



Ecole Nationale Supérieure Louis-Lumière

Mémoire de Master

Promotion 2015 – Section Son

UTILISATION DE LA TECHNOLOGIE
WFS DANS LA CRÉATION SONORE
CINÉMATOGRAPHIQUE : POSSIBILITÉS
ET LIMITES

Rémi CARREAU, Thibaut MACQUART

Directeurs internes : Alan, BLUM, Claude GAZEAU

Directeurs externes : Etienne CORTEEL, Steven GOUTHY

Rapporteurs : Jean-Pierre HALBWACHS, Etienne HENDRICKS

REMERCIEMENTS

Nous tenons en premier lieu à remercier Etienne Corteel, pour son accueil, son investissement, ses conseils et l'aide précieuse qu'il nous a apporté dans la mise en pratique de notre installation. Nos remerciements sincères vont à l'équipe de Sonic Emotion Lab, Réda Frauly et Raphaël Foulon, pour leur disponibilité, leur écoute, leurs conseils avisés. Nous avons pu bénéficier grâce à eux d'outils remarquablement bien conçus et performants.

Nous remercions également Steven Gouthi. Son apport professionnel a guidé notre démarche et nous a permis de prendre du recul sur notre travail.

Nous remercions bien évidemment Alan Blum et Claude Gazeau d'avoir accepté de diriger ce mémoire, et de l'avoir fait avec la générosité, l'investissement, la disponibilité et la rigueur que nous leur connaissons. Leur attention portée à notre mémoire aura été d'une grande aide.

Nous remercions nos rapporteurs Jean Pierre Halbwachs, et Etienne Hendricks dont l'aide et les conseils ont été très précieux pour nos tests. Nos remerciements vont également à Laurent Millot qui nous a aidés à construire les bases de notre démarche.

Nous tenons par ailleurs à remercier particulièrement la société SNO, en la personne de Jean-François Fougeras, qui nous a permis de concrétiser et de mener dans de bonnes conditions notre mémoire. Nous remercions également Franck Gillardeaux et Mathieu Delquignies qui ont réussi à apporter des solutions technologiques et matérielles à nos exigences.

Nous remercions également le personnel de l'ENS Louis Lumière, en particulier Philippe Simonet, Mohammed Elliq, Taïeb Keraoun, et Laurent Stehlin, pour leur aide dans la mise en place de notre installation.

Nous remercions les vingt et une personnes présentes à nos tests perceptifs, pour leur patience et leurs remarques pertinentes et intéressées.

Enfin, nous tenons sincèrement à remercier chacune de nos familles, qui nous ont permis de mener sereinement nos études, et de les mener jusqu'au bout...

RESUME

La *Wave Field Synthesis*, ou synthèse de front d'onde, est une technique de spatialisation et de restitution visant à reproduire les caractéristiques physiques d'un champ sonore. Cette technologie a pour principale caractéristique de proposer une zone d'écoute étendue, chose qui est en pratique impossible avec les systèmes courants. Restée pendant longtemps dans les laboratoires de recherche, la WFS est aujourd'hui commercialisée, ce qui en fait un dispositif potentiellement applicable à la diffusion sonore en salle de cinéma. A l'heure où des systèmes disposant d'un nombre important d'enceintes voient le jour, tel que l'*Atmos* de Dolby, il nous semble intéressant de considérer la WFS comme une source de possibilités supplémentaires à la création sonore cinématographique.

Ce mémoire a donc pour nous été l'occasion d'appréhender ces nouvelles techniques de spatialisation sous le prisme de la création, et en particulier du point de vue du mixeur. Cette teinte applicative a accompagné notre démarche tout au long de cette étude, ce qui nous a amené à ne pas cloisonner les éléments théoriques de nos expérimentations. Notre problématique a été abordée sous différents angles : technologique, en essayant d'adapter le système aux contraintes de diffusion sonore en salle de cinéma ; pratique, en confrontant les outils actuels de la WFS de la postproduction ; perceptif, en tentant d'avancer puis de valider certaines tendances sur des effets que nous avons nous même constatés.

Mots-clés : Waves Field Synthesis, WFS, synthèse de front d'onde, holophonie, postproduction sonore, son au cinéma, spatialisation, mixage, perspective sonore, cohérence audiovisuelle.

ABSTRACT

Wave Field Synthesis is a sound spatialization and reproduction technique that aims at reproducing the physical characteristics of an acoustic field. This approach mainly creates an extended listening area which is almost impossible with the current systems. Kept for a long time for laboratories and research projects, WFS is now available on the market, and makes a feasible technique for sound reproduction in the cinema industry. As systems with a lot of loudspeakers are being installed, like *Atmos* developed by Dolby, we think it's interesting to consider WFS as opening new possibilities for cinema sound creation.

This project therefore was the opportunity for us to understand this new technique through the prism of creation, and in particular from the sound mixer's point of view. This practical aspect guided us through our study, in which theoretical principles and experimentations intertwine. The issues were approached from different angles: technologically, by trying to adapt the system to the sound reproduction constraints in cinema rooms; practically, by bringing the WFS tools and the postproduction requirements face to face; perceptively, by discussing and trying to validate some tendencies about effects that we noticed.

Keywords: Wave Field Synthesis, WFS, holophony, sound postproduction, cinema sound, spatialization, mixing techniques, sound perspective, audiovisual coherence.

TABLE DES MATIERES

INTRODUCTION	2
1. LE SON SPATIALISE AU CINEMA	9
1.1. UNE HISTOIRE DU SON SPATIALISE	10
1.2. LE 5.1, UNE ESTHETIQUE CODIFIEE	19
1.3. DEUX SYSTEMES EMERGENTS DE SPATIALISATION	25
1.3.1. ATMOS, DEVELOPPE PAR DOLBY	26
1.3.2. AURO 3D, PROMOTIONNE PAR BARCO	35
1.3.3. NOUVELLES TECHNOLOGIES : RESUME ET OUVERTURE	42
1.4. LA WFS : UN CONCEPT INNOVANT	45
1.4.1. LES PRINCIPES PHYSIQUES FONDATEURS	47
1.4.2. DE LA THEORIE A LA PRATIQUE : SIMPLIFICATIONS ET CONSEQUENCES	53
1.4.3. APPORTS, A PRIORI, DE LA WFS	61
1.4.4. HYPOTHESES : LA WFS EN CONTEXTE DE CREATION SONORE CINEMATOGRAPHIQUE	67
2. LA SPATIALISATION EN WFS : ETUDE PRELIMINAIRE	77
2.1. PREMIERE APPROCHE DU SYSTEME A SYNTHESE DE FRONT D'ONDE	79
2.1.1. UN OUTIL DE MISE EN « ESPACE WFS » : LE PERFORMER	79
2.1.2. COMPORTEMENT BASIQUE DU WAVE 1	84
2.2. LA WFS FACE AUX MATERIAUX D'UNE BANDE SON CINEMATOGRAPHIQUE	92
2.2.1. TRAITEMENT DE LA VOIX, MISE EN ESPACE	93
2.2.2. LE CAS DES AMBIANCES FRONTALES	107

2.2.3.	LA SPATIALISATION DE LA MUSIQUE "OFF"	122
2.3.	PREMIERES CONCLUSIONS SUR L'UTILISATION DE LA WFS AU CINEMA	139
3.	<u>MISE EN PRATIQUE DE LA WFS DANS LE MIXAGE CINEMA</u>	<u>146</u>
3.1.	LES OUTILS DEVELOPPES PAR SONIC EMOTION	148
3.2.	PARTICULARITES TECHNIQUES DE LA POST PRODUCTION CINEMATOGRAPHIQUE	158
3.2.1.	UNE CONFIGURATION HYBRIDE ET MODULABLE	159
3.2.2.	UN SYSTEME ELECTRO-ACOUSTIQUE ADAPTE	162
3.3.	PROPOSITION DE CONFIGURATION POUR UN MIXAGE CINEMA	164
3.3.1.	CHOIX ET DISPOSITION DES ENCEINTES	166
3.3.2.	ORGANISATION DE SESSION ET ROUTING	171
3.3.3.	CALIBRATION	180
3.4.	EXPERIENCE DE MIXAGE	187
3.4.1.	ERGONOMIE, METHODE ET ESPACE(S) DE TRAVAIL EN WFS	188
3.4.2.	POSSIBILITES DE SPATIALISATION EN WFS	192
3.4.3.	PLUSIEURS EFFETS DISCORDANTS	196
3.4.4.	CONCLUSIONS PRATIQUES SUR L'UTILISATION DE LA WFS EN MIXAGE CINEMA	201
4.	<u>EVALUATION PERCEPTIVE DE LA WFS AU CINEMA</u>	<u>206</u>
4.1.	PROPOS LIMINAIRES A LA MISE EN PLACE DES TESTS	207
4.1.1.	CONTEXTE D'ETUDES	207
4.1.2.	CRITERES D'ETUDES	210
4.1.3.	CONSTITUTION ET DESCRIPTION DU CORPUS	215
4.2.	PROTOCOLE, ANALYSE ET DISCUSSION	224

4.2.1. PROTOCOLE DE TESTS	224
4.2.2. ANALYSE STATISTIQUE DES RESULTATS	229
4.2.3. INTERPRETATION ET DISCUSSION	235

CONCLUSION

244

BIBLIOGRAPHIE

248

ANNEXES

254

ANNEXE 1 : DETAILS DES PROTOCOLES MIS EN PLACE EN PARTIE 2

255

ANNEXE 1.1 : TRAITEMENT DE LA VOIX, MISE EN ESPACE – PROTOCOLE 1 : VOIX-OFF SEULE - DETAILS
255

ANNEXE 1.2 : TRAITEMENT DE LA VOIX, MISE EN ESPACE – PROTOCOLE 2 : VOIX-OFF ASSOCIEE A UNE
REVERBERATION - DETAILS 258

ANNEXE 1.3 : CAS DES AMBIANCES FRONTALES – PROTOCOLE 1 – AMBIANCE STEREOPHONIQUE DE
CONTENU FREQUENTIEL LARGE - DETAILS 264

ANNEXE 1.4 : CAS DES AMBIANCES FRONTALES – PROTOCOLE 2 – AMBIANCE STEREOPHONIQUE DE
CONTENU FREQUENTIEL RESSERRE - DETAILS 270

ANNEXE 1.5 : CAS DES AMBIANCES FRONTALES – PROTOCOLE 3 – AMBIANCE MONOPHONIQUE -
DETAILS 273

ANNEXE 1.6 : SPATIALISATION DE LA MUSIQUE OFF – PROTOCOLE 1 - DETAILS 276

ANNEXE 1.7 : SPATIALISATION DE LA MUSIQUE OFF – PROTOCOLE 2 – DETAILS 280

ANNEXE 2 : DOCUMENTS RELATIFS AUX TESTS

282

ANNEXE 2.1 : EXTRAIT DU QUESTIONNAIRE DE TESTS 282

ANNEXE 2.2 : ORDRE DES SEQUENCES TIRE AU SORT POUR CHAQUE SEANCE 286

ANNEXE 2.3 : APERÇU DES TABLEAUX DE DONNEES ET DES RESULTATS DE L'ANOVA 291

ANNEXE 3 : DOCUMENTATIONS TECHNIQUES

294

ANNEXE 3.1 : ENCEINTES DX8 – SPECIFICATIONS TECHNIQUES 294

ANNEXE 3.2 : AMPLIFICATEURS UTILISES POUR LES RAMPES LATERALES – SPECIFICATIONS TECHNIQUES	297
ANNEXE 3.3 : PROCESSEUR WAVE 1 – SPECIFICATIONS TECHNIQUES	299
ANNEXE 3.4 : CONVERTISSEUR <i>MADI TO ANALOG</i> – SPECIFICATIONS TECHNIQUES	300

TABLE DES FIGURES

- Fig. 1.1 : Vue des microphones de scène du Théatrophone. Source [Hugonnet, 2000] 12
- Fig. 1.2 : Disposition des microphones et leurs liaisons aux écouteurs. Source [Hugonnet, 2000] 12
- Fig. 1.3 : Al Jolson dans The Jazz Singer, de Alan Crosland, Warner Bros, 1927. 13
- Fig. 1.4 : Rideau de microphones relié un banc de haut-parleur reconstituant un front d'onde. Source [Hugonnet, 2000]..... 14
- Fig. 1.5 : Garity et Hawckins devant une partie de l'installation du Fantasound. Source [Journal of the Society of Motion Picture Engineers, 1941]. 16
- Fig. 1.6 : Configuration d'une salle équipée d'un système 5.1. Source [Dolby, 2012]. 18
- Fig. 1.7 : Logo de la technologie Dolby Atmos. Source www.dolby.com. 26
- Fig. 1.8 : La salle 1 du Pathé Wepler, première installation Atmos en France..... 27
- Fig. 1.9 : Configuration de haut-parleurs recommandée pour un système Atmos.. Source [Dolby, 2012]. 29
- Fig. 1.10 : Les canaux de Beds, additionnés aux objets sonores. Source [Dolby, 2012]. 31
- Fig. 1.11 : L'application Monitor. Source [Dolby, 2013]..... 32
- Fig. 1.12 : Le plugin Dolby Atmos Panner. Source [Dolby, 2013]..... 33
- Fig. 1.13 : Le panner de la DFC Gemini..... 34
- Fig. 1.14 : Logo de la technologie Auro 3D. Source : www.auro-3d.com. 35
- Fig. 1.15 : Illustration de la configuration Auro 3D maximale, en 13.1. Source : www.auro-3d.com..... 37
- Fig. 1.16 : Configuration Auro 3D 13.1 dans une salle de cinéma. Source : www.auro-3d.com..... 38
- Fig. 1.17 : Fenêtre principale du Auro-Panner. Source : www.auro-3d.com..... 39

- Fig. 1.18 : Fenêtre principale du Auro-Matic. Source : www.auro-3d.com..... 40
- Fig. 1.19 : Illustration originale du principe de Huygens. 48
- Fig. 1.20 : A gauche se trouve la « traduction mentale » des sources en petits haut-parleurs proches..... 49
- Fig. 1.21 : Géométrie pour l'intégrale de Kirchhoff-Helmholtz..... 50
- Fig. 1.22 : Illustration de l'effet de fenêtrage. Source [Corteel, 2004]. 54
- Fig. 1.23 : Mise en évidence de l'effet de diffraction pour un banc de haut-parleurs de taille finie. Source [Caulkins, 2007]. 55
- Fig. 1.24 : Contribution des différentes rampes à la synthèse d'une source latéralisée. Source [Corteel, 2004]. 57
- Fig. 1.25 : Mise en évidence du repliement spatial par diagramme du champ de pression d'un système WFS pour deux fréquences différentes..... 58
- Fig. 1.26 : Principe optique de la parallaxe..... 64
- Fig. 1.27 : Types de sources synthétisables par WFS. Source [Corteel, 2004]. 66
- Fig. 1.28 : Illustration de la distorsion angulaire induite par la distance holophonique. 74
- Fig. 2.2 : Fenêtre principale du Performer, interface de positionnement et de visualisation des objets sonores..... 81
- Fig. 2.3 : Onglet « Routing » du Performer, permettant de décider du sort des sources sonores. 83
- Fig. 2.4 : Illustration de la zone de fenêtrage frontal..... 86
- Fig. 2.5 : Avec le Wave 1, la synthèse de sources à l'intérieur de la zone d'écoute est rendue impossible. 90
- Fig. 2.6 : Positions testées des objets WFS..... 94
- Fig. 2.7 : Mise en évidence du phénomène d'aliasing spatial. Source [Corteel, 2004]. 98
- Fig. 2.8 : Fenêtre principale du plugin Reverence. 100
- Fig. 2.9 : Première configuration : objet WFS à hauteur du premier cercle du Performer, groupe de réverbération (stéréo) derrière la source et relativement proche de celle-ci. 101

- Fig. 2.10 : Deuxième configuration : objet WFS au niveau du deuxième cercle du Performer, groupe de réverbération (stéréo) derrière la source et relativement proche de celle-ci, largeur du groupe « réverbération » reconsidérée. 102
- Fig. 2.11 : A gauche, la sensation d'espace créée en WFS ; à droite, la profondeur créée ne nous semble pas rendre de perspective..... 103
- Fig. 2.12 : Position 1 : groupe bipolaire "ambiance" au niveau des enceintes (un pôle dans chaque enceinte). 111
- Fig. 2.13 : Position 3 : groupe bipolaire "ambiance" au niveau du second cercle du Performer ; largeur du groupe reconsidérée..... 112
- Fig. 2.14 : En WFS, pour une position d'écoute excentrée, une image sonore large semble être préservée..... 114
- Fig. 2.15 : Contribution du banc de haut-parleurs à la synthèse d'une source ponctuelle, d'après [Husson, 2011]. 121
- Fig. 2.16 : Fonction d'alimentation du banc de haut-parleur d'après [Holophony]. 121
- Fig. 2.17 : En restitution discrète, une ligne horizontale décrit la zone d'écoute où une bonne balance avant/arrière est perçue 129
- Fig. 2.18 : Fenêtre principale du plugin Reverence. Un preset d'une réverbération de type « Small Hall » est chargé. 130
- Fig. 2.19 : Anymix, l'un des plugins d'upmix utilisés par les mixeurs 133
- Fig. 2.20 : Illustration de l'effet de précedence 136
- Fig. 3.1 : Interface du plugin TosCA. « ID #7 » fait référence à la souce n°7 dans le Performer. 154
- Fig. 3.2 : Intégration réseau des interfaces pour l'exploitation du Wave 1. 156
- Fig. 3.3 : Synoptique de l'installation faite dans l'auditorium de l'ENS Louis-Lumière..... 165
- Fig. 3.4 : Vue de l'auditorium..... 169
- Fig. 3.5 : Vue de l'auditorium 2 170
- Fig. 3.6 : Configuration des Mix Busses dans la console. 172
- Fig. 3.7 : Configuration des Aux Busses dans la console..... 173
- Fig. 3.8 : Patch d'entrée de Nuendo 175

- Fig. 3.9 : Timeline du séquenceur Nuendo 176
- Fig. 3.10 : Routing dans le Performer, utile à notre session de mixage 179
- Fig. 3.11 : Preset de l'enceinte DX8, avec bass management..... 181
- Fig. 3.12 : Preset de l'enceinte DX8, sans bass management..... 182
- Fig. 4.1 : Mundos Immundos, de Charlotte BAYER-BROC, photogramme. 216
- Fig. 4.2 : Petite Bête, de Margaux REMAURY, photogramme..... 217
- Fig. 4.3 : La Menace d'une Rose, de Thomas LHERMITTE, photogramme. 218
- Fig. 4.4 : Couper le voile, de Maximilien WILLMOCH, photogramme. 220
- Fig. 4.5 : Louis Le Manchot, de Camille JAULENT, photogramme 221
- Fig. 4.6 : Gabriela, de Anna LEYSENS, photogramme. 223

INTRODUCTION

L'audition humaine est un sens particulièrement utile à la perception de l'espace. L'idée de « sphère auditive »¹ témoigne d'ailleurs du fait que le son soit entendu dans toutes les directions. Il en est tout autre pour la vue, par exemple, qui amène le corps à se mouvoir pour continuer à percevoir. L'ouïe est d'ailleurs parfois le moteur de ce mouvement permettant d'amener la vue à se focaliser sur l'objet à percevoir. Nombreux de ces constats ont aujourd'hui été validés par les neurosciences, qui avancent notamment que le colliculus supérieur se distingue des autres centres auditifs de par sa faculté à embarquer une « véritable topographie de l'espace sonore »²

Par conséquent, les recherches dans le domaine sonore se sont assez rapidement intéressées à la captation et à la reproduction de l'espace. Aujourd'hui, le format 5.1 est devenu un standard, amené dans les salles obscures par des sociétés comme Dolby ou DTS. L'objectif commun de ces techniques est le suivant : augmenter la sensation d'immersion du spectateur dans un espace différent de celui où il se trouve réellement et, de par cette illusion, espérer l'impliquer toujours plus dans la narration proposée par le « spectacle audiovisuel »³.

¹ Terme emprunté à Etienne Corteel, dans *Caractérisation et extension de la Wave Field Synthesis en conditions réelles*, Thèse de doctorat de l'université de Paris VI, faisant office de document de référence dans les applications actuelles de la WFS [Corteel, 2004].

² Laurent Millot & Gérard Pelé dans « La sensation d'espace dans le spectacle audiovisuel », contribution à l'ouvrage *Une architecture du son*, [Millot & Pelé, 2006].

³ Notion développée dans [Millot & Pelé, 2006].

Cependant, la plupart des techniques de restitution développées aujourd'hui se fondent sur une illusion, un « trompe l'oreille » [Corteel, 2004], qui n'est appréciable que pour une toute petite partie de la zone d'écoute, souvent centrale : le célèbre *sweet spot*. Si une position d'écoute différente est adoptée par l'auditeur, l'illusion est biaisée, conduisant à une déformation de l'image sonore pouvant aller jusqu'à la perception d'un unique canal pour les positions d'écoute les plus excentrées. Ces défauts, inhérents à ces systèmes dérivés de la stéréophonie, amènent le mixeur à considérer ces distorsions avec précaution, de telle sorte à ne pas perturber le spectateur dans sa situation d'écoute. Ce phénomène conduit bien souvent à une inhibition de la création sonore, déjà fortement « cadrée » par la présence de l'image cinématographique.

D'autres systèmes ont pour principale ambition de reconstruire les propriétés physiques du champ sonore. On compte parmi eux la WFS et l'Ambisonie. Pendant les années 80, le laboratoire en Sismologie et en Acoustique de l'Université de Delft, aux Pays-Bas, s'est intéressé à la synthèse de champ sonore par modélisation de fronts d'onde : l'holophonie voyait le jour. Ces recherches ont abouti à la publication d'un document, aujourd'hui considéré comme étant le socle fondamental de toute étude ultérieure : *A holographic approach to acoustic control*, écrit par le professeur A. J. Berkhout [Berkhout, 1988]. Ce n'est qu'en 1993, avec [Berkhout et al., 1993], que le terme Wave Field Synthesis est apparu.

La WFS et les autres méthodes de rendu mathématiquement associées (comme l'ambisonie d'ordre élevé, ou HOA pour *High Order Ambisonic*) sont réalisables sur n'importe quel système de restitution dès lors que celui-ci contient plus 2 haut-parleurs. C'est d'ailleurs au début des années 90 que la plupart des principes techniques ont été avancés. Ils visent à approximer la théorie de sorte à rendre réalisable en pratique le procédé. Il faudra attendre le début des années 2000 pour que la WFS sorte des laboratoires universitaires, et soit appréhendée selon

INTRODUCTION

un angle plus technologique, plus applicatif. Ce fut l'idée centrale du projet CARROUSO (cf. [Brix et al., 2001]).

L'un des avantages significatifs de la WFS est de passer d'une méthode d'encodage basée sur le système de reproduction (le format est défini par le nombre de haut-parleurs), à un encodage basé sur la source sonore (chaque source peut être encodée séparément, et être reproduite d'une manière dépendante du système de restitution). Ainsi, a priori, plus le système est "défini" (plus l'intégration en haut-parleurs est importante), plus la source pourra être reproduite dans les meilleures conditions, ou du moins, dans des conditions proches de celles définies lors de l'encodage, à savoir lors du mixage. De fait, la restriction de la zone d'écoute sur les systèmes dérivés de la stéréophonie, directement liée à l'organisation figée des haut-parleurs, peut être détournée : le *sweet spot* s'efface, au profit d'une zone d'écoute élargie [Berkhout et al., 1993].

Dans [Brix et al., 2001], document recensant les fondements et les aboutissants du projet CARROUSO, les auteurs avançaient que « les progrès en micro-électronique, avec le coût de plus en plus faible de la puissance de calcul indique que dans un future proche, des applications dans le monde professionnel pourront voir le jour. » Un constat similaire avait également été réalisé par Raphaël Mouterde, dans le cadre de son étude à l'ENS Louis-Lumière [Mouterde, 2004]. Le procédé technique, décrit dans [Corteel, 2004], est aujourd'hui commercialisé par la société franco-suisse *Sonic Emotion*. Son concurrent allemand, *losono*, a également proposé pendant une dizaine d'années des solutions technologiques de synthèse de front d'onde, notamment destinées aux professionnels du cinéma. Cependant, la complexité et le coût des installations, induits entre autres par un nombre de haut-parleurs démesuré, ont pendant longtemps rebuté les exploitants et les sociétés de l'industrie cinématographique.

INTRODUCTION

Aujourd'hui, le système proposé par *Sonic Emotion* intègre une algorithmie capable de synthétiser un champ d'ondes avec un nombre réduit de haut-parleurs. A l'heure où le procédé spectaculaire *Atmos* développé par Dolby commence à faire son apparition dans les auditoriums et les salles de cinéma, il semble légitime de se demander si la WFS ne peut pas s'y faire une place, et ainsi offrir d'autres possibles à la création sonore cinématographique. Sans prétendre redéfinir le langage du cinéma, cette technologie, en plus d'être à même de solutionner quelques-uns des problèmes intrinsèques aux systèmes actuels, pourrait permettre d'envisager un autre rendu de l'espace sonore, de questionner autrement le rapport qu'entretient le cinéma avec les lieux qu'il investit. De plus, choisir d'étudier la WFS aujourd'hui nous donne accès à des outils commercialisés, chose importante pour nous, afin d'imaginer concrètement une chaîne de postproduction complète, et s'adresser ainsi directement aux utilisateurs, aux professionnels. Nous souhaitons en effet envisager l'étude de la WFS au cinéma autrement, et éviter le cadre des études uniquement perceptives, en nous appropriant le système en tant que nouvelle technique potentiellement capable de créer et de restituer une bande son cinématographique dans son ensemble.

La postproduction sonore dans l'industrie du cinéma sera donc notre contexte d'étude. Les installations spécifiques (Géode, Planétarium, ...), ainsi que le contenu moins courant (cinéma expérimental, installations dites « d'art numérique », ...), ne feront pas partie de notre étude. La question d'une telle diffusion sonore pour une image en relief reste également une piste possible, mais nous a semblé difficilement envisageable en raison de contraintes techniques évidentes. D'autre part, nous ne limiterons pas notre étude à une façade d'enceintes placées derrière l'écran, comme cela a pu être le cas dans [Mouterde, 2004]. Imaginer un système dans le contexte de l'industrie cinématographique implique, selon nous, de répondre à un certain nombre d'attentes, à la fois du public, mais aussi des professionnels. De la même manière que les formats actuels (5.1, Atmos, ...)

INTRODUCTION

proposent d'entourer le spectateur avec des canaux de restitution disposés sur un plan, nous envisagerons par conséquent un système WFS « entier », composé d'une ceinture de haut-parleurs entourant le spectateur.

La question globale à laquelle tentera de répondre cette étude est la suivante : *Dans quelle mesure la WFS peut-elle s'inscrire dans la postproduction sonore cinématographique ?* Cette question contient de multiples ramifications : technologique d'abord, en évaluant comment un système WFS peut s'intégrer dans les auditoriums de mixage et les salles obscures ; pragmatique ensuite, en nous mettant à la place du mixeur, et en confrontant les outils propres à la WFS aux méthodes de mixage usuelles ; perceptive enfin, en essayant de voir si les effets permis par la synthèse de front d'onde respectent le contrat sous-jacent du « spectacle audiovisuel », qui vise à offrir à chaque « spectateur-consommateur » un contenu identique [Millot & Pelé, 2006].

Ce document retrace notre démarche personnelle. Nous tenterons de faire en sorte de ne pas cloisonner la théorie de la pratique. En effet, les éléments dits « théoriques » prendront parfois la forme d'encadrés explicatifs, apparaissant là où ils se feront nécessaires, nous permettant ainsi de nous adresser d'avantage aux acteurs de la postproduction sonore. D'autre part, le mémoire s'articulant principalement autour des possibilités et des limites d'un système WFS, nous pensons qu'il est intéressant de s'attarder sur ce point précis de la méthode : obtenir un effet prouve les possibilités du système, mais expliquer comment obtenir cet effet apporte des réponses directes aux utilisateurs. Ce travail répond également aux demandes des sociétés partenaires, particulièrement soucieuses des retours d'utilisation de leurs outils.

Notre étude sera structurée comme suit. Une première partie proposera une approche contextuelle de notre travail. Après un bref historique des techniques de

restitution « spatialisantes » au cinéma, et un tour d'horizon des systèmes actuellement émergents, nous reviendrons sur les notions physiques de la WFS. Bien qu'il existe doré et déjà un corpus important sur le sujet, il nous semble important de ré-aborder cette étape en essayant d'en proposer une approche particulière : nous envisagerons en effet le système WFS dans le cadre spécifique du cinéma, entendu ici en termes d'espace de diffusion sonore. Cette première partie nous permettra enfin de poser les premières hypothèses sur l'utilisation d'un dispositif WFS en postproduction sonore cinématographique.

Une deuxième partie sera pour nous l'occasion de relater nos premières expérimentations autour du système WFS proposé par Sonic Emotion. Effectuée sous forme de stage court, cette période de tests nous a permis de nous familiariser avec les outils et les effets proposés par la synthèse de front d'onde. Nous présenterons une série de matériaux sonores usuels des bandes sons au cinéma, confrontés aux traitements du processeur WFS, en comparaison avec les systèmes standards usuels (stéréo, ou 5.1). Bien que les résultats ne soient, dans un premier temps, soumis à aucune validation scientifique, nous trouvons intéressant de proposer au lecteur notre approche personnelle du système, débouchant sur des possibilités et des limites, technologiques et perceptives, que nous aurons constatées. Cette première approche pratique aura par ailleurs la particularité d'être réalisée en l'absence d'image.

La troisième partie de notre étude s'attachera à rendre compte de nos expérimentations « taille réelle » dans l'auditorium de l'ENS Louis-Lumière. Véritable mise en situation de mixage, cette expérience comportait un objectif double : mettre en relation nos constats perceptifs précédents avec la présence d'une image ; appréhender la chaîne de postproduction que nous avons conçue autour du système WFS. A travers le mixage de courtes séquences, nous proposerons une analyse à la fois pratique et perceptive, en comparant les

INTRODUCTION

résultats obtenus sur un système 5.1 et le système à synthèse de front d'onde. Nous resterons cependant dans l'auditorium et, en ce sens, nous ne chercherons pas à définir un moyen de transporter le contenu issu du mixage, ou de l'encapsuler dans quelconque format. Le but de cette partie ne sera pas non plus de modéliser des techniques de mixage depuis les exemples étudiés, les problèmes d'expression sonore se posant au cas par cas pour chaque film. Nous nous attarderons plutôt sur les méthodes de mixage amenant aux résultats perçus, pour dégager un peu plus les possibilités du système et éventuellement ses limites. Cette partie nous permettra enfin de lier nos perceptions auditives avec la présence d'une image cinématographique 2D.

Une dernière partie nous permettra de confronter notre travail à une réalité extérieure, en analysant les résultats des tests perceptifs mis en place. Cette partie sera l'occasion pour nous de prendre un peu de distance sur les nombreux constats effectués au préalable, et de récolter les impressions et les sensations d'autres professionnels. Le test réalisé se restreindra volontairement à l'étude de 3 critères. Ces tests perceptifs ne prétendront pas offrir des résultats absolus quant à la pertinence de l'usage d'un système WFS en postproduction sonore. Ils mettront en avant, nous l'espérons, certaines tendances quant à la validité de nos critères d'étude. Si la finalité perceptive de notre étude est importante, il ne faudra pas oublier qu'il s'agit également de documenter une méthode de travail nouvelle, autour d'un système très peu utilisé à l'heure actuelle dans le cinéma.

Ce document est construit de telle sorte à guider le lecteur vers des résultats intermédiaires. Chaque partie est en effet ponctuée d'analyses, de conclusions, de discussions. Nous invitons donc le lecteur à consulter librement chacune de ces parties.

1. Le son spatialisé au cinéma

Cette première partie vise à déterminer le cadre et le contexte dans lequel notre étude s'inscrit. La volonté de recréer un espace fut au cœur des premières recherches sonores. Un rapide rappel historique de l'histoire du son au cinéma nous amènera à envisager le 5.1 sous son aspect technique mais surtout de porter un regard sur les codes esthétiques qui se sont construits autour de son utilisation. Dès lors, il s'agira de questionner le pourquoi de ces codes et de savoir ce qui tient d'avantage aux limites du système de diffusion et ce qui fait d'avantage l'objet de limites induites par la perception audiovisuelle elle-même.

Avec l'arrivée sur le marché de nouveaux systèmes prétendant redéfinir la spatialisation du son au cinéma, la conception du son spatialisé sera sans doute amenée à évoluer. Nous ne pouvons prétendre ici qu'à une explication de ces technologies et leurs principes, sans pouvoir apporter d'éléments de réflexion sur leur utilisation et leurs apports perceptifs ou créatifs.

Nous proposerons enfin une première approche théorique de la WFS. Ces éléments, que nous avons voulu complets mais abordables, nous amèneront à justifier pourquoi la WFS nous semble intéressante dans un cadre cinématographique. Nous émettrons enfin plusieurs hypothèses, a priori, sur l'utilisation de ce système au cinéma.

1. Le son spatialisé au cinéma

1.1. Une histoire du son spatialisé

« C'était en octobre 1987, à la convention de la SMPTE. Les gens se demandaient « combien de canaux doit on avoir pour le format standard du son au cinéma ? » Certains disaient deux...d'autres quatre...un autre disait huit. J'ai levé ma main et dit « cinq point un ». Tout le monde m'a regardé en se demandant « Mais de quoi parle t-il ? ». »

Tomlinson Holman, ingénieur du son et inventeur de la norme THX, d'après
[Kerins, 2001].

L'histoire de la diffusion du son au cinéma est une succession de procédés, d'inventions éphémères, d'échecs ou de succès, technologiques et/ou commerciaux. Là où d'autres arts sonores ont très vite fixé leur format de travail et de diffusion, le cinéma a toujours cherché et cherche encore une autre solution. Si cette recherche semble ne pas trouver de point final, c'est sans doute que *« Le cinématographe est une écriture avec des images en mouvements et des sons »*¹. C'est une trivialité certes, mais elle semble pourtant être la raison de cette recherche permanente du format qui permettrait d'adapter le son à la puissance du médium visuel.

Cependant, il ne faut pas oublier que le cinéma, s'il est un art, est aussi une industrie, qui cherche d'avantage à répondre aux lois de son marché qu'à s'encombrer des réflexions fondamentales sur ce qu'elle produit. On pourrait nous

¹ Bresson Robert, *Note sur le cinématographe*, Gallimard, Paris, 1975, p.18.

1. Le son spatialisé au cinéma

répondre que les films évoluant, il est nécessaire d'adapter le format de diffusion du son en fonction de cette évolution. On pourrait également formuler une réponse qui inverserait la cause et la conséquence, invoquant le fait que les inventions techniques ont donné naissance à une autre forme narrative ou scénographique. En effet, du cinéma primitif au cinéma contemporain, les évolutions technologiques ont permis une amélioration qualitative du son et de l'image, « *mais il n'y a jamais eu de relation mécanique entre une avancée technique et l'invention d'une forme filmique conséquente* »². Les liens entre l'invention des outils et l'évolution esthétique de l'art cinématographique sont trop complexes pour se résumer à un sens unique.

Dès la fin du XIX^{ème} siècle, les premières expérimentations autour de la spatialisation du son furent mises en place. En 1881, lors de l'exposition internationale de l'électricité, Clément Ader installa dix microphones sur les bords de scène de l'Opéra de Paris qui alimentaient des récepteurs téléphoniques. Les quarante spectateurs pouvaient suivre les mouvements des comédiens en écoutant les micros 1-6, 2-7... Ce système d'écoute stéréophonique, quoique intéressant, fut considéré comme trop coûteux et trop complexe pour qu'on lui envisage un avenir industriel.

² Canonville Christian, « Dimension sonore au cinéma : variations imaginatives », *Médiamorphoses* n°18, 2006, p. 46.

1. Le son spatialisé au cinéma

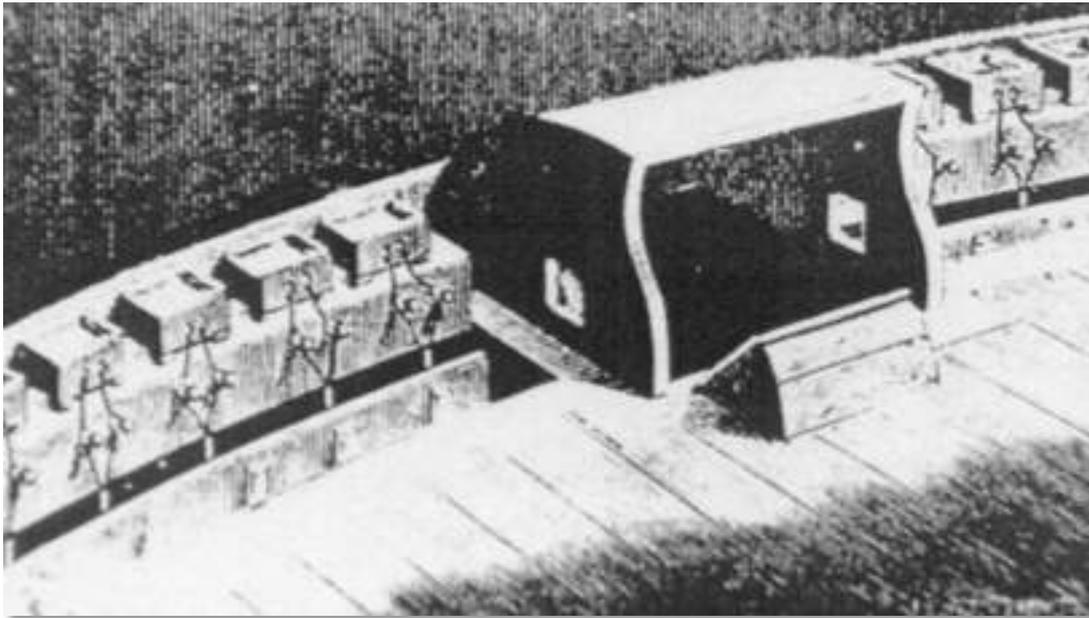


Fig. 1.1 : Vue des microphones de scène du Théatrophone. Source [Hugonnet, 2000]

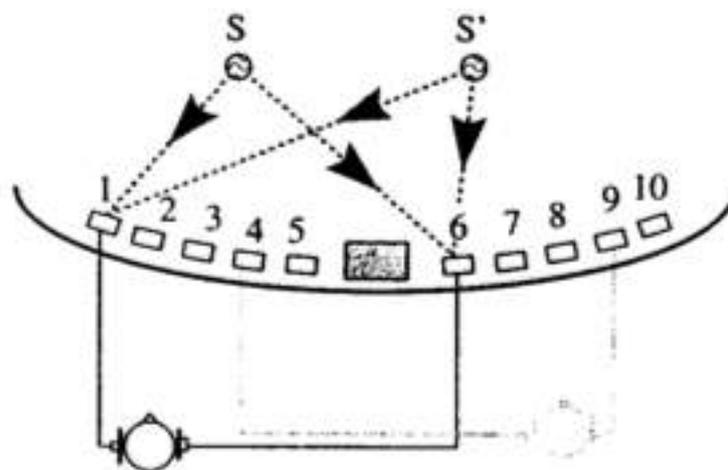


Fig. 1.2 : Disposition des microphones et leurs liaisons aux écouteurs. Source [Hugonnet, 2000]

En 1905, Edwar H. Amet, synchronise pour la première fois une source sonore à une image et, grâce à un commutateur, il peut envoyer le son sur les différents haut-parleurs placés derrières l'écran. Les dialogues des acteurs sont déplacés

1. Le son spatialisé au cinéma

dans l'espace frontale et suivent les déplacements des comédiens. Quelques années plus tard, en 1927, sur l'écran du Warner Theater de New York, Al Jolson, le visage peint de noir, les lèvres exagérément dessinées, caricature un chanteur de jazz noir américain. Ce qui sera considéré comme le premier film parlant de l'histoire du cinéma n'est finalement qu'un film muet à l'éthique douteuse, agrémenté de quelques moments musicaux et d'un seul et unique dialogue synchrone entre le chanteur et sa mère. On parle ici d'un synchronisme absolument approximatif mais ce film consacre un des premiers procédés commercialisés de diffusion d'un son synchrone à une image : le *Vitaphone*.



Fig. 1.3 : Al Jolson dans *The Jazz Singer*, de Alan Crosland, Warner Bros, 1927.

Si *The Jazz Singer* est diffusé en monophonie, c'est sans doute qu'à l'époque, les recherches sur la spatialisation sonore en sont encore à leur balbutiement. Ces

1. Le son spatialisé au cinéma

recherches sont le terrain d'une véritable guerre industrielle entre deux compagnies anglosaxonnes, la compagnie Bell (USA) dirigée par Fletcher et la firme EMI, dirigée par Blumlein.

Fletcher travaille sur le concept de l'holophonie, pensant qu'il faut recréer physiquement le champ acoustique pour restituer l'espace correctement. Pour cela, il met en place un rideau de microphones omnidirectionnels pour capter le champ d'une scène acoustique. Ces microphones sont reliés directement à un banc de haut-parleurs qui restituent le champ sonore capté. Dans la théorie et dans le système proposé par Fletcher, il n'y a pas de point d'écoute privilégié. Cependant là aussi, les impératifs commerciaux freine sa démarche. Pour que le système soit envisageable à l'échelle industrielle, on demande à la compagnie Bell de réduire le nombre de canaux ce qui dégrade considérablement le champ restitué, la notion de front d'onde n'étant plus vraiment respectée.

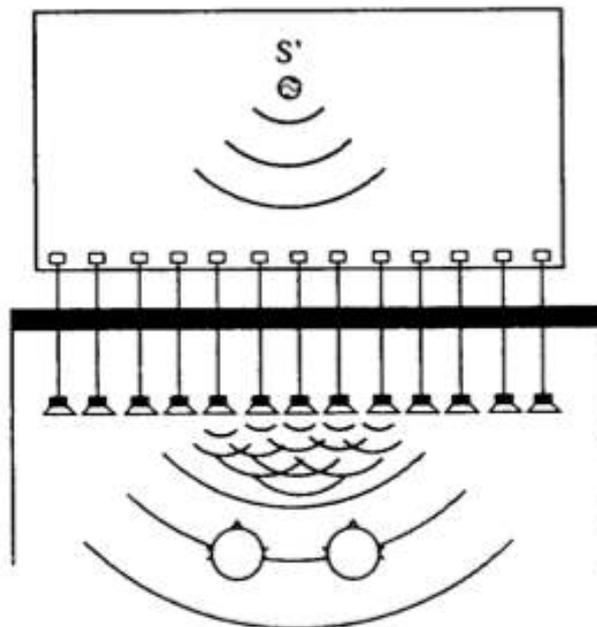


Fig. 1.4 : Rideau de microphones relié un banc de haut-parleur reconstituant un front d'onde. Source [Hugonnet, 2000].

1. Le son spatialisé au cinéma

De son côté Blumlein considère qu'une approche psycho-acoustique du phénomène est plus intéressante que les fronts d'ondes pour reconstruire le champ acoustique. Pour cela il doit définir une position de l'auditeur, au sommet d'un triangle équilatéral formé avec les haut-parleurs. C'est par la différence d'intensité des signaux émis par les haut-parleurs qu'il entend recréer l'illusion perceptive du positionnement d'une source entre les deux enceintes. En 1931, Blumlein dépose les brevets de base de l'enregistrement et de la restitution stéréophonique qui deviendront un standard dans de multiples domaines du son. Seulement, la perception et la localisation des sources virtuelles ne sont valables que pour un espace d'écoute réduit. Dès qu'un auditeur se trouve en dehors de cette zone, il introduit des différences d'intensité et de temps supplémentaires qui ne correspondent plus aux différences initiales entre les canaux : les sources ne sont plus entendues à l'endroit où elles avaient été placées au départ.

Si ce système ne semble pas être adaptée à l'expérience d'écoute collective, c'est pourtant lui qui s'imposera et sera commercialisée. Moins coûteux et moins complexe, il est à la source de beaucoup d'autres systèmes dérivant de ce principe.

Par la suite, l'industrie cinématographique s'intéresse assez logiquement à ces recherches pour trouver un remplaçant à la monophonie. Plusieurs procédés se succèdent sans qu'aucun ne parvienne vraiment à remplir les critères financier et pratique imposés par les exploitants tout en proposant une qualité sonore suffisante. Le plus célèbre d'entre eux, le *Fantasound* (1941), inventé par William Garity, alors ingénieur des studios Disney, est utilisé pour le film *Fantasia* de

1. Le son spatialisé au cinéma

Stokowski. Dans son article sur l'histoire de *Fantasia*³, John Aldred relate le contenu du journal de la SMPTE d'Aout 1941, dans lequel on trouve des détails sur la complexité de l'installation nécessaire à l'enregistrement de l'orchestre et au mixage du film. Il est certain que ce système n'avait été conçu et envisagé que pour ce film « orchestre », produit avec un budget bien supérieur à la moyenne des productions américaines de l'époque. Très vite, Garity et le mixeur du film J. N. A. Hawckins ont compris que le *Fantasound* ne serait qu'une parenthèse dans la recherche de la diffusion sonore au cinéma, tant sa complexité semblait rédhibitoire devant les impératifs commerciaux.



Fig. 1.5 : Garity et Hawckins devant une partie de l'installation du *Fantasound*. Source [Journal of the Society of Motion Picture Engineers, 1941].

³ John Aldred, *Fantasia !*, consulté le 2 mai 2015 sur <http://www.filmsound.org/film-sound-history>

1. Le son spatialisé au cinéma

Le cinéma peine donc à instaurer un format de diffusion de manière pérenne, combinant à la fois une mise en œuvre simple, un coût modéré, et une bonne qualité sonore. Se suivent alors diverses inspirations de la théorie de Blumlein et de la stéréophonie, cette dernière ayant déjà fait sa place dans les mondes musical et radiophonique. En augmentant le nombre de canaux, le problème du stockage des informations apparaît. Le son synchronisé à l'image est en effet dépendant de la place qu'il occupe sur la pellicule et ne dispose que d'un espace limité : il est en effet possible d'inscrire uniquement deux pistes en marge des photogrammes de l'image. Pour bénéficier malgré tout de plusieurs canaux de restitution, des recherches sont menées sur le matriçage, afin de coder plusieurs canaux sur les deux pistes audio disponibles. En 1976, la firme Dolby met sur le marché le Dolby SR, à la fois procédé de réduction de bruit et technique de matriçage. Ce format sera le premier après la monophonie à s'imposer à grande échelle dans l'industrie cinématographique. Le Dolby SR permet de matricer un canal gauche, centre, droit et *surround* sur deux canaux. Pourtant, le matriçage reste un système délicat à exploiter, puisqu'il introduit des problèmes de phase et de diaphonie entre les canaux.

Il faudra attendre les années 90 pour que des systèmes proposant des canaux discrets, indépendants les uns des autres, par un codage de voie et non de signal, voient le jour. Ce sont les normes toujours en vigueur à l'heure actuelle (Dolby SR-D pour *Spectral Recording Digital* en 1992, suivi en 1993 de son concurrent le DTS pour *Digital Theater Sound*) gravitant autour d'un dispositif de diffusion commun, le 5.1. Aujourd'hui les trois voies frontales (gauche, centre, droite), les haut-parleurs latéraux (plusieurs pour le canal arrière gauche et le canal arrière droit), et le caisson de grave, se retrouvent dans la quasi-totalité des salles obscures et la totalité des auditoriums de mixage cinéma.

1. Le son spatialisé au cinéma

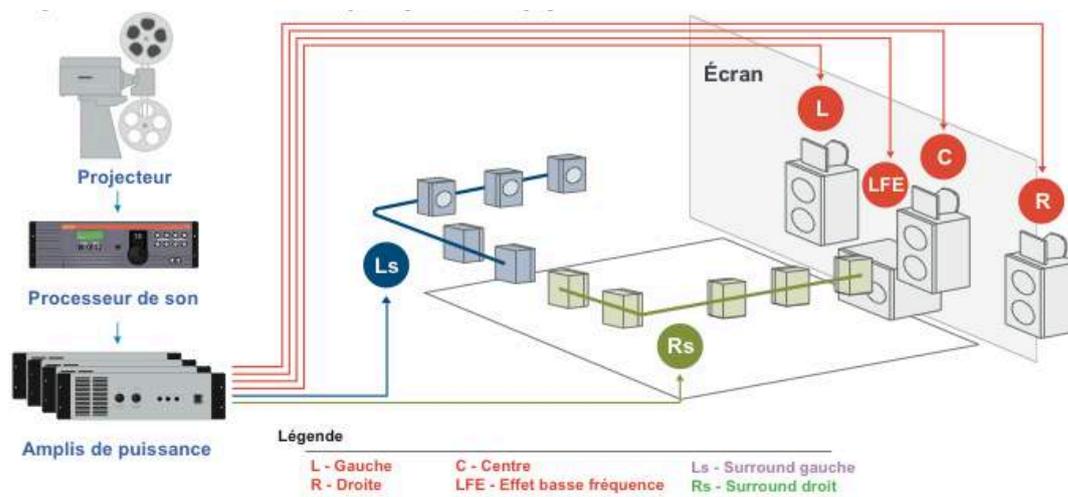


Fig. 1.6 : Configuration d'une salle équipée d'un système 5.1. Source [Dolby, 2012].

1. Le son spatialisé au cinéma

1.2. Le 5.1, une esthétique codifiée

Les formats multicanaux sont en effet devenus les systèmes majoritaires de diffusion du son. Le plus répandu d'entre eux, le 5.1 apparaît nous l'avons dit, depuis plusieurs années déjà, comme le format standard de la diffusion du son au cinéma. Il a ouvert une autre voie de réflexion autour la construction de l'identité sonore d'un film, sa conception de l'espace et, par-là, son rapport au temps. Cette faculté ne lui est pourtant pas réservée et, bien avant son apparition, ces questions semblaient déjà trouver une forme de réponse à travers d'autres formats antérieurs. Son apparition a permis cependant de résoudre des problèmes techniques liés aux précédentes technologies et a fourni une autre palette expressive au mixeur, un autre chemin d'exploration du rapport entre l'image et le son qui peut désormais s'étendre au delà de l'écran. Ce format s'avère malgré tout être le lieu d'un conflit scénographique lorsqu'il est associé à une projection visuelle.

L'origine de ce conflit tient sans doute dans la conception même du dispositif de restitution. Le 5.1 peut être considéré, dans une première approche, comme un format dérivé de la stéréophonie : à la paire stéréophonique frontale, on ajoute un canal de centre et deux canaux arrières, eux même assimilés dans leur fonctionnement à une autre paire stéréophonique. Cependant, Millot et Pelé nous rappellent que si ce système dérive bien technologiquement de la stéréophonie, la situation perceptive qu'il induit dans sa manière de considérer l'espace lui est propre (cf. [Millot & Pelé, 2006], p. 284). Les auteurs postulent également que « *tout système qui crée des sources sonores virtuelles et possède en même temps de sources réelles, dispersées et de caractéristiques déterminées, est susceptible*

1. Le son spatialisé au cinéma

d'engendrer des conflits de la perception spatiale ». Car, s'il faut considérer le 5.1 indépendamment de la stéréophonie en termes de perception, ce dernier conserve de son format originaire, la notion de *sweet spot*. Nous rappelons que cet anglicisme désigne la position d'écoute de référence où l'auditeur serait pleinement en mesure d'entendre le mixage tel qu'il a été conçu par le mixeur. Cette notion a trait à toute situation d'écoute sur un système qui, pour placer des objets sonores dans un espace virtuel, utilise des valeurs d'intensité différentes entre les signaux qui alimentent les haut-parleurs du dispositif. Ce sont ces différences qui permettent de situer un objet sonore entre deux points de diffusion. L'effet de précedence nous indique que la localisation d'une source sonore dépend de la direction d'arrivée du premier front d'onde. On comprend donc assez rapidement que ces valeurs différentielles ne peuvent être conservées que pour un seul point, défini par une triangulation entre l'auditeur et le dispositif : un autre auditeur, placé plus près d'une source de diffusion que d'une autre, ne bénéficiera pas du même équilibre de niveau et localisera le son à un autre endroit. Le système fabrique des images sonores que l'on qualifiera de « fantômes » dans le sens où elles n'existent que par une forme d'illusion, efficiente seulement pour un point d'écoute et qui se rompt dès le moindre déplacement de l'auditeur. Les sources placées entre l'avant et l'arrière sont plutôt instables, ce qui les rend d'autant plus dépendantes de la position d'écoute. Dans les salles de cinéma, là où 2 haut-parleurs arrières ne suffisent pas pour délivrer un niveau sonore satisfaisant, il devient nécessaire d'utiliser d'autres haut-parleurs de surround, latéraux. En conséquence, pour de nombreux auditeurs/spectateurs, certains des haut-parleurs n'apparaîtront plus derrière, mais devant, perturbant l'image sonore frontale. De plus les sources placées entre les haut-parleurs voient leur timbre modifiés et la position des haut-parleurs doit être la même que lors de la production du contenu.

1. Le son spatialisé au cinéma

Lorsque l'on associe ce système imparfait à la projection d'une image frontale sur un écran, ces conflits spatiaux, inhérents au système, se conjuguent aux problématiques de la perception d'un objet audiovisuel. On pourrait alors voir les défauts du 5.1 comme rédhibitoires à sa mise en place dans un dispositif cinématographique, tant les distorsions de localisation qu'il semble impliquer pour une position excentrée par rapport au sweet spot semblent à première vue incompatibles avec ce que Michel Chion appelle le « *contrat audiovisuel* » [Chion, 1991]. Ce contrat définit les principes de coexistence d'un son et d'une image au sein d'un dispositif audiovisuel qui s'attache à nous faire associer ce que nous entendons à ce que nous voyons. Si ces rapports audiovisuels sont culturels et historiques, ils reposent malgré tout sur des phénomènes psycho-physiologiques.

S'étendant au-delà de l'audiovision, la « synchrèse » consiste à associer dans un seul phénomène « *l'évènement sonore ponctuel et l'évènement visuel ponctuel* » [Chion, 1998]. L'auteur définit dans un autre temps un autre phénomène d'assimilation, « l'aimantation spatiale », phénomène qui tend à regrouper dans un espace commun un objet visuel et un objet sonore, localisé pourtant sur un autre point ou provenant d'une autre direction. Pour Claude Baiblé, « *la reprojexion vers la source de l'excitation sensorielle est la caractéristique fondamentale des systèmes perceptifs à distance* » [Baiblé, 1990]. C'est « *la force éjaculatrice de l'œil* »⁴, selon Robert Bresson. En définissant ainsi ces phénomènes, nous voyons qu'ils introduisent la possibilité pour le 5.1 d'exploiter notre faculté à associer dans une démarche réflexive, un son à une image et que les distorsions géométriques perçues par les auditeurs excentrés pourraient s'atténuer ou du moins devenir tolérables. Ce que nous avançons ici présente pourtant une limite, et la codification esthétique des bandes sons construites sur un système 5.1 en est la preuve irréfutable.

⁴ Bresson Robert, op. cite, p. 24.

1. Le son spatialisé au cinéma

En distinguant ce qui tient aux « *conflits de la perception spatiale, qui résultent de la géométrie des systèmes* » (cf. [Millot & Pelé], p. 285) c'est-à-dire à leur technologie elle-même, et l'utilisation qui en est faite, il semble que le 5.1 ait depuis longtemps trouvé un cadre à son expression, instaurant des codes que rarement il transgresse, lorsqu'il s'est affranchi du "pittoresque" et du "spectaculaire" dont on l'affuble encore trop souvent. C'est en lui fixant des limites que les acteurs de la bande son au cinéma ont donné au 5.1 la possibilité de devenir un format compatible, malgré ses défauts, avec le « spectacle audiovisuel ». La spatialisation des différents matériaux constituant la bande sonore est « moralisée », régie par des principes qui trouveront certes toujours des figures d'exceptions, mais qui semblent s'être érigés comme des lois presque inébranlables. Les voix sont confinées dans le canal de centre, c'est l'image et l'*aimantation audiovisuelle* qui les placent. Les ambiances s'étalent du centre aux bords de l'écran. La musique trouve sa place dans un espace parfois sensiblement plus large et certains effets sont envoyés à l'arrière.

Si le 5.1 a construit une esthétique sonore frontale, il reste à définir la cause de cette construction qui semble tenir à la fois à la nature de la projection d'un objet audiovisuel et aux imperfections du système précédemment décrites, dans un équilibre que nous ne prétendons pas pouvoir définir ici mais qui constitue finalement la profonde motivation de notre travail. Cette frontalité semble s'imposer et ce quelque soit le format de diffusion sonore. L'exaltation sensorielle grossière proposée dans certains films, par une utilisation outrancière de la spatialisation, ne fait que conforter cette idée que l'engagement frontal du regard, appelle à une orientation sonore identique.

Bien que cet « *engagement frontal [...] suppose une implication limitée du schéma corporel dans le dispositif* » [Baiblé, 1998]. l'ouverture spatiale du son au delà de

1. Le son spatialisé au cinéma

l'écran n'est, fort heureusement, pas proscrite et semble au contraire être à même de questionner autrement notre rapport à une œuvre cinématographique et notre degré d'implication dans l'espace de la narration. La légitimité, s'il en est, des sons spatialisés à l'arrière est régie par un principe de cohérence, qui voudrait que l'espace diégétique ne soit jamais éclaté par leur mise en application. Rick Altman explique dans [Altman, 1995] que les spectateurs ont une attente bien différente des canaux surround. Là où une forme de cohérence spatiale est attendue sur la façade, lieu de projection de l'image, les canaux surrounds eux, s'ils respectent une logique géométrique déterminée par la narration, détournent très vite l'attention du spectateur. « Les mouvements d'attention auditive » provoqués par une utilisation excessive des canaux arrières, marquent une rupture inadaptée avec l'espace de projection et la narration (cf. [Bailblé, 1998], p. 240). Pourtant, le cinéma offre des exemples d'utilisations mesurées des surrounds, certains films pariant sur le fait que l'auditoire associera les sons diffusés par ces canaux au même espace diégétique que ceux provenant de l'écran. Mark Kerins relate les propos du *sound designer* Garry Rystrom, à propos des canaux arrières : « *Les surround dans un film ne sont pas efficaces s'ils sont constamment actifs. Vous devez choisir ces moments* » [Kerins, 2011], joignant son propos à celui de Walter Murch, qui définit le *could* et le *should* dans [Murch, 1992], parlant du devenir des objets sonores dans l'espace de la salle de cinéma. L'utilisation des arrières peut maintenant être dictée par ce qui semble approprié à l'histoire d'un film et non par une contrainte technologique comme c'était le cas avec le Dolby Stéréo.

Cependant il ne faut pas considérer l'apport du 5.1 sous le seul angle des surrounds déconnectés de l'ensemble du système dans lequel ils s'implantent, ni considérer comme une finalité la possibilité de déplacer plus librement les sons dans l'espace. L'utilisation de largeur de l'espace disponible à une influence esthétique qui dépasse le cadre de la simple localisation des éléments sonores et de l'enveloppement du spectateur.

1. Le son spatialisé au cinéma

En ouvrant la perspective, le 5.1, via des canaux discrets indépendants les uns des autres, a permis un démasquage des sources. Cette possibilité a entraîné une complexification des bandes sons matérialisée par la densification des éléments sonores simultanés, ou du moins regroupé dans le temps. Dans une démarche presque paradoxale, on utilise alors encore plus de sons qui se combinent les uns aux autres pour créer une autre matière. On réintroduit alors une forme de masquage pleinement conscient, en maîtrisant à divers degrés le phénomène. Là où le système semble ne plus poser de limite, il se confronte pourtant à la réalité psycho-acoustique de l'oreille humaine, qui ne peut prêter attention qu'à un nombre limité de sons à la fois. Murch dépeint une règle du « *deux et demi* » [Murch, 1992] pour limiter le nombre de sons constituant un objet sonore, rejoignant ainsi Beck dans [Beck, 2003], pour qui une utilisation limitée du nombre de sons permet à l'auditeur d'entendre à la fois « *la forêt et les arbres* ». Cette densification conduit également à une autre tendance, celle des niveaux forts, trop forts, trop souvent et de manière injustifiée. Randy Thom dans [Thom] décrit cette tendance qui pour lui, tient d'avantage aux volontés des réalisateurs qu'à la pratique des mixeurs. Elle est sans doute le reflet de la faiblesse de certains films qui abusent de la dynamique offerte par le 5.1, pensant combler par l'excès sonore, leurs lacunes scénaristiques, narratives et la vacuité de leur propos.

1.3. Deux systèmes émergents de spatialisation

A l'image du 5.1 dans les années 90, le monde du cinéma garde encore aujourd'hui une place pour les innovations sonores. La question du « son 3D » n'a jamais été aussi récurrente qu'à l'heure actuelle, avec toutes les interprétations que cela peut impliquer. Parle-t-on de « son 3D » dès lors que les canaux de diffusion soient étagés à des hauteurs différentes ? Associe-t-on ce terme lorsque le système considéré est capable de donner l'illusion d'un espace crédible autour de l'auditeur, comme en sont capables les technologies binaurales ? Ou enfin, parle-t-on de « son 3D » à partir du moment où une dimension nouvelle, située derrière l'écran, est virtuellement restituée par le système d'écoute, comme cela peut être le cas en WFS ?⁵

Aujourd'hui, il semblerait que les 2 technologies commercialement en lice pour suppléer au 5.1 se basent sur une approche zénithale du « son 3D ». L'*Auro 3D*, développé par Barco et récemment plébiscité par Georges Lucas, puis l'*Atmos* de Dolby, déjà implanté dans de nombreuses salles, se font peu à peu une place dans la course aux normes de diffusion sonore. Il reste bien entendu à déterminer où se situe la limite, pour ces systèmes à canaux « de plafond », entre un simple argument de vente, et un réel apport créatif. Une étude complète et détaillée serait à mener. Cela dit, nous entreverrons dans cette partie que l'apport de ces systèmes ne se limite pas à une « simple » diffusion verticale des sources sonores.

⁵ Nous développerons ces aspects de la WFS plus tard dans notre étude, en particulier dans les parties 1.4 et 2.

1. Le son spatialisé au cinéma

1.3.1. Atmos, développé par Dolby



Fig. 1.7 : Logo de la technologie Dolby Atmos. Source www.dolby.com.

Contexte

L'Atmos est le nouveau format de mixage, d'encodage, et de restitution sonore développé par Dolby. Apparue en 2012, il compte aujourd'hui une vingtaine de salles en France, contre un peu plus d'une trentaine en Europe. En France, le premier cinéma à abriter le système fut le Pathé *Wepler*, situé à Paris. *Rebelle*⁶, des studios Disney-Pixar fut le premier film mixé en Atmos, tandis que *Le Hobbit : Un Voyage Inattendu*⁷, fut le premier film de cette technologie diffusé en France. 2013 fut une année particulièrement productive pour le nouveau format de Dolby. Après *En solitaire*⁸, premier film français mixé en Atmos, vint le spectaculaire *Gravity*⁹ qui, de part son contenu interstellaire, fut la première diffusion Atmos convaincante à nos yeux. A l'heure actuelle, bien que la technologie semble bien assise aux Etats-Unis, les salles françaises n'accueillent que très peu de contenu Atmos, et ce, même si le mixage a été réalisé de cette façon.

⁶ *Rebelle*, titre original *Brave*, réalisé par Mark Andrews et Brenda Chapman, sorti en France le 1 août 2012.

⁷ *Le Hobbit : Un Voyage Inattendu*, réalisé par Peter Jackson, sorti en France le 12 décembre 2012.

⁸ *En Solitaire*, réalisé par Christophe Offenstein, 97min, sorti en France le 25 août 2013.

⁹ *Gravity*, réalisé par Alfonso Cuarón, 91min, sorti en France le 23 octobre 2013.

1. Le son spatialisé au cinéma



Fig. 1.8 : La salle 1 du Pathé Wepler, première installation Atmos en France.

Configuration globale

Les quelques lignes suivantes n'ont pas pour objectif de détailler au lecteur les caractéristiques techniques de l'installation. Nous nous contenterons de décrire les éléments technologiques qui démarquent l'Atmos des dispositifs « classiques » 5.1. Pour de plus amples informations, nous invitons le lecteur à consulter les documents suivants : [Dolby, 2012], [Dolby 2, 2012] & [Besse, 2013].

Chez Dolby, « son 3D » est synonyme d'une nouvelle dimension sonore, verticale, proposée au spectateur. Cette caractéristique est en effet l'argument de vente principal des systèmes Atmos par rapport aux systèmes discrets plus anciens, qui étaient déjà pour la plupart, développés par Dolby... Cependant l'ajout de haut-parleurs au plafond n'est pas la seule plus value technologique, a priori, du

1. Le son spatialisé au cinéma

dispositif. Là où les systèmes actuels restent dans une approche « orientée canal », le nouveau format de Dolby est le premier à proposer un concept « orienté objet ».

Le mixage « orienté objet » est une notion qui fut développée pendant le projet CARROUSO. L'objectif premier était en effet de « transférer un champ sonore, généré dans un espace réel ou virtuel, dans un autre espace habituellement éloigné » [Brix et al., 2001]. La norme MPEG-4, développée dans le cadre du projet CARROUSO, permet entre autres d'abriter un contenu audio, séparé de ses informations spatiales associées. Dans Atmos, le principe est le suivant :

- Le mixeur produit une bande son dans un auditorium, d'une configuration de haut-parleurs donnée.
- Les informations de spatialisation sont enregistrées séparément des sources sonores, telles des coordonnées géométriques à associer aux éléments sonores.
- La bande son ainsi produite, faite d'un flux audio séparé d'un flux de métadonnées de spatialisation, arrive dans une nouvelle salle, la salle de restitution, disposant d'une autre configuration de haut-parleurs.
- Le processeur de restitution se charge de calculer « la meilleure façon » de reproduire le positionnement des sources, tel que l'a renseigné le mixeur en auditorium.
- En d'autres termes, le *processing* de traitement sur le système est différent selon les salles, le nombre de haut-parleurs ou le niveau émis par chacun d'entre eux également, mais le rendu final est potentiellement identique, du moins en termes de placement de sources sonores.

1. Le son spatialisé au cinéma

Notons bien qu'Atmos ne prévoit pas une « bonne » restitution pour des configurations de haut-parleurs complètement différentes de la configuration globale préconisée, illustrée par la figure 1.9 ci-dessous. De plus, il va de soi que plus la salle est dotée d'un nombre important de haut-parleurs, plus elle sera capable de restituer la position d'une source avec précision.

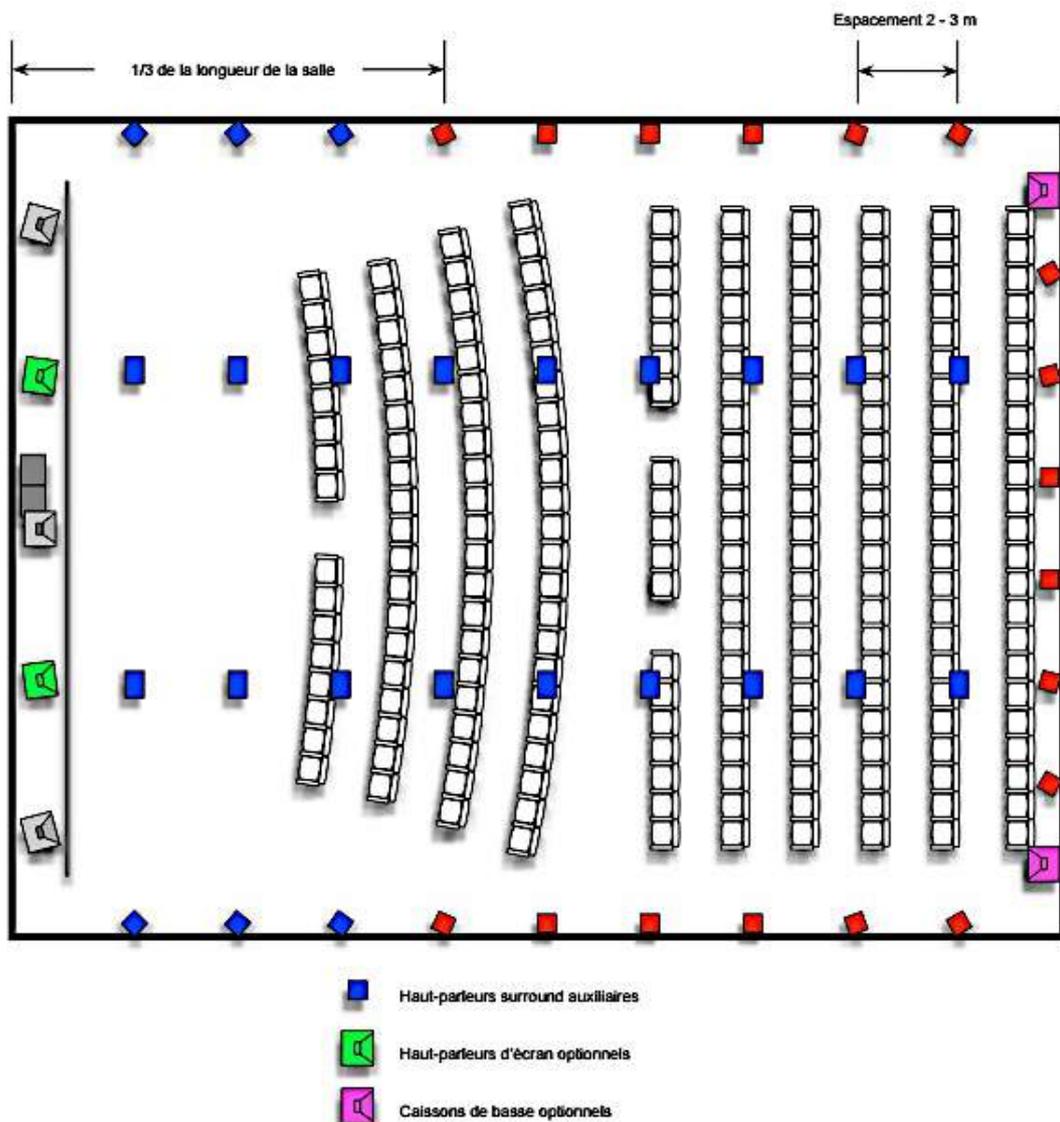


Fig. 1.9 : Configuration de haut-parleurs recommandée pour un système Atmos. En bleu apparaissent les nouveaux haut-parleurs de surround ; en vert sont indiqués les haut-parleurs d'écran optionnels (écran de plus de 12m de large) ; en rose apparaissent les subwoofers utilisés comme renfort de basse des canaux de surround. Source [Dolby, 2012].

1. Le son spatialisé au cinéma

Les canaux de surround « classiques » doivent habituellement être capables de délivrer, pour chacun d'entre eux, 82 dB_{SPL}(C). Pour sa part, Atmos prévoit un niveau sonore de 85dB_{SPL}(C) pour chaque canal, c'est à dire pour chaque enceinte acoustique présente dans la salle. Outre la débauche de niveau sonore que cela représente, chaque haut-parleur doit également disposer d'une étendue spectrale élargie par rapport aux canaux de surround « classiques ». On passe en effet d'une largeur spectrale très restreinte (dimensionnée par la norme ISO 2969 X), à un registre fréquentiel quasi « pleine bande », pouvant aller de 50Hz à 19kHz. Enfin, de sorte à étendre le spectre de ces canaux jusque dans les très basses fréquences, Atmos prévoit l'intégration de *subwoofers* à l'arrière de la salle, ce qui permet un contrôle des fréquences graves par *bass management*.

Le mixage

Avec Dolby Atmos, Le mixage *orienté objet* s'implémente dans l'espace de travail de l'auditorium. Le mixeur dispose de canaux de mixages principaux (appelés *Beds*) comparables aux actuels *stems*, ainsi que de pistes objets (ou *Objects*). 128 pistes d'objets sonores sont exploitables, tandis que 10 d'entre elles tout au plus, sont réservées pour représenter les *Beds* 5.1, 7.1, ou 9.1 (configuration 7.1 avec 2 voies de plafond). Il reste ainsi 118 objets sonores pouvant être positionnés n'importe où par le mixeur, et donc être potentiellement, diffusés sur n'importe quel point de restitution de la salle. Notons que ces objets sonores peuvent être diffusés un par un ou groupés, cela dépendant de l'effet recherché.

1. Le son spatialisé au cinéma

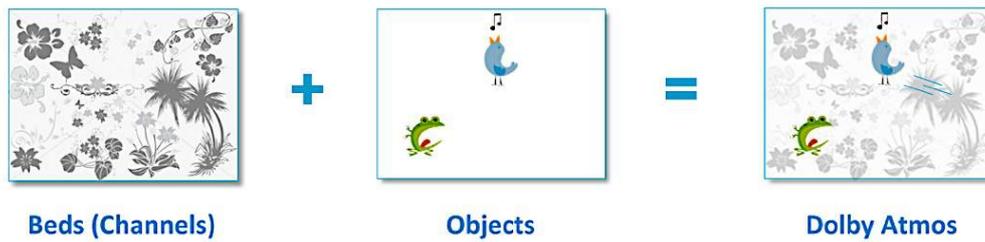


Fig. 1.10 : Les canaux de Beds, additionnés aux objets sonores, forment l'intégralité du flux Atmos qui sera ensuite transporté jusqu'au serveur de projection. Source [Dolby, 2012].

Les systèmes Atmos incluent un processeur de rendu et de restitution, le RMU (Rendering Mastering Unit), qui comme nous l'avons évoqué précédemment, se charge de calculer « la meilleure manière » de restituer les intentions de spatialisation du mixeur. De base, le processeur est accompagné de deux logiciels, l'un sous forme d'application, l'autre sous forme de plugin intégrable dans Pro Tools, permettant respectivement un contrôle visuel global et le positionnement des sources. Ci-après suit une brève présentation de ces deux logiciels. Pour plus de détails, nous invitons le lecteur à consulter [Dolby, 2013].

- *Dolby Atmos Monitor* application :

C'est cette application qui permet le contrôle audio et visuel du processeur RMU, pendant que celui-ci est en train de calculer les métadatas liées aux sources. Le Dolby Atmos Monitor permet principalement :

- o La bascule entre le mode Atmos *orienté objet*, et le mode traditionnel *orienté canal*, assurant ainsi une rétrocompatibilité du système ;

1. Le son spatialisé au cinéma

- L'enregistrement du mixage principal, composé des *Beds* et des objets, qui sera embarqué par le serveur lors de la projection en salle ;
- L'écoute du mixage enregistré, via la fonction *playback*.

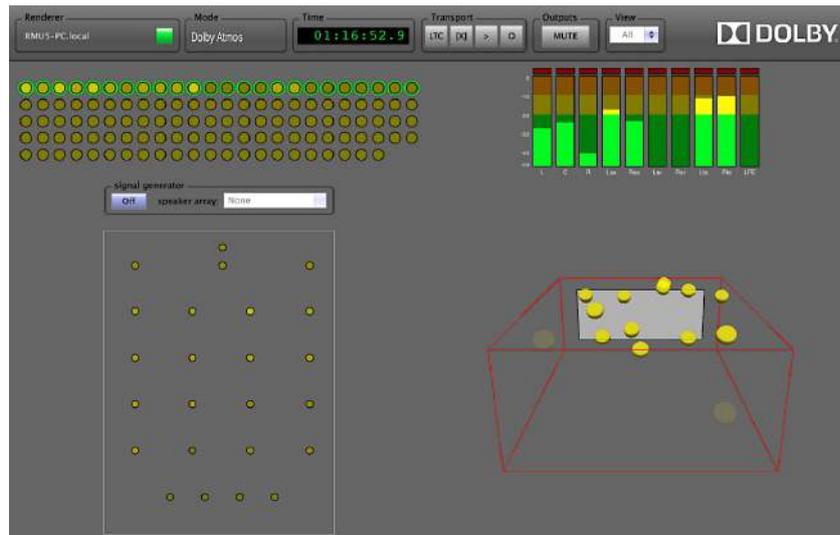


Fig. 1.11 : L'application Monitor inclue un crête-mètre, des indicateurs de signal pour chaque objet, ainsi que des visualisations de positionnement d'objets. Source [Dolby, 2013].

- *Dolby Atmos Panner* plugin :

Ce plugin permet donc le positionnement des objets sonores dans l'espace tridimensionnel défini par la configuration Atmos. Il est inséré sur chaque piste mono ou stéréo de Pro Tools, qui seront des pistes labélisées « objet ». Les positions d'objet ainsi que d'autres métadonnées deviennent alors des paramètres d'automation comme les autres. Pro Tools se charge ensuite d'envoyer ces métadonnées au RMU, via liaison ethernet, alors que

1. Le son spatialisé au cinéma

le signal audio est acheminé via liaison MADI. Enfin, le rendu est réalisé par le RMU, lors du mixage ou de la phase d'enregistrement.

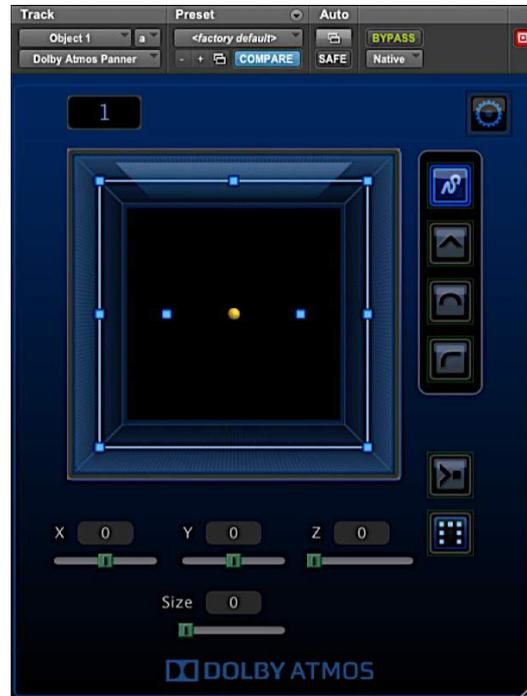


Fig. 1.12 : Le plugin Dolby Atmos Panner se présente comme une fenêtre de pan tridimensionnelle, visuellement une simple extension des panners « classiques » dans Pro Tools. Source [Dolby, 2013].

Pour finir, nous pourrions noter une collaboration intéressante entre Dolby et le constructeur de consoles AMS-Neve qui, avec leur modèle DFC Gemini, sont parvenus à intégrer un panner « 3D » à leurs machines. Le résultat en est plutôt attirant, d'autant que cette console permet le contrôle d'autres logiciels tels que Nuendo ou Pyramix (via le protocole *Star Command*). Enfin, la DFC Gemini permet également le mixage en format Auro 3D.

1. Le son spatialisé au cinéma

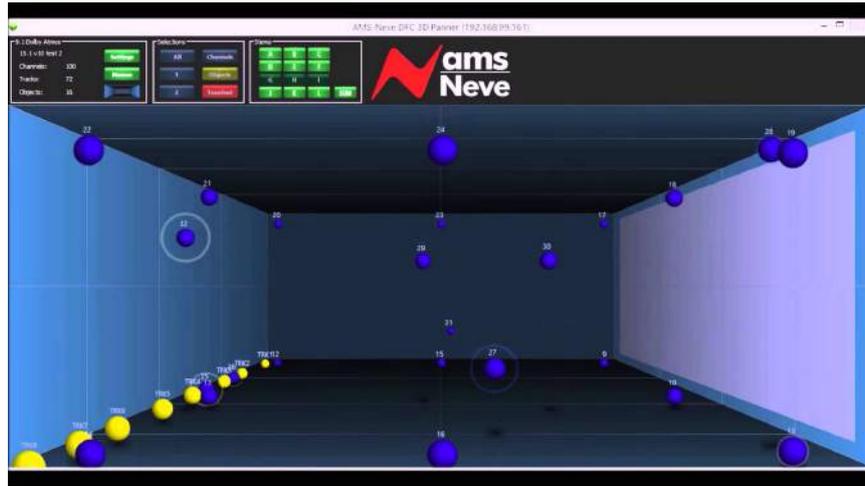


Fig. 1.13 : Le panner de la DFC Gemini, de part sa clarté et son intégration dans une console, forme une solution intéressante aux formats de mixage « 3D ».

1. Le son spatialisé au cinéma

1.3.2. Auro 3D, promotionné par Barco



Fig. 1.14 : Logo de la technologie Auro 3D. Source : www.auro-3d.com.

Contexte

L'*Auro 3D* est un système de spatialisation sonore inventé en 2006, mis en avant par la société belge Barco, entreprise aux multiples facettes également active sur le marché du vidéoprojecteur cinéma. A l'heure actuelle, bien qu'aucune salle française n'abrite un tel système, une implantation assez importante est à noter aux Etats-Unis et en Inde. Le premier film mixé en Auro 3D fut *Red Tails*, de la société *Lucas Film* (2012). Trivialement, aucun film français n'a été mixé grâce à cette technologie, excepté un certain *Lucy*, réalisé par Luc Besson en 2014, mais mixé... aux Etats-Unis. La plupart du contenu Auro 3D touche bien entendu les productions américaines, mais également le marché indien, issu de l'industrie *Bollywood*. Tout comme chez Atmos, Auro Technologies propose un développement du concept dans la branche « home cinema », décidément un marché fortement prisé par les principales sociétés concurrentes.

Configuration globale

1. Le son spatialisé au cinéma

De la même manière que pour l'Atmos, nous n'embarquerons pas le lecteur dans une description détaillée du système, bien que pour Auro 3D, les différences avec un système 5.1 standard soient assez minces. Nous invitons le lecteur à consulter les documents suivants, pour plus de précisions : [Van Baelen et al.] & [Besse, 2013].

Chez Barco, l'approche « son 3D » est différente de chez Dolby. Le principal apport tient de la configuration du système, qui double, dans le sens de la hauteur, les canaux gauche, centre et droit installés derrière l'écran, ainsi que les deux canaux surrounds placés sur les côtés de l'auditeur. Cependant, *exit* le mixage objet : Auro 3D reste un format de mixage discret, par canal, qui ne permet pas une description virtuelle du positionnement des sources. En effet, le système embarque « simplement » des lois de panoramique standard, étendues à la dimension verticale, comme en témoignent les plugins présentés ci-après.

Différents formats

On notera différents formats, différentes configurations de haut-parleurs, exploitables par le système. La présence d'un canal central élevé marque la différence principale entre les systèmes « home cinéma » et les systèmes pour salles obscures.

- En home cinéma :
 - o 9.1 : ajout de 4 canaux au dessus de L, R, Ls et Rs
 - o 10.1 : ajout des 4 canaux précédents, ainsi qu'un canal zénithal (le « Voice of God »)

1. Le son spatialisé au cinéma

- Au cinéma :

- 11.1 : par rapport aux configurations home cinéma, elle prévoit un ajout d'un canal zénithal frontal, au dessus du canal central déjà existant en 5.1.
- 13.1 : c'est la configuration maximale, qui prévoit une séparation des *sides* surround, et des *rear* surround, comme en 7.1.

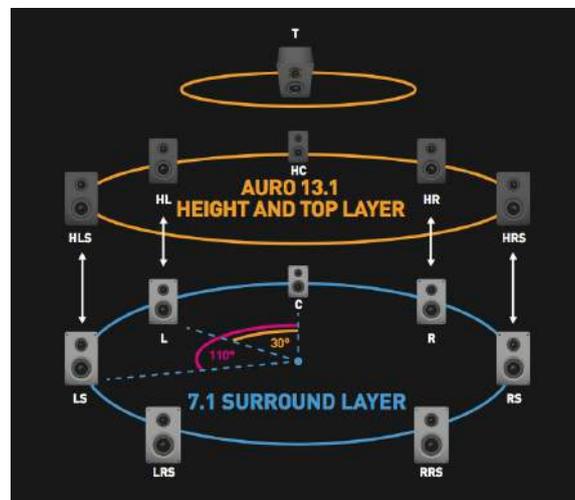


Fig. 1.15 : Illustration de la configuration Auro 3D maximale, en 13.1. Les canaux de surround « side » et « rear » sont séparés. Source : www.auro-3d.com.

L'idée d'Auro 3D est la suivante : on double « simplement » le nombre d'enceintes en ajoutant une ceinture plus haute à la configuration déjà existante en 5.1 (cf. figure ci-dessus). Ceci permet, selon Barco, d'obtenir un « son 3D avec une configuration des plus réduites » [Van Baelen et al.].

1. Le son spatialisé au cinéma



Fig. 1.16 : Configuration Auro 3D 13.1 dans une salle de cinéma. On imagine que l'installation des haut-parleurs situés en hauteur peut devenir complexe pour les salles basses de plafond. Source : www.auro-3d.com.

Cependant, Barco ne semble pas prévoir de normes particulières concernant le registre fréquentiel de ses canaux surround, comme en témoignent leurs recommandations : « pour passer d'un son surround à un vrai son 3D, il suffit simplement d'ajouter 8 haut-parleurs à la configuration existante. » [Van Baelen et al.]. Rappelons que pour Atmos, Dolby prévoit une étendue spectrale de toutes les enceintes jusqu'à 50Hz.

Le mixage

En plus de son processeur de rendu et de restitution, Auro 3D propose des plugins permettant le « positionnement » des sources sonores. Nous souhaitons préciser au lecteur que nous préférons réserver le terme « positionnement » pour les systèmes de diffusion *orientés objets*. En effet, pour les systèmes *orientés canaux*, comme c'est le cas pour Auro 3D, le positionnement des sources est

1. Le son spatialisé au cinéma

dépendant de la configuration des haut-parleurs, chose qui est a priori détournée sur un système comme Atmos par exemple.

Les deux plugins nécessaires à l'exploitation d'un système Auro 3D sont les suivants :

- Un panner, *Auro 3D Authoring Tools* :

C'est ce plugin qui permet le « positionnement » des sources décrit précédemment. Il dispose des caractéristiques principales suivantes :

- o exploitation de la verticalité
- o compatibilité préservée avec des systèmes standards « non-3D », grâce à un algorithme de *downmix* développé par Barco
- o possibilité de contrôler les *downmixs* générés



Fig. 1.17 : Fenêtre principale du Auro-Panner. En bref, un panner gérant la verticalité, avec une interface apparemment agréable. Source : www.auro-3d.com.

1. Le son spatialisé au cinéma

- Un plugin d'upmix, *Auro-Matic* :

Ce plugin permet de réaliser un *upmix*, technique décrite dans notre étude, partie 2.2.3, p. 122. L'idée est de « convertir » un contenu plus restreint (stéréophonique, quadriphonique, 5.1, ...), en un contenu diffusable sur tous les haut-parleurs du système Auro 3D. Ses caractéristiques principales sont les suivantes :

- o conversion d'un contenu monophonique, stéréophonique, ou tout autre contenu « 2D » (5.1, 7.1), en contenu Auro-3D
- o intégration dans le plugin *Auro-3D Authoring Tools* décrit précédemment



Fig. 1.18 : Fenêtre principale du Auro-Matic. En bref, un upmix gérant la verticalité, avec une interface tout aussi agréable. Source : www.auro-3d.com.

1. Le son spatialisé au cinéma

Barco présente sa technologie comme étant capable de restituer du « son 3D », tout en s'adaptant aux systèmes plus anciens et aux autres formes de diffusion. En pratique, les étapes de fabrication sont les suivantes :

- Un mixage « master » est tout d'abord réalisé, au format Auro-3D maximal (11.1 ou 13.1), prévu à la diffusion en salle de cinéma.
- Par suite, un downmix est automatiquement réalisé, mais peut tout de même être contrôlé par le mixeur.
- Le downmix 5.1 obtenu peut servir, comme usuellement, à la déclinaison de la bande son pour d'autres supports de diffusion (stéréophonique, ...), ou à la diffusion du contenu dans les salles de cinéma standards (« non-3D »).

1. Le son spatialisé au cinéma

1.3.3. Nouvelles technologies : résumé et ouverture

Les années 2010 marquent un renouveau dans la diffusion du son au cinéma. Deux sociétés, Auro Technologies et Dolby, mettent en avant deux systèmes qui rivalisent vers une normalisation de leur procédé. Il est difficile d'entrevoir, à notre échelle, dans quelle mesure ces technologies ont réussi à se faire un chemin pour tenter de supplanter le standard 5.1, ancré depuis 20 ans dans les salles obscures. En revanche, il nous semble intéressant de nous demander en quoi ces dispositifs représentent un apport vis à vis des systèmes plus anciens.

Le Dolby Atmos, tout d'abord, marque une avancée importante dans la manière de penser le mixage. En passant du côté des formats *orientés objets*, la technologie offre, sur le papier, une solution aux problèmes de transportabilité des mixages. Avec Atmos, la spatialisation apportée est censée pouvoir être transportée d'un lieu à un autre, sur des installations potentiellement composées d'un nombre de haut-parleurs différent. Cet argument va également dans le sens des exploitants de salle, qui ne pourront réellement pas tous s'équiper d'un système composé de 64 haut-parleurs...

Perceptivement, le système Atmos semble à même de contourner l'un des problèmes propres à la restitution 5.1. Grâce à un dimensionnement des haut-parleurs capable d'agir dans les basses fréquences, et grâce à l'ajout de *subwoofers* arrières en vue d'un *bass management*, les nombreux canaux de l'Atmos sont a priori capables d'opérer un enveloppement efficace, jouant sur le *cône de confusion* décrit en 1.4.3, p. 61. En effet, la diffusion de contenu

1. Le son spatialisé au cinéma

fréquemment bas autour de l'auditeur est, selon nous, un apport potentiel en vue d'une immersion réussie. Par ailleurs, le prolongement des canaux de surround jusqu'à l'écran semble permettre une transition plus douce entre le champ sonore frontal et le champ sonore diffusé depuis la salle, ce qui offrirait par exemple d'avantage de fluidité de mouvement entre l'image et la zone d'écoute.

Cependant, l'apport d'une telle définition sur les canaux de surround reste à évaluer. En effet, il nous semble légitime de nous demander si la précision de localisation en Atmos, amenée par la possibilité de spatialiser un son dans n'importe quel haut-parleur, est une réelle plus value vis à vis du « spectacle audiovisuel » proposé au spectateur. Selon Claude Baiblé, « il y a une risque scénographique à transformer les sons off (contigus, semi-proches ou lointains) en sources localisées, susceptibles d'attirer l'attention hors écran, ce qui reviendrait à découpler l'attention auditive de l'attention visuelle – ordinairement coaxiales – en deux systèmes distincts de localisation » [Baiblé, 2013]. Pour le moment, ce sera au mixeur, en collaboration avec le réalisateur, de déterminer de la pertinence d'un tel potentiel de discrimination en dehors de l'écran.

Une autre question est de savoir si les canaux zénithaux, proposés à la fois par Atmos et par Auro 3D, sont à même de créer des effets perceptifs intéressants dans le cadre d'une diffusion cinématographique. Est-ce qu'entendre « la pluie tomber sur nous » ne ferait pas plutôt office d'effet « gadget », potentiellement dérangeant à notre implication narrative ? « Peut-on dissocier aussi facilement le point de regard du point d'écoute, sans entraîner des réaction d'orientation quasi-automatiques, réactions ici sans objet, puisque l'écran rectangulaire reste frontalement posé ? » [Baiblé, 2013].

Enfin, que dire d'Auro 3D, qui en plus de garder un format de mixage et de restitution *orienté canal*, mise une innovation sur l'ajout de canaux surround sur

1. Le son spatialisé au cinéma

une dimension verticale. Notre expérience d'écoute sur ce type de système est nulle. Cependant, nous nous demandons si l'apport de verticalité (déjà peu perçue par l'être humain) sur des canaux placés derrière l'auditeur peut prétendre à des effets vraiment intéressants en termes d'enveloppement ou d'immersion.

Au vu du nombre important de canaux requis par les deux dispositifs décrits précédemment, et au vu de leurs différences assez fines, si ce n'est en termes de nombre de haut-parleurs, avec les systèmes « classiques » de restitution (5.1, 7.1, ...), il nous semble intéressant de nous plonger dans l'étude d'une autre technologie, différente dans ses fondements théoriques : la WFS. Précisons au lecteur que le système WFS que nous étudierons ne disposera pas de canaux zénithaux. Même si cela est aujourd'hui possible avec l'intégration d'une dimension verticale dans les produits proposés par Sonic Emotion, nous nous bornerons volontairement à l'étude d'un système disposé sur le plan horizontal. Les canaux zénithaux forment, selon nous, une étude à part entière comme cela a déjà pu être le cas dans [Dupas, 2007].

1. Le son spatialisé au cinéma

1.4. La WFS : un concept innovant

Dans cette sous-partie, l'objectif ne sera pas d'amener le lecteur à une compréhension détaillée de la théorie sous-jacente à la WFS. Bien qu'aborder les principes fondamentaux nous semble important pour comprendre les phénomènes en découlant, nous n'amènerons pas le lecteur à une explication mathématique ni même algorithmique du fonctionnement du système. En se plaçant du point de vue du mixeur cinéma, les *teintes* perceptives et technologiques nous semblent plus appropriées quant à la problématique que nous nous sommes fixés (cf. Introduction).

La sous-partie suivante ne contiendra par conséquent qu'une couche primaire de principes théoriques. Cependant, tout au long de ce mémoire, et en particulier dans la partie 2, nous proposerons au lecteur des encadrés explicatifs, lui permettant s'il le souhaite, de rentrer plus finement dans les principes de la synthèse de front d'onde et de ses éventuels phénomènes impliqués. Nous invitons d'autre part le lecteur à se référer à la liste de documents ci-dessous, non-exhaustive, mais tout de même très fournie en éléments théoriques sur le sujet :

- [Berkhout, 1988] : **BERKHOUT A. J.**, *A holographic approach to acoustic control*, Journal of the Audio Engineering Society, vol. 36, 1988
- [Berkhout et al., 1993] : **BERKHOUT A. J., DE VRIES D., VOGEL P.**, *Acoustic control by wave field synthesis*, Journal of the Acoustical Society of America, vol. 93, 1993

1. Le son spatialisé au cinéma

- [Corteel, 2004] : **CORTEEL Etienne**, *Caractérisation et extensions de la Wave Field Synthesis en conditions réelles*, Thèse de doctorat : traitement du signal, Université Paris VI – Pierre et Marie Curie, 2004

Dans les sous-parties 1.4.3 et 1.4.4, nous tenterons d'analyser, a priori, les apports et les limites de la WFS dans un contexte de postproduction cinématographique. Les apports technologiques du système s'adaptent-ils perceptivement au « spectacle audiovisuel » proposé ? Les contraintes technologiques de la synthèse de front d'onde sont-elles cohérentes avec l'industrie cinématographique actuelle ? Autant d'hypothèses que nous formuleront, dans le but de les valider, ou non, lors de nos travaux expérimentaux (cf. parties 2, 3, 1).

1. Le son spatialisé au cinéma

1.4.1. Les principes physiques fondateurs

La WFS, ou synthèse de front d'onde en français, se base sur un principe physique destiné en premier lieu à l'étude des ondes optiques : c'est le principe de Huygens, découvert en 1678, traitant de la reproduction de front d'ondes, base théorique de l'holographie (BDP : procédé permettant la restitution de l'image d'un objet en 3 dimensions). La compréhension de ce principe permet de bien appréhender le fonctionnement d'un système de restitution sonore par synthèse de front d'onde. Il peut s'énoncer comme suit :

« Chaque élément d'un front d'onde peut être considéré comme étant le centre d'une perturbation qui génère à son tour un front d'onde secondaire. »

Ce principe fut complété plus tard par Fresnel :

« Chaque point M d'une surface S atteinte par la lumière peut être considéré comme une source secondaire émettant une onde sphérique dont l'amplitude et la phase sont celles de l'onde incidente au point M . »

Si un point M reçoit une onde d'amplitude $E(M, t)$, alors on peut considérer qu'il réémet une onde sphérique de même fréquence, même amplitude, et même phase. En somme, au lieu de considérer que l'onde progresse de manière continue, on décompose sa progression en imaginant qu'elle avance de proche en proche.

1. Le son spatialisé au cinéma

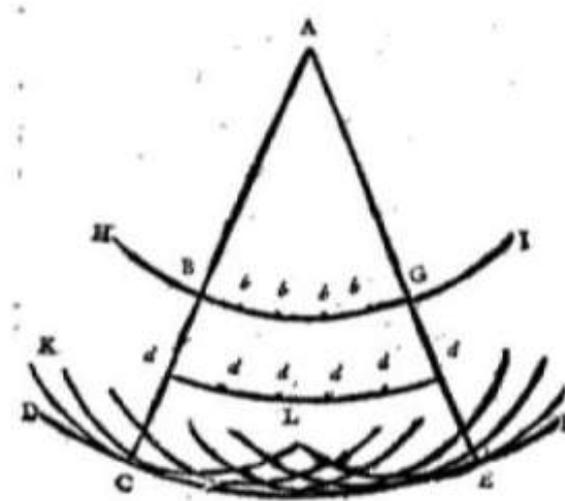


Fig. 1.19 : Illustration originale du principe de Huygens. La source primaire A produit une onde qui se propage en HI puis en DF. Situé sur HI, le point B peut être une source secondaire, centre d'une perturbation KCL, de même que tous les autres points b de la sphère de centre A et de rayon AB. La somme de toutes ces perturbations recréeront l'onde originale.

La superposition des fronts d'onde secondaires ainsi considérés permet alors la création d'un front d'onde principal, théoriquement identique au front d'onde original. Si chaque front d'onde secondaire est émis par une source, alors on peut mentalement remplacer ces dites sources par des haut-parleurs. Un grand nombre de petits haut-parleurs proches forment la dénommée *rampe de haut-parleurs*.

1. Le son spatialisé au cinéma

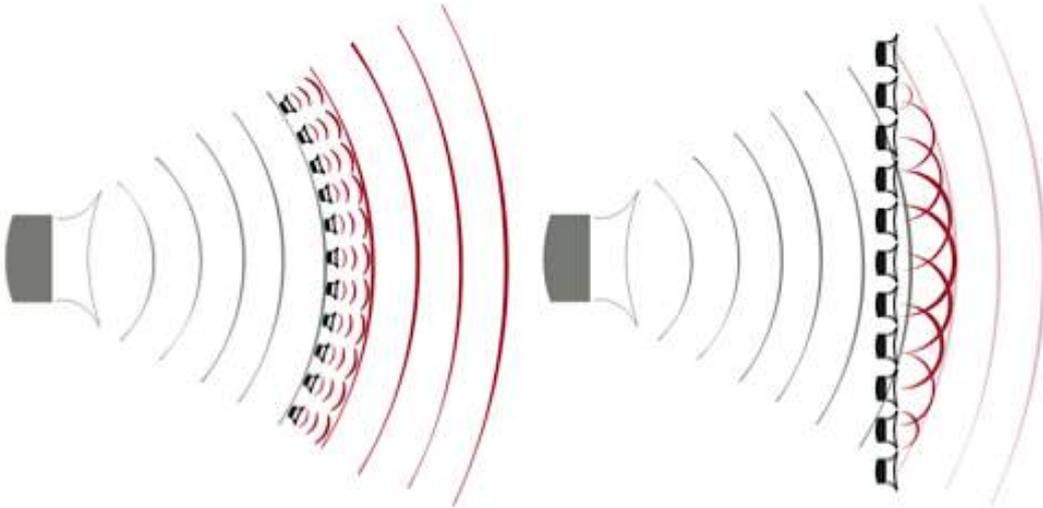


Fig. 1.20 : A gauche se trouve la « traduction mentale » des sources en petits haut-parleurs proches. A droite, l'ensemble des haut-parleurs est réorganisé en rampe linéaire, cela facilitant l'intégration et la gestion de l'installation.

Chaque haut-parleur de la rampe est attaqué par un signal lui étant propre, correspondant entre autres aux résultats donnés par les algorithmes issus des intégrales de Kirchhoff-Helmholtz, et des représentations du théorème de Rayleigh. Alors qu'une intégrale vise à exprimer une grandeur avec un suivi continu de celle-ci, nous verrons par la suite que la WFS s'appuie sur un ensemble d'approximations de ces formules (cf. partie 1.4.2).

L'intégrale de Kirchhoff-Helmholtz permet d'exprimer, et donc de connaître, la pression contenue en un point placé dans un volume délimité par une surface. La pression est issue de la contribution de sources extérieures à cette surface, avec pour hypothèse de connaître la pression sonore sur une distribution de sources secondaires situées sur la surface délimitant le volume.

On comprend assez rapidement que le volume ci-décrit correspondra réellement à la zone de reproduction (ou *zone d'écoute*), que la surface délimitant ce volume (contenant les sources secondaires) correspondra aux limites de la zone d'écoute

1. Le son spatialisé au cinéma

(contenant les haut-parleurs), et que les sources dont on pourra connaître la pression dans le volume seront des sources a priori extérieures à la zone d'écoute. Les sources primaires, synthétisées par les sources secondaires, seront appelées les *sources virtuelles*, car faisant référence à un espace virtuelle situé derrière les haut-parleurs.

Soit p_0 le champ de pression sur $\delta\Omega$, k le nombre d'onde, et \vec{n} le vecteur unité dirigé vers l'extérieur. \vec{R} est le vecteur entre la source secondaire et le point d'écoute, tel que $\vec{R} = \vec{r} - \vec{r}_0$.

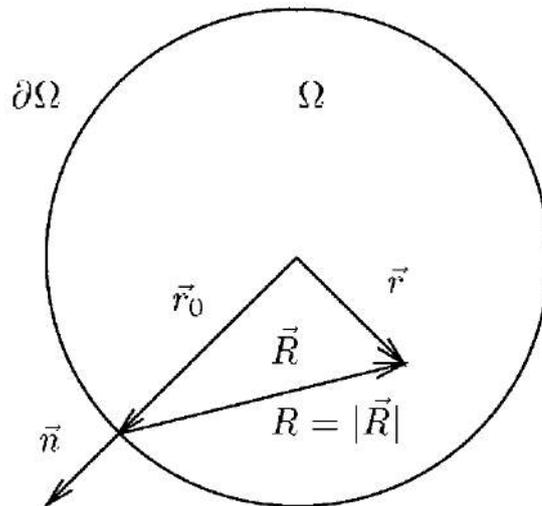


Fig. 1.21 : Géométrie pour l'intégrale de Kirchhoff-Helmholtz. Ce schéma est à lire dans le cas général en 3 dimensions.

Sur la frontière $\delta\Omega$ de la surface Ω , le champ acoustique est donné par l'intégrale suivante :

$$p(\vec{r}) = \iint_{\partial\Omega} \left[\vec{\nabla} p_0 \cdot \vec{n} - \frac{\vec{R}}{R} \cdot \vec{n} (1 + jkR) \frac{p_0}{R} \right] \frac{e^{-jkR}}{4\pi R} dS_0 \quad \forall \vec{r} \in \Omega$$

1. Le son spatialisé au cinéma

Dans l'expression comprise entre les crochets, on peut identifier une combinaison linéaire de termes **monopolaires** (1^{er} terme), et de termes **dipolaires** (2nd terme). Les monopoles sont assimilables technologiquement à des sources omnidirectionnelles, tandis que les dipôles donnent suite à des sources bidirectionnelles.

$$\frac{e^{-jkR}}{4\pi R} \qquad \frac{\vec{R}}{R} \cdot \vec{n} (1 + jkR) \frac{e^{-jkR}}{4\pi R^2}$$

Ces principes physiques impliquent l'usage d'une infinité de haut-parleurs pour pouvoir imaginer une synthèse parfaite. Ils impliquent également l'utilisation conjointe de monopoles et de dipôles pour restituer le champ sonore. Aujourd'hui, la plupart des systèmes se base sur l'usage de monopoles, ce qui rend un champ sonore correct à l'intérieur de la zone d'écoute [Sporer, 2004] : seul le champ sonore externe à la zone d'écoute, censé être non-nul, obtiendra une valeur non-négligeable de par l'usage de haut-parleurs omnidirectionnels.

Enfin, la conception de haut-parleurs omnidirectionnels pour toute la bande audible restant très complexe, l'intégration d'un degré supérieur de directivité dans la rampe est a priori inévitable. Etienne Corteel, dans [Corteel, 2004] p. 36, avance que « la synthèse de figure de directivité est correcte en dessous de la fréquence d'aliasing et se dégrade légèrement en basses fréquences quelque soit la source considérée ». Pour cela, l'auteur prévoit des méthodes « d'égalisation multicanale » censées compenser les problèmes, dans le plan horizontal, induits par la directivité des haut-parleurs : « la transparence de la reproduction n'est pas

1. Le son spatialisé au cinéma

parfaite mais bien meilleure que dans le cas d'une égalisation individuelle »¹⁰.

¹⁰ Dans [Corteel, 2004], p. 116, il est distingué deux types d'égalisation visant à compenser les défauts d'un système. L'égalisation individuelle considère individuellement les haut-parleurs du dispositif et vise à compenser leurs défauts en moyenne. On cherche dans ce cas à compenser les « défauts » des transducteurs acoustiques, pour une petite zone privilégiée de l'auditoire. L'égalisation multicanale tend quant à elle à assurer la synthèse correcte de l'objectif de reproduction par l'ensemble des haut-parleurs du dispositif, au moins en un nombre de points finis appelés *points de contrôle*. On cherche ici à assurer la validité du « résultat final » sur les points de contrôle, cette validité étant ensuite étendue à la zone d'écoute toute entière. Le document renseigne également sur les méthodes « d'inversion » mises en jeu (p. 118).

1. Le son spatialisé au cinéma

1.4.2. De la théorie à la pratique : simplifications et conséquences

Si l'on s'en tient à l'approche académique, la construction d'un champ sonore par WFS fait appel à un trop grand nombre de haut-parleurs, au vu des applications envisagées dans cette étude. En effet, il est irréaliste d'imaginer une salle de cinéma composée d'un nombre d'enceintes à celui attendu par la théorie, et donc recouverte sur toute la surface située autour du spectateur. Au delà du coût financier que cela représenterait, les difficultés d'installation et de maintenance ne seraient pas à même d'intéresser les exploitants dans le contexte d'une industrie cinématographique. Il y a donc nécessité de se restreindre à un dispositif plus réaliste. Les limites en découlant, largement explicitées dans la littérature jusqu'ici dans [Corteel, 2004] & [Caulkins, 2007] par exemple, feront l'objet d'un bref résumé dans les lignes qui suivront.

Passage d'un plan à une ligne infinie de sources secondaires

Les principes précédemment évoqués nous ont montré la synthèse d'un champ sonore via la contribution de sources secondaires placées sur une surface, elle-même composée de plans infinis. Dans les situations d'écoute quotidiennes, et en particulier dans une salle de cinéma, les spectateurs sont situés dans un même plan horizontal. Selon Etienne Corteel, on peut « montrer que les sources secondaires situées dans le plan horizontal produisent la contribution la plus importante du champ sonore perçu par les auditeurs » [Corteel, 2004], p. 18.

1. Le son spatialisé au cinéma

Il est alors possible de réduire le plan de diffusion à une ligne infinie de haut-parleurs. Cette approximation n'est pas sans perturber le champ sonore synthétisé, qui ne correspond plus exactement au champ primaire : une symétrie cylindrique autour de la ligne de haut-parleurs en découle, ce qui modifie la loi d'atténuation du système, passant d'une dispersion sphérique à une dispersion cylindrique. Selon Etienne Corteel, cette modification n'affecte « que très peu la transparence sauf à proximité immédiate du banc de haut-parleurs » [Corteel, 2004], p. 77.

Passage d'une ligne infinie à un segment

De manière assez évidente, il nous semble impossible de concevoir un système disposant d'une rampe de haut-parleurs de taille infinie. Un retour à la réalité nécessite alors une troncature de la rampe : on passe d'une ligne à un segment. Ce segment peut être assimilé à une fenêtre au travers de laquelle l'auditeur « voit » un espace acoustique. A l'inverse, on peut définir une zone de visibilité pour la source primaire à travers le segment de sources secondaires.

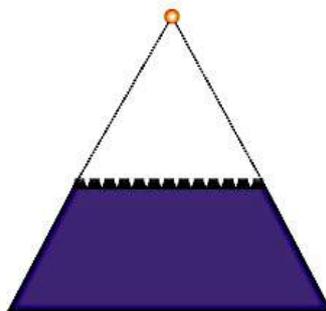


Fig. 1.22 : Illustration de l'effet de fenêtrage. La source primaire (en jaune), peut être synthétisée au travers du segment de sources secondaires dans la zone appelée zone de visibilité (en violet). Source [Corteel, 2004].

1. Le son spatialisé au cinéma

Le fenêtrage implique une restriction de la zone d'écoute, là où la synthèse reste correcte. Le problème est que cette zone dépend de la position de la source virtuelle à synthétiser. On comprend par exemple que si la source jaune se déplace à gauche sur la Fig X, la zone de visibilité, i.e. la zone d'écoute, change de géométrie.

Outre la réduction de la zone couverte, les dimensions finies de la rampe de haut-parleurs entraînent l'apparition d'ondes de diffraction qui suivent le front d'onde principal. Ce phénomène a été mis en évidence par Terrence Caulkins dans [Caulkins, 2007], en simulant la synthèse d'une onde plane de basse fréquence par un banc de haut-parleurs de taille finie. Le résultat a été mesuré par une ligne de microphones situés dans la zone de restitution.

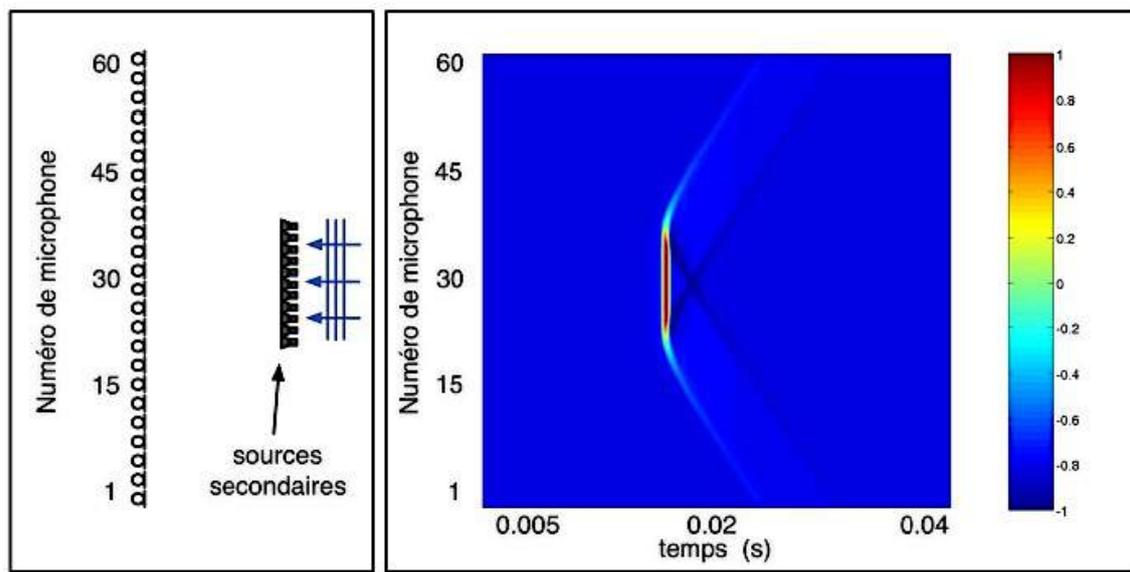


Fig. 1.23 : Mise en évidence de l'effet de diffraction pour un banc de haut-parleurs de taille finie. Source [Caulkins, 2007].

1. Le son spatialisé au cinéma

Le graphique de droite montre l'évolution temporelle du champ capté par la ligne de microphones. On observe l'apparition de deux fronts d'onde parasites (en noir) qui suivent le front d'onde principal et paraissent provenir de sources secondaires qui seraient situées à chaque extrémité du banc de haut-parleurs. Terrence Caulkins cite les travaux de [Nicol, 1999] et [Start, 1997] pour l'étude détaillée de ces effets de diffraction. La méthode usuellement employée pour réduire ces aberrations consiste à atténuer la contribution des haut-parleurs situés aux extrémités. Cette méthode présenterait toutefois le désavantage de réduire le niveau des basses fréquences pour des positions d'écoute excentrées [Corteel, 2004].

Fermeture de la zone d'écoute par plusieurs segments

Une organisation des haut-parleurs en segment, i.e. en rampe (ou en banc), ne permet à l'auditeur que d'accéder à une fenêtre réduite sur la scène sonore, potentiellement frontale. Bien qu'une scène frontale semble être au centre de notre propos en postproduction sonore cinématographique, nous souhaitons offrir la possibilité aux créateurs de disposer de toute la périphérie de la salle de projection. De fait, il faut penser à un système constitué de plusieurs segments, refermant la zone d'écoute. Une organisation rectangulaire de ces segments nous semble être appropriée compte tenu de la géométrie générale des salles de cinéma à l'heure actuelle.

1. Le son spatialisé au cinéma

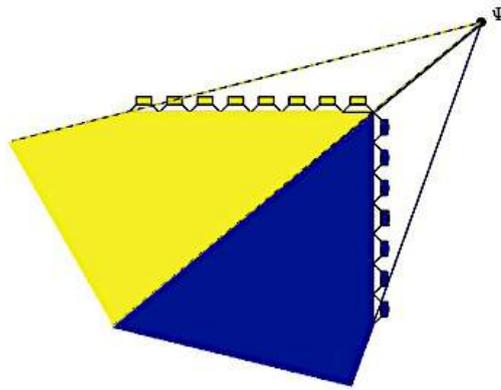


Fig. 1.24 : Contribution des différentes rampes à la synthèse d'une source latéralisée. Source [Corteel, 2004].

Comme le montre la figure ci-dessus, la position de la source virtuelle décidera de la contribution de telle ou telle rampe à la synthèse de la source. Technologiquement, il est difficile de savoir précisément comment ces différentes contributions se mettent en place. En revanche, il est possible d'avancer que dès qu'une source virtuelle « voit » une rampe de haut-parleurs, cette rampe participera à la synthèse de la source (cf. partie 2.1.2).

Echantillonnage de la rampe

Une simplification supplémentaire concerne la discrétisation du banc, ce dernier étant censé être un segment continu constitué d'une infinité de sources secondaires. En effet, quelque soient les haut-parleurs utilisés, ils ne sont ni infiniment petits, ni infiniment proches. Cette discrétisation « équivaut à effectuer un échantillonnage spatial régulier de la distribution de sources secondaires », affirme Etienne Corteel dans [Corteel, 2004], p. 24. De même qu'un échantillonnage numérique amène à un repliement spectral des fréquences les

1. Le son spatialisé au cinéma

plus aigues¹¹, un échantillonnage spatial aura pour conséquence « un phénomène de recouvrement des composantes spatiales du champ sonore synthétisé à partir d'une fréquence limite, la fréquence d'aliasing spatial » [Corteel, 2004], p. 24.

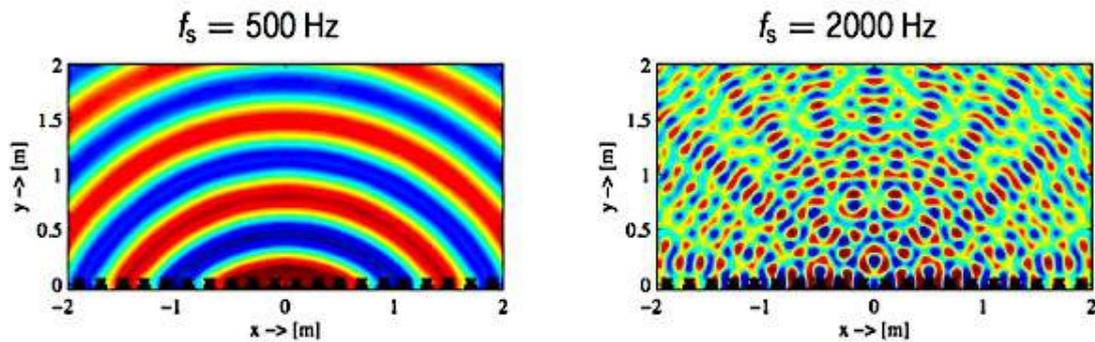


Fig. 1.25 : Mise en évidence du repliement spatial par diagramme du champ de pression d'un système WFS pour deux fréquences différentes.

On peut comprendre qu'en dessous de cette fréquence d'aliasing spatial, le champ sonore restitué par une rampe discrétisée de haut-parleurs sera identique au champ sonore restitué par une ligne continue de sources secondaires. Au dessus de cette fréquence, les ondes issues des sources secondaires ne se superposent plus correctement, et « un phénomène de répétition des composantes spatiales du champ [sonore] apparaît » [Corteel, 2004]. La synthèse n'est plus correcte (cf. ci-dessus). En bref, cette fréquence d'aliasing est fonction de l'écartement entre les haut-parleurs, de la position de la source, et de la position d'écoute. Les conséquences perceptives seraient spatiales (localisation biaisée) et fréquentielles (altération du timbre). Pour plus de précisions sur le sujet, nous invitons le lecteur à consulter l'encadré i p. 96 traitant de l'aliasing spatial.

¹¹ Le théorème de Shannon indique que les fréquences du signal à situées au delà d'une fréquence maximale F_{\max} , avec $F_e = 2 \times F_{\max}$ (F_e : Fréquence d'échantillonnage), sont « repliées » dans le bas du spectre, ce qui a pour conséquence des artefacts audibles (gonflements dans les basses fréquences par exemple).

Introduction du champ réverbérant : la salle d'écoute

Les principes théoriques précédemment évoqués, visant la synthèse d'un champ sonore par une infinité de sources secondaires placées sur un plan infini, sont tous considérés dans un espace de type « champ libre », autrement dit un espace ne contenant ni obstacle, ni paroi, bref, un espace théoriquement « infini ». Outre les diverses simplifications résumées lors des paragraphes précédents, un « retour à la réalité » passera forcément par la prise en compte du lieu de restitution. Tout lieu de restitution contiendra très probablement obstacles et/ou parois, ce qui est d'autant plus certain pour une salle de cinéma.

On comprend bien que la mise en place de parois, souvent placées proches des haut-parleurs, pourra gravement affecter la validité de la synthèse mise en jeu. Les ondes réfléchies par les différents murs de la pièce entreront a priori elles aussi dans la superposition complexe des sources secondaires. Le champ réverbérant ainsi créé complexifie considérablement les modèles ondulatoires permettant le contrôle des sources virtuelles. On s'attardera, dans la plupart des cas, à soigner l'empreinte acoustique du lieu de restitution, afin de la rendre la plus « mat » possible. Les salles de cinéma se positionnent plutôt bien en termes de traitements acoustiques, par rapport à d'autres lieux de diffusion, comme certaines salles de concerts par exemple.

Une compensation par égalisation pourra également être envisagée. Dans [Corteel, 2004], p. 130, l'auteur propose une méthode originale de compensation permettant de « contrôler la synthèse de sources virtuelles dites *de compensation* par la WFS ». Cette technique est chargée de « synthétiser une source de compensation dont les caractéristiques de rayonnement sont données par le

1. Le son spatialisé au cinéma

champ acoustique réfléchi au premier ordre sur la paroi considérée ». D'après l'auteur, cette méthode est algorithmiquement peu coûteuse, bien qu'elle mette en action tous les haut-parleurs du dispositif simultanément. Des études plus poussées montrant l'intérêt de telles égalisations sont à mener, particulièrement dans le domaine de l'industrie cinématographique où les « profils acoustiques » des salles de diffusion sont apparemment moins variables que pour des salles de concert par exemple.

1. Le son spatialisé au cinéma

1.4.3. Apports, a priori, de la WFS

Nous avons vu dans la sous-partie précédente que la translation du concept de la WFS dans la réalité n'était pas sans causer un certain nombre de limites physiques, introduisant des effets indésirables, plus ou moins perceptibles. Cette sous-partie projette de présenter rapidement, au lecteur et à l'utilisateur, ce qu'il est possible de réaliser, a priori, avec un système à synthèse de front d'onde. Cette présentation, également récurrente dans la littérature, sera complétée plus loin dans notre étude, par les constats que nous amènerons découlant de nos pratiques instrumentales (cf. partie 2 & 3).

Séparation de la captation et de la restitution : concept de source virtuelle

Tout d'abord, l'un des concepts fondamentaux de la WFS prévoit une séparation entre le dispositif de captation, et le dispositif de restitution [Berkhout et al., 1993]. Là où la plupart des systèmes actuels lient étroitement la manière de capter la pression sonore à la manière de la restituer, la WFS permet une description de la scène sonore par des sources dites *notionnelles*. Issue d'un système de prise de son quelconque, cette source sera ensuite synthétisée en calculant les signaux d'alimentation de chacun des haut-parleurs du système de restitution. Le terme de *source virtuelle* est plus communément utilisé.

Ainsi, on comprend bien où peut nous mener une telle manière d'aborder le traitement du son. Une captation, faite de 2 microphones, de 3 microphones, de X microphones... Est tout à fait synthétisable sur un système WFS, chose qui est

1. Le son spatialisé au cinéma

invisageable sur un système discret comme le 5.1 sans *downmix* (réduction du nombre de canaux via mixage) préalable.

Elargissement de la zone d'écoute : une posture « exocentrée »

Les technologies issues de la stéréophonie, les technologies *discrètes*, fonctionnent sur un « trompe l'oreille » qui n'est valable que pour une position figée, autrement appelée le *sweet spot*. Pour une position d'écoute différente, l'illusion disparaît, et l'auditeur est amené à percevoir à nouveau les sources de restitution, i.e. les haut-parleurs.

Outre les constats empiriques effectués lors de nos diverses expériences d'écoute (cf. Introduction), la littérature fait largement état d'un certain élargissement de la zone d'écoute (cf. [Corteel, 2004], [Berkhout et al., 1997], etc). Les positions d'auditeur pour lesquelles les sources sont correctement localisées, et donc pour lesquelles l'illusion reste valable, semblent se multiplier et ne plus se résumer à une toute petite portion de la salle de diffusion. En effet, le concept de source virtuelle amène à une position certes virtuelle de la source, mais à une position « virtuellement exacte » de celle-ci. Ainsi, si l'auditeur se déplace dans la zone d'écoute, la direction associée à la source restera identique : un « hologramme sonore » est proposé à l'auditeur.

Selon les termes d'Etienne Corteel dans [Corteel, 2004], la posture d'écoute « égocentrée » sur les systèmes discrets, devient « exo-centrée » en WFS. Cet élargissement de la zone d'écoute a particulièrement attiré notre attention dans le cadre d'une bande son associée à un film. Il semblerait que tout mixeur soit intéressé par la création d'un contenu sonore dont les propriétés spatiales soient partagées par tous les auditeurs de la zone d'écoute, et pas uniquement par

1. Le son spatialisé au cinéma

quelques spectateurs situés proches du point central privilégié. Le simple fait d'imaginer une ambiance stéréophonique, perçue comme étant « stéréophonique » par tous les auditeurs quelque soit leur position dans la zone d'écoute, reste une idée très attrayante.

Distance holophonique de la source

Avec les paramètres de *pan*, le mixeur dispose, depuis l'invention de la stéréophonie, d'un pouvoir de latéralisation, lui permettant de disposer les éléments sonores à différents emplacements de l'image sonore. En plus de ce paramètre, la WFS propose un paramètre de distance virtuelle de la source par rapport au centre du dispositif : c'est la *distance holophonique* [Noguès et al., 2003]. La profondeur associée à une source devient un outil de plus dans la palette du mixeur. Notons bien que cet aspect de la spatialisation ne permet pas à lui seul d'apprécier la distance d'une source [Blauert, 1999]. Dans l'encadré ii (p. 104), le lecteur pourra constater que cette distance holophonique est tout à fait associable aux paramètres « classiques » permettant de jouer sur la *présence de la source* (niveau, filtrage, rapport son direct/son réverbéré, ...) ¹².

Perceptivement, selon Etienne Corteel, « La manipulation de la distance holophonique est rendue possible grâce à la variation cohérente des indices binauraux de localisation auditive perçue lors des mouvements de l'auditeur » [Corteel, 2004], p. 235. Grâce à la manipulation de la distance holophonique, un effet de parallaxe (cf. figure ci-dessous), en relation avec une perspective sonore

¹² *Présence de la source* : terme plus exacte pour parler du degré de présence de la source, et donc par conséquent, de l'éloignement subjectif de celle-ci. Ce terme est issu des études menées par l'IRCAM dans les années 90 sur "l'impression spatiale". Ce paramètre prend en compte plus d'éléments que l'unique rapport son direct/son réverbéré.

1. Le son spatialisé au cinéma

créée par la position des sources virtuelles, est alors créé. Cette perspective est donc a priori cohérente pour l'ensemble des « points de vue » de la scène.

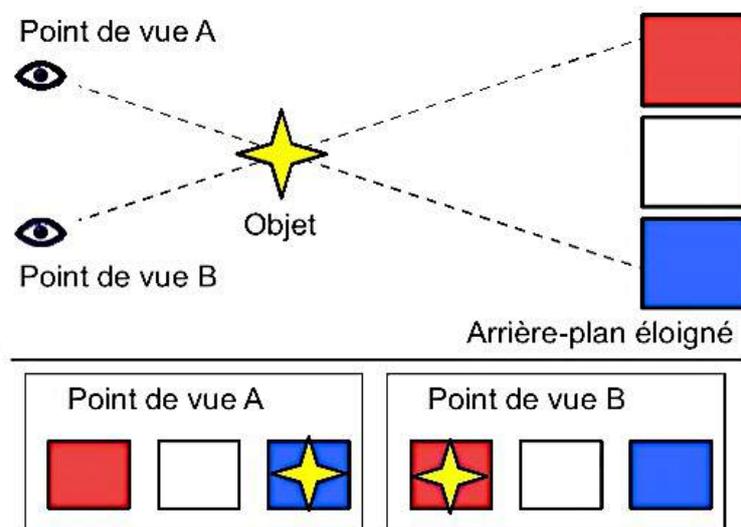


Fig. 1.26 : Principe optique de la parallaxe : Le sujet placé en A percevra l'objet sur le fond bleu, alors que le sujet placé en B percevra l'objet sur fond rouge. Il en est de même pour un phénomène de parallaxe acoustique.

Types de sources synthétisables

Nous avons vu qu'à l'instar des systèmes discrets, la WFS permet de reproduire des sources dont les propriétés de localisation sont partagées par tous les auditeurs de la zone d'écoute. Par ailleurs, une distinction peut être faite entre les différents types de sources synthétisables, directement fonction de la distance holophonique.

Dans *La sensation d'espace dans le spectacle audiovisuel* [Millot & Pelé, 2006], Laurent Millot et Gérard Pelé distinguent trois possibilités pour la création de perspective en WFS :

1. Le son spatialisé au cinéma

- Les **sources ponctuelles** sont des sources situées à l'intérieur de la surface émissive et synthétisées par un front d'onde sphérique sur un plan contenant les oreilles et les haut-parleurs. Ce type de source permet la perception de sources stables en position absolues.
- Il est également possible de synthétiser une approximation **d'onde plane** dans le plan horizontal, en formant des fronts d'onde de rayon de courbure infini, ce qui correspond à une distance holophonique élevée de la source virtuelle. C'est en fait le cas limite des sources ponctuelles. Les ondes planes sont perçues comme provenant d'une direction fixe et permettent la perception d'une scène « à l'infini » quelque soit la position de l'auditeur dans la zone d'écoute.
- Enfin, certains systèmes WFS permettent la synthèse de sources virtuelles situées à l'intérieur du sous-espace de reproduction et désignées par le terme **sources focalisées**. Le processeur que nous utiliserons par la suite ne permet pas la synthèse de ce type de source qui demande une puissance de calcul importante, ainsi qu'un nombre de haut-parleurs tendant vers ce que préconise la théorie. De plus, ce type d'effet est difficilement envisageable dans le cadre d'une diffusion audiovisuelle pour des questions évidentes de rupture entre les espaces de projection respectifs de l'image et du son. Comment assurer une « aimantation audiovisuelle » si le son est géométriquement placé « devant » l'image ?

1. Le son spatialisé au cinéma

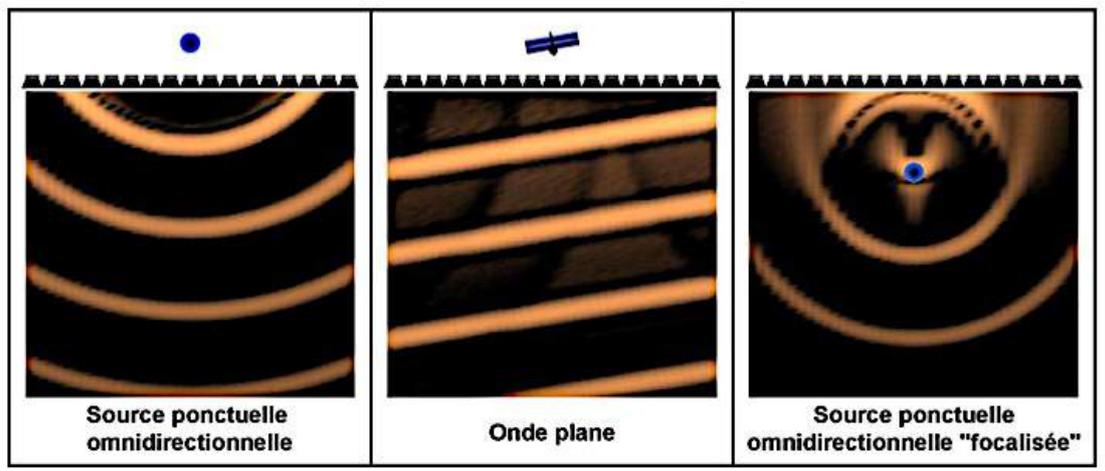


Fig. 1.27 : Types de sources synthétisables par WFS. Source [Corteel, 2004].

1. Le son spatialisé au cinéma

1.4.4. Hypothèses : la WFS en contexte de création sonore cinématographique

Après une première lecture historique nous permettant d'appréhender un peu plus les enjeux de la spatialisation sonore au cinéma, suivi d'une présentation des techniques innovantes dans le domaine du « son 3D », nous terminons cette partie en présentant la technologie qui sera au centre de notre étude : *la Wave Field Synthesis*. L'utilisation de ce type de système, confronté aux matériaux actuels de la postproduction sonore cinématographique, forme donc notre domaine d'étude. Nous souhaitons préciser au lecteur que les parties suivantes de ce mémoire ne s'attarderont pas à essayer de déterminer quel type d'éléments sonores ou quelle organisation de ces éléments conviendra le mieux dans la création d'une bande son cinématographique. Notre sujet se bornera bien à l'étude de l'apport de la WFS du point de vue du mixeur, c'est à dire en prenant en compte les matériaux déjà existants. Cela dit, quelques brèves remarques relatives au montage son pourront être distillées de-ci de-là. Elles doivent être considérées comme autant de portes ouvertes à d'autres études plus poussées.

Imaginer un système comme la WFS dans le contexte de la création sonore cinématographique nécessite d'appréhender dans quelles mesures l'utilisation de cette technologie pourra s'adapter à ce contexte bien particulier. « L'adaptation » ainsi formulée doit être considérée à la fois d'un point de vue technologique (utilisation des outils, installation technique), mais aussi d'un point de vue perceptif. En effet, le premier objectif est de savoir si les nombreuses contraintes technologiques liées à la WFS pourront permettre son intégration en auditorium de mixage et en salle de cinéma. Le second objectif serait de voir si les effets nouveaux permis par la synthèse de front d'onde s'accordent avec le dispositif

1. Le son spatialisé au cinéma

particulier du « spectacle audiovisuel », grandement dirigé par la présence d'un écran.

Les approximations précédemment évoquées dans notre étude (cf. 1.4.2) permettent, au prix d'un certain nombre de problèmes, d'intégrer le concept de la WFS dans des conditions réelles. Comme l'indiquent Laurent Millot et Gérard Pelé en citant Yves Rocard, « même si, dans ces conditions, le champ sonore synthétique s'écarte du champ réel, ce champ synthétique peut pour autant correspondre à une approximation acceptable du champ réel » [Millot & Pelé, 2006]. De plus, la WFS permet de s'écarter de plusieurs biais intrinsèques aux techniques discrètes :

- **Placement et nombre des haut-parleurs débridés** : Contrairement aux standards discrets comme le 5.1, la WFS ne nécessite pas une organisation équivalente des haut-parleurs entre la production et la restitution du contenu. Outre le fait que le nombre de haut-parleurs ait pour conséquence un plan grande « définition » de restitution, l'étape de mixage n'a plus nécessité d'être effectuée dans une salle rigoureusement comparable, dans le placement des haut-parleurs, à la salle de cinéma. Le principe de séparation entre la création du contenu et sa restitution forme les bases du mixage *orienté objet*, dont les fondements ont été amenés par le projet CARROUSO au début des années 2000 [de Brix et al., 2001]. La plupart des formats commerciaux actuels, excepté l'Atmos, sont des formats « orientés canaux », en ce sens que le format de codage (5.1, 7.1, ...) définit la configuration de restitution. En WFS, comme en Atmos, le signal audio de chaque source est envoyé au processeur de restitution, qui se charge ensuite de calculer le signal à envoyer à chaque haut-parleur. Précisons à nouveau que contrairement à la synthèse de front d'onde, le

1. Le son spatialisé au cinéma

nouveau standard de Dolby ne permet pas la reconstruction d'un champ sonore suivant les principes de l'holophonie (cf. 1.3.1).

Par ailleurs, contrairement à ce qu'a pu préconiser la littérature pendant plusieurs années, le nombre de haut-parleurs nécessaires à la reconstruction d'un champ sonore correcte tend à diminuer. On s'éloigne de l'écartement de 15 cm à 20 cm requis par la WFS traditionnelle, qui donnait lieu à des installations démesurées nécessitant des dizaines voire des centaines de haut-parleurs, comme celle du studio Todd-AO à Burbanks. Le processeur dont nous avons disposé, développé par la société *Sonic Emotion*, intègre les algorithmes propres à la WFS en exploitant un nombre réduit de haut-parleurs. Cet aspect de l'installation sonore nous semble être en adéquation avec une industrie cinématographique plutôt tournée vers l'économie de moyens. Etienne Corteel, aujourd'hui chef du département développement chez Sonic Emotion, préconise un espacement entre les haut-parleurs comparable à ce que propose Dolby avec *Atmos* aujourd'hui. De plus, cet espacement plus important permet l'utilisation de transducteurs qui correspondent, de par leurs dimensions, aux niveaux sonores requis dans une salle de cinéma.

- **Positionnement libéré de l'auditeur** : L'un des désavantages principaux de la stéréophonie concerne la zone d'écoute très limitée, comparable à la zone de travail du mixeur dans l'auditorium. Ceci a pour résultat une mauvaise stabilité de localisation des sources fantômes, et d'autre part, une balance avant/arrière sensiblement dépendante de la position d'écoute. L'effet de précedence d'un haut-parleur sur l'autre peut avoir des conséquences perceptives désagréables pour des positions d'écoute excentrées : l'image sonore frontale n'est plus valable. Pour des positions d'écoute reculées dans la salle de cinéma, la balance

1. Le son spatialisé au cinéma

avant/arrière peut facilement attirer l'attention sur les canaux surround, lorsque ceux-ci prennent de l'importance dans le mixage.

En WFS, nous avons vu que la localisation correcte d'une source sonore était indépendante de la position du spectateur dans la zone d'écoute. Cette caractéristique est plutôt intéressante au vu du dispositif cinématographique qui prévoit d'accueillir un nombre élevé de spectateurs, avec pour certains d'entre eux, une position tout à fait excentrée dans la salle de cinéma. Raphaël Mouterde, dans [Mouterde, 2004], mettait déjà en valeur cet apport de la WFS. Cette « bonne » localisation est valable à la fois pour des sources ponctuelles issues d'une prise de son monophonique, mais également pour les sources issues d'une prise de son plus complexe, allant de la stéréophonie, aux dispositifs ambisoniques, en passant par les arbres multicanaux. Ici encore, les idées quelque peu dépassées visant à limiter la WFS à une captation rigoureuse du champ sonore, nécessitant un très grand nombre de microphones, doivent être et sont déjà réévaluées au profit d'un regard plus réaliste et tourné vers la pratique.

Dans [Theile et al., 2003], les auteurs prévoient déjà l'intégration du contenu stéréophonique dans la restitution WFS. L'idée est de capter une source à l'aide d'un dispositif stéréophonique « classique », ou de tout autre dispositif multicanal, afin d'extraire une image de la scène sonore contenant déjà des informations spatiales (réflexions précoces, champ diffus, ...). Ces canaux enregistrés sont alors considérés comme des sources à part entière par le système WFS, comme des haut-parleurs virtuels situés dans l'espace à synthétiser. Cette technique se nomme le *Virtual Panning Spot*, et est explicitée dans l'encadré iii, p. 109. Ainsi, la localisation correcte obtenue pour des sources ponctuelles peut être étendue aux sources multicanales, considérées alors comme un ensemble de sources

1. Le son spatialisé au cinéma

ponctuelles. Cette caractéristique, nous permettant a priori d'apprécier une image sonore large quelque soit notre position dans la salle de cinéma, nous paraît également très intéressante.

- **Nouvelle définition des canaux arrières** : Dans les salles de cinéma actuelles, là où 2 haut-parleurs arrières ne suffisent pas pour délivrer un niveau sonore satisfaisant, il devient nécessaire d'utiliser d'autres haut-parleurs de surround, placés sur les côtés de la salle. En conséquence, pour de nombreux spectateurs, certains haut-parleurs ne sont plus situés derrière, mais devant eux, perturbant ainsi l'image sonore frontale. En WFS, le principe de sources virtuelles explicité précédemment permet aussi le placement d'objets sonores à l'arrière de la zone d'écoute, valable pour tous les spectateurs de la salle. Peu importe son positionnement, le spectateur préservera la source sonore placée derrière lui, et non pas à côté de lui, ou devant lui, tel qu'il est possible de le percevoir dans les configurations discrètes actuelles (5.1, 7.1, ...).

Par ailleurs, la WFS nécessitant des caractéristiques identiques, ou du moins comparables, de chacun des canaux de restitution, les haut-parleurs situés à l'arrière doivent nécessairement être dimensionnés de telle sorte à développer de forts niveaux sonores, et doivent être d'une étendue spectrale équivalente aux haut-parleurs de façade. Ces caractéristiques permettraient de mieux appréhender les éléments sonores de surround. En effet, dans le but d'éviter le réflexe d'attention, principalement causé par la diffusion d'éléments « précis » ou aigus à l'arrière, les mixeurs font usuellement appel à du contenu chargé en basses fréquences pour alimenter les canaux de surround. Cette technique se base perceptivement sur le *cône de confusion*, traduisant une ambiguïté de localisation avant/arrière pour les fréquences situées en dessous de 1500Hz (cf.

1. Le son spatialisé au cinéma

[Corteel, 2004], p. 215). Ainsi, on comprend bien tout l'intérêt pour le mixeur de disposer de basses fréquences pour les canaux arrières, lui offrant des possibilités efficaces d'enveloppement. Bien que ces caractéristiques ne soient pas propres à la WFS, puisqu'étant partagées avec l'*Atmos* de Dolby, nous avons tout même trouvé intéressant d'en faire part à notre lecteur.

Outre l'évaluation des multiples apports que pourrait fournir la synthèse de front d'onde dans un contexte de création cinématographique, nous trouvons intéressant d'amener le lecteur à réfléchir sur les potentielles limites du système, qu'elles soient, encore une fois, technologiques ou perceptives :

- **Présence de l'écran** : Dans [Millot & Pelé, 2006], Laurent Millot et Gérard Pelé mettent régulièrement l'accent sur les potentiels conflits existants entre la projection visuelle et la diffusion sonore dans le « spectacle audiovisuel ». En admettant que l'écran reste effectivement le procédé de projection de l'image en salle de cinéma, l'utilisation de la WFS pourrait induire un certain nombre de problèmes. Le premier, serait lié à ce que les auteurs appellent la « dissymétrie entre son et image », p. 290. En effet, le plan vertical dans lequel est contenu l'image est orthogonal au plan horizontal du domaine sonore. Problème déjà connu en diffusion 5.1, les mixeurs ont su s'approprier les canaux arrières en les considérant principalement comme une extension et un étalement de la façade. La « gadgétisation » des surrounds, qui pouvait probablement être tentante aux prémices du Dolby Digital (et qui peut encore l'être pour une catégorie réduite de films dits « à effets »), semble avoir été raisonnée depuis plusieurs années au profit d'une utilisation plus « fine » de ces canaux. Ainsi, la WFS ne doit pas être vue, selon nous, comme des possibilités supplémentaires de localisation arrière, mais plutôt comme une

1. Le son spatialisé au cinéma

amélioration de l'outil déjà bien connu des mixeurs, permettant enveloppement et élargissement.

Par ailleurs, l'effet de parallaxe précédemment évoqué (cf. partie 1.4.3), induisant un phénomène de perspective sonore, peut entraîner, a priori, des distorsions angulaires suivant la position du spectateur dans la zone d'écoute. Dans [Melchior et al., 2003], les auteurs décrivent largement ce phénomène. Géométriquement, on comprend assez rapidement que l'erreur de position d'une source sonore par rapport au stimulus visuel dépend de la position de la source, de sa profondeur, et de la position du spectateur (cf. figure ci-dessous). Au plus la source sera éloignée de l'auditeur -en azimut et en profondeur- au plus la distorsion angulaire sera importante. Etienne Corteel dans [Corteel, 2004] évoque « l'effet ventriloque »¹³, connu comme étant capable de fusionner des stimuli visuels et des stimuli auditifs placés à des positions différentes. Il reste cependant à évaluer comment « l'aimantation audiovisuelle » pourrait s'opérer pour des sources sonores virtuellement placées en profondeur : la nature de l'effet ventriloque est pour le moment réservée, en général, à la caractérisation du phénomène pour des personnages de part et d'autre de l'écran, et ayant leur voix diffusée dans le canal central. Dans tous les cas, le mixeur aura le choix d'éloigner ou non les sources sonores en profondeur. Un placement des sources au plus proche de l'écran l'affranchira complètement de tout problème de distorsion angulaire lié à la perspective. Par ailleurs, comme l'indiquent les auteurs dans [Corteel et al., 2014], la distance des sources peut être considérée comme un paramètre d'homogénéité pour l'auditoire. En effet, pour une distance holophonique importante, la source est

¹³ Dans [Corteel et al., 2014], les auteurs avancent qu'une marge d'erreur de 5 à 15 degrés est permise par l'effet ventriloque. L'auteur cite pour ce sujet [Perrott, 1993].

1. Le son spatialisé au cinéma

virtuellement plus large, et cela peut permettre un certain *lissage* des erreurs de position introduites par la perspective.

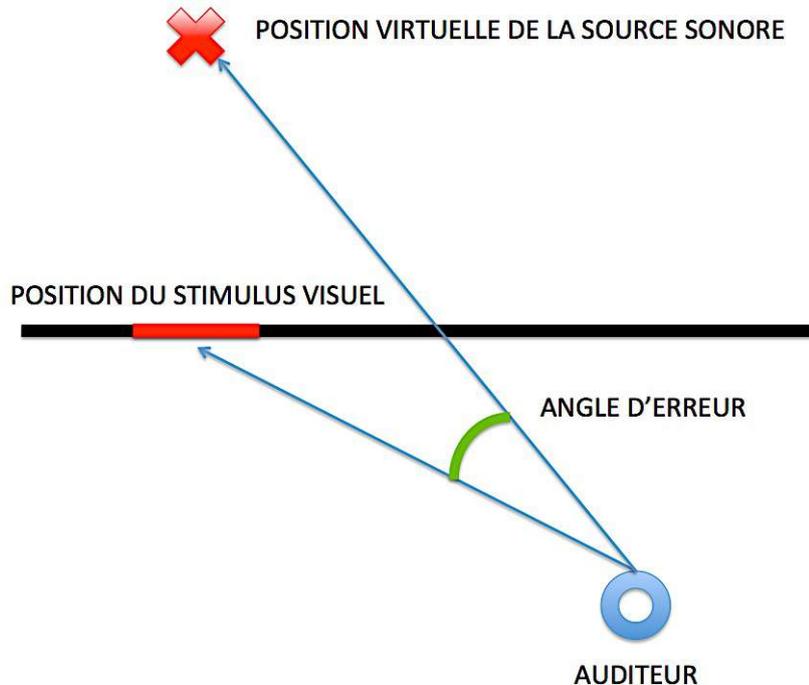


Fig. 1.28 : Illustration de la distorsion angulaire induite par la distance holophonique.

- **Transportabilité et compatibilité** : Comme déjà évoqué précédemment, l'hypothèse est faite qu'en WFS, la création du contenu soit indépendante du dispositif de restitution, à la « précision » du système près (nombre de haut-parleurs). C'est le principe du mixage *orienté objet* décrit précédemment. Cependant, cette idée non-validée d'une certaine *transportabilité*¹⁴ du mixage ne peut nous satisfaire. En 5.1, la transportabilité de la bande son est bien souvent jugée sur des caractères spectraux, variables d'une installation à une autre. Quelques critères

¹⁴ La transportabilité d'un mixage peut être défini par sa capacité à être perçu de manière identique (ou du moins équivalente) d'une salle à l'autre.

1. Le son spatialisé au cinéma

peuvent également concerner la largeur stéréophonique, ou encore le niveau des canaux surrounds, mais le nombre réduit de canaux et de haut-parleurs semble permettre une transportabilité satisfaisante, régie par un certain nombre de normes de calibrations¹⁵. Dans l'hypothèse où une salle de cinéma WFS serait soumise à un processus de calibration, comment évaluer objectivement, ou même perceptivement, la transportabilité du mixage d'une installation à une autre ? Cette question semble s'éloigner de notre sujet, qui dans un environnement réduit d'étude (une seule installation, dans un seul auditorium de mixage), ne nous portera pas directement préjudice. Cependant, cette question pourrait tout à fait faire l'objet d'une étude à part entière.

Par ailleurs, la « compatibilité descendante » est un concept aujourd'hui devenu récurrent dans l'industrie cinématographique. Les mixages 5.1 ou 7.1 sont très souvent soumis à des *downmix* leurs permettant une diffusion sur des supports contenant un nombre plus réduit de canaux : télévision, internet, appareils mobiles. Le processus de « downmix stéréo », permettant le passage des 6 canaux du 5.1 aux 2 canaux stéréophoniques, est un exemple bien connu des mixeurs, ces derniers automatisant souvent cette opération via l'utilisation d'un plugin dédié. Comment imaginer une compatibilité descendante entre un format de type objet telle que la WFS, et une diffusion discrète stéréophonique, dispositif restant le plus courant dans les installations domestiques ? Les solutions proposées par *Atmos* aujourd'hui semblent se contenter d'un *downmix* des *Beds* pour délivrer un contenu discret, rapidement compatible avec les installations les plus modestes. Bien qu'étant en dehors de notre étude, notons que ces problématiques de transport et de compatibilité font en ce moment même

¹⁵ En France, ces calibrations sont pour la plupart menées par la société Dolby.

1. Le son spatialisé au cinéma

l'objet d'un programme de recherche nommé *Edison 3D*, visant à développer un consensus sur les outils, les formats, et les modes de diffusion de la « 3D sonore ». Les grandes lignes de ce projet sont consultables sur le site de l'Agence Nationale de Recherche (cf. [Edison] dans la bibliographie).

En résumé, la WFS nous semble être un outil puissant mis à disposition du mixeur dans la conception de scènes sonores. L'une des propriétés les plus importantes est cette capacité du système à fournir une localisation stable des sources virtuelles, quelque soit la position du spectateur dans la zone d'écoute, i.e. dans la salle de cinéma. Les problèmes typiques et les contraintes liées à l'image stéréophonique sont a priori détournés par l'usage de la WFS. Cependant, un certain nombre de limitations transparaissent déjà de par la nature même du dispositif holophonique, mettant en jeu des conflits entre le champ visuel et le champ sonore. Dans la suite de notre étude, nous aurons pour objectif d'évaluer si les apports de la WFS sont suffisants pour lui imaginer une intégration dans un processus de création cinématographique. En effet, les réticences qu'attire *Atmos* aujourd'hui, liées à la « lourdeur » de l'installation, peuvent très bien être associées au système WFS, composé d'un nombre équivalent de haut-parleurs. Il faudra donc que la WFS soit vraiment convaincante perceptivement pour prétendre à une intégration standardisée dans les salles de cinéma.

2. La spatialisation en WFS : étude préliminaire

La deuxième partie de ce mémoire vise à développer notre première approche du système, que ce soit d'un point de vue pratique, par la prise en main des outils de spatialisation, ou d'un point de vue perceptif, par la mise en place de petits tests comparatifs visant à nous faire prendre conscience du potentiel et des limites du dispositif. Nous tenons à préciser au lecteur que les observations développées dans cette partie ne pourront en aucun cas prétendre à la même valeur que des résultats obtenus dans le cadre rigoureux d'un test perceptif soumis à un panel de sujets. En effet, nous nous sommes dans un premier temps basés sur nos sensations, elles-mêmes issues d'une culture du son et du mixage qui, nous étant propres, ne pourraient induire une généralisation de ces constats en « lois » perceptives, s'il en est.

Ainsi, nous nous attacherons d'abord à une brève description des outils mis à notre disposition pour cette première approche de la WFS. Le lecteur trouvera une description plus détaillée de ces outils en 3.1. Ces premiers paragraphes nous permettront également de poser le comportement basique du dispositif, par le biais de constats effectués lors de nos premières écoutes. Nous proposerons

ensuite une série de tests plus précis, ayant un objectif double : celui, dans un premier temps, de nous familiariser avec l'outil de spatialisation ; puis celui, dans un second temps, de commencer à appréhender l'usage de la WFS pour le traitement spatial des matériaux courants dans la création sonore cinématographique : la voix, les ambiances, ou encore la musique.

2.1. Première approche du système à synthèse de front d'onde

2.1.1. Un outil de mise en « espace WFS » : le Performer

Nos premiers pas pratiques dans ce travail de mémoire se sont déroulés dans les locaux de Sonic Emotion. Grâce à l'accueil d'Etienne Corteel et de ses collaborateurs, nous avons eu accès au matériel nécessaire à l'utilisation de la WFS pour du contenu sonore lié à l'image. Le laboratoire¹, situé en sous-sol au 42 bis rue de Lourmel, regroupait le matériel suivant :

- Un processeur *Wave 1*, nerf central du procédé de synthèse de front d'onde : cet ensemble rackable, composé de cartes DSP² du constructeur *RME*, et d'un système d'exploitation à noyau *Linux*, permet de mettre en place toutes les fonctions relatives à la synthèse de front d'onde. Ce système d'exploitation a la particularité d'être stocké sur une clé USB, regroupant ainsi tout le traitement audio, i.e. la synthèse de front d'onde, à l'intérieur du boîtier.

¹ Notons que c'est sur cette installation que nous avons réalisé tous les tests évoqués dans cette partie 2.

² DSP (*Digital Signal Processing*) : unité de calcul spécialisée dans le traitement du signal.



Fig. 2.1 : Le processeur Wave 1.

- Un système électro-acoustique, composé d'une trentaine de haut-parleurs de petite taille, chacun relié à un canal d'amplification : bien que de modèle réduit, ces enceintes semblaient nous offrir une précision et un niveau suffisants au vue des dimensions du laboratoire (environ 4 mètres de longueur, pour 3 mètres de largeur). Leurs lacunes dans le registre fréquentiel bas sont compensées par l'utilisation d'un *subwoofer*, configuré en *bass management* dans le Wave 1³.
- Un software de type DAW⁴ : *Nuendo*, du développeur *Steinberg*, allait nous permettre de lire du contenu sonore, avant de le router dans le processeur Wave 1 via liaison MADI.
- Un software, le *Performer*, permettant le contrôle du Wave 1 : ce software est considéré comme le GUI (Graphical User Interface) de la chaîne. Véritable plateforme de commande, elle permet un placement visuel et aisé des sources sonores, ce même placement étant retranscrit en fonctions de calcul par le Wave 1.

³ Explications sur le comportement du bass management en 3.1.2.

⁴ DAW (*Digital Audio Workstation*) est un terme qui regroupe les séquenceurs audio comme Pro Tools, Pyramix, Nuendo, etc.

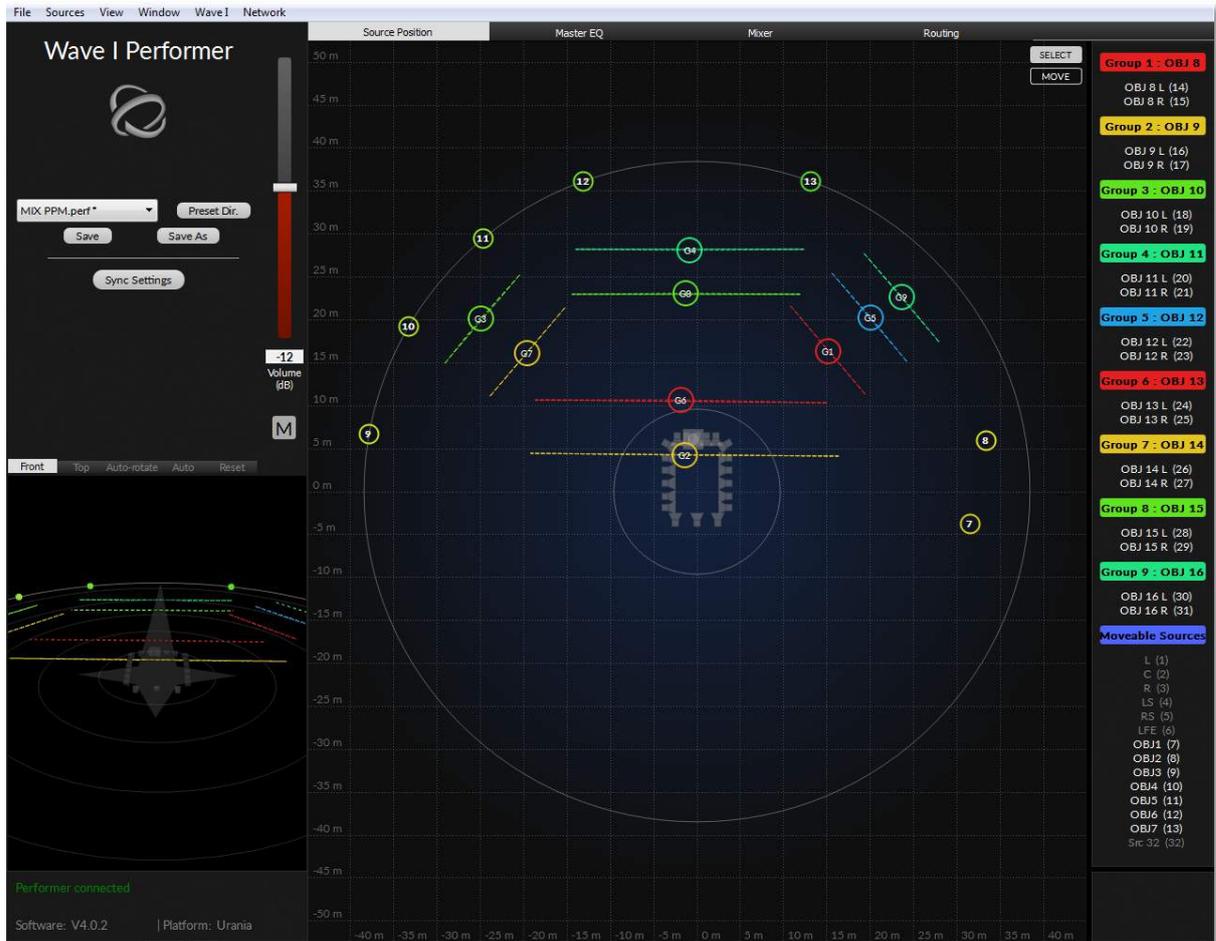


Fig. 2.2 : Fenêtre principale du Performer, interface de positionnement et de visualisation des objets sonores.

Dès lors qu'une source sonore sera positionnée dans la zone interne au plus grand cercle du Performer, on pourra dire que la source est placée dans « l'espace WFS », par abus de langage. Nous appellerons « objets sonores » ou « objets » les sources ainsi synthétisées. Cette terminologie permettra par la suite de bien différencier les sources restituées par synthèse de front d'onde, et les sources directement routées dans un ou plusieurs canaux : pour ces dernières, on parlera alors de sources « discrètes ». Précisons au lecteur qu'une rigueur terminologique tentera d'être préservée tout au long de ce document, permettant par exemple une

distinction aisée entre une source stéréophonique discrète, et une source stéréophonique restituée via le dispositif WFS (cf. partie 2.2.2, « Cas des ambiances frontales »).

Le Performer est donc une interface visuelle de positionnement des sources sonores. Il offre une vision 2D et 3D de la scène. Afin de rendre la scène sonore plus claire, les sources peuvent être organisées en groupe, cachées, renommées, et sont présentées dans une liste située sur le côté droit de l'écran.

Par ailleurs, chaque source peut être routée directement vers une sortie physique (un ou plusieurs haut-parleurs, ou un sous-système⁵ entier via une module de routing propre au Performer (onglet "Routing"). C'est grâce à ce mode « discret » du processeur que nous aurons la possibilité, dans une même session de travail du Performer, de comparer la même source sonore via restitution discrète et via restitution WFS.

⁵ Pour un détail du concept de sous-système, voir la partie sur la présentation du Designer, partie 3.1.

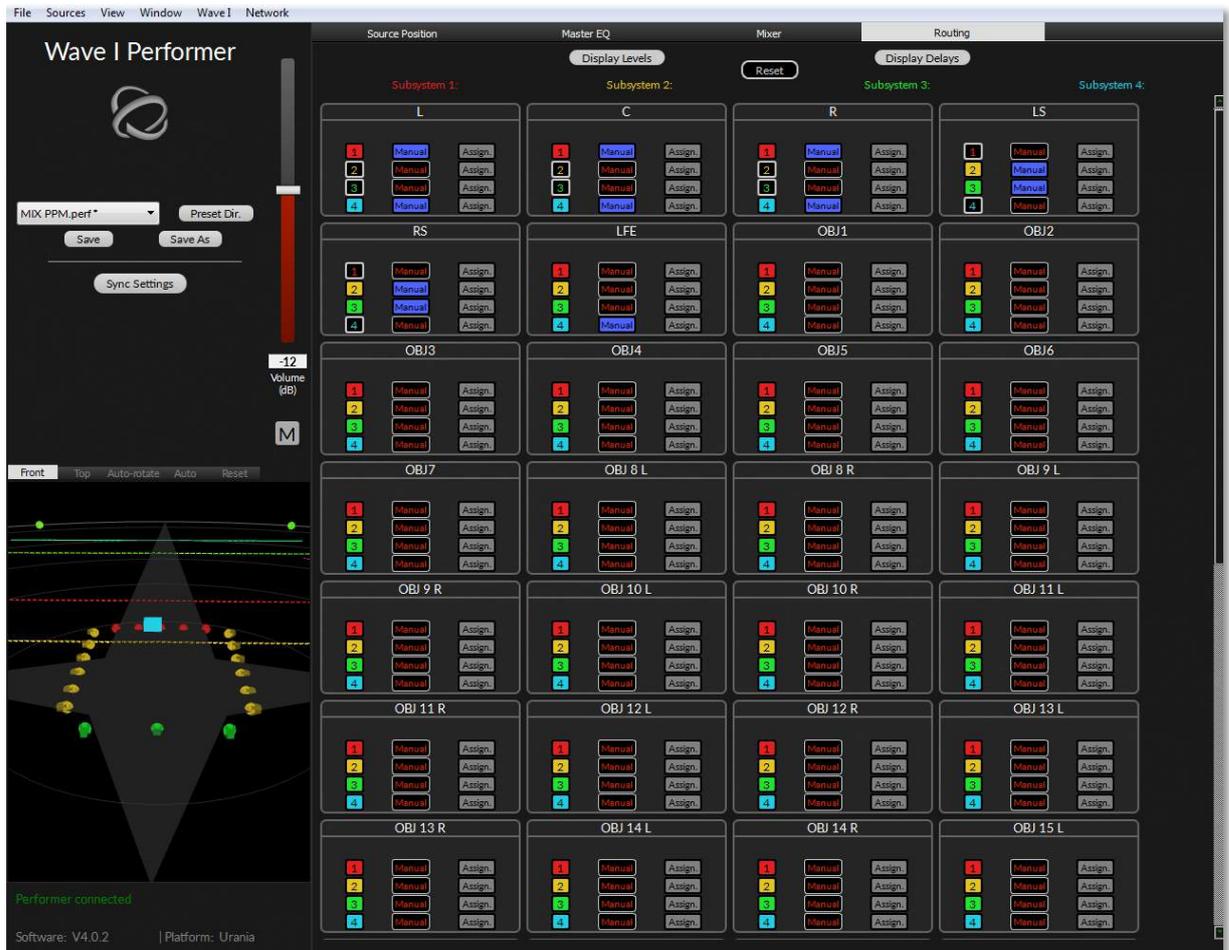


Fig. 2.3 : Onglet « Routing » du Performer, permettant de décider du sort des sources sonores.

Enfin, le Performer inclut une table de mixage permettant d'ajuster le niveau de chaque objet source (onglet "Mixer"). Cette fonction ne nous sera utile que pour contrôler l'arrivée du signal dans chacun des objets.

Chaque configuration peut alors être sauvegardée comme un fichier de preset au format .xml, et peut être facilement rappelée (voir fenêtre de sauvegarde, en haut à gauche de la page, cf. figure ci-dessus).

2.1.2. Comportement basique du Wave 1

Ayant pour objectif final de nous intéresser au mixage, c'est à dire au mélange de plusieurs sources, sur un système à synthèse de front d'onde, il nous a semblé nécessaire dans un premier temps de nous restreindre à l'étude d'un contenu plus réduit, plus minimaliste. C'est ainsi que nous avons abordé les matériaux de base d'une bande son liée à une image, constats relatés plus tard dans ce document (cf. partie 2.2). Avant même de nous rapprocher de ces matériaux, il nous a semblé pertinent d'évaluer le comportement du système pour une source sonore simple. Le bruit rose aurait pu faire l'objet de cette étude : de par son contenu fréquentiel large, sa constance énergétique dans le temps, il aurait formé un matériau intéressant pour permettre l'évaluation spatiale et fréquentielle du système. Or, puisque nous savions dès le départ que notre démarche ne mettait en jeu aucune mesure acoustique, nous avons fait le choix de nous rabattre sur un contenu moins « calibré », et potentiellement moins désagréable à l'écoute. Ainsi, une ambiance monophonique de type « fond d'air urbain » a été sélectionnée, nous permettant de juger du comportement du système sur une zone fréquentielle assez large, mais proposant tout de même un contenu moins agressif et fatigant que le bruit rose.

Nous avons débuté ces premiers tests par l'évaluation d'une seule rampe de haut-parleurs : la rampe frontale. Celle-ci nous semble effectivement déterminante, puisque le contenu proposé au spectateur est de forme scénique, s'agissant dans notre cas d'un spectacle audiovisuel (cf. Introduction). Voici ce qui a pu être extrait de l'usage de cette rampe dans la synthèse de l'ambiance « fond d'air urbain » :

La zone de fenêtrage frontal

Les sensations perçues pour différentes positions de l'objet nous permettent de distinguer trois zones de traitements appliqués par le Wave 1 à la rampe frontale :

- Une zone frontale d'une étendue angulaire approximative de 60° , où les traitements semblent se faire de manière progressive (en "fondue") si un déplacement latéral de source est envisagé. La source devient effectivement localisable là où elle est virtuellement placée sur le Performer.
- Une zone d'inactivité « Ouest », où aucun traitement, sinon celui implémenté de manière constante, semble être associé à une position de source. En effet, que l'objet soit placé en haut de la zone « Ouest », ou en bas de celle-ci, l'effet perceptif obtenu, et par conséquent le traitement du Wave 1 sur la rampe frontale, nous semblent identiques. Cette zone d'inactivité entre en jeu de manière progressive dès lors que la source commence à quitter la zone frontale précédemment définie. Plus simplement, cette zone d'inactivité a pour résultat une position de source fixe et extrême dans le haut-parleur le plus à gauche de la rampe.
- Une zone d'inactivité « Est », est un simple symétrique par rapport à l'axe central de la zone « Ouest » décrite précédemment, aux différences spectrales près.

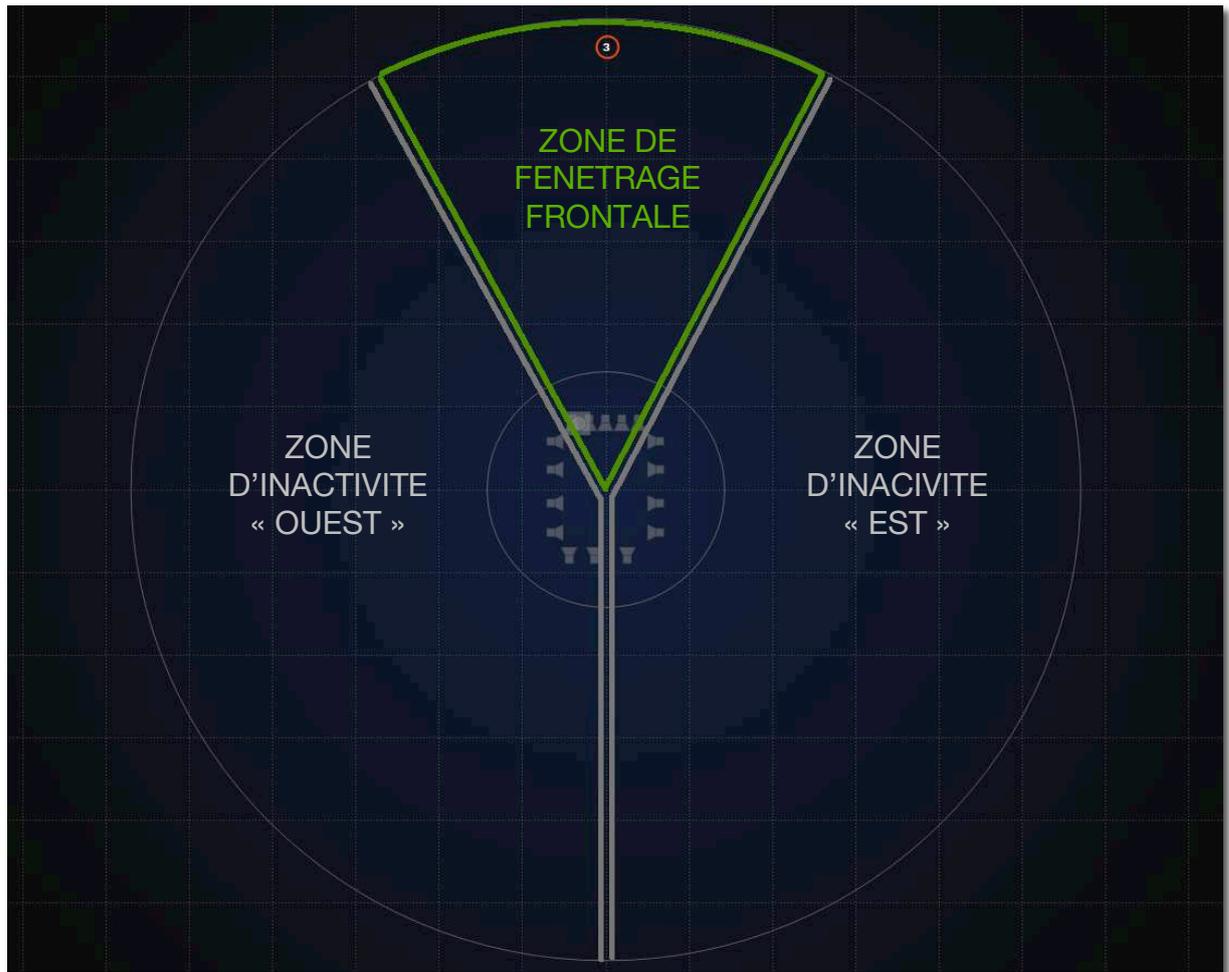


Fig. 2.4 : Illustration de la zone de fenêtrage frontal. C'est dans cette zone que les sources sont correctement synthétisées par la rampe frontale de haut-parleurs.

Nous pouvons donc relever que la zone correspondant aux traitements dynamiques effectués sur la rampe frontale ne concerne qu'une zone polaire restreinte, délimitée par un angle approximatif de 60° , passant par le repère central⁶, et par les haut-parleurs situés aux extrémités de la rampe. Nous nommerons cette zone « la zone de fenêtrage frontal » ou plus simplement la « fenêtre frontale », ou encore la « fenêtre WFS » dans le cas plus général ou la rampe de haut-parleurs peut se situer sur un côté, ou à l'arrière de l'auditeur. C'est

⁶ Ce point n'est en aucun cas à considérer comme un sweet spot. Il n'est qu'un repère central à notre échelle géographique.

effectivement à l'intérieur de cette fenêtre que les objets pourront être correctement synthétisés par le processeur, ce dernier calculant alors les signaux à envoyer à la rampe. La WFS peut alors être vue comme une fenêtre (ou un ensemble de fenêtres) ouverte sur une scène sonore.

La contribution des autres rampes

Nous pouvons par suite conjecturer que la zone d'action de la rampe frontale est totale : quelle que soit la position de la source dans la salle, la ligne frontale continuerait à contribuer à la synthèse de cette source. Cela dit, il reste à vérifier que cette rampe reste active si elle en vient à être associée à d'autres rampes, latérales par exemple. Après un second test réalisé, pour une configuration « en ceinture » du système, et pour une position de source à l'arrière de la salle, nous constatons que la rampe frontale n'émet aucun son. Ainsi, la zone d'action de la rampe frontale n'est pas totale dès lors que celle-ci est associée à d'autres rampes, latérales et arrière dans notre cas.

Par ailleurs, si la structure du système ne se limite plus à une façade de haut-parleurs mais à une ceinture (ou 4 rampes) autour de l'auditeur, le positionnement d'une source en dehors de la fenêtre frontale prend tout son sens. Un effet perceptif d'éloignement par élargissement est alors créé. A noter que cet effet d'élargissement est assez rapidement obtenu. A priori, dès que la source vient à sortir de la zone de fenêtrage frontal, les haut-parleurs latéraux commencent à contribuer à la synthèse de cette source, et donc par conséquent à la rendre plus large, en plus de la rendre localisable sur un côté. Plus généralement, dès lors qu'une source sera géométriquement en regard d'une rampe de haut-parleurs, cette rampe participera à la synthèse de la source.

La distance de la source : modification de la structure de l'onde

Nous constatons tout d'abord qu'une source repoussée à "l'infini" (proche du second cercle du Performer) n'entraîne pas une perception lointaine de celle-ci. On se rapproche d'une diffusion en onde plane, comme le ferait une source lointaine. Par contre, l'effet de salle, le filtrage associé, ainsi que l'atténuation en niveau, ne sont pas appliqués à la source par défaut. Ce sera au mixeur de contrôler ces paramètres pour rendre crédible l'effet d'éloignement. Dans le cas où la source possède un effet de salle provenant de la prise, cet effet, mêlé à son tour au comportement "lointain" de l'onde, devient alors cohérent avec ce que nous pourrions obtenir en écoute naturelle. Les artifices standards (utilisation de réverbération stéréo, typiquement) pour élargir la source, ne sont plus nécessaires pour créer cet effet. La forme du front d'onde devient un nouvel artifice de travail très puissant.

Dans le cas inverse, celui du rapprochement d'une source, un nouveau traitement est appliqué par le processeur, difficilement exprimable, assimilable à un resserrement. Notons qu'avant ce rapprochement, la source était perçue comme plus large lorsqu'elle était placée proche du second cercle. Un fondu assez bref est alors perceptible lorsque la source transite d'une position éloignée, à une position proche du premier cercle du Performer. Notons bien qu'un rapprochement de la source au plus proche des enceintes ne permettra en aucun cas de simuler un effet de proximité, impliquant modifications spectrales et éventuellement dynamiques. Là encore, c'est sur la forme du front d'onde que le processeur agit, en rendant, dans le cas d'une source proche, un certain effet de ponctualité.

L'effet abandonné : la source focalisée

Une fois le premier cercle du Performer franchi, la source est alors théoriquement et géométriquement "dans la zone d'écoute"⁷. Peu de changements sont alors perceptibles sur la rampe frontale de haut-parleurs lorsque la source est promenée dans la "zone d'écoute". On notera une "aberration" : le passage de la source au travers de la normale verticale passant par le repère central se fait assez brutalement, et des changements spectraux sont audibles.

Après discussion avec Etienne Corteel sur le sujet, celui-ci nous a indiqué que la synthèse de sources focalisées n'était plus vraiment dans leurs lignes de développement, du fait de l'intégration très complexe de celle-ci en vue d'un résultat fiable. En effet, ce type de synthèse nécessiterait un nombre trop important de haut-parleurs, ainsi qu'une calibration très rigoureuse des uns par rapport aux autres. Après une série de tests effectués par Sonic Emotion, une distance de 30cm était déjà trop élevée pour tenter la synthèse de sources focalisées. Hors, pour toute autre effet situé cette fois-ci en dehors de la zone d'écoute, une intégration moins fine des haut-parleurs permettrait de recréer environ 90% des effets attendus. L'implémentation des algorithmes à l'origine de la création « d'holophones » à l'intérieur de la zone d'écoute a alors été abandonnée.

⁷ Cf. 1.4.2 : Un système WFS est censé être capable de synthétiser des sources à l'intérieure même de la zone d'écoute.

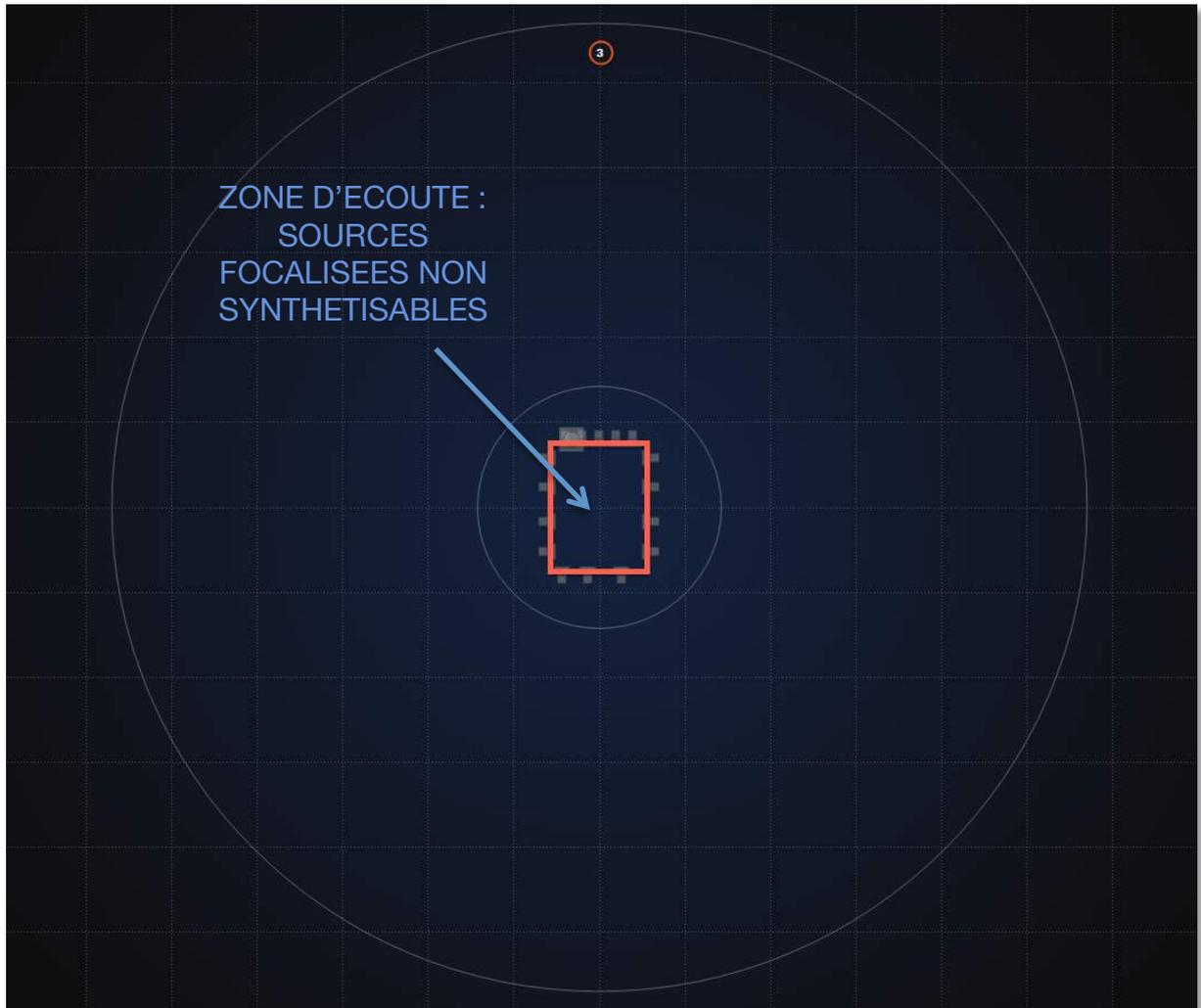


Fig. 2.5 : Avec le Wave 1, la synthèse de sources à l'intérieur de la zone d'écoute est rendue impossible.

La compensation via bass management

Pour un objet sonore donné, et quelque soit son positionnement dans le Performer, le système prévoit de déléguer la reproduction des basses fréquences à l'usage d'un *subwoofer* : c'est la technique du *bass management*. Dans notre cas de figure, cette compensation était nécessaire au vu de la taille réduite des enceintes placées sur les rampes WFS. Le *bass management* peut être activé dans le Wave 1 à condition qu'un *subwoofer* soit déclaré lors de la configuration de

l'installation⁸. Pour plus d'informations sur le *bass management*, le lecteur pourra se référer aux documents [Dolby, 2000] et [Richard, 2002].

La perception des fréquences graves étant supposé omnidirectionnelle dans de nombreux cas (ou du moins, moins directives que les fréquences aigues), on comprend assez vite que pour un objet placé derrière l'auditeur, cette source pourra disposer d'une étendue spectrale suffisante, le *bass management* faisant son effet. Cette caractéristique est d'autant plus intéressante si l'on considère que l'enveloppement créé par les éléments sonores placés à l'arrière (ambiances, réverbérations, ...) est plus efficace dès lors que la balance spectrale penche vers le registre grave (cf. 1.4.4, p. 67). Notons que cette particularité n'est pas propre à la WFS : l'Atmos permet également une gestion du registre grave via *bass management*, à la différence prêt que les *subwoofers* sont placés à l'arrière de la salle (cf. 1.3.1).

⁸ Pour plus de précisions, voir partie 3.1 sur la configuration du Designer.

2.2. La WFS face aux matériaux d'une bande son cinématographique

Informés du comportement basique du système, nous avons par suite trouvé nécessaire de nous pencher sur les éléments de base d'une bande son cinématographique, mis en regard de la synthèse de front d'onde. Cette nouvelle mise en pratique nous a permis successivement de manipuler l'outil de spatialisation (le Performer), puis d'en appréhender les effets perceptifs conséquents, enchaînement tout à fait courant voire essentiel dans le travail du mixeur. En plus de nous faire découvrir la palette d'effets inhérents à la WFS, ces situations de test nous ont amené à réellement comparer le rendu d'un dispositif discret (stéréophonique ou 5.1) avec celui d'un système à synthèse de front d'onde.

La partie suivante fera état de nos analyses globales par rapport aux tests effectués. S'il le souhaite, le lecteur pourra accéder aux détails de ces analyses en Annexe 1.

2.2.1. Traitement de la voix, mise en espace

Objectifs : Ecouter le rendu de la sensation d'espace d'une voix parlée, sur un système WFS et sur un système discret.

Sur un système discret, l'espace autour d'une voix est construit avec la modification du niveau, du filtrage, et un processus de réverbération (éventuellement multicanal). Nous nous interrogeons ici sur le placement d'une voix dans l'espace WFS. Pour cette première série de tests nous avons choisi d'utiliser une voix off, principalement pour deux raisons : premièrement, l'usage de la voix-off reste un procédé récurrent dans la narration cinématographique, et elle peut être étudiée en tant que telle ; deuxièmement, une voix-off possède des propriétés acoustiques intéressantes puisqu'étant enregistrée en proximité, i.e. sans effet de salle. Ainsi, si nous envisageons de replacer cette source dans un espace acoustique, nous n'aurons pas à composer avec un éventuel effet de salle issu de la prise de son.

Protocole 1 : voix-off seule (détails en Annexe 1.1)

Une voix-off (masculine, anglosaxone) est lue depuis un séquenceur (Nuendo) puis spatialisée par le Performer. Deux pistes contiennent chacune la même voix-off. La première piste est routée et assignée à un objet "WFS" du Performer : son placement sera donc synthétisé en front d'onde par le processeur. La seconde piste est routée vers un autre objet, configurée sans synthèse WFS, en "direct out": elle alimente alors directement et uniquement l'enceinte centrale, de manière similaire à l'utilisation standard d'un système 5.1. Cette configuration,

disposant de deux sources identiques lues dans Nuendo, nous permet lors du test de simuler un accès à l'un ou l'autre des systèmes (WFS ou 5.1), cela en cliquant simplement sur le "solo" de la piste "voix WFS" ou "voix 5.1" dans notre séquenceur.

Pour la diffusion en WFS, nous avons choisi 3 positions pour la voix-off, toutes centrales : la première au plus près des enceintes, la deuxième un peu plus reculée sur l'axe vertical (au niveau du premier cercle du Performer), la troisième quasiment tout au fond de la l'espace frontal du Performer (au plus proche du deuxième cercle).

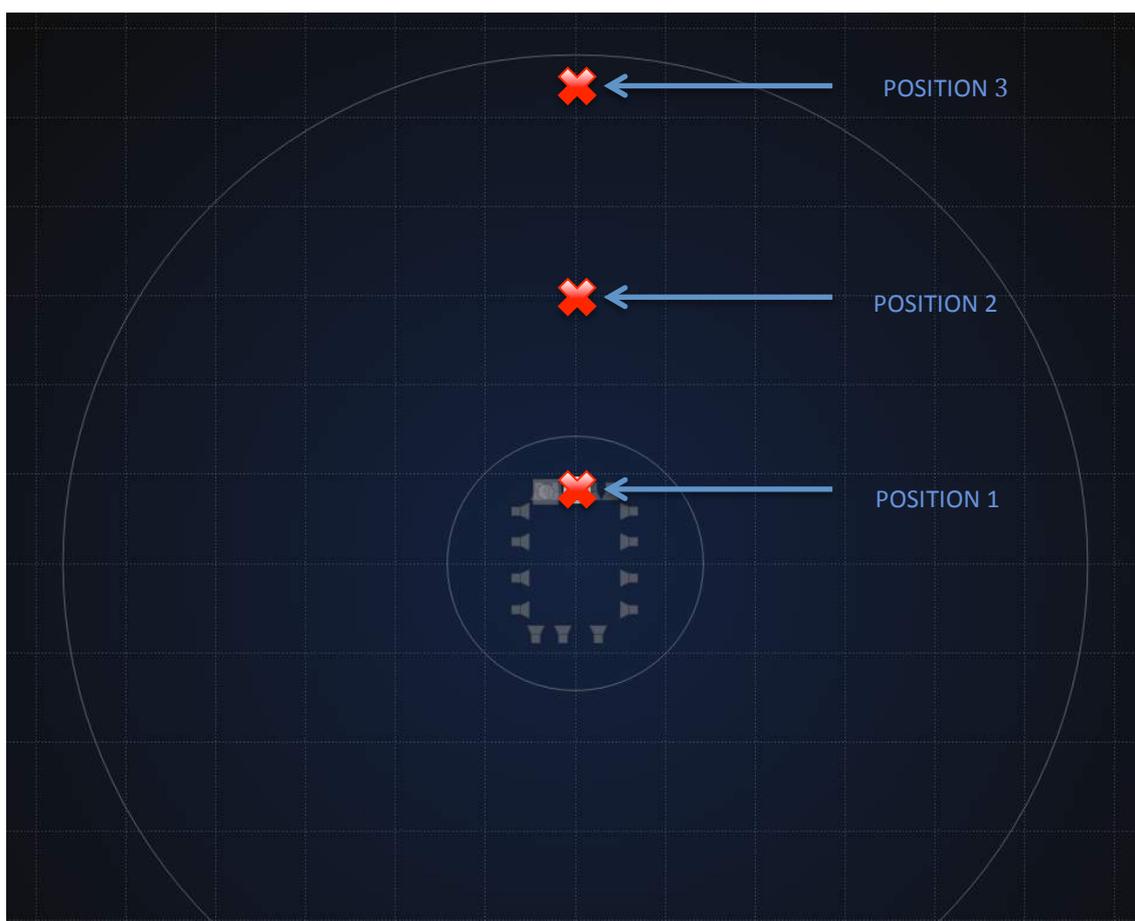


Fig. 2.6 : Positions testées des objets WFS.

Analyse protocole 1

Au vu de ces 3 positions successives de source, la voix-off semble subir dans tous les cas une modification spectrale plus ou moins marquée. Tandis que la position 1 et la position 3 affectent le haut du spectre de manière audible, la position 2 semble proposer une balance spectrale plus proche de celle de la voix-off directement diffusée dans l'enceinte centrale (l'encadré 1 suivant fait état des artefacts spatiaux propres à la WFS).

i. ENCADRE : DIFFERENCES SPECTRALES ET ALIASING SPATIAL EN WFS

L'application de la théorie de la synthèse de front d'onde demande en pratique d'effectuer un certain nombre d'approximations. L'une d'entre elles concerne la discrétisation du banc, censé être en théorie un segment continu constitué d'une infinité de haut-parleurs. En effet, quelque soit les enceintes utilisées, elles ne sont ni infiniment petites, ni infiniment proches. Cette discrétisation entraîne un phénomène d'aliasing spatial ou repliement spatial, comparable dans son principe au repliement spectral lors d'un échantillonnage numérique.

Cet aliasing entraîne une dégradation du timbre et de la localisation. Comme pour l'échantillonnage fréquentiel numérique dont la fréquence "Fe" va déterminer la fréquence maximale que l'on pourra reproduire, l'échantillonnage spatial du segment WFS (le nombre de haut-parleurs par unité de longueur) va déterminer la fréquence maximale que l'on pourra restituer avec précision. Ce phénomène est en partie dû à l'incapacité d'un système WFS à reproduire avec précision des fronts d'onde pour des fréquences élevées, et/ou ayant des angles d'incidence élevés. Il est donc fonction du nombre de haut-parleurs utilisés mais également de la forme du front d'onde synthétisé et donc de la position de la source virtuelle.

La figure Fig. 2.7 ci-après décrit un processus de mesure de cet aliasing via un banc de microphone. Configuration utilisée : Le banc de 60 haut-parleurs mesure 6m de long, la distance entre chaque haut-parleur est de 10cm, et on mesure sa réponse sur un banc de 60 microphones espacés de 15cm chacun.

La source virtuelle synthétisée est un monopole situé 3m derrière le banc de haut-parleurs, 3m sur la droite, et sa zone de visibilité est délimitée par des traits pointillés mauves sur l'ensemble des figures.

(b) : Magnitude (dB) de la réponse fréquentielle mesurée au niveau des microphones. La fréquence de repliement est indiquée par une courbe jaune.

(c) : Niveau (dB) de la réponse impulsionnelle mesurée au niveau des microphones, filtré passe-bas en dessous de 1.5kHz.

(d) : Niveau (dB) de la réponse impulsionnelle mesurée au niveau des microphones et filtrée passe-haut au dessus de 1.5kHz.

Cependant l'amélioration de "l'algorithmie WFS" a permis de palier à ce problème, les processeurs de Sonic Emotion permettant aujourd'hui d'utiliser un nombre restreint de haut-parleurs tout en évitant ou en minimisant l'aliasing spectral. De plus, si un nombre élevé de haut-parleurs permet d'augmenter la fréquence à partir de laquelle l'aliasing se produit, les artéfacts résultants d'une telle configuration peuvent devenir intolérables.

En effet, avec une rampe fournie de haut-parleurs, le moindre mouvement de tête peut amener l'auditeur à percevoir des modifications spectrales importantes. A l'inverse, l'utilisation d'un nombre restreint de haut-parleurs conduit à une fréquence d'aliasing plus basse, mais dont les effets sont beaucoup moins perceptibles.

L'étude *Wave Field Synthesis in the Real World* [Sporer & Klehs, 2004] analyse l'impact du nombre de haut-parleurs et de la distance de la source sur la qualité audio perçue. Les tests menés sur un panel d'experts utilisent deux stimuli sonores (voix parlée et bruit), écoutés chaque fois à deux distances différentes sur plusieurs configurations d'espacement de haut-parleurs.

Cette étude affirme que le nombre de haut-parleurs n'a pas d'influence perceptible sur la qualité audio de la source perçue. Cependant, une différence de qualité est perçue en fonction de la distance virtuelle de la source : la plus grande distance est perçue comme ayant la meilleure qualité audio comparée à la source placée plus près des haut-parleurs.

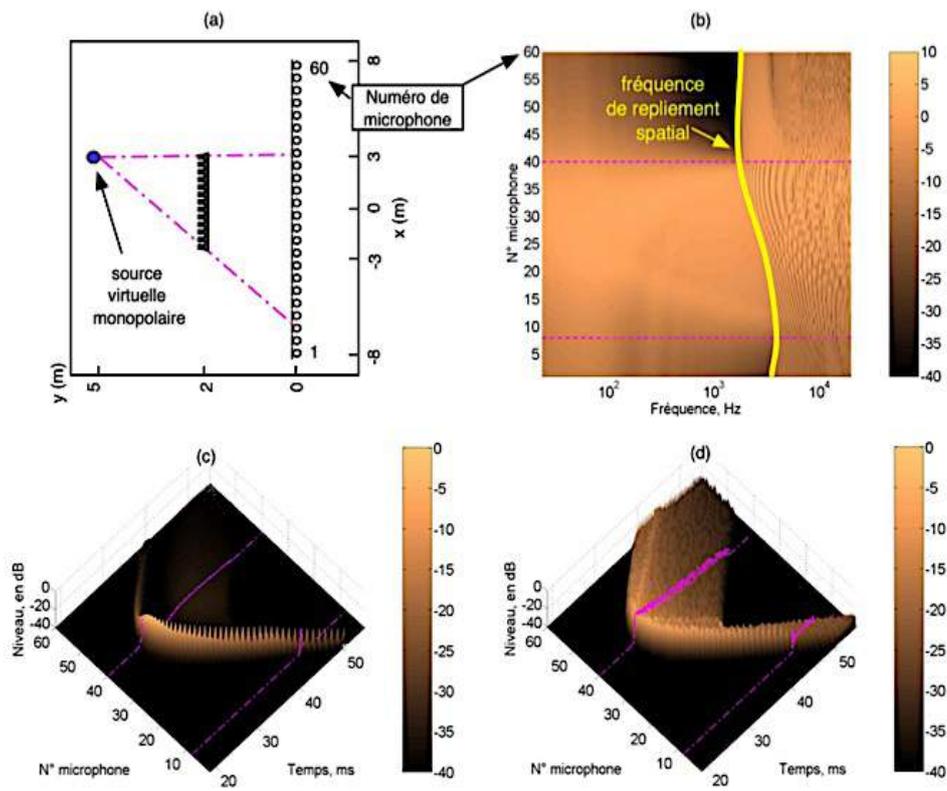


Fig. 2.7 : Mise en évidence du phénomène d'aliasing spatial. Source [CorteeL, 2004].

Spatialement, la position 2 nous semble être la plus intéressante dans le cas d'une voix-off. Outre le respect du timbre de la source, elle semble offrir un élargissement de la source tout en gardant une localisation de celle-ci au centre de l'image frontale, i.e. a priori au centre de l'écran.

Notons qu'une technique parfois utilisée en 5.1 permet de reproduire un élargissement via l'un des paramètres du panoramique : la divergence. Une certaine quantité du signal central (souvent exprimée en %) est envoyée à la fois dans le canal gauche et le canal droit. L'un des désagréments de cette méthode réside dans le fait de percevoir une diffusion "multipoints", c'est-à-dire de percevoir la même source diffusée par 2 enceintes différentes au même moment⁹. Cette aberration n'a effectivement pas été perçue via synthèse de front d'onde.

⁹ De fait, par effet de précedence, il peut en résulter une distorsion de localisation pour des positions d'écoute latérales. Par exemple, l'auditeur placé sur le côté gauche de la salle pourrait percevoir le signal en provenance du canal gauche avant même le signal en provenance du canal central : la voix-off serait alors localisée à gauche.

Protocole 2 : voix-off associée à une réverbération (détails en Annexe 1.2)

La même voix-off est maintenant associée à une réverbération de petit volume. Cette réverbération est obtenue grâce au plugin natif de Nuendo, "Reverence", utilisant la convolution de réponses impulsionnelles¹⁰.



Fig. 2.8 : Fenêtre principale du plugin Reverence.

Nous avons testé la source WFS en deux points du Performer ; sa réverbération associée, stéréophonique, était placée juste derrière, arrangée en groupe, et d'une

¹⁰ Une réverbération de type "petite pièce vide" a été utilisée. Notons qu'un grand volume aurait pu être choisi, mais ne nous aurait pas facilité le placement de source car, dans un souci d'éviter l'effet "présence + réverbération", nous aurions été contraints d'augmenter le niveau du son réverbéré, ce qui aurait quelque peu brouillé l'intelligibilité de la voix. De plus, puisque la perception de l'effet de salle d'une source sonore -de sa "spatialité"- est fortement influencée par la quantité de premières réflexions qui lui sont associées [Rumsey, 2001] (p. 35), nous avons plus de chance d'obtenir un effet satisfaisant avec un volume réduit de réverbération.

largeur apparente que nous avons fait varier¹¹. L'effet obtenu, à base de réverbération, de filtrage, et surtout de synthèse de front d'onde, a par suite été approché pour la source discrète, uniquement par filtrage et réverbération. Le but était d'obtenir un résultat comparable pour les deux sources, en termes de timbre, de niveau, de rapport direct/réverbéré, et de largeur apparente¹². Comme pour le protocole précédent, les constats amenés mettront en vis-à-vis la source avec WFS, et la source en sortie directe, sans WFS.

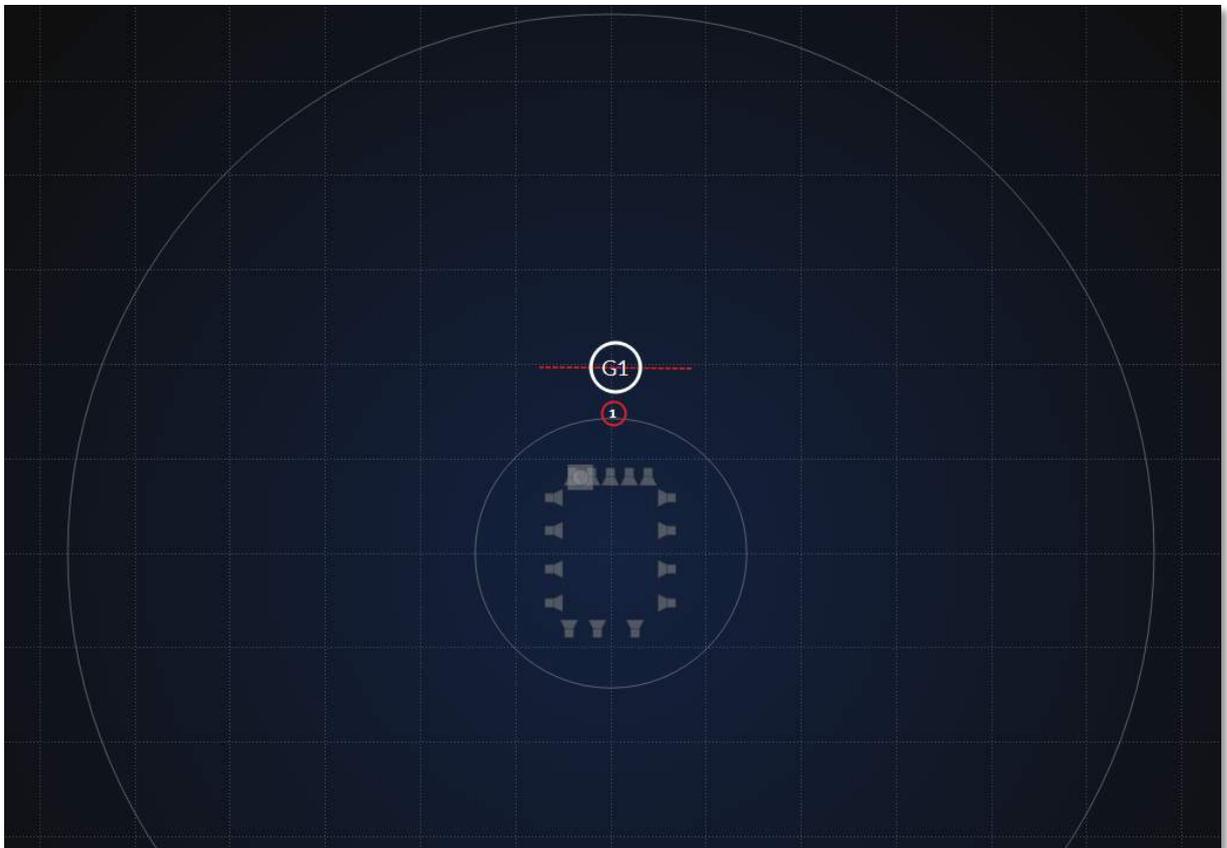


Fig. 2.9 : Première configuration : objet WFS à hauteur du premier cercle du Performer, groupe de réverbération (stéréo) derrière la source et relativement proche de celle-ci.

¹¹ Cf. 2.2.3, p. 122, pour plus de précisions sur le fonctionnement en groupe.

¹² Concept de largeur apparente explicité notamment dans [Rumsey, 2001], p. 36.



Fig. 2.10 : Deuxième configuration : objet WFS au niveau du deuxième cercle du Performer, groupe de réverbération (stéréo) derrière la source et relativement proche de celle-ci, largeur du groupe « réverbération » reconsidérée.

Analyse protocole 2

Dans un contexte de postproduction sonore cinématographique, le niveau, la balance spectrale, et le rapport son direct/son réverbéré sont des paramètres habituellement considérés pour "placer" une source dans un espace acoustique donné. Dans l'utilisation de la WFS, on peut observer l'émergence de 3 nouveaux paramètres supplémentaires visant à créer une sensation d'espace associé à une source.

1) Le placement de la source dans "l'espace WFS", qui pour une source mono se résume à choisir une position de l'objet sur le plan du Performer, permet de modifier le comportement dispersif de cette source. Opter pour une distance holophonique faible amènera le système à associer à la source un mode "sphérique" de diffusion ; à l'inverse, une distance holophonique importante amènera à la synthèse d'un comportement en ondes planes. Nous avons constaté que ce paramètre, propre à la WFS, permettait de percevoir une perspective dans un espace qui serait créé derrière la rampe frontale de haut-parleurs. En restitution discrète, la sensation de profondeur semble être créée, mais se limiter à l'éloignement d'une "ligne sonore", comme si l'espace situé en amont de cette ligne lointaine n'était pas perçu (cf. figure ci-dessous).



Fig. 2.11 : A gauche, la sensation d'espace créée en WFS ; à droite, la profondeur créée ne nous semble pas rendre de perspective.

ii. ENCADRE : EFFET DE DISTANCE EN WFS

Dans un mixage multicanal classique, la reproduction des effets de salle et d'espace est amenée par l'utilisation de réverbérations et de filtrages éventuels. C'est en jouant sur le niveau, le rapport champ direct/champ réverbéré et sur le choix de cette réverbération, que le mixeur construit une distance et une sensation d'espace réaliste autour de la source. Sur un système WFS, la courbure de l'onde apporte une information spatiale supplémentaire, et permet a priori d'accroître la sensation d'espace associée à une source. En tant que technique de reproduction sonore holophonique, la WFS permet de gérer indépendamment deux paramètres liés à la distance d'une source : premièrement la forme du champ acoustique associé au son direct de la source virtuelle (la structure de la forme d'onde) ; secondement la distribution de l'énergie dans l'espace lié à la source sonore.

Dans *Monitoring distance effect with Wave Field Synthesis* [Noguès et al., 2003], les auteurs définissent deux types de phénomènes liés à la perception de distance via la terminologie suivante :

- La **distance holophonique**, qui se réfère à la position de la source virtuelle, créée par synthèse de front d'onde. Sa cohérence perceptive peut dépendre de la précision du système (taille et caractéristiques des enceintes, repliement spatial).
- La présence de la source, qui se réfère à l'impression subjective de distance liée au temps et à la distribution spatiale de la réponse de la pièce. On assimile ce paramètre au **rapport champ direct/champ réverbéré**.

Les auteurs affirment également que la courbure du front d'onde renvoie directement l'auditeur à la sensation d'une direction et d'une distance de la source sonore perçue. Cependant, la forme du front d'onde ne constitue pas en soit une information suffisante pour la perception de la distance. La construction d'un espace en WFS pourrait donc combiner l'utilisation des outils d'ors et déjà à disposition du mixeur, associée aux différentes structures de formes d'onde synthétisables.

2) Le placement relatif, en profondeur, de la réverbération par rapport à la source est un autre paramètre important quant à l'illusion d'espace créée. En admettant une position centrale de la réverbération -i.e. les deux pôles sont placés symétriquement de part et d'autre de l'axe verticale du Performer-, une position trop lointaine de la réverbération par rapport à la source nous a amené à percevoir une certaine décorrélation entre le champ direct et le champ réverbéré. La fusion nous a semblé croître pour une position de réverbération proche, et en retrait de la source lui étant associée. Comme évoqué précédemment, ce constat a entre autres été fait pour une réverbération de type petit volume. Il est probable que, pour un volume plus important, dont la densité de premières réflexions est moins importante [Rumsey, 2001], la fusion entre champ direct et champ réverbéré nous serait apparue moins déterminante quant à la perception d'un placement réaliste de source sonore dans un espace acoustique. Nuancions également en précisant que la fusion entre champ direct et champ réverbéré n'est pas forcément un effet voulu, et s'il le devient, peut être fortement dépendant de la nature de la source sonore originale.

3) Lors de ces tests, nous avons constaté une probable corrélation entre la position de la source sur l'axe verticale, et la largeur de la réverbération lui étant associée : plus la source est loin, plus la réverbération demande à être élargie pour proposer un effet de perspective cohérent. De plus, il semblerait qu'une relation puisse exister entre la "fenêtre WFS" précédemment décrite (cf. 2.1.2, « zone de fenêtrage frontal ») et la largeur des réverbérations. Dans notre cas, on constate qu'une largeur satisfaisante de réverbération était obtenue pour un placement des pôles stéréophoniques proches des bords de la zone de fenêtrage frontal.

Enfin, ces paramètres établis sont à apprécier avec beaucoup de précautions, et demanderaient à eux seuls une étude complète. Ils ne forment qu'une porte

d'entrée vers un champ de possibilités techniques, et d'effets perceptifs, dans le traitement de l'espace associé à une voix. Notons également qu'en plus d'avoir utilisé un seul type de source, un seul type de réverbération, et deux positions de sources, nous n'avons pas envisagé l'utilisation de réverbérations multicanales. La gestion des pôles arrières de celles-ci dans une synthèse de front d'onde formerait également une étude à part entière.

2.2.2. Le cas des ambiances frontales

Objectifs : Comparer la perception d'une ambiance "bipolaire" sur un système standard 5.1 et sur un système WFS

Sur un système de restitution 5.1, la gestion des ambiances au mixage est souvent influencée par le contenu du montage son et la construction de celui-ci : une série de pistes est réservée aux ambiances frontales (stéréophoniques pour les canaux "Left" et "Right", monophoniques pour le canal "Center"), tandis qu'une autre sert à accueillir les ambiances surround (souvent stéréophoniques). Qu'advient-il des ambiances frontales, composantes importantes d'un mixage 5.1 standard, lorsque celles-ci sont placées dans "l'espace WFS" ?

Protocole 1 : ambiance stéréophonique de large bande spectrale
(détails en Annexe 1.3)

Une ambiance stéréophonique, spectralement large, de type "fond d'air campagne", est envoyée dans 2 sources du Performer. Avec ces sources est formé un groupe composé de 2 pôles (cf. encadré iii suivant). Dans la suite de ce document, nous appellerons sources "bipolaires" les sources stéréophoniques diffusées via synthèse de front d'onde. Par extension, on pourra parler de sources "tripolaires", "quadripolaires", ou plus généralement "multipolaires".

L'ambiance sélectionnée comporte par ailleurs quelques faibles événements distinctifs, tels qu'un léger vent dans les arbres, des passages de véhicules lointains, et quelques oiseaux. A l'écoute en stéréophonie, cette ambiance nous semblait déjà être d'une largeur apparente importante, et nous apparaissait enveloppante (pour une position d'écoute au sweet spot). Nous avons reproduit la même méthode de tests que ceux effectués sur la voix : la configuration dans le séquenceur nous permettait de basculer entre une écoute 5.1 et une écoute WFS.

NB : Par abus de langage, il est probable de trouver dans la littérature le terme source "stéréo" pour désigner une source WFS contenant deux pôles. De fait, par "stéréo", il est considéré que la source diffusée est la représentation de 2 haut-parleurs virtuels placés dans l'espace WFS [Theile et al., 2003].

iii. ENCADRE : LE *VIRTUAL PANNING SPOT* ET LA GESTION DES SOURCES COMPLEXES

A l'heure actuelle, la puissance de calcul limitée des processeurs de diffusion WFS restreint le nombre de sources synthétisables. Or, la littérature un ancienne sur le sujet nous préconise entre autres la prise de son monophonique de proximité afin d'éviter tout effet de salle qui sera lui aussi spatialisé dans l'espace WFS. Ce fractionnement absolu est en pratique difficilement envisageable (en particulier au cinéma), et parfois peu recommandable, ce pour plusieurs raisons.

La première, nous l'avons évoquée, tient aux limites de calcul des processeurs WFS. Comment envisager par exemple la reconstruction d'une foule d'une centaine de personnages dans un théâtre, qui nécessiterait dès lors une puissance de calcul considérable pour permettre la synthèse d'une centaine de sources. On devrait de plus envisager un dispositif microphonique permettant l'enregistrement séparé de chaque personnage.

Une autre raison réside dans la nature même de la prise de son en production cinématographique. Le cadre, fenêtre sur l'action, forme également une limite pour tout dispositif de prise de son, ce dernier étant censé se faire discret et même invisible. De fait, le placement parfois lointain de la perche, ne s'accorde pas avec le principe préconisé de réaliser une prise de son de proximité.

Une solution pouvant être amenée par les micros « cravate » (communément appelés « HF »), ne semble pas être viable à l'heure actuelle : l'aspect qualitatif de ce type de dispositif n'est pas à même de substituer la perche.

A ces différents problèmes s'ajoute le fait que certaines sources ne peuvent être captées par une prise de son monophonique, et rendue par la synthèse d'un monopole sans subir de dégradation. Le cas plusieurs fois évoqué du piano, instrument beaucoup trop large spatialement et spectralement pour être rendu en monophonie, est un exemple qui nous invite à envisager les choses autrement.

Partant du principe que nous percevons des groupes d'objets sonores, délimités en zones, une technique particulière est explicitée dans *Potential wavefield synthesis application in the multichannel stereophonic world* [Theile et al., 2003] afin de limiter le nombre de canaux nécessaires à la transmission d'une source, tout en permettant de restituer sa largeur spatiale et spectrale. Cette technique se nomme le *Virtual Panning Spot*.

Précisons au lecteur que dans notre étude, nous avons utilisé le terme « pôle » pour désigner les « spots » explicités dans [Theile et al., 2003].

Notons que le choix d'une ambiance spectralement large pour débiter cette série de test n'est pas totalement hasardeux. Suite aux constats effectués lors de notre première approche du système WFS, nous avons déjà pu noter que plus la source disposait d'un spectre large, plus les effets d'aliasing fréquentiel amenés par la synthèse de front d'onde allaient être problématiques. Nous avons alors, pour cette deuxième série de tests, fait le choix de classer les ambiances en 2 catégories : les ambiances à large spectre, telles que des fonds d'air, et les ambiances à timbre plus resserré, telles que des chants d'oiseau en forêt (en l'absence de quelconque fond d'air), des "chants" d'insectes, une foule humaine, ...

Comme pour les expérimentations relatives à la voix, 3 positions dans "l'espace WFS" ont été testées : le placement 1 restera la position au plus proche des haut-parleurs (les deux pôles de l'ambiance stéréo sont chacun placés sur les haut-parleurs gauche et droit) ; la configuration 2 correspondra toujours à un placement intermédiaire du groupe (le groupe bipolaire est placé au niveau du premier cercle du Performer) ; la configuration 3 restera la position la plus éloignée (le 2 pôles du groupe bipolaire touchent quasiment au second cercle du Performer).



Fig. 2.12 : Position 1 : groupe bipolaire "ambiance" au niveau des enceintes (un pôle dans chaque enceinte). Les positions 2 et 3 seront donc les mêmes qu'en figure Fig. 2.6 p. 94.

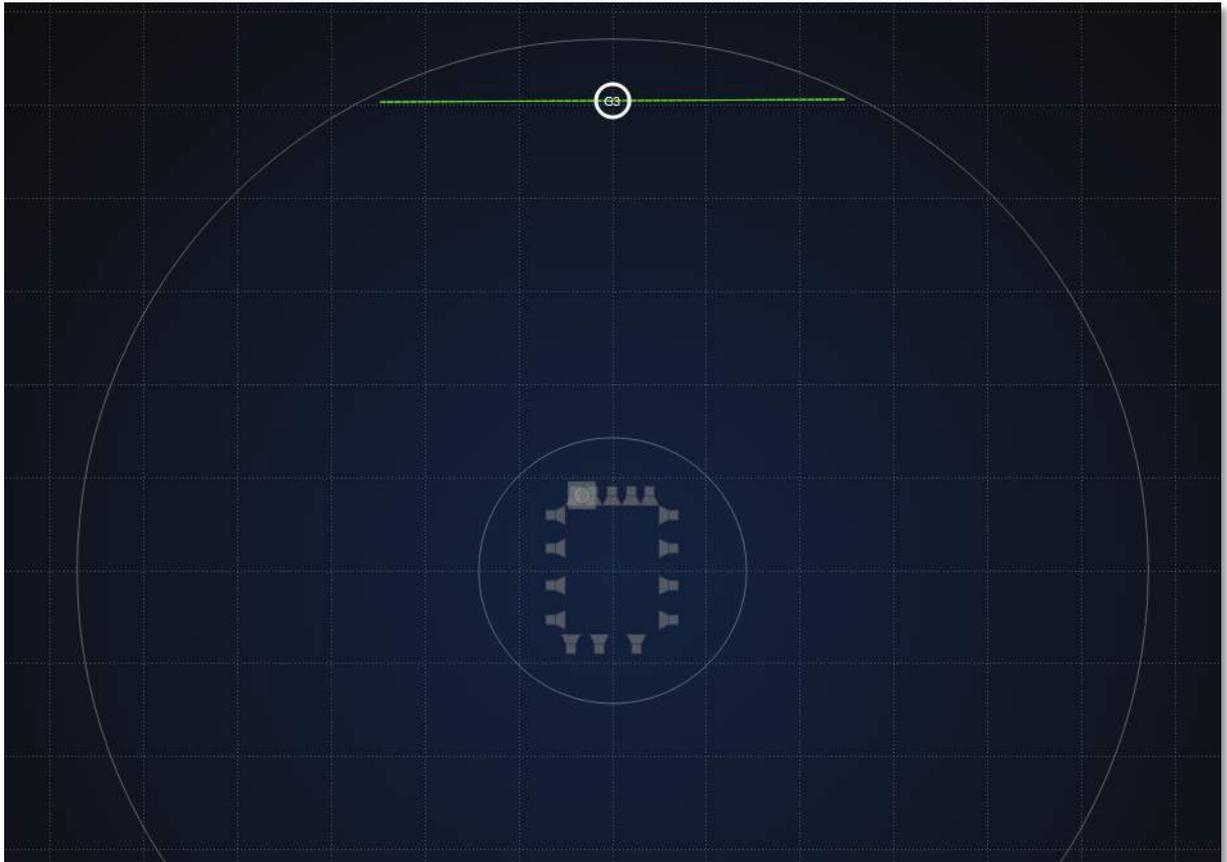


Fig. 2.13 : Position 3 : groupe bipolaire "ambiance" au niveau du second cercle du Performer ; largeur du groupe reconsidérée.

Analyse protocole 1

Pour cette ambiance sélectionnée, et pour la position 3 particulièrement, les comparaisons effectuées lors de ces tests amènent à constater des différences notables entre les deux systèmes de restitution. Au sweet spot, la stéréophonie semble nous offrir des résultats convaincants en termes de largeur apparente et de sensation d'immersion. Cela dit, ces quelques caractéristiques sont assez radicalement remises en question dès lors que la position d'écoute se latéralise, chose plutôt fréquente dans une salle de cinéma...

En WFS, outre le fait qu'une largeur apparente et une sensation d'immersion semblent être atteintes, l'un des biais induits par la diffusion stéréophonique discrète semble être évité : quelque soit notre placement d'écoute dans la salle, la sensation d'une image sonore large est préservée, et ce même pour des placements extrêmes. La position la plus lointaine de l'ambiance (position 3) nous apparaît comme étant la plus adaptée à ce type de contenu puisqu'elle apporte cette caractéristique a priori intéressante. De plus, ce placement d'ambiance semble lui apporter profondeur et perspective, choses difficilement perceptibles lors d'une restitution stéréophonique, et ce même pour une position d'écoute centrale. Pour la suite de nos manipulations, c'est donc un positionnement que nous privilégierons lorsqu'il s'agira de spatialiser une ambiance fréquemment large.

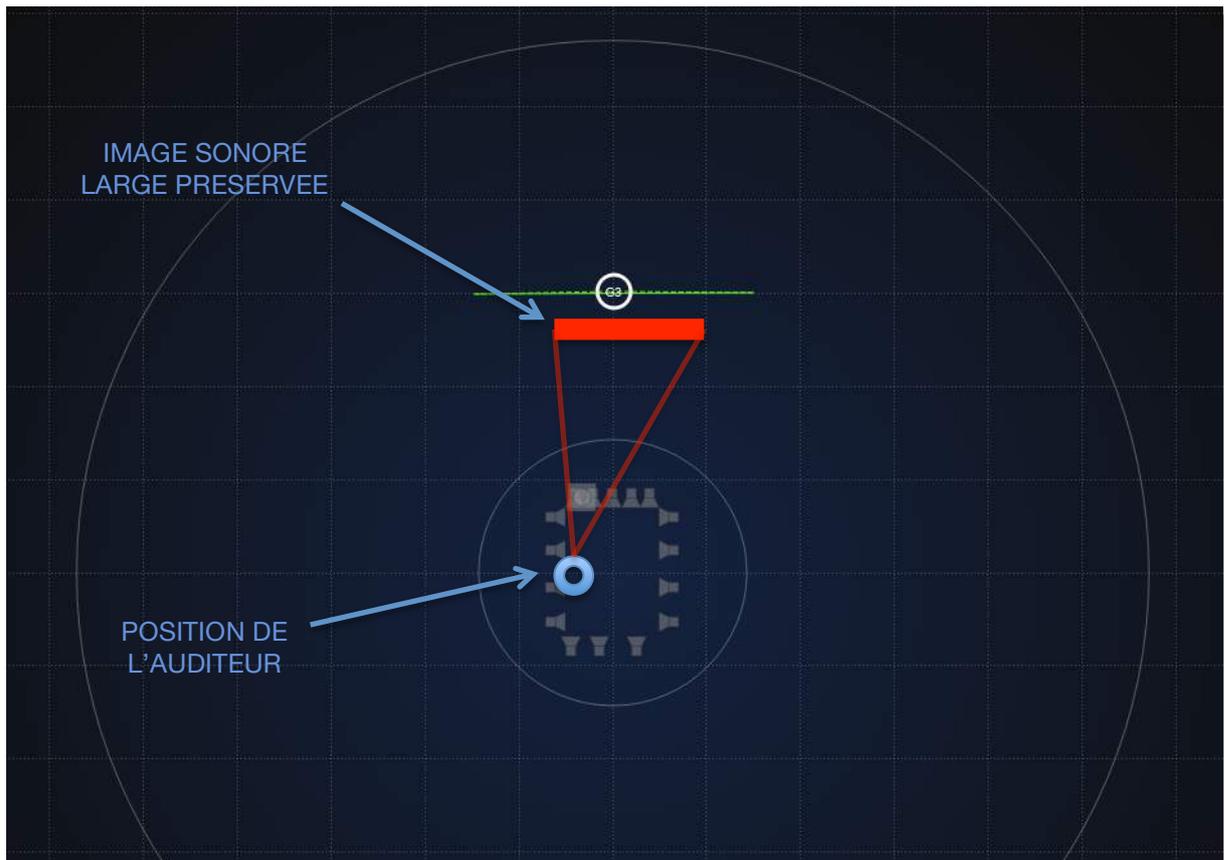


Fig. 2.14 : En WFS, pour une position d'écoute excentrée, une image sonore large semble être préservée.

Par ailleurs, même si les artefacts fréquentiels restent minimes pour un placement d'ambiance en position 3, nous trouvons qu'il est préférable de laisser le choix au mixeur de disposer des caractéristiques de la synthèse de front d'onde, ou de les éviter le cas échéant. Ainsi, le respect du timbre original de la source, et la cohérence de l'image sonore quelque soit la position d'écoute formeraient deux bornes entre lesquelles le mixeur choisirait la WFS, ou pas. Quant à l'impression de profondeur et de perspective créée au delà de la ligne des haut-parleurs, n'est-il pas au final intéressant d'accéder à ces effets, quitte à altérer légèrement le contenu fréquentiel de la source ?

Protocole 2 : ambiance stéréophonique à contenu fréquentiel resserré (détails en Annexe 1.4)

Une nouvelle ambiance, privée de fond d'air, est composée d'éléments ponctuels et distincts : des chants d'oiseaux relativement nombreux, sont accompagnés d'un oiseau unique, plus lointain, amené à "exciter" l'acoustique du lieu. Un groupe bipolaire est à nouveau formé avec cette ambiance stéréophonique. Du fait de l'absence de quelconque fond d'air, l'ambiance contient un timbre plus resserré, caractérisé par la tessiture même des oiseaux.

En écoute stéréophonique, la sensation d'espace associée à l'ambiance nous paraît déjà assez importante. La réverbération de montagne/forêt semble participer, pour beaucoup, à la "spatialité" ainsi créée. Malgré cette impression d'espace, on notera une certaine frontalité propre à l'ambiance.

Sur ce test, les 3 positions de groupe seront testées, identiques aux positions présentées pour le protocole 1.

Analyse protocole 2

Les tests effectués sur cette ambiance contenant des évènements distincts (oiseaux, absence de fond d'air) nous offrent, jusqu'à maintenant, les constats les plus convaincants en termes de restitution WFS.

La contenu spectralement resserré de l'ambiance semble dans un premier temps nous éviter nombre des soucis provenant du filtrage issu du la synthèse de front d'onde. Bien qu'en position 2 une très légère décoloration du registre aigu puisse

se faire ressentir, elle tend à complètement disparaître pour les positions 1 et 3, ce qui libère un peu plus les possibilités de placement pour ce type d'ambiance.

Dans un second temps, une direction de source cohérente, quelque soit notre position dans la zone d'écoute, semble être préservée. Ceci se manifeste par une largeur apparente conservée, et ce même pour des positions d'écoute latérales extrêmes. On constate notamment qu'en WFS, l'image sonore frontale est d'avantage respectée avec cette ambiance fréquemment restreinte, là où avec le fond d'air nous pouvions plus facilement obtenir de légères distorsions d'image suivant notre position d'écoute dans la salle.

Le relief, pour ce type d'ambiance, est un effet pour lequel nous n'attendions pas forcément un rendu si évident. En particulier pour la position 3, les éléments constituant l'ambiance (oiseaux proches nombreux ; oiseau lointain unique) semblent être placés sur des plans différents après synthèse de front d'onde. Cette perception nouvelle de la perspective est effectivement une plus value intéressante par rapport à la diffusion stéréophonique, qui comme pour tous les autres critères étudiés, serait à vérifier avec d'autres types d'ambiance.

Enfin, la création de l'espace au delà de la rampe frontale, ainsi que la contribution d'un plus grand nombre de haut-parleurs à la synthèse de la source, tend à nous faire ressentir un certain réalisme par rapport à la restitution stéréophonique de l'ambiance. En effet, la ligne d'enceintes n'est plus aussi audible : une source est perçue, mais il devient difficile de déterminer le dispositif la restituant.

Protocole 3 : ambiance monopolaire placée au niveau de l'enceinte centrale (détails en Annexe 1.5)

Lorsque nous avons entamé le travail sur les ambiances, l'effet d'élargissement constaté précédemment sur la voix-off (cf. 2.2.1) nous a assez naturellement dirigé vers l'étude d'une ambiance mono positionnée dans "l'espace WFS". Comme auparavant, nous comparerons le rendu du système WFS avec celui du système discret (ambiance mono placée dans le canal central). En synthèse de front d'onde, les 3 positionnements précédemment décrits seront à nouveau utilisés.

L'ambiance monophonique choisie est issue d'une scène de combat dans un hangar, contenant encouragements et hurlements du public. C'est une ambiance très fournie, très vivante, d'un caractère assez profond, ou du moins disposant d'un champ réverbéré important (acoustique de hangar). Le choix de cette ambiance s'est fait dans le but de diversifier un peu plus les matériaux sonores étudiés. N'étant ni de nature urbaine, ni composée d'éléments naturels (oiseaux, vent, ...), elle a pour caractéristique notable de contenir une grande quantité d'éléments ponctuels, d'éclatements de voix, finissant presque par se noyer entre eux.

Analyse protocole 3

Le traitement réservé aux ambiances stéréophoniques frontales (cf. 2.2.2, protocoles 1 et 2) nous a naturellement conduit à évaluer le rendu d'une ambiance monophonique sur un système WFS. Matériau récurrent de la postproduction sonore cinématographique, l'ambiance monophonique est principalement utilisée pour deux raisons :

- **Permettre le raccord** : Autrement appelée "ambiance raccord", ce type d'utilisation de l'ambiance monophonique a pour but de « lisser » les cisaillements opérés lors du montage son dit "des directs". Ambiance la plupart du temps issue du tournage, le "raccord" accompagne souvent les perches dans le canal central, ce qui le rend plutôt "transparent" par rapport au reste des matériaux sonores. En WFS, nous conjecturons que l'ambiance raccord pourra prendre deux directions opposées. Dans une scène silencieuse, là où très peu d'éléments accompagnent les voix, le raccord sera probablement nécessaire. La largeur apparente constatée lors de ces tests, ainsi que la fusion amenée par la WFS (cf. 2.2.1) pourront potentiellement conduire à un raccord satisfaisant, dépendant lui même du sort des sources "voix" et "directs". Dans une scène plus bruyante, il sera intéressant d'évaluer l'apport de l'ambiance raccord. En effet, du fait du remplissage du canal central, induit par la contribution de ce même canal à la synthèse d'autres ambiances, nous nous demandons si le raccord présentera toujours un intérêt. Les autres ambiances ne suffiront-elles pas au « lissage » précédemment évoqué ?
- **Jouer l'ambiance au centre** : Dans certains cas de figure, la stéréophonie habituellement utilisée pour les ambiances ne correspond pas à l'espace sonore requis par le champ pictural. Les endroits clos, les gros plans, les éléments sonores extérieurs entendus de l'intérieur (exemple : oiseaux extérieurs entendus depuis l'intérieur d'une maison), sont des exemples de configuration où la stéréophonie et sa largeur potentielle ne sont pas tout à fait adaptées à la scène jouée à l'écran. Dans ces cas bien précis, le retour à la monophonie devient souvent une solution. En WFS, nous avons constaté que le système présentait plusieurs apports quant au rendu spatial des ambiances monophoniques. Là où la largeur sur le système discret se résume à un point, le dispositif WFS conduit à une

restitution plus diffuse de la source. Plus la distance holophonique est importante, plus l'ambiance paraîtra large, ce qui selon nous est un effet intéressant pour ce type de source : une source monopolaire habituellement resserrée, peut maintenant prendre de la largeur, et devenir presque aussi large qu'une ambiance stéréophonique. D'un autre côté, la divergence, outil utilisé sur un système multicanal pour perceptivement élargir une source ponctuelle, amène à localiser la source dans une autre direction dès lors que la position de l'auditeur n'est plus centrale. Grâce à la synthèse de front d'onde, nous avons effectivement constaté que même si l'ambiance monophonique dispose d'une largeur importante (distance holophonique importante sur l'axe vertical du Performer), elle continue à être perçue dans la même direction quelque soit la position de l'auditeur dans la salle. Cette cohérence de localisation, décidément récurrente dans toutes nos observations faites jusqu'ici, forme une plus value très intéressante quant au traitement des ambiances monopolaires.

iv. ENCADRE : CONTRIBUTION DES HAUT-PARLEURS ET ALIMENTATION D'UN SYSTEME WFS

Les sources ponctuelles sont synthétisées par le processeur qui applique un retard et un gain à certains haut-parleurs de la rampe. Le premier front d'onde provient du haut-parleur situé directement devant la source (selon la normale au banc de haut-parleurs). C'est ce premier front d'onde qui permet la localisation de la source. Comme on le voit sur la figure Fig. 2.15 qui suit, les deux haut-parleurs entourant le premier concerné émettent à leur tour à l'instant $T+1$, et viennent contribuer à l'énergie du front d'onde constitué. A l'instant $T+2$, les deux haut-parleurs suivants contribuent à la création de ce front d'onde, et ainsi de suite jusqu'à ce que toutes les enceintes concernées participent à la synthèse de la source. Notons qu'il n'est pas obligatoire que tous les haut-parleurs contribuent à la création du front d'onde.

On le comprend donc ici, un système WFS va travailler essentiellement avec des retards. On va tout d'abord indiquer les informations relatives à la localisation de la source virtuelle, à savoir l'azimut (ou l'angle d'incidence de la source), et sa distance. Comme nous l'avons vu dans une partie précédente (cf. 2.1.1), ces informations de spatialisation sont renseignées par l'utilisateur via des interfaces. Le processeur d'écoute va interpréter ces données et les traduire en retard pour que chaque haut-parleur contribue au front d'onde, et ce au bon moment.

La dernière étape consiste en un filtrage du signal : il correspond à une calibration du système d'écoute opérée par le processeur.

La différence fondamentale entre un "élargissement" par divergence, et un "élargissement" via WFS, réside dans les traitements requis pour parvenir à l'effet escompté. Là où la divergence consiste en une "simple" distribution du signal, suivant les lois de panoramique du système, l'élargissement en WFS (amené par un éloignement virtuel de la source) fera intervenir délais (« dT » sur la figure Fig. 2.16) et gains (flèches montantes sur la même figure Fig. 2.16). Dans le cas de la simulation d'une source en onde sphérique, la première contribution provenant du haut-parleur situé sur la normale de l'axe de la source est primordiale pour la localisation. La contribution suivante des haut-parleurs juxtaposés est retardée et atténuée. Dans le cas d'une simulation d'une onde plane, en fonction de sa position, l'atténuation et le retard apportés sont moindres, d'où la sensation d'une source plus large.

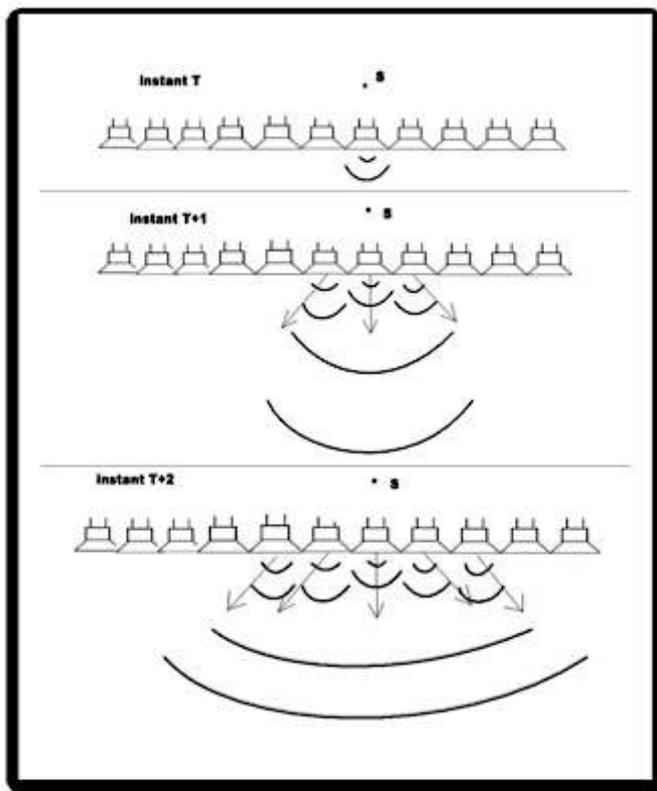


Fig. 2.15 : Contribution du banc de haut-parleurs à la synthèse d'une source ponctuelle, d'après [Husson, 2011].

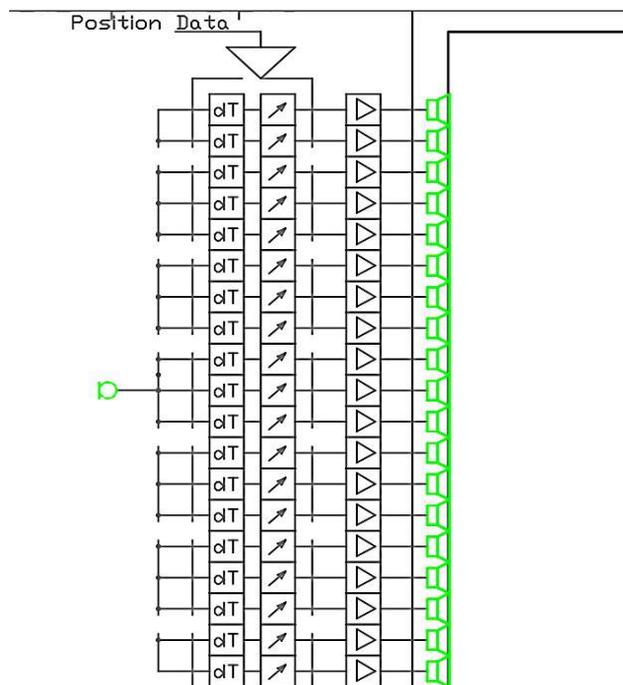


Fig. 2.16 : Fonction d'alimentation du banc de haut-parleur d'après [Holophony].

2.2.3. La spatialisation de la musique "off"

Objectifs : comparer le rendu d'extraits musicaux "off" sur un système WFS et un système 5.1

Quasi-exclusivement issue du monde stéréophonique, la musique au cinéma est souvent sujette à une spatialisation elle-même induite par les technologies discrètes. Lorsque le mixeur n'a pas accès aux éléments séparés de la musique (on parle également de "stems"), il est "contraint" de composer avec les canaux mis à sa disposition : dans la meilleure configuration au nombre de 5 ou 6 -cas où la musique est issue d'un mixage multicanal-, ils sont le plus souvent limités au nombre de 2 -cas standardisé du mixage musical-. C'est sur ces sources musicales stéréophoniques que nous nous pencherons lors de ces tests.

En WFS, la musique pourrait, au vu de ce que nous avons constaté sur les ambiances, s'approcher d'une organisation en bipole. La possibilité d'une structure en groupe (d'éloigner ce groupe, de l'élargir) nous semble être un paramètre intéressant quant au sort des sources musicales subissant une synthèse de front d'onde. Nous évaluerons dans un premier temps l'apport de la WFS par rapport à une source stéréophonique, sans avoir recours à quelconque effet "surround". Par suite, nous comparerons le rendu WFS avec la restitution multicanale faisant appel aux canaux arrières.

Nous disposons de 4 extraits musicaux :

- Wolfgang Amadeus Mozart, *Requiem*, "Lacrimosa"
- Hans Zimmer, *Time*, extrait de la bande originale du film *Inception*

- Cliff Martinez, *Bride of Deluxe*, extrait de la bande originale du film *Drive*
- Eddie Vedder, *Hard Sun*, extrait de la bande originale du film *Into the Wild*

Protocole 1 : groupe « musique » placé dans l'espace WFS ; source discrète stéréophonique (surrounds inactifs) (détails en Annexe 1.6)

Le premier protocole de cette série de tests mettait en comparaison la diffusion des 4 extraits musicaux, à la fois en WFS (via le groupe bipolaire créé), puis en stéréophonie (via les haut-parleurs gauche et droit de la rampe frontale). Dans la suite de cette partie, tout comme précédemment, nous avons jugé accessoire de proposer au lecteur une vision précise de nos constats relatifs aux protocoles mis en place (voir Annexes 1.6 et 1.7 pour les résultats détaillés relatifs aux protocoles 1, 2). Les effets perçus pour les sources musicales ont été très proches de ceux obtenus pour les ambiances, à la différence près que chacun de ces effets nous ont semblé plus marqués :

- Le resserrement de l'image frontale et l'atténuation du registre aigu se sont à nouveau faits ressentir dans l'usage de la WFS, et pouvaient même nous paraître plus audibles du fait de la concentration en transitoires des extraits musicaux.
- Par ailleurs, la WFS s'est avérée être efficace en vue d'une cohérence de l'image sonore frontale quelque soit la position d'écoute de l'auditeur dans la salle. Nous noterons quand même un constat nouveau : là où les sensations de profondeur et de perspective semblaient être obtenues assez facilement pour les ambiances (cf. 2.2.2, « position 3 »), cet effet nous a paru plus discret pour les extraits musicaux sélectionnés.
- Nous avons même noté une césure entre le contenu musical dit « orchestral », et le contenu dit « actuel » : dans le premier cas, l'unicité de l'espace (ensemble instrumental compact, placé dans une acoustique unique), semblait être à même de restituer la profondeur ; dans le second

cas, le placement complexe des sources dans l'espace « stéréo » (ajout de nombreuses réverbