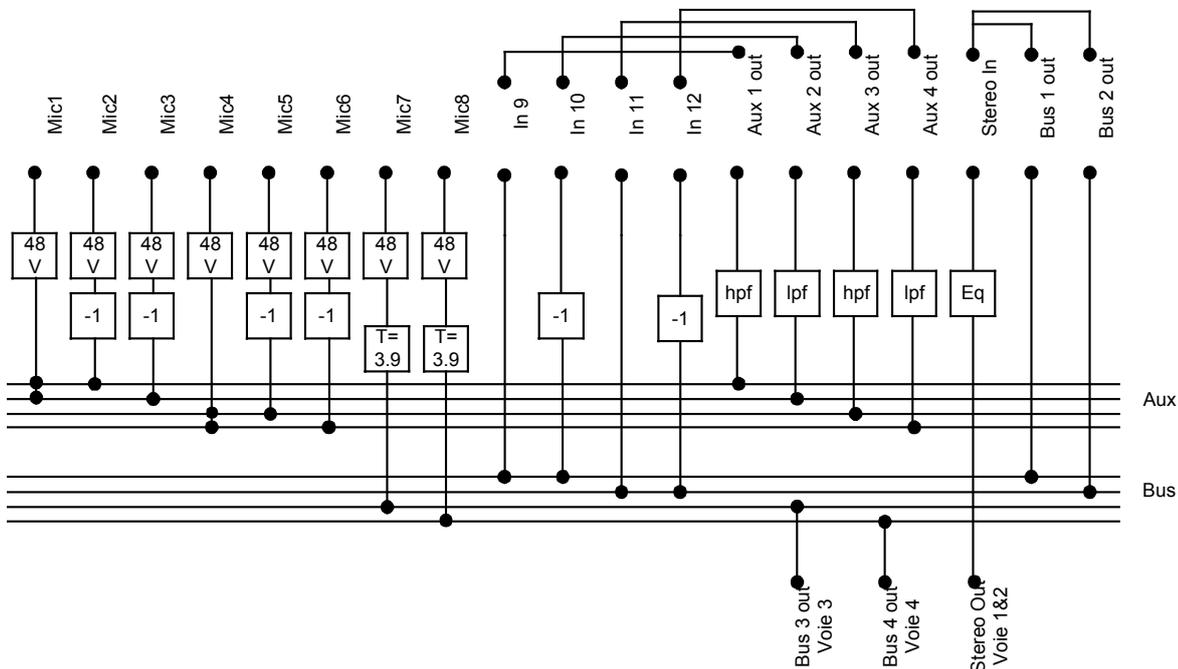
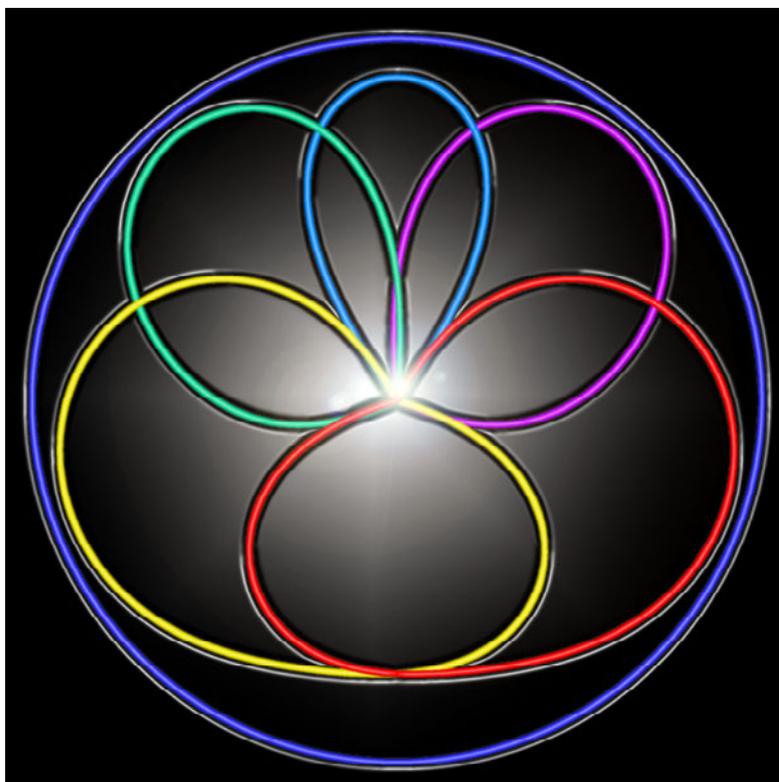


figure 69. Synoptique du câblage de la console Yamaha 03D

On trouvera sur la figure 69, le synoptique complet d'un câblage possible de la console O3-D correspondant à deux microphones du 2nd ordre pour la prise de son des tests subjectifs finaux. Le traitement numérique et les réinsertions induisent des retards de traitement qu'il est nécessaire de compenser pour les voies 3 & 4 qui ne comportent qu'un seul microphone. Le retard mesuré par MLSSA en mode autocorrélation est de 3.9ms. Les inversions de phase sont représentées par des symboles de label -1. Les filtres passe-haut et passe-bas du second ordre sont représentés respectivement par les cases hpf et lpf. L'égalisation finale est assurée par le bloc Eq sur une voie stéréo. Ce câblage utilise presque toute la ressource d'entrée et de sortie de la console.



La prise de son 5.0 en Haute Résolution Spatiale



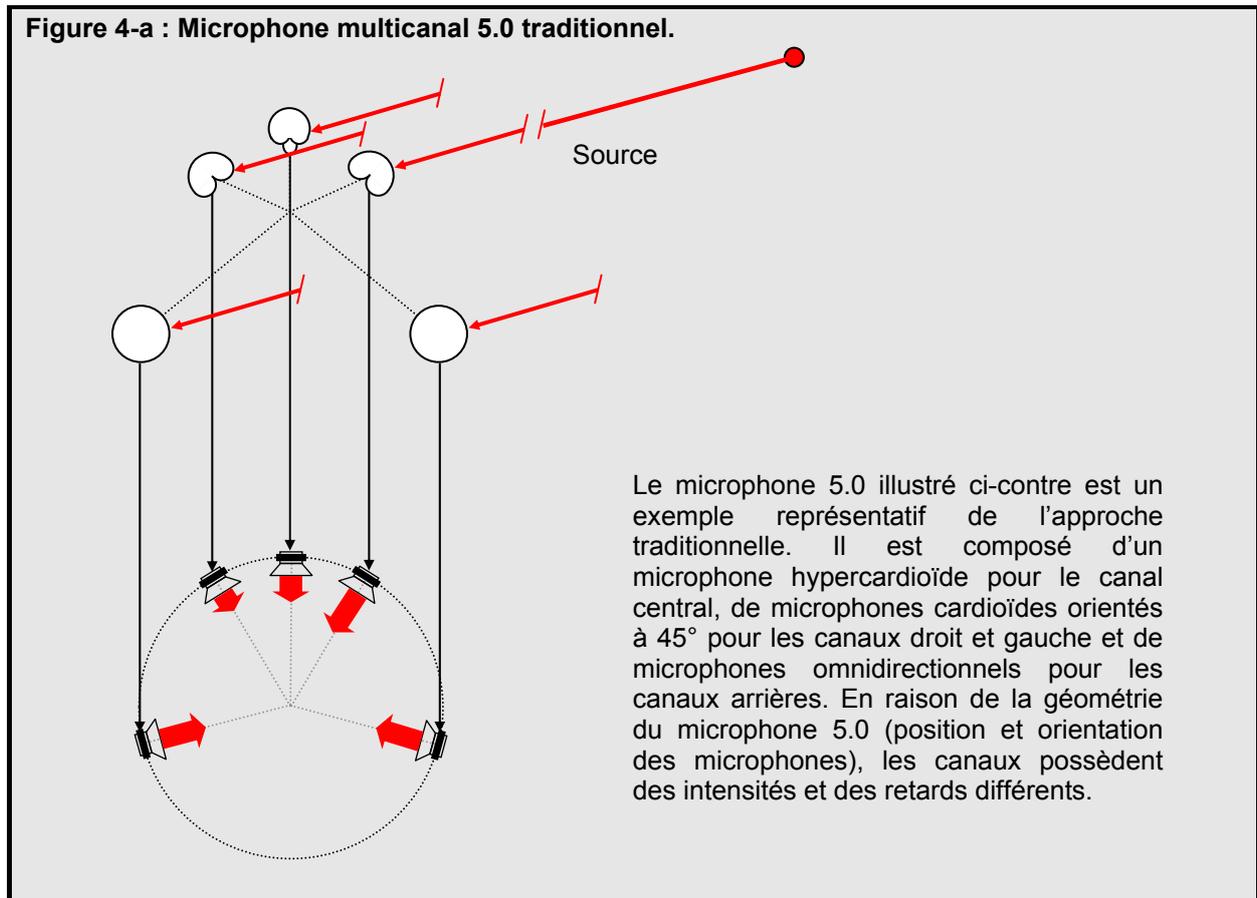
La prise de son 5.0 en Haute Résolution Spatiale

1 La haute résolution spatiale

1.1 Les limites de la prise de son multicanale traditionnelle

Actuellement, tous les systèmes de prise de son multicanal cherchent à exploiter au mieux les directivités des microphones actuels. En effet, la démarche adoptée est directement issue de la stéréophonie et consiste à produire des différences d'intensité et de retard entre les signaux captés en agissant respectivement sur l'orientation des microphones et sur leurs distances respectives. La position des sources sonores autour du microphone est « enregistrée » grâce à ces différences d'intensité (ΔI) et de temps (ΔT). Dans cette approche, toute la difficulté consiste à déterminer la combinaison optimale d'orientation et d'écartement des microphones. Cette approche est illustrée par la figure 4-a.

Figure 4-a : Microphone multicanal 5.0 traditionnel.



Or cette démarche est fondamentalement limitée car les microphones actuels n'offrent qu'un nombre restreint de directivités. Il s'agit de la palette de directivités allant de la figure omnidirectionnelle à la figure bidirectionnelle (figure-8) en passant par les intermédiaires (infracardiïde, cardioïde, hypercardioïde). Ainsi, les systèmes de prise de son traditionnels ne peuvent pas reproduire toute la diversité des signaux que les systèmes de diffusion multicanal permettent de reproduire. Aussi, il semble légitime de se demander si cette approche conduit réellement à une prise de son multicanale optimale.

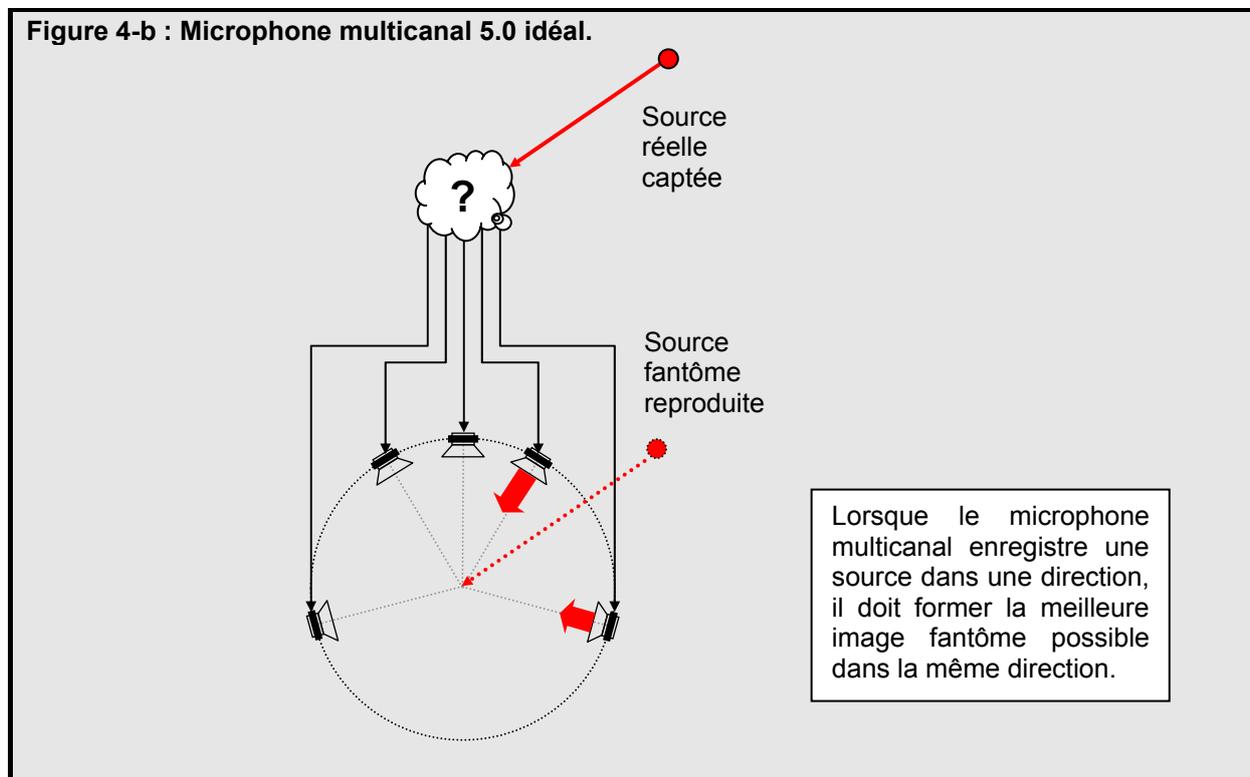
Par ailleurs, bien que fondamentalement limitée, l'approche traditionnelle permet de définir un nombre colossal de microphones multicanaux dont les différences seraient parfaitement audibles. En effet, en

jouant sur seulement 5 types de directivité, 10 orientations et 50 positions (5 angles et 10 rayons), on dénombre déjà plus de 15 milliards de systèmes de prise de son différents ! Il est donc illusoire d'espérer pouvoir les répertorier et les tester. Signalons qu'avec la même liberté, il n'existe que 2500 systèmes stéréophoniques. Par conséquent le passage de la stéréophonie à 2 canaux au multicanal à 5 canaux se traduit par une explosion des combinaisons possibles en matière de système de prise de son. Dès lors, l'approche empirique d'essais/erreurs sur laquelle s'est fondée la stéréophonie semble également révéler ses limites.

1.2 Quelle est la prise de son multicanale optimale

Face à la complexité de la prise de son multicanale, **Trinnov Audio** aborde le problème sous un angle totalement nouveau. Faisons temporairement abstraction des directivités des microphones actuels et essayons de répondre à la question : **quel serait le comportement idéal d'un système permettant une prise de son 5.0 optimale sur 360° ?**

Le réponse est triviale : Lors d'un enregistrement idéal d'une source sonore dans une direction donnée, le système de restitution 5.0 doit former la meilleure source fantôme possible dans la même direction que la source enregistrée. Idéalement, la seule limitation tolérée est la limitation intrinsèque du système de restitution 5.0. Le système de prise de son ne doit pas apporter de limitation supplémentaire. Le concept de microphone multicanal idéal est illustré par la figure 4-b.

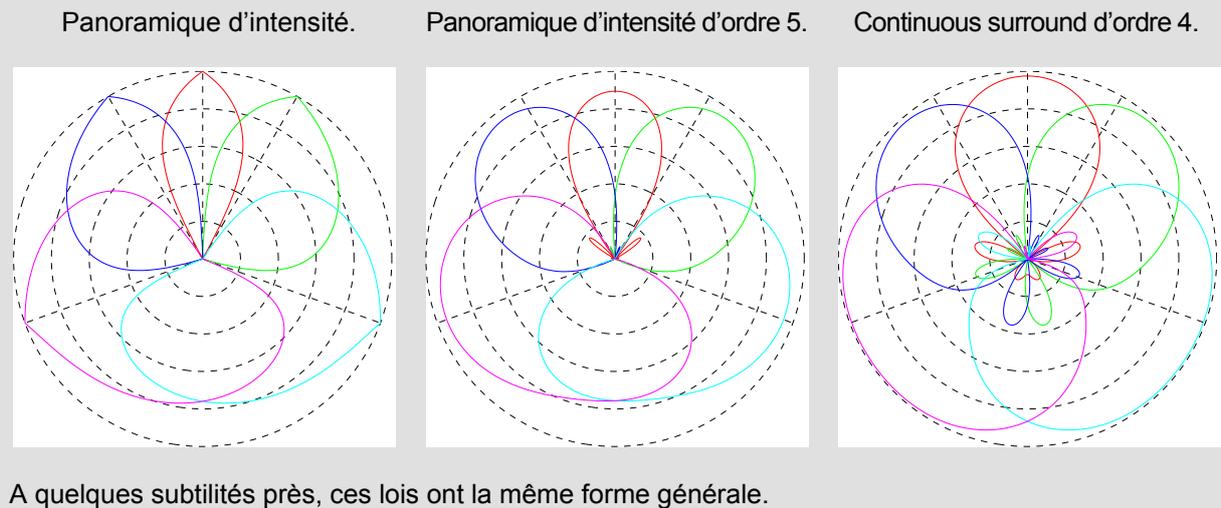


Ainsi définie, la prise de son multicanal optimale correspond à une loi de panoramique (également appelée loi de pan-pot ou encore loi de monophonie dirigée). En effet, ces lois ont été précisément étudiées pour former des images fantômes optimales sur 360°. Une loi de panoramique 5.0 est définie par 5 figures de directivités qui expriment l'intensité des canaux pour former une image fantôme optimale dans chaque direction. **Ainsi, les directivités d'une loi de panoramique peuvent s'interpréter comme les directivités idéales d'une prise de son multicanale optimale.** Il existe de nombreuses lois de panoramique, mais à quelques subtilités près, ces lois ont toujours la même forme générale. A titre d'exemple citons les lois suivantes :

- Le panoramique d'intensité.
- Le panoramique d'intensité d'ordre 5. Le concept « d'ordre » sera défini ultérieurement.
- La loi continuous surround panning d'ordre 4 (P. Craven, 2003).

Ces 3 exemples sont illustrés ci-dessous.

Figure 4-c : Directivités idéales pour une prise de son 5.0 optimale



1.3 Concept de prise de son multicanale à haute résolution spatiale

Le concept de haute résolution spatiale s'inspire directement du concept de haute résolution temporelle (également appelé haute fidélité).

On parle d'enregistrement à haute résolution temporelle lorsque le signal est enregistré avec une très grande précision même dans ses variations les plus rapides. Pour pouvoir suivre ces variations rapides, le système d'enregistrement doit être capable d'enregistrer les fréquences élevées et donc avoir une bande passante étendue (jusqu'à au moins 20kHz). Par conséquent, la haute résolution temporelle impose une large bande passante.

Par analogie, on parle d'enregistrement à haute résolution spatiale lorsque le champ acoustique est enregistré avec une très grande précision même dans ses variations spatiales les plus rapides. Les variations spatiales correspondent à la répartition angulaire du champ acoustique autour du microphone. Pour pouvoir suivre ces variations spatiales rapides, le système d'enregistrement doit être capable d'enregistrer les fréquences spatiales élevées et donc avoir une bande passante spatiale étendue.

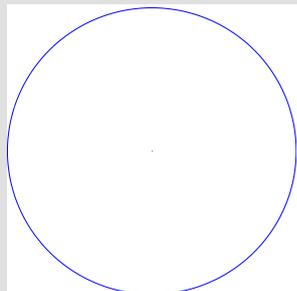
Une fréquence spatiale correspond en fait à une figure de directivité dont la finesse des lobes détermine directement la précision spatiale. Plus la figure de directivité est sélective, plus la résolution spatiale qu'elle offre est élevée. En effet, une figure de directivité sélective possède une meilleure aptitude à discriminer les sources sonores dans l'espace. Afin de définir un spectre spatial, les fréquences spatiales sont classées en fonction de leur précision spatiale appelée « ordre ». La figure 4-d illustre certaines fréquences spatiales jusqu'à l'ordre 5.

Comme nous l'avons déjà souligné, les microphones actuels n'offrent qu'un nombre restreint de directivités. En réalité ces directivités ne sont que des mélanges de directivité omnidirectionnelles et bidirectionnelles. Par conséquent, la résolution spatiale atteignable est toujours limitée à l'ordre 1.

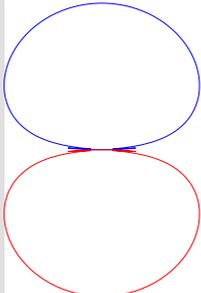
Au contraire, les directivités idéales pour une prise de son 5.0 optimale correspondent à un mélange de fréquences spatiales qui atteint l'ordre 5. Ainsi, une prise de son 5.0 optimale requiert une précision spatiale d'ordre 5. Autrement dit, la précision spatiale optimale est 5 fois supérieure à celle obtenue avec les microphones actuels. Le concept de haute résolution spatiale repose donc sur un fondement scientifique et technique extrêmement profond. Le figure 4-e illustre la différence de sélectivité entre l'ordre 1 et l'ordre 5 pour les directivités hypercardioïdes et cardioïdes.

Figure 4-d : Fréquences spatiales jusqu'à l'ordre 5.

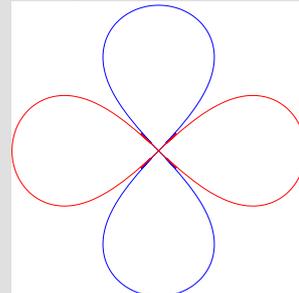
Fréquence spatiale d'ordre 0
(une directivité omnidirectionnelle)



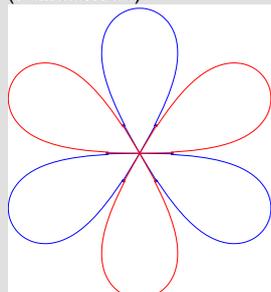
Fréquences spatiale d'ordre 1
(3 directivités bidirectionnelles)



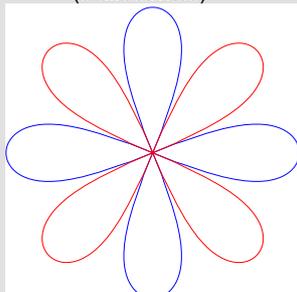
Fréquences spatiale d'ordre 2
(5 directivités):



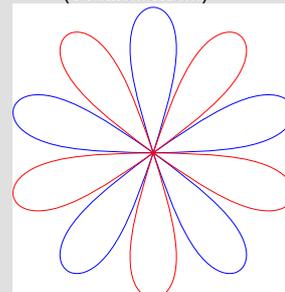
Fréquence spatiale d'ordre 3
(7 directivités)



Fréquence spatiale d'ordre 4
(9 directivités)



Fréquence spatiale d'ordre 5
(11 directivités)

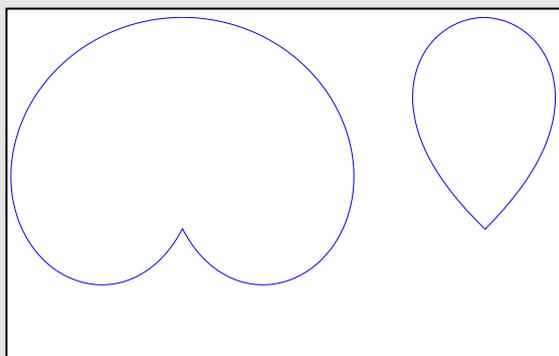


Dans ces figures, les lobes positifs des fréquences spatiales sont représentés en bleu et les lobes négatifs en rouge.

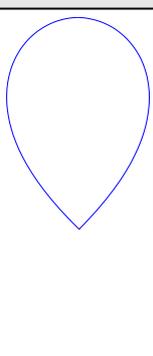
Figure 4-e : Comparaison des résolution spatiale d'ordre 1 et d'ordre 5.

Cardioïde :

Ordre 1

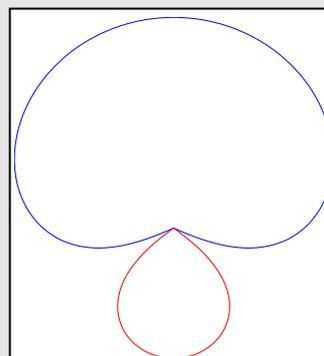


Ordre 5

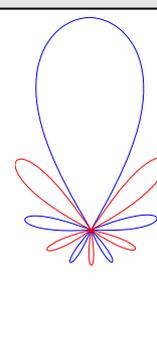


Hypercardioïde :

Ordre 1



Ordre 5



Dans ces figures, les lobes positifs des fréquences spatiales sont représentés en bleu et les lobes négatifs en rouge.

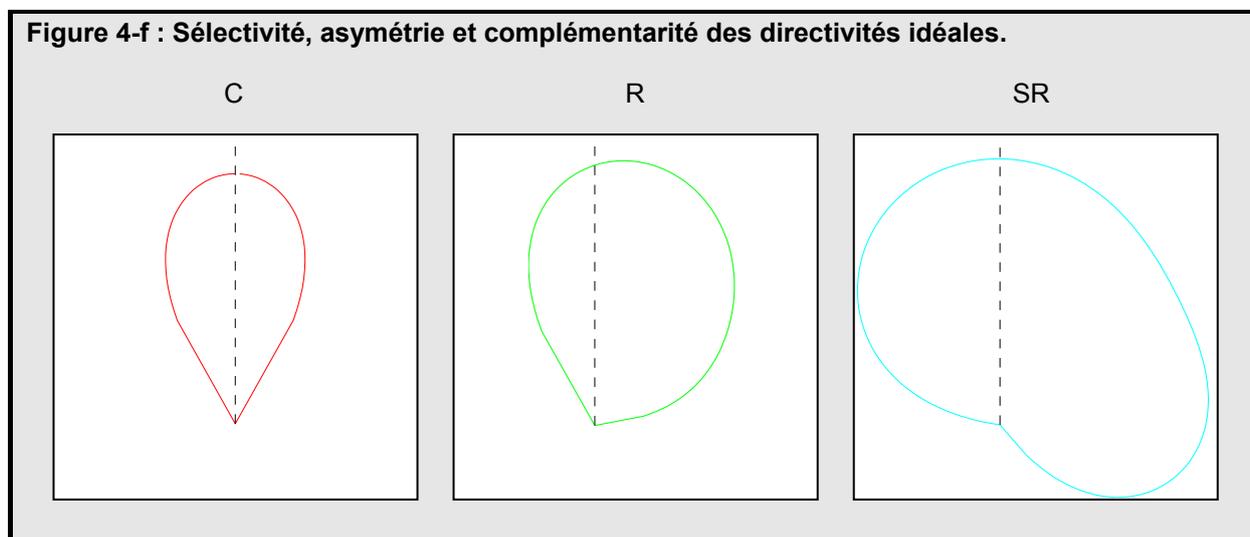
Si des microphones très sélectifs étaient disponibles, une prise de son à haute résolution spatiale serait obtenue de manière très simple. Il suffirait de placer 5 microphones coïncidents dont les directivités correspondraient aux directivités idéales définies précédemment.

1.4 La haute résolution spatiale requiert une nouvelle approche

Les directivités nécessaires pour obtenir un prise de son 5.0 à haute résolution spatiale présentent les caractéristiques suivantes :

- Elles sont très sélectives. C'est cette sélectivité spatiale qui est à l'origine de la haute résolution spatiale. Celle-ci est évidente pour les 3 directivités frontales qui offrent une sélectivité d'ordre 5 pour toute la zone frontale (+30° / -30°). Comme cela a déjà été souligné, la sélectivité requise est 5 fois plus élevée que la sélectivité des microphones actuels. Signalons également que les directivités arrières présentent également une sélectivité spatiale élevée (bande passante spatiale d'ordre 3).
- Elles présentent des asymétries. Hormis la directivité associée au canal central, les 4 autres directivités sont asymétriques. Cette caractéristique est indispensable si l'on souhaite une adaptation optimale à l'irrégularité de la configuration 5.0. Si les haut-parleurs avaient été disposés selon un pentagone régulier (0°, +/-72° et +/-144°) les directivités idéales auraient été symétriques. Les microphones actuels sont incapables de produire des directivités asymétriques.
- Elles sont complémentaires. Les directivités qui produisent des images fantômes entre 2 haut-parleurs ont un comportement parfaitement complémentaire de sorte que les images fantômes soient produites avec la même intensité dans toute la zone comprise entre les deux haut-parleurs.

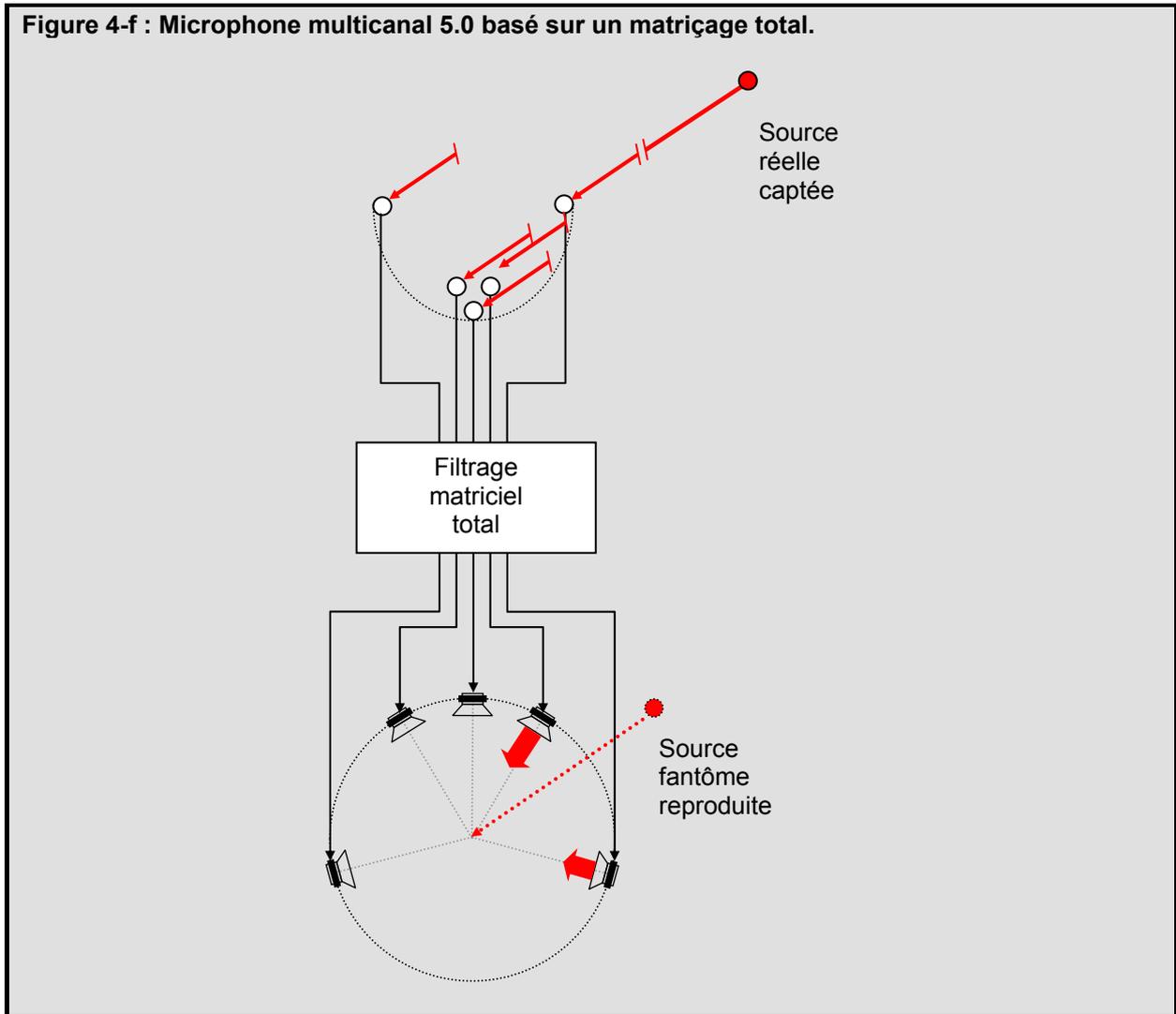
Ces caractéristiques sont illustrées par la figure 4-f.



Les directivités nécessaires pour un enregistrement à haute résolution spatiale ne sont donc pas obtenues naturellement par les microphones actuels. Pour atteindre les performances requises, il est nécessaire de **renoncer à l'approche traditionnelle qui consiste à capter chaque canal avec un unique microphone. Il faut s'orienter vers une approche plus complète basée sur un matricage.**

Afin d'atteindre une qualité spatiale maximale, **Trinnov Audio** a mis en œuvre la solution ultime en matière de traitement linéaire : le **filtrage matriciel total**. Chaque canal est obtenu en tirant profit de l'ensemble des microphones. Le signal de chaque microphone est filtré de manière spécifique puis l'ensemble des signaux ainsi formés est combiné de manière à produire un canal optimum. Seul un système aussi évolué permet d'atteindre la haute résolution spatiale. La figure 4-g illustre un microphone multicanal à haute résolution spatiale basé sur cette architecture.

Figure 4-f : Microphone multicanal 5.0 basé sur un matricage total.



2 Avantages de la haute résolution spatiale

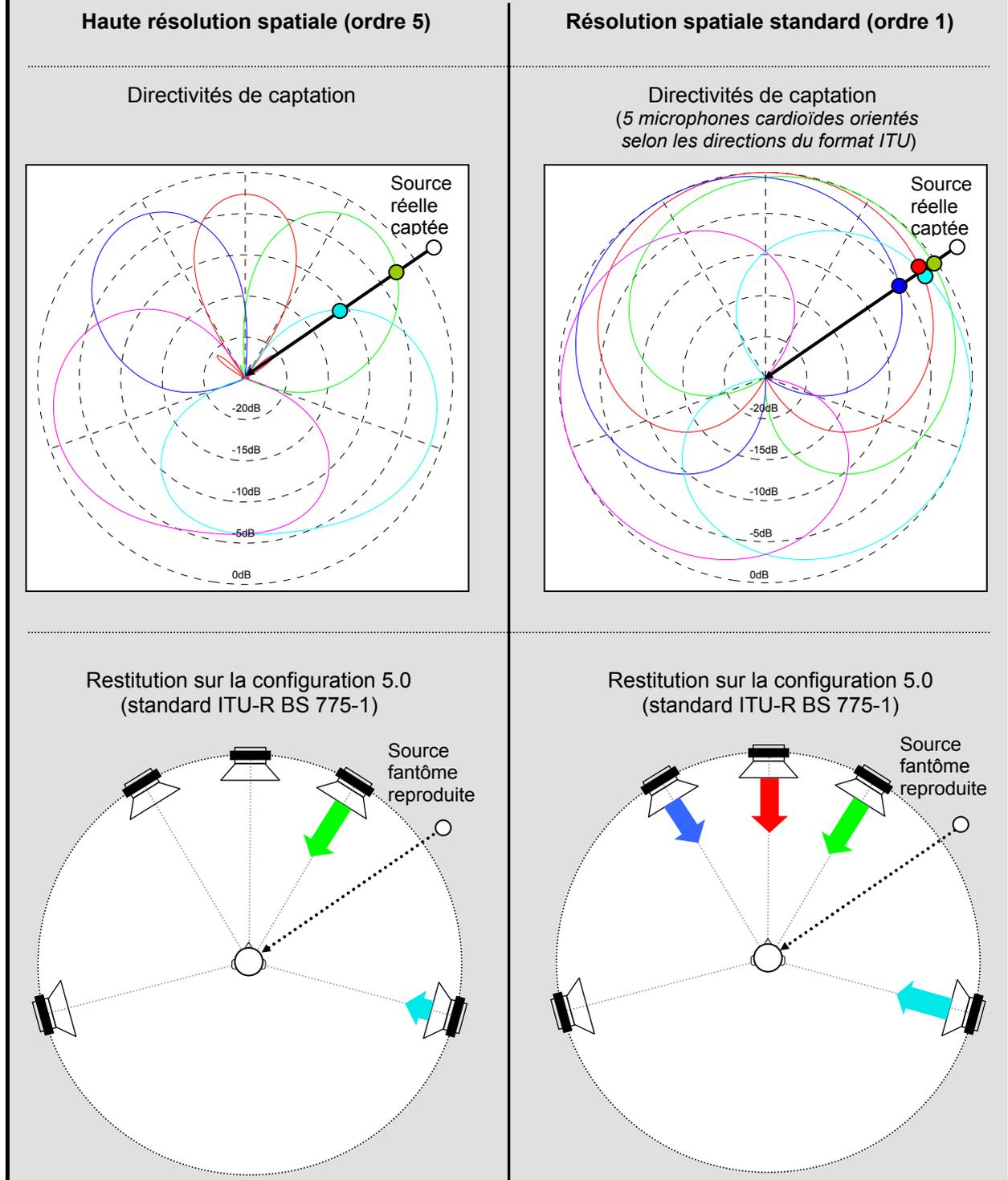
Par rapport à un enregistrement multicanal à résolution standard, un enregistrement à haute résolution spatiale apporte une sélectivité 5 fois supérieure, notamment sur la scène frontale. Dans un enregistrement à haute résolution spatiale, seuls les deux haut-parleurs au voisinage immédiat de chaque source fantôme sont actifs. Au contraire, dans le cas d'un enregistrement à résolution standard, l'ensemble des haut-parleurs est actif. Il en résulte des interactions complexes qui dégradent la qualité spatiale de l'enregistrement. La figure 4-g compare les enregistrements à haute résolution spatiale et à résolution spatiale standard.

Une prise de son à haute résolution spatiale présente de nombreux avantages par rapport à une prise de son à résolution standard. En particulier :

- Elle optimise la **séparation** des canaux.
- Elle augmente de la **précision** (la ponctualité) des sources.
- Elle augmente de la taille de la zone d'écoute (**sweet spot**)
- Elle respecte les **timbres** quelle que soit la direction des sources.
- Elle produit des images fantômes **sur 360°**,
- Elle permet un **Down-mix** facile et de bonne qualité, sans modification de timbre ni distorsion d'angle.

- Elle offre une **compatibilité** totale avec les lois de pan-pot habituelles, ce qui facilite l'ajout de microphones d'appoint.
- Elle contrôle la spatialisation des sources en dehors du plan azimutal.
- Elle permet un meilleur contrôle du **facteur de distance**.
- Elle offre un meilleur contrôle de l'angle utile de prise de son grâce à un contrôle précis des directions des images reproduites.
- Elle offre une meilleure robustesse au **traitement** : le traitement d'un canal n'affecte les images fantômes que dans son voisinage.

Figure 4-g : Comparaison des prises de son 5.0 à haute résolution spatiale et à résolution spatiale standard.



Les avantages de la prise de son à haute résolution sont présentés plus en détail dans chacune des parties suivantes.

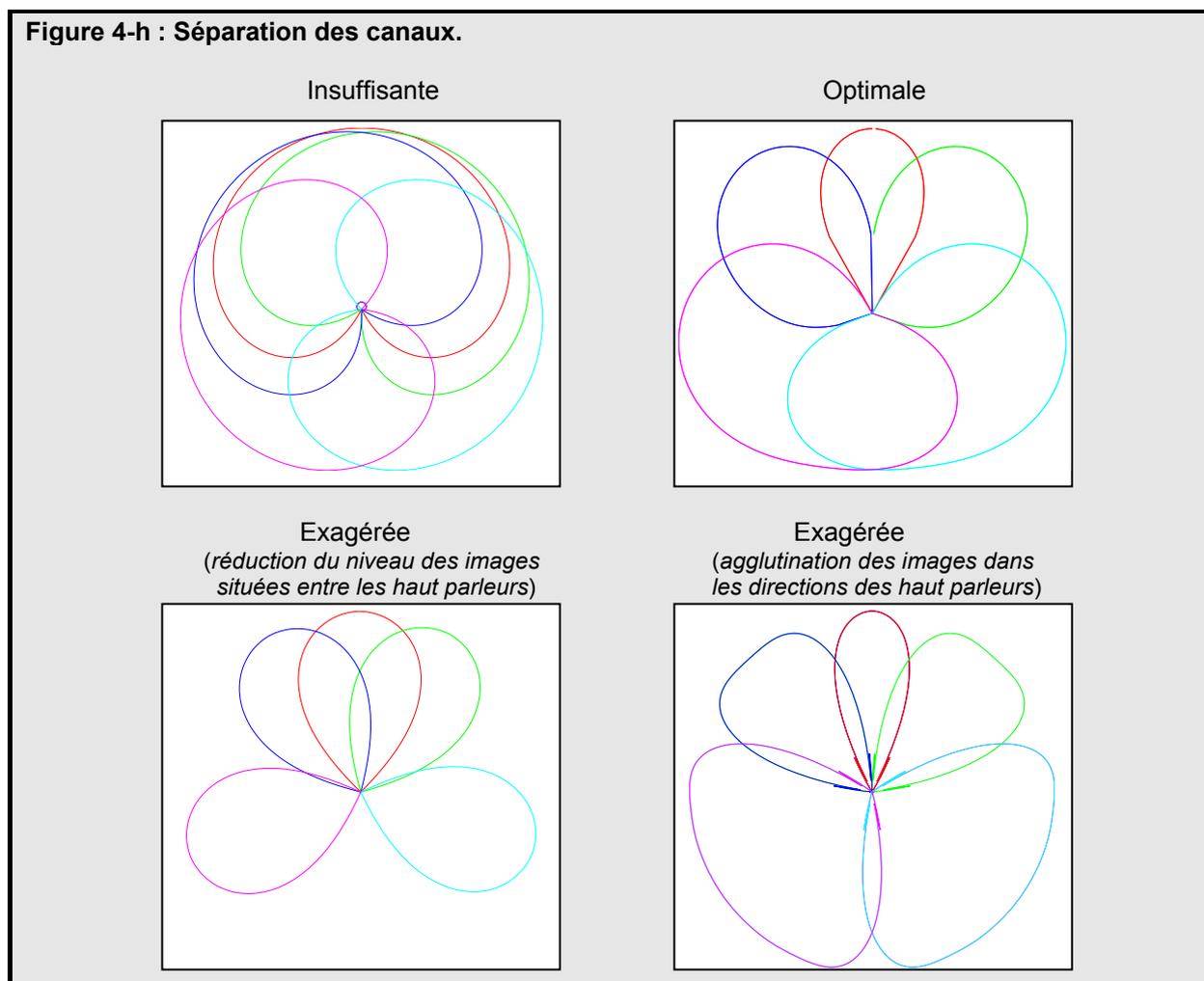
2.1 Séparation optimale des canaux

Dans un signal multicanal, la qualité spatiale globale est directement liée à la séparation (ou décorrélation) des canaux. **Une séparation insuffisante conduit à des performances spatiales limitées.** Lorsque la séparation augmente, la qualité spatiale globale augmente également. Cependant, il existe une limite au-delà de laquelle toute augmentation de la séparation nuit aux performances. En fait, la corrélation est indispensable pour produire des images fantômes. Si l'on augmente trop la séparation, deux phénomènes indésirables apparaissent :

- Les images fantômes situées entre les haut-parleurs réduisent en niveau sonore et des « trous » se forment dans l'environnement sonore.
- Les images fantômes qui apparaissent entre les haut-parleurs se déplacent vers les haut-parleurs les plus proches. Les images s'agglutinent sur les haut parleurs et des « trous » se forment dans l'environnement sonore.

En l'absence totale de corrélation, les 5 canaux véhiculent des signaux totalement différents et l'espace sonore restitué correspond à 5 sources dans les directions des 5 haut-parleurs. Donc sans corrélation, il n'y a plus d'images fantômes. **Ainsi, une séparation exagérée dégrade les images fantômes.** Par conséquent, il existe un niveau optimum de séparation des canaux. **Le but du microphone 5.0 à haute résolution spatiale est justement de réaliser une séparation optimale** (et non maximale) des canaux. La figure 4-h illustre le concept de séparation optimale des canaux.

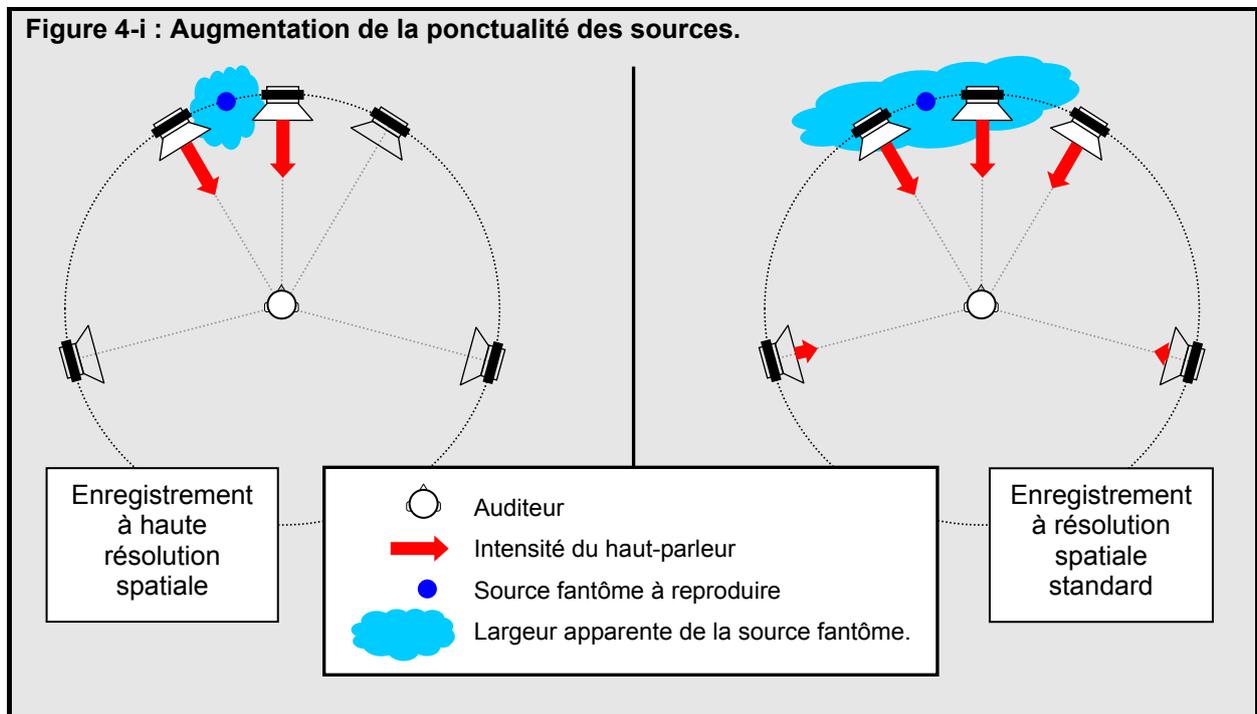
Figure 4-h : Séparation des canaux.



2.2 Augmentation de la ponctualité des sources

Dans le cas d'un enregistrement à résolution standard, le signal émit par une source se trouve répercuté sur quasiment tous les canaux. Par conséquent, l'image fantôme correspondante est reproduite par un nombre élevé de haut-parleurs. En réalité, la contribution d'un grand nombre de canaux pour reproduire une source fantôme est en général problématique. Par exemple, lorsque les 5 canaux sont actifs, ils recréent simultanément 10 images fantômes. En effet, un ensemble de 5 haut-parleurs permet de définir 10 couples qui produisent chacun une image fantôme. Ainsi, le son multicanal ne peut pas être réduit à la juxtaposition de 5 images qui apparaîtraient entre les paires juxtaposées de haut-parleurs (par exemple entre les canaux C et R). Il faut également prendre en compte les 5 autres images fantômes « parasites » qui se forment entre les paires non juxtaposées de haut-parleurs (par exemple entre les canaux L et SR). Or ces 20 images fantômes apparaissent dans des directions différentes et sont à l'origine d'informations spatiales contradictoires. Il en résulte un étalement spatial des sources fantômes. Les conflits d'images fantômes sont des problèmes spécifiques aux systèmes multicanaux et n'apparaissent pas en stéréophonie puisqu'il n'y a que 2 canaux et 1 image fantôme.

Dans le cas d'un enregistrement à haute résolution spatiale, seuls 2 canaux contribuent majoritairement à la reproduction de l'image fantôme. Il existe donc une seule et unique image fantôme il n'y a aucun problème de contradiction. Il en résulte que la source reproduite est bien localisée et très ponctuelle.



Signalons que la haute résolution spatiale supporte que plus de deux haut-parleurs soient actifs pour produire des images fantômes précises. Dans ce cas, le système est optimisé pour que la contribution globale des 5 canaux aboutisse à une reconstruction correcte des différentes fréquences spatiales du champ acoustique dans la zone d'écoute.

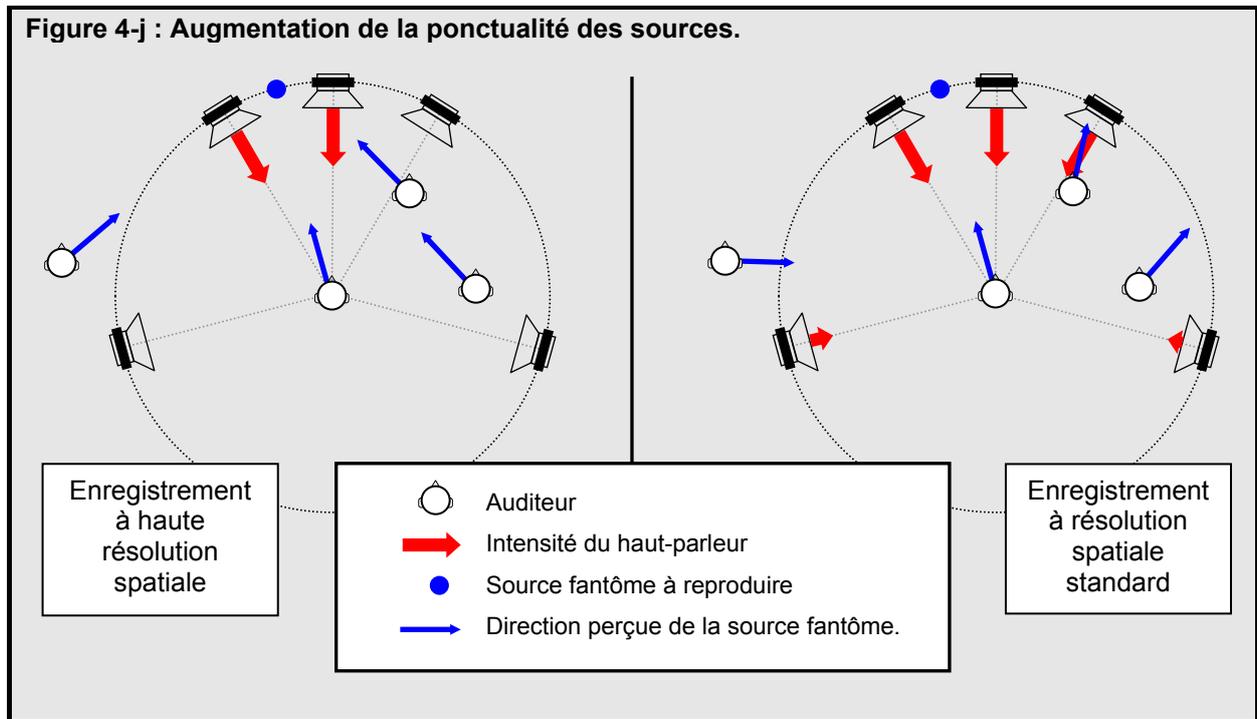
2.3 Augmentation de la taille de la zone d'écoute

Dans un enregistrement à résolution spatiale standard, le signal émit pas une source se trouve répercuté sur quasiment tous les canaux. Lorsque l'auditeur se déplace dans le local de restitution, il perçoit une intensité dominante provenant du haut-parleur le plus proche. La source fantôme reproduite se trouve attirée dans la direction du haut-parleur dominant. Il en résulte une instabilité

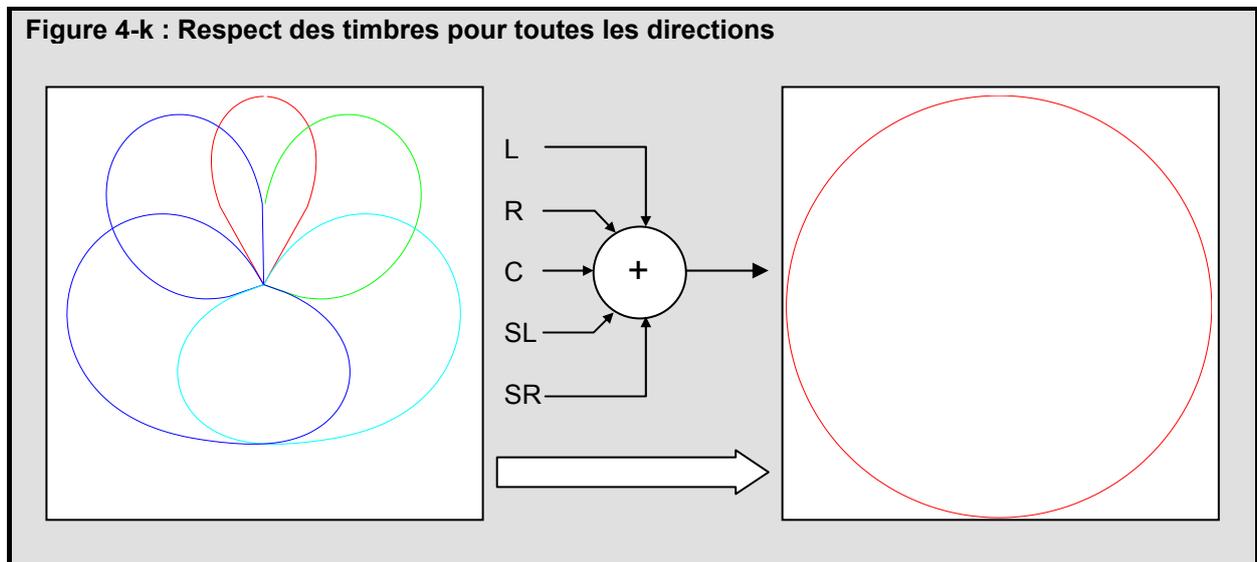
marquée des sources fantômes avec les déplacements de l'auditeur. La zone pour laquelle la spatialisation est convenablement maintenue est très restreinte.

Au contraire, **dans un enregistrement à haute résolution spatiale, seuls les haut-parleurs au voisinage immédiat de la source fantôme sont actifs.** Quelle que soit la position de l'auditeur, l'image fantôme reste localisée entre les deux haut-parleurs les plus proches. Il en résulte une grande stabilité des sources fantômes avec les déplacements de l'auditeur. Bien entendu, la stabilité est également liée à l'écart entre les haut-parleurs et les images les plus stables seront obtenues pour la zone frontale (+30°/-30°). Rappelons que la configuration ITU 5.0 a justement été étudiée pour produire une scène sonore frontale très stable. **Seul un microphone à haute résolution spatiale parvient à exploiter de manière optimale cette configuration afin de fournir une zone d'écoute maximale.**

Le figure 4-j illustre la direction perçue pour différentes positions d'écoute.



2.4 Meilleur respect des timbres dans toutes les directions



Le respect des timbres est une caractéristique garantie dans un enregistrement à haute résolution spatiale. Un système de prise de son multicanal respecte les timbres si un microphone omnidirectionnel placé au centre d'un système de restitution idéal délivre le même signal qu'un microphone omnidirectionnel placé au point d'enregistrement. Or, au centre du système de restitution, les 5 canaux se somment exactement. Donc, les timbres sont respectés si, en sommant les 5 canaux du signal 5.0, on retrouve exactement le signal qui aurait été capté par un microphone omnidirectionnel. Autrement dit, **une prise de son 5.0 préserve les timbres si le down-mix vers le format mono est parfait** (voir partie 4.6). De manière plus scientifique, cela signifie que **la fréquence spatiale d'ordre 0 du champ acoustique (composante omnidirectionnelle) doit absolument être captée et reproduite avec une grande précision pour toutes les fréquences.** Cette caractéristique est assurée par les microphones à haute résolution spatiale et est illustrée sur la figure 4-k.

La résolution spatiale peut également varier avec la fréquence, sans affecter le timbre. Il s'agit ici d'un résultat nouveau apporté par la haute résolution spatiale et qui mérite quelques explications. Habituellement, un microphone dont la directivité varie avec la fréquence introduit inévitablement des modifications de timbre. En effet, pour une direction donnée, un tel microphone capte le son avec une sensibilité qui varie en fonction de la fréquence, ce qui introduit une modification de timbre. Ce problème ne peut pas être résolu dans le cas où un unique microphone est considéré. Au contraire, dans un système multicanal il est indispensable de considérer simultanément l'ensemble des canaux. Or, des systèmes de résolution différentes peuvent tout à fait former des images fantômes dont la direction et le timbre sont très précisément reproduites. Dans ce cas, les seules différences apparaissent au niveau de la spatialisation. Il est donc tout à fait possible d'imaginer un système très performant présentant une résolution élevée dans la bande 300Hz – 5000Hz et une résolution qui diminue progressivement aux extrêmes de la bande passante (20Hz – 20000Hz). Pour s'en convaincre, rappelons que ce concept est déjà utilisé de manière caricaturale dans le format 5.1. Dans ce format, les très basses fréquences (< 80Hz) sont transmises par un unique canal « sub. » et ne comportent pas de spatialisation. Malgré une transition brutale dans la résolution spatiale, les timbres sont convenablement respectés sur toute la bande passante, y compris pour les plus basses fréquences. Bien entendu, il en est de même lorsque les transitions entre résolutions sont plus douces. Par conséquent, un système dont la résolution varie avec la fréquence peut être envisagé et conduire à des performances élevées sans pour autant introduire de modification de timbre des sources fantômes sur toute la bande passante. Bien entendu, les performances spatiales globales dépendent fortement de la résolution spatiale atteinte pour chaque fréquence. Celle-ci doit absolument être élevée dans la bande 300 Hz – 5000 Hz.

2.5 Présence d'images fantômes sur 360°

La haute résolution spatiale garantit la présence d'images fantômes sur 360°. Il s'agit d'une conséquence directe du **respect de la fréquence spatiale d'ordre 0 (composante omnidirectionnelle) du champ acoustique.** Toutes les sources sonores entourant le microphone seront reproduites avec un niveau correct et forment des images fantômes. En revanche, cette condition est insuffisante pour garantir un contrôle sur la répartition de ces sources au moment de la restitution.

2.6 Down-mix facile et de bonne qualité

Les problèmes de down-mix ont été étudiés dès l'apparition du format stéréophonique afin d'assurer une compatibilité descendante avec le format monophonique. Pour les formats multicanaux, le problème de la compatibilité devient nettement plus délicat puisque le format 5.0 doit être compatible avec l'ensemble des formats dont les plus utilisés sont les formats 4.0 « Quad », 4.0 « Surround », 3.0 « Surround », 2.0 « Stéréo » et 1.0 « Mono ».

L'aptitude d'une prise de son à supporter le down-mix est liée au principe mis en œuvre.

- Dans le cas d'une prise de son coïncidente, un down-mix optimum est obtenu de manière simple en recombinaison des différents canaux par un matriçage très simple. En effet, la spatialisation est obtenue uniquement au moyen de variations d'intensités (ΔI) entre les canaux. Ainsi le mélange des canaux ne modifie que la spatialisation, sans affecter le timbre.

- Dans le cas d'une prise de son non coïncidente, le down-mix est extrêmement délicat et ne conduit jamais à des résultats optimaux. En effet, ces systèmes introduisent des différences de temps (ΔT) entre les canaux ce qui rend impossible leur mélange. Un matricage simple de microphones non coïncidents introduit systématiquement des déformations de timbres (effets de filtre en peigne) et des déformations de la spatialisation (effets d'aliasing spatial : apparition de nombreux lobes secondaires incontrôlés).

Le signal fourni par un système à haute résolution spatiale repose sur un principe de coïncidence. La spatialisation est obtenue uniquement au moyen de variations d'intensités (ΔI) entre les canaux et aucune variation de temps n'est introduite (ΔT). Par conséquent, **un enregistrement à haute résolution spatiale permet des down-mix faciles et de haute qualité.**

Par rapport à une prise de son à résolution standard, la haute résolution spatiale apporte un avantage décisif : **le down-mix d'une prise de son à haute résolution spatiale conserve les propriétés d'une prise de son à haute résolution spatiale. Simplement, la résolution de l'enregistrement s'adapte automatiquement à la résolution maximale que peut supporter le nouveau format.** Ainsi, le down-mix d'un enregistrement à haute résolution spatiale sera très proche du résultat qui aurait été obtenu en appliquant le principe de la haute résolution spatiale directement au nouveau format. On parle de prise de son à haute résolution spatiale correspondante. Toutefois, les performances atteintes seront légèrement en retrait car l'optimisation est moins poussée. Bien entendu, le down-mix mono ne comporte aucune spatialisation et le résultat obtenu est identique à celui d'une prise de son monophonique.

En particulier, **l'opération de down-mix n'introduit pas de modification de timbres ni de distorsion d'angle dans les images fantômes.** Les images fantômes formées par les formats 4.0 et 3.0 sont localisés dans les mêmes directions que celles du format 5.0. Seules les performances en terme de spatialisation (précision des sources, taille de la zone d'écoute, ...) sont réduites. Dans le cas du format 2.0, les directions des images fantômes sont préservées pour toute la zone frontale ($+30^\circ$ à -30°). En dehors de cette zone frontale, les images sont repliés sur la zone frontale, tout en préservant parfaitement la latéralisation et les timbres. Le tableau 4-I illustre l'opération de down-mix d'une prise de son à haute résolution spatiale pour les principaux formats multicanaux.

Tableau 4-I : Down-mix d'une prise de son à haute résolution spatiale.

Format	Configuration	Matrice de Down Mix	Prise de son à haute résolution spatiale correspondante.
4.0 "Quad" Résolution spatiale: Ordre 3			
4.0 "Surround" Résolution spatiale: Ordre 5			

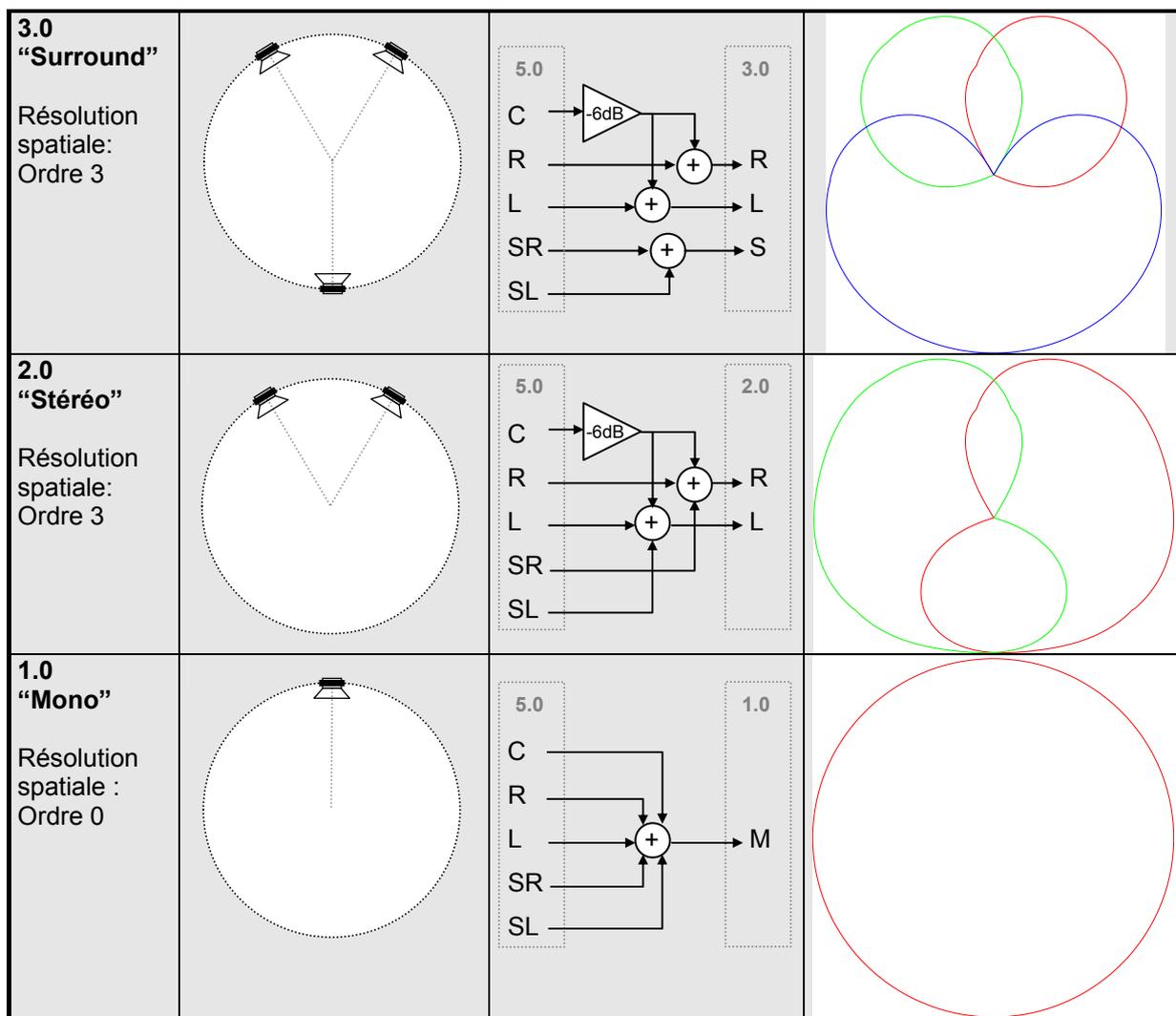
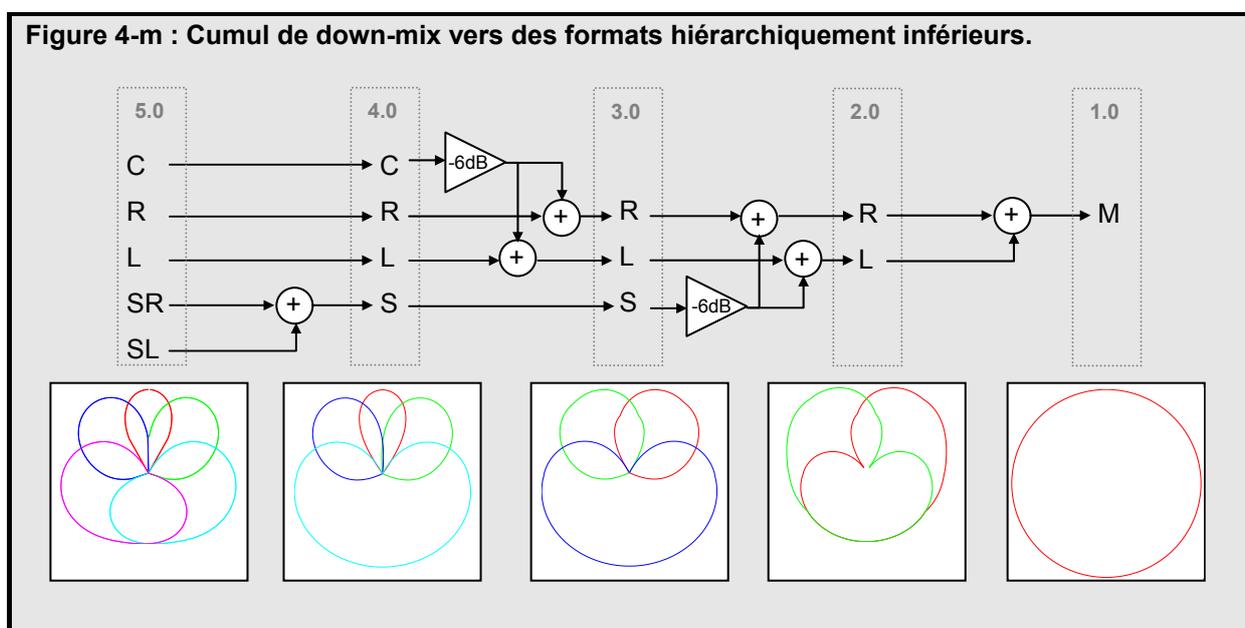


Figure 4-m : Cumul de down-mix vers des formats hiérarchiquement inférieurs.



Sous réserve que la hiérarchie des formats soit convenablement respectée, **la haute résolution spatiale autorise le cumul des down-mix au travers de formats différents**. Dans ce cas, le down-mix final sera identique à un down-mix direct vers le format final et préservera donc toutes les qualités

de la haute résolution spatiale. Par exemple, le format 5.0 initial peut être down-mixé dans le format 4.0 Surround, lui-même down-mixé dans le format 3.0 Surround, lui-même down-mixé dans le format 2.0 Stéréo et lui-même finalement down-mixé dans le format mono. Le signal mono ainsi obtenu sera identique à celui obtenu au moyen d'un down-mix mono directement appliqué au format 5.0 initial.

2.7 Compatibilité totale avec les lois de pan-pot habituelles

Le comportement d'un microphone à haute résolution spatiale est similaire à celui des lois de panoramique (voir partie 1.2). Par conséquent, **une prise de son à haute résolution spatiale se mélange parfaitement avec des sons mono spatialisés artificiellement à l'aide d'une loi de panoramique**. Dans les deux cas, le rendu sonore est comparable et le mélange est très homogène.

Il s'agit ici d'une souplesse nouvelle offerte par la haute résolution spatiale. En effet, les prises de son à résolution standard sont, par définition, limitées à une précision d'ordre 1 alors qu'une source spatialisée avec une loi de panoramique est à haute résolution spatiale. Il apparaît des différences flagrantes en terme de qualité de spatialisation et la fusion ne conduit pas à un résultat homogène.

Dans le cas particulier où la source supplémentaire provient d'un microphone de proximité, la haute résolution spatiale en facilite l'intégration dans l'enregistrement final. Toutefois, dans ce cas, certaines précautions doivent être prises :

- Les sources d'appoints supplémentaires doivent être dirigées dans les mêmes directions que les images fantômes correspondantes produites par le microphone. Dans le cas contraire, il en résulte une diminution de la précision spatiale liée à un étalement spatial de l'image fantôme. Toutefois, cette précaution n'est plus nécessaire lorsque le niveau sonore de la source supplémentaire dépasse largement celui de la source (> 15dB).
- Les retards entre les microphones d'appoint et le microphone multicanal doivent être ajustés afin de compenser la distance qui les séparent ainsi que la latence de traitement du microphone multicanal. Une erreur dans le choix du retard produira des différences de temps qui dégraderont les timbres au moment du mixage (effets de filtre en peigne).

En revanche, il n'y a pas de recommandations particulière concernant le dosage des niveaux entre des microphones d'appoint et le microphone multicanal. Les deux utilisations extrêmes sont :

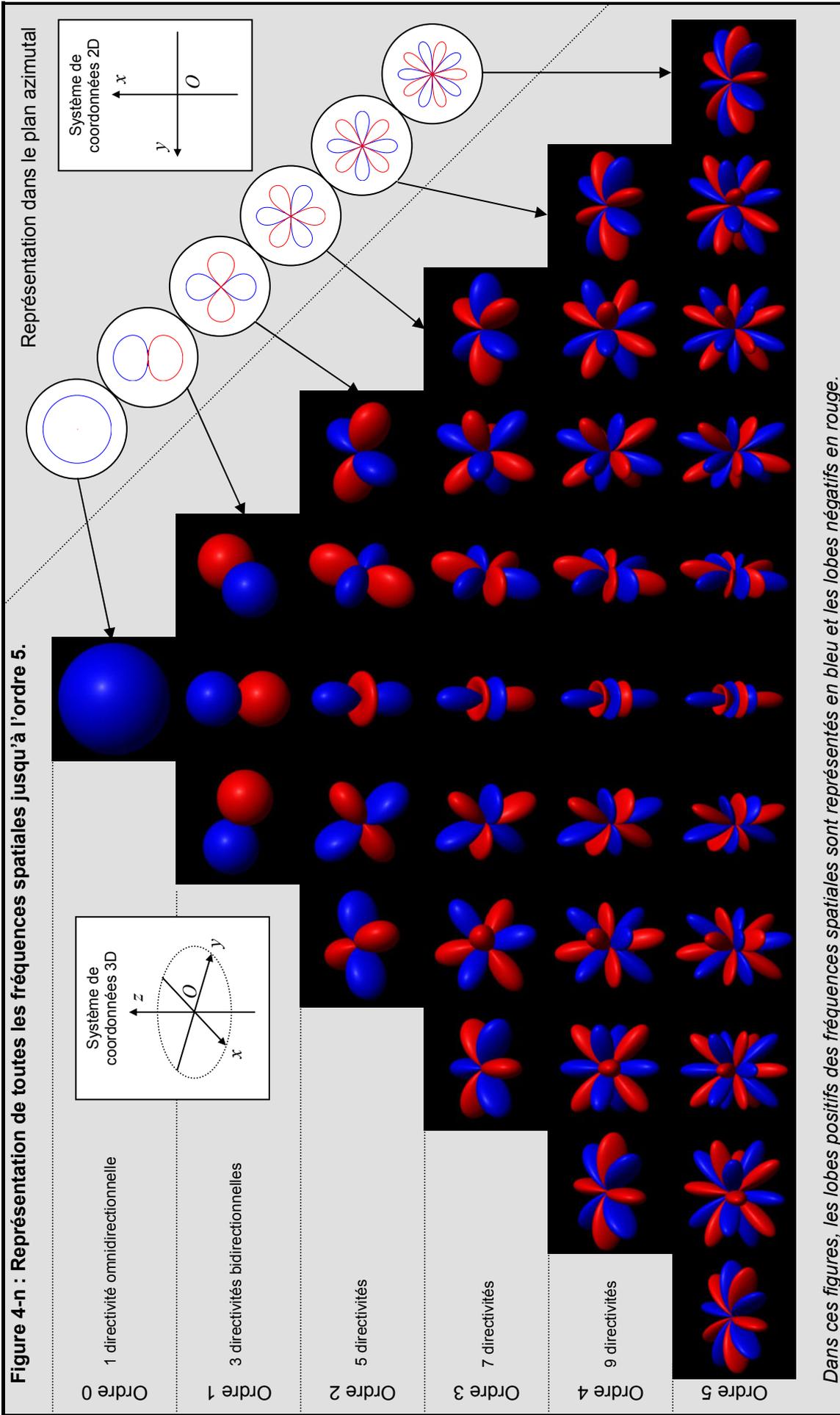
- Le microphone multicanal est dominant et les microphones de proximité apportent juste une discrète touche de présence.
- Les microphones de proximité sont dominants et le microphone multicanal apporte juste un discret complément d'ambiance.

Entre ces deux utilisations extrêmes, tous les réglages sont envisageables. Ils permettent notamment de contrôler finement, lors du mixage, le rapport entre le champ direct et le champ réverbéré, tout en préservant fidèlement l'acoustique du lieu d'enregistrement.

2.8 Contrôle de la spatialisation en dehors du plan azimutal

Jusqu'à présent, seules les sources placées dans le plan azimutal ont été considérées. Or dans un environnement sonore réel, les sources sonores sont placées dans toutes les directions de la sphère (et ne sont pas limitées aux seules directions du cercle azimutal). Par exemple, une source sonore qui survole le microphone doit être correctement enregistrée tout au long de la trajectoire. De même, un effet de salle se traduit par d'innombrables réflexions contre les murs produisant des sources images réparties dans toutes les directions autour du microphone. **Par conséquent, un microphone multicanal doit être optimisé pour toutes les directions de la sphère.**

La haute résolution spatiale permet de contrôler finement le comportement en dehors du plan azimutal. **En effet, les fréquences spatiales permettent de discriminer les sons dans l'espace avec une même finesse, quelle que soit la direction des sources.** Les 36 fréquences spatiales jusqu'à l'ordre 5 sont illustrées par la figure 4-n.



Dans un environnement sonore réel, les sons proviennent de toutes les directions. Or, dans la configuration 5.0, les haut-parleurs sont tous situés dans le plan azimutal. Les sources reproduites ne peuvent donc être localisées que dans le plan azimutal. Face à cette limitation du format 5.0, il est nécessaire de déterminer le comportement idéal du microphone en dehors du plan azimutal. De manière caricaturale, le comportement souhaité est qu'une source située hors du plan azimutal lors de l'enregistrement se trouve restituée avec la même direction azimutale mais sans élévation. Evidemment, un tel comportement est trop caricatural. En effet, une source qui survolerait le microphone d'avant en arrière serait reproduite uniquement en avant puis elle basculerait brutalement en arrière lorsqu'elle dépasse le microphone. Un tel comportement n'est pas réaliste par rapport au mouvement réel de la source.

Un comportement plus réaliste serait que :

- Toute source dans le plan azimutal est enregistrée avec les figures de directivités idéales.
- Toute source située exactement au dessus (ou en dessous) du microphone multicanal est restitué par les 5 haut-parleurs de sorte qu'aucune direction ne soit privilégiée. Une telle captation ne présente aucune séparation et correspond à une prise de son au moyen de 5 microphones omnidirectionnels coïncidents.
- Toute source située à une élévation intermédiaire entre ces deux cas extrêmes adopte un comportement intermédiaire. Autrement dit, le comportement devient progressivement omnidirectionnel au fur et à mesure que l'élévation augmente.

Un tel comportement produit bien l'effet souhaité, lorsque la source survole le microphone : l'image fantôme restituée est ramenée dans le plan mais elle passe progressivement d'avant en arrière. Ainsi, le concept de directivités idéales, initialement défini dans le seul plan azimutal, peut s'étendre à toutes les directions autour du microphone. Le figure 4-0 illustre **les directivités idéales d'une prise de son multicanal optimale pour toutes les directions de l'espace.**

Les directivités nécessaires pour obtenir un prise de son 5.0 à haute résolution spatiale présentent les caractéristiques suivantes :

- Elles sont très sélectives pour les directions proches du plan azimutal, c'est à dire pour des élévations de +/-45° par rapport au plan azimutal. C'est toujours cette sélectivité spatiale qui est à l'origine de la haute résolution spatiale.
- Elles présentent des asymétries. L'absence de symétrie par rapport aux directions des haut-parleur avait déjà été mise en évidence dans le plan azimutal. L'absence de symétrie axiale apparaît de manière encore plus évidente lorsque l'on compare le comportement en azimut et en élévation des directivités.

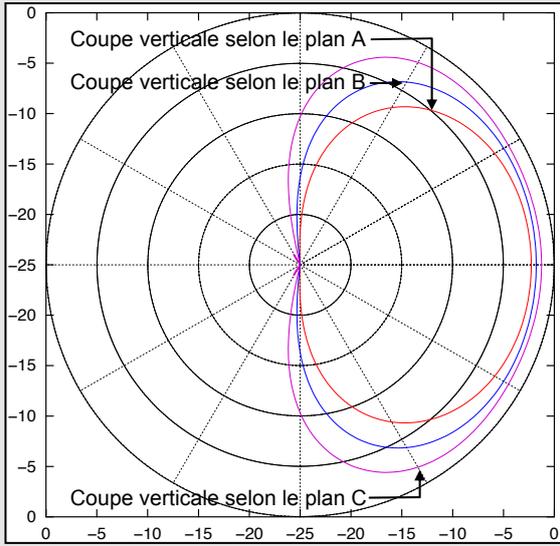
Bien entendu, de telles directivités idéales préservent les propriétés essentielles de la haute résolution spatiale pour toutes les directions de la sphère :

- Séparation optimale des canaux. En effet, une séparation exagérée des canaux pour les sources en élévation produit des trajectoires de sources peu réalistes.
- Augmentation de la ponctualité des sources.
- Augmentation de la taille de la zone d'écoute.
- Meilleur respect des timbres dans toutes les directions.
- Présence d'images fantômes sur 360°.
- Down mix facile et de bonne qualité quelle que soit la direction des sources.
- Facilite l'ajout de microphones d'appoint.
- Meilleur contrôle de l'angle utile de prise de son dans le plan horizontal mais aussi vertical.

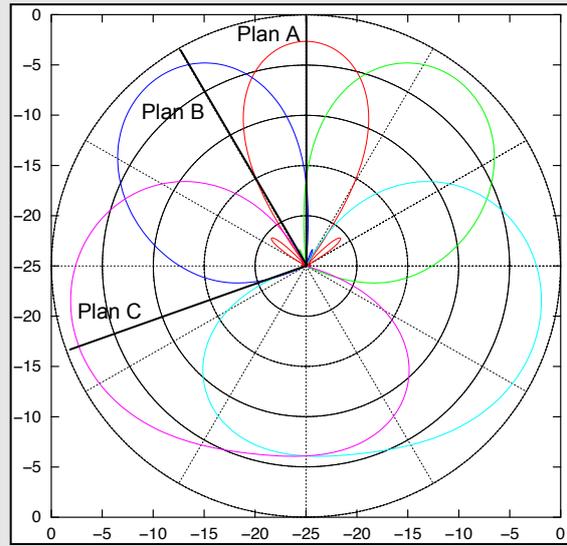
Les systèmes de prise de son à haute résolution spatiale développés par **Trinnov Audio** sont basés sur les 36 fréquences spatiales et permettent un contrôle très fin des directivités idéales pour toutes les directions de l'espace. **La haute résolution spatiale permet notamment d'adopter un comportement totalement différent entre l'azimut et l'élévation.**

Figure 4-o : Directivités idéales d'une prise de son 5.0 pour toutes les directions de l'espace.

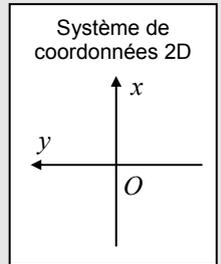
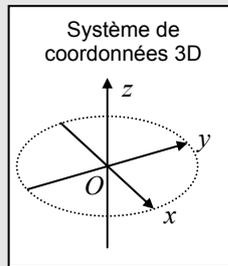
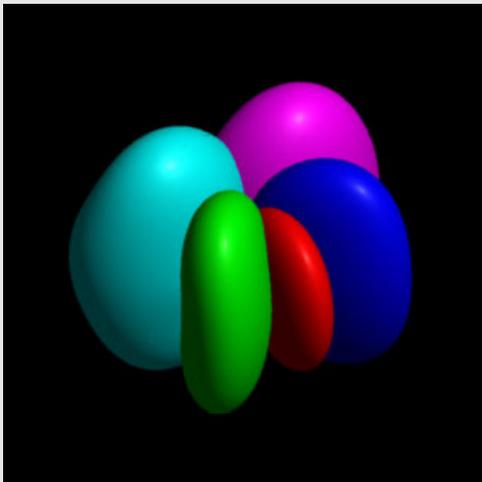
Coupes verticales selon les directions du format 5.0



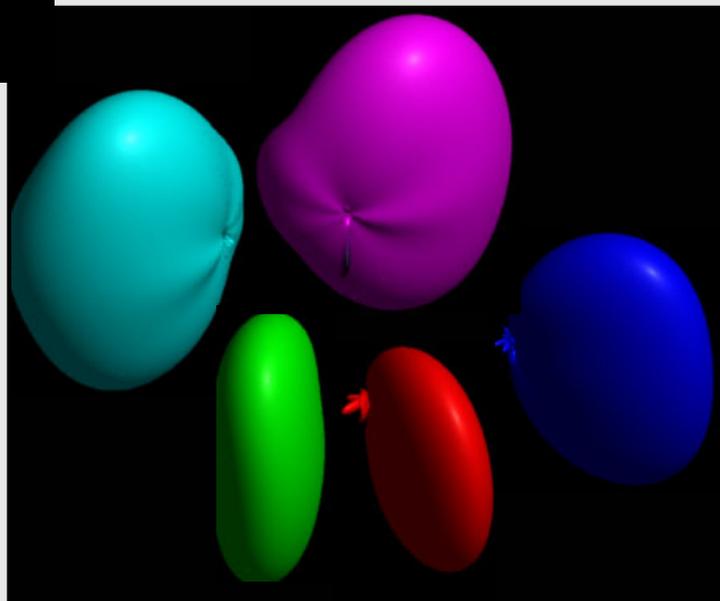
Coupe dans le plan horizontal



Représentation des directivités pour toutes les directions.



Vue en « éclaté » des directivités idéales



2.9 Maîtrise optimale du facteur de distance

Lorsqu'une source sonore réelle rayonne dans un environnement réel, on considère habituellement que le champ acoustique résultant est un mélange :

- de champ direct, formé par la première onde qui parvient au microphone et éventuellement par les premières réflexions contre les parois du lieu d'enregistrement (principalement celle provenant du sol). L'intensité sonore diminue lorsque la distance entre la source et le microphone augmente. A la position du microphone, l'intensité sonore du champ direct est toujours orientée dans la direction de la source.
- de champ réverbéré, constitué de la superposition des innombrables réflexions du son contre les parois du lieu d'enregistrement. L'intensité sonore du champ réverbéré est identique en tout point du lieu d'enregistrement. En chacun des points, l'intensité sonore est uniformément répartie entre toutes les directions.

Ainsi, le rapport d'intensité entre le champ direct et le champ réverbéré varie en fonction de la distance entre la source et le microphone. La distance pour laquelle, le champ direct et le champ réverbéré ont même intensité est appelée distance critique. En général, une prise de son analytique dans laquelle les sons sont détaillés et les sources sonores sont localisées avec précision est obtenue en plaçant le microphone à une distance inférieure à la distance critique.

Malheureusement, nombreuses sont les conditions d'enregistrement pour lesquelles la distance critique n'excède pas quelques mètres. Dans le cadre d'une prise de son monophonique, le rapport d'intensité entre le champ direct et le champ réverbéré peut être modifié en utilisant un microphone directif. Par exemple, un microphone cardioïde orienté dans la direction de la source présente une sensibilité au champ direct 1,73 fois supérieure (4.8dB) comparée à la sensibilité au champ réverbéré. Ainsi, la distance critique se trouve artificiellement augmentée d'un facteur 1.73 (on parle de facteur de distance).

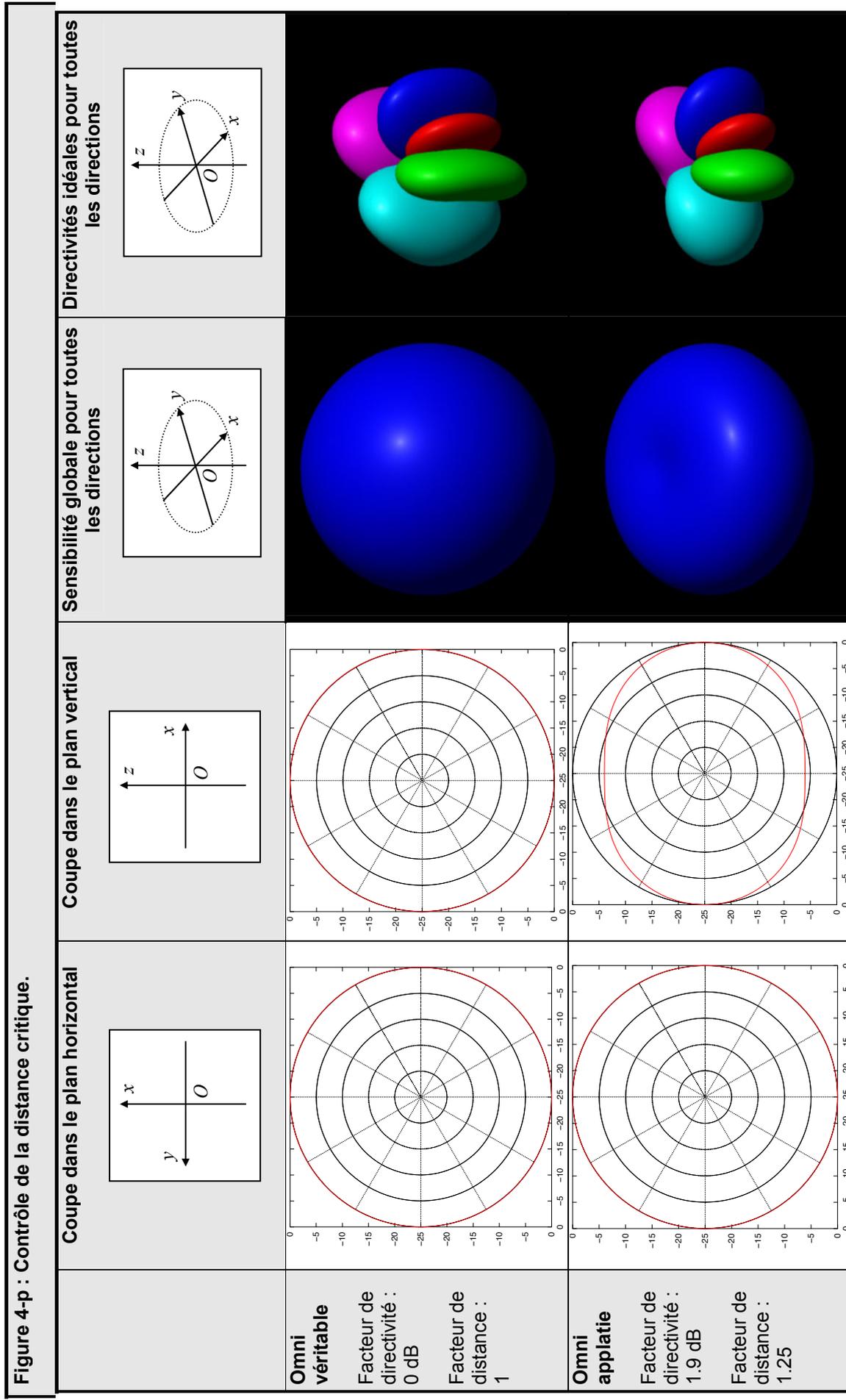
Bien entendu, les considérations relatives à la distance critique s'appliquent également à la prise de son multicanal. Dans ce cas, les 5 canaux doivent être considérés simultanément afin de déterminer la sensibilité globale du système vis à vis du champ direct et du champ réverbéré. Dans une situation d'enregistrement classique où les sources sont positionnées sur la zone frontale, les canaux frontaux (L,R,C) reproduisent un mélange de champ direct et de champ réverbéré alors que les canaux arrières (SL,SR) reproduisent principalement du champ réverbéré. Ainsi, une solution immédiate pour augmenter la distance critique consisterait à réduire le niveau des canaux arrière. Malheureusement, cette solution dégrade la fidélité spatiale puisque les images latérales viennent s'agglutiner sur les haut-parleurs avant droit et avant gauche.

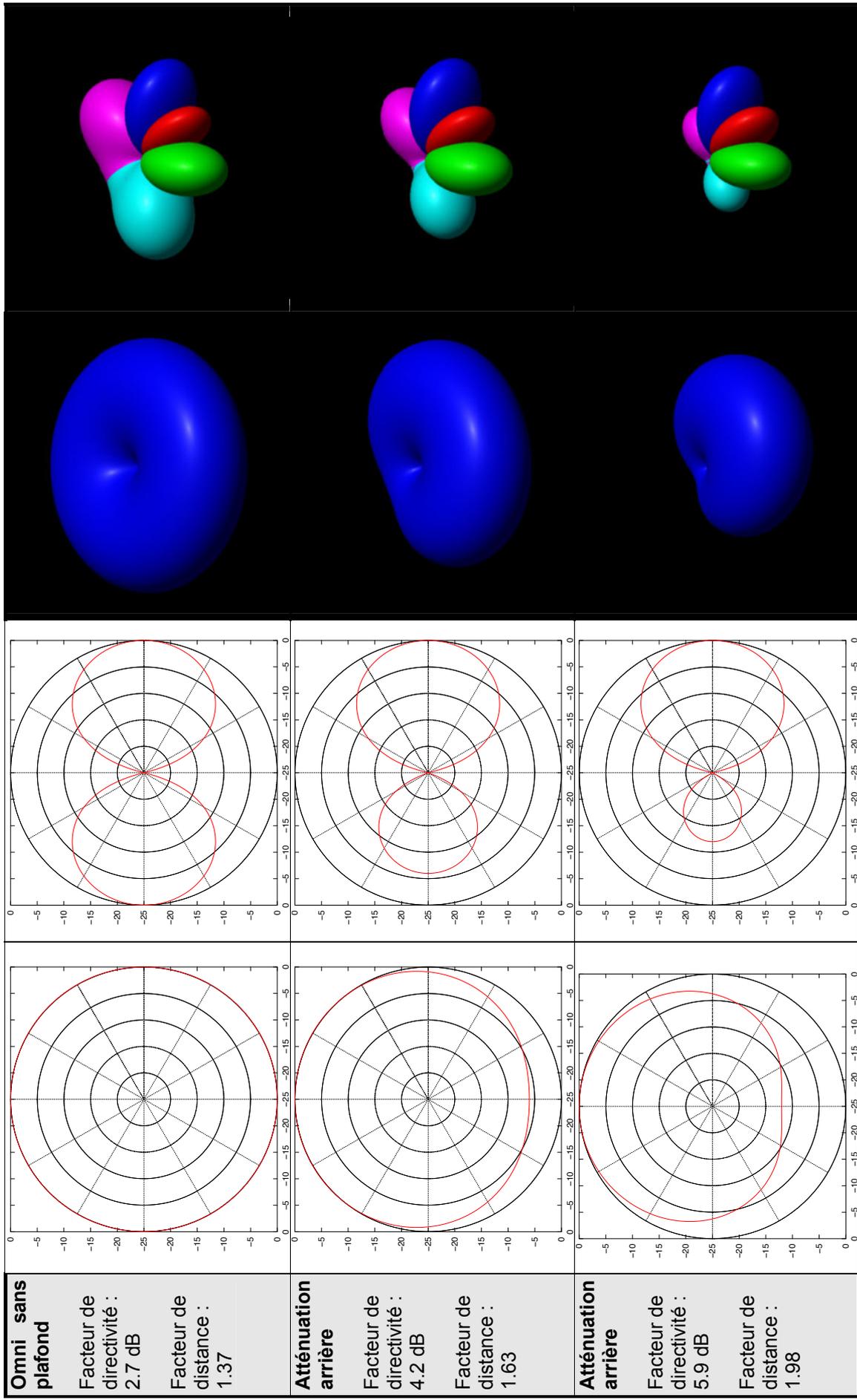
La technologie Trinnov Audio permet de contrôler la sensibilité globale du microphone multicanal en fonction de la direction des sources, sans introduire de déformations de timbres ni de distorsion d'angle. Le contrôle est obtenu en préservant toutes les propriétés de la haute résolution spatiale. Ce contrôle est possible en exploitant les 36 fréquences spatiales jusqu'à l'ordre 5.

Le principe consiste à réduire la sensibilité globale du microphone dans les directions pour lesquelles il n'y a pas de sources mais seulement du champ réverbéré. Cette diminution de sensibilité affecte de manière identique l'ensemble des directivités idéales, de sorte qu'il n'y ait aucune influence sur la position des images fantômes, mais seulement sur leur niveau sonore. Ainsi, l'intensité du champ direct est préservée alors que celle du champ diffus diminue : le facteur de distance est supérieur à 1. Plusieurs stratégies sont envisageables en fonction de la répartition des sources autour du microphone :

- Les sources entourent le microphone sur les 360° du cercle azimutal et leur élévation n'est pas trop importante (+/- 40° par rapport au plan azimutal). Dans ce cas, le facteur de distance peut être augmenté en réduisant la sensibilité globale du microphone pour les directions dont l'élévation est importante (au dessus et au dessous du microphone).
- Les sources sont localisés dans l'hémisphère avant (+/-90°) et leur élévation n'est pas trop importante (+/- 40° par rapport au plan azimutal). Dans ce cas, le facteur de distance peut être augmenté en réduisant la sensibilité pour les sources situées en élévation (au dessus et au dessous) et situées dans l'hémisphère arrière du microphone.

La figure 4-p illustre le contrôle de la distance critique offert par la haute résolution spatiale.





Comme cela a déjà été souligné précédemment, les 5 canaux d'une prise de son 5.0 à haute résolution spatiale sont en phase et se somment parfaitement au centre du système de restitution. Cette opération est équivalente à un down-mix mono. Dans le cas où la distance critique n'est pas modifiée artificiellement, le down-mix mono correspond exactement à une prise de son mono omnidirectionnelle. Dans les autres cas, le down-mix mono correspond à une prise de son qui aurait été effectuée par un microphone dont la directivité correspondrait à la sensibilité globale du microphone multicanal. Une telle directivité présente une bande passante spatiale étendue et ne possède pas de symétrie axiale. Pour des raisons comparables à celles avancées dans la partie 1.4, cette directivité ne peut pas être obtenue avec un unique microphone. Seule la technologie à haute résolution spatiale **Trinnov Audio** permet de modéliser de cette manière les directivités. Avec un facteur de directivité de 5.9 dB, les performances peuvent être supérieures à celles obtenues avec un cardioïde (4.8 dB) alors que les sons de l'hémisphère arrière restent audibles (atténuation de 12 dB).

2.10 De l'angle utile de prise de son à la distorsion angulaire

Jusqu'à présent, nous avons décrit des microphones multicanaux qui présentent une parfaite fidélité spatiale. Les sources fantômes restituées sont localisées exactement dans les directions des sources enregistrées. Or il est parfois souhaitable d'obtenir une répartition différente. Par exemple dans le cas d'une prise de son rapprochée d'une large scène sonore, il peut être souhaitable de resserrer la scène frontale.

En stéréophonie, le concept d'angle utile de prise de son a été développé afin d'exprimer l'ouverture angulaire pour laquelle les sources enregistrées se retrouvent entre les deux haut-parleurs. Cependant, l'angle utile de prise de son ne donne aucune information sur la manière dont les sources se répartissent dans l'espace stéréophonique. Le concept de distorsion angulaire reprend cette idée mais l'exprime de manière plus complète. En effet, une courbe de distorsion angulaire précise pour chaque direction de l'espace, la direction de l'image fantôme correspondante.

La notion de distorsion angulaire se base sur les fréquences spatiales d'ordre 1. En effet, selon la méta-théorie de la perception auditive de Michael Gerzon :

- la fréquence spatiale d'ordre 0 (composante omnidirectionnelle) du champ acoustique est représentative du timbre des sources.
- les 3 fréquences spatiales d'ordre 1 (composantes bidirectionnelles) du champ acoustique sont représentatives de la position des sources dans l'espace. En revanche elles ne donnent aucune information sur les autres caractéristiques des sources comme la taille, la forme, l'orientation, ...

Comme cela a déjà été évoqué dans la partie 2.4, la comparaison des composantes omnidirectionnelles entre le champ enregistré et le champ reproduit permet de déterminer les modifications de timbres apportées par le microphone multicanal. De manière analogue, la comparaison des composantes bidirectionnelles entre le champ enregistré et le champ reproduit permet de déterminer les modifications de positions apportées aux sources sonores. Autrement dit, **l'analyse des fréquences spatiales d'ordre 1 permet de caractériser très finement la distorsion angulaire.**

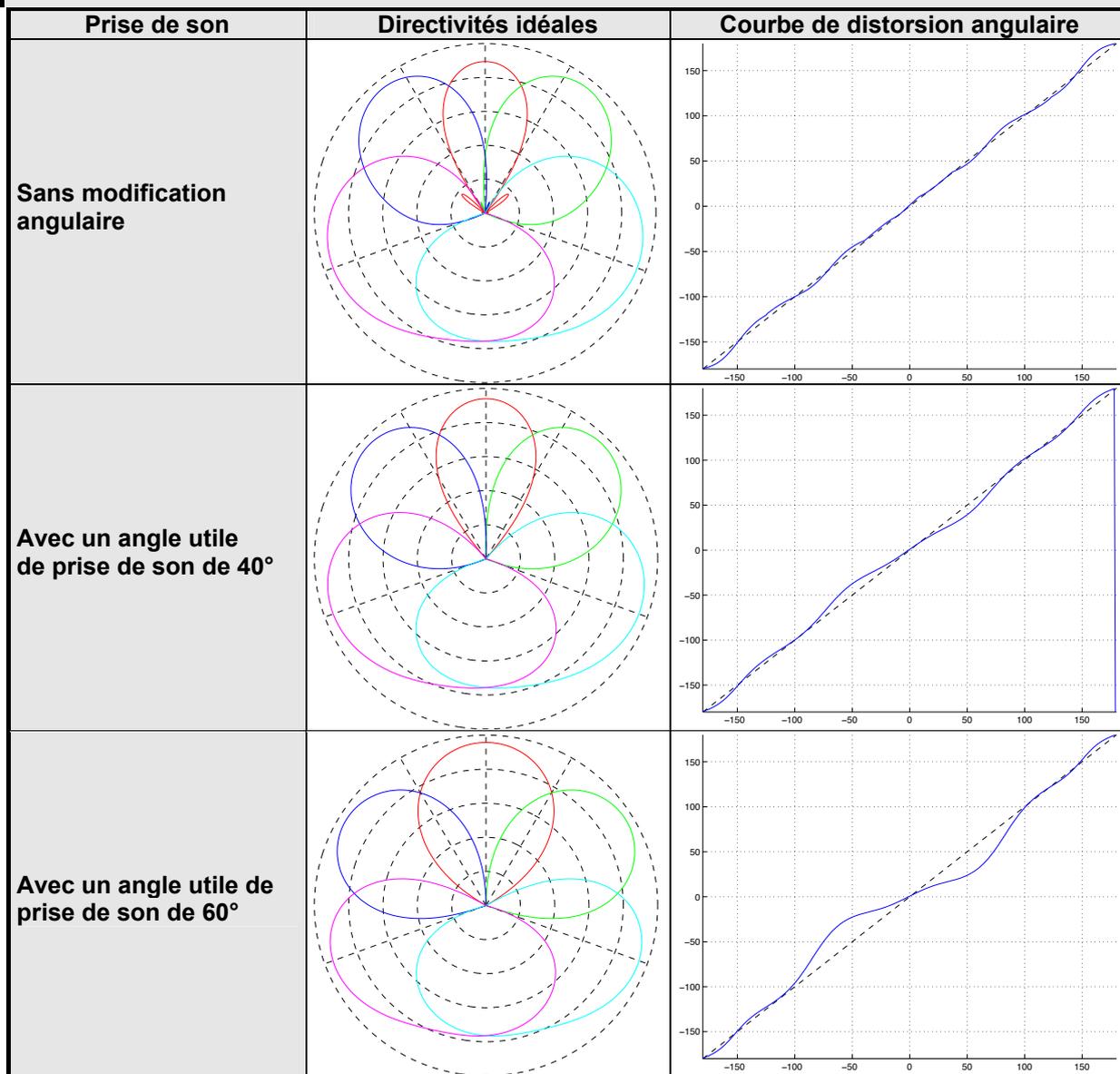
La technologie développée par **Trinnov Audio** permet, à l'inverse, de contrôler les fréquences spatiales, notamment les fréquences spatiales d'ordre 1. Ainsi, cette technologie permet de contrôler finement la distorsion angulaire des microphones multicanaux à haute résolution spatiale. Le contrôle de la distorsion angulaire est illustré par la figure 4-q, pour des angles de couverture de 30° (pas de distorsion), 40° (légère distorsion) et 60° (resserrement marqué de la scène frontale).

2.11 Meilleure robustesse au traitement

Il s'agit d'une conséquence immédiate de la séparation élevée entre les canaux. La modification d'un canal n'affecte les images fantômes que de part et d'autre du haut-parleur correspondant au canal modifié. Le reste de la scène sonore n'est pas touchée par cette modification. Ainsi, la haute résolution spatiale offre une robustesse accrue aux traitements usuels, tels que la correction de

niveaux, l'égalisation, la compression de dynamique ou encore le codage au moyen de matrices actives (du type pro-logic).

Figure 4-q : Contrôle de la distorsion angulaire.



3 Le fondement scientifique de la haute résolution spatiale : la maîtrise du champ acoustique

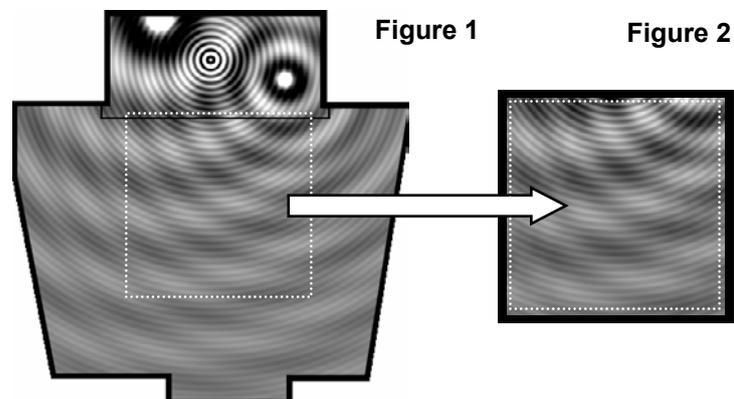
3.1 Origine scientifique de la technologie

Parallèlement aux techniques traditionnelles issues de la stéréophonie, certains laboratoires de recherche explorent de nouvelles approches, notamment l'*IRCAM*, l'*Université de Delft*, *France Télécom R&D* ou encore *Trinnov Audio*. L'idée maîtresse consiste à considérer la nature physique du son, c'est à dire le champ acoustique. Les technologies développées s'appellent *Wave Field Synthesis (WFS)* ou *High Order Ambisonics (HOA)*. Ces technologies trouvent leurs origines dans l'acoustique fondamentale : le principe de Huygens (1690) et les séries de Fourier-Bessel (1870). Il s'agit d'une sorte d'approche absolue qui permet d'enregistrer le champ acoustique du lieu d'enregistrement dans une large zone et de le reproduire de manière parfaitement identique dans le lieu de restitution. Ainsi, que l'on soit dans le lieu d'écoute ou dans le lieu d'enregistrement, il n'y a aucune différence, l'expérience auditive est absolument identique. Malheureusement, ces systèmes ne sont pas utilisés car les microphones et les haut-parleurs se comptent par centaines.

Figure 4-r : La maîtrise du champ acoustique.

Figure 1 : Dans cette salle de concert vue du dessus, trois sources sonores créent un champ acoustique dans lequel sont plongés les auditeurs.

Figure 2 : Dans cet auditorium, le champ acoustique capté pendant le concert est parfaitement reproduit. Les auditeurs sont alors plongés dans le même champ acoustique que celui du concert.



Toutefois, ce fondement scientifique valide de manière profonde et objective le concept de haute résolution spatiale. En effet, la technologie HOA se base précisément sur la décomposition du champ acoustique en fréquences spatiales appelées fonctions de Fourier-Bessel (constituées par les fonctions de Bessel et les harmoniques sphériques). La puissance de cet outil est identique à celle offerte par la décomposition d'un signal en ses fréquences constituantes : les timbres peuvent être totalement transformés ou au contraire parfaitement préservés. Par analogie, HOA est un puissant outil de contrôle des champ acoustique : les timbres et les positions des sources peuvent être totalement transformées ou au contraire parfaitement préservés.

3.2 Enregistrement à haute résolution spatiale - Principe d'échantillonnage

Comme cela a été évoqué dans 4.1.4, les directivités nécessaires pour un enregistrement à haute résolution spatiale ne sont pas offertes naturellement par les microphones actuels. Pour atteindre les performances requises, il est nécessaire de les recomposer grâce à un traitement linéaire complexe : le filtrage matriciel total. Un microphone multicanal à haute résolution spatiale fonctionne donc en deux étapes :

- Une étape d'acquisition d'informations sur le champ acoustique à l'aide d'un réseau de microphones usuels.
- Une étape de traitement, basée sur un filtrage matriciel total, permettant de synthétiser les directivités souhaitées

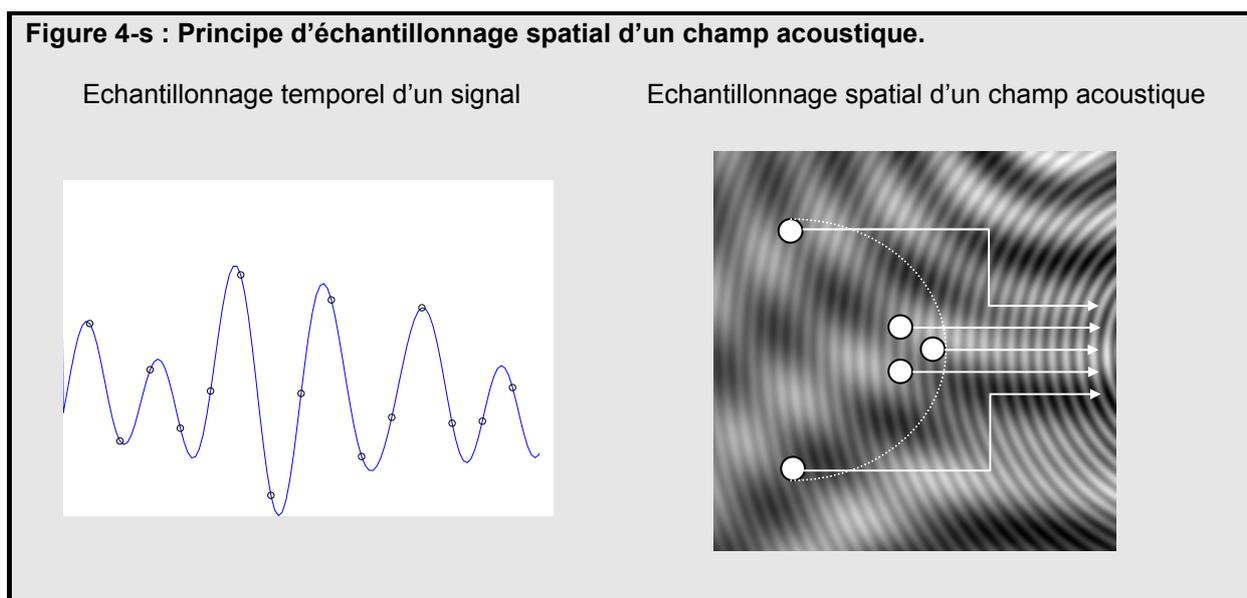
L'étape d'acquisition peut s'interpréter comme une étape d'échantillonnage spatial du champ acoustique et s'inspire directement de la méthode d'échantillonnage temporel d'un signal.

En effet, un signal (dont la bande passante est limitée) peut être observé en seulement certains instants sans qu'il n'y ait de perte d'information. Le signal temporel initialement continu se trouve alors parfaitement représenté par une série de valeurs appelées échantillons. On parle d'échantillonnage temporel du signal.

Par analogie, un champ acoustique peut être observé en seulement certaines positions sans qu'il n'y ait de perte d'information. Le champ acoustique initialement continu se trouve alors parfaitement représenté par une série de signaux appelés échantillons. On parle d'échantillonnage spatial du champ acoustique.

En réalité, un réseau de microphones réalise précisément des mesures sur le champ acoustique à des positions différentes et correspond donc à un échantillonneur de champ acoustique. L'ensemble des mesures réalisées caractérise les variations spatiales du champ acoustique. Ce principe est illustré par la figure 4-s.

Figure 4-s : Principe d'échantillonnage spatial d'un champ acoustique.



3.3 Enregistrement à haute résolution spatiale - Principe d'encodage

L'encodage est une opération qui consiste à déterminer les fréquences spatiales du champ acoustique. Lorsque le réseau de microphones est exposé à un champ acoustique connu en tout point, les lois de la physique permettent de calculer précisément les échantillons, c'est à dire les signaux délivrés par les microphones du réseau.

Or dans une situation d'enregistrement, on se trouve exactement dans la situation inverse : le réseau de microphones délivre des signaux qui sont connus et l'on souhaite obtenir des informations concernant le champ acoustique initial. Dans le cas particulier d'un enregistrement à haute résolution spatiale, on souhaite mesurer les fréquences spatiales du champ acoustique jusqu'à l'ordre 5. **Ainsi exprimée, l'opération d'encodage correspond exactement à l'inverse du processus d'échantillonnage spatial du champ acoustique. Ainsi, un enregistrement à haute résolution spatiale est obtenu en inversant le processus d'échantillonnage du champ acoustique au moyen d'un filtrage matriciel total.**

Un exposé détaillé de cette technique est développé dans les publications « *A New Comprehensive Approach of Surround Sound Recording* » (114^{ème} Convention de l'AES, preprint 5717) et « *High Spatial Resolution Multichannel Recording* » (116^{ème} Convention de l'AES).

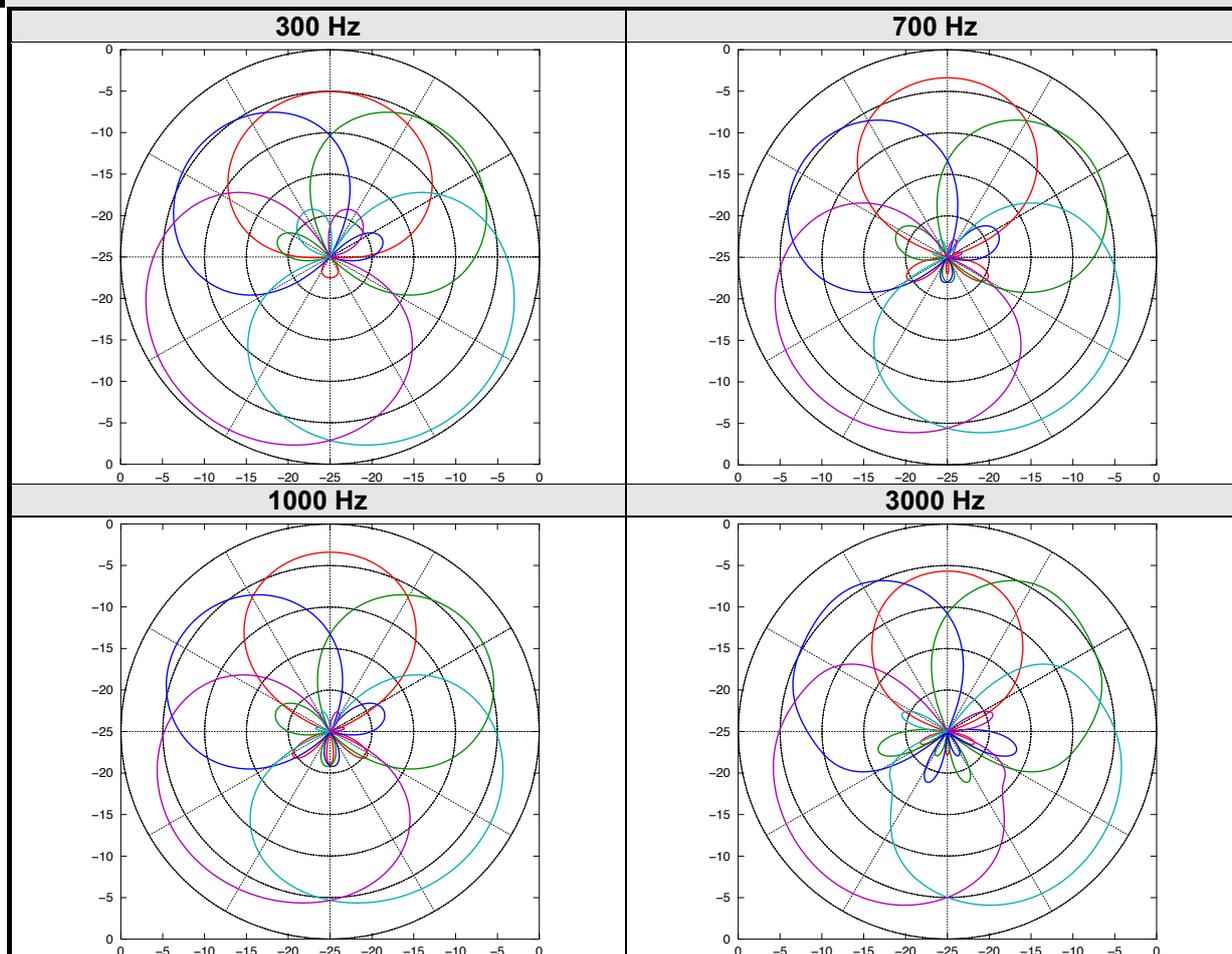
3.4 Capture en ΔT et transmission en ΔI

Grâce aux travaux de **Trinnov Audio**, la haute résolution spatiale peut être atteinte avec un réseau de **8 microphones omnidirectionnels et une unité de traitement**.

Dans ce système, la capture du champ acoustique est réalisée selon une approche en pure ΔT . **L'information est prélevée uniquement grâce à l'écartement des capteurs**, sans exploiter de caractéristiques directionnelles. L'agencement des microphones dans le réseau est optimisé pour prélever efficacement une grande quantité d'informations sur le champ acoustique. Les travaux de **Trinnov Audio** ont mis en évidence que, compte tenu de la faible sélectivité spatiale des microphones actuels, un réseau non coïncident réalise une meilleure captation qu'un réseau coïncident (*114^{ème} Convention de l'AES, preprint 5717*). En d'autres termes, l'échantillonnage du champ acoustique par un réseau non-coïncident de microphones est plus efficace qu'un échantillonnage par un réseau coïncident.

Grâce à un traitement basé sur le contrôle du champ acoustique, l'unité de traitement détermine les 5 canaux du format 5.0 qui permettent une reproduction optimale du champ acoustique. Le traitement est basé sur 40 filtres à 1024 fréquences permettant une reconstitution optimale de chaque canal à partir des 8 microphones omnidirectionnels. L'ensemble du système (réseau + unité de traitement) se comporte comme 5 microphones coïncidents réalisant une prise de son à haute résolution spatiale. Les performances typiques du microphone **Trinnov Audio HSR 5.0** sont présentées sur la figure 4-t. Dans le signal multicanal délivré, l'information est uniquement codée en variation d'intensité sonore (ΔI). Ce type de codage est particulièrement bien adapté à la transmission et la restitution de son multicanal.

Figure 4-t : Performances typiques du microphone Trinnov Audio HSR 5.0



3.5 Trinnov SRP : une plate-forme dédiée à la haute résolution spatiale.

La technologie à haute résolution spatiale offre de nouvelles perspectives en matière de prise de son multicanale. En exploitant une architecture originale composée d'un réseau de microphone et d'un traitement numérique spécifique, cette technologie est à l'origine d'une nouvelle génération de microphones multicanaux.

Trinnov Audio propose une solution souple et évolutive : *Trinnov SRP (Surround Recording Platform)*. Plus qu'un simple microphone multicanal il s'agit d'une véritable plate-forme évolutive dédiée à la prise de son à haute résolution spatiale. Actuellement, le *SRP* accueille le premier microphones multicanal 5.0 à haute résolution spatiale. La plate-forme à été conçue pour recevoir les futurs microphones à haute résolution spatial qui seront proposés sous forme de packages *HSR (High Spatial Resolution)*. Un package est constitué d'un disque optique pour la mise à jour logicielle du *SRP* et éventuellement d'un nouveau réseau de microphones.

Le premier package disponible est le *HSR 5.0/Omni-8* composé :

- du réseau *Omni-8* permettant d'accueillir 8 microphones omnidirectionnels de studio. Grâce à un jeu de bagues, vous adaptez vos microphones de studio favoris.
- de la banque de filtres matriciels total *HSR 5.0* permettant d'exploiter le réseau *Omni-8*.
- d'outils de monitoring dédiés à la prise de son à haute résolution spatiale.
- d'outils de gestion du facteur de distance.

Trinnov SRP équipé du package *HSR 5.0/Omni-8*



Pour plus d'information sur la plate-forme Trinnov SRP :

Trinnov Audio

30-32 avenue de la République
94815 Villejuif Cedex
France

Tel : +33 1 46 81 27 56
Fax : +33 1 46 81 77 27
Mob : +33 6 16 81 62 87
www.trinnov.com – contact@trinnov.com

Sebastien Montoya, Product Manager

sebastien.montoya@trinnov.com
Tel : +33 6 74 09 95 38

Arnaud Laborie, General Manager

arnaud.laborie@trinnov.com
Tel : +33 6 16 81 62 87
Fax: +33 1 46 81 77 27

2003



COMMUNICATIONS ET AFFICHES

Samedi 22 mars

Session A : MICROPHONES

- **Double microphone autodirectif** — *Alexander A. Goldin*
- **Réseau de microphones circulaires pour l'enregistrement audio multicanal discret** — *Edo Hulsebos, Thomas Schuurmans, Diemer de Vries, Rinus Boone*
- ➔ **Une nouvelle approche globale de la prise de son surround** — *Arnaud Laborie, Rémy Bruno, Sébastien Montoya*
- **Les mécanismes créant le bruit du vent dans les microphones** — *Stuart Bradley, Tao Wu, Sabine von Hünerbein, Juha Backman*
- **Circuits intégrés pour microphones à électret hautes performances** — *Arie van Rhijn*
- **Susceptibilité aux radiofréquences des microphones à condensateur** — *Jim Brown, David Josephson*
- **Formation de faisceaux de microphones pour l'enregistrement multicanal** — *Juha Backman*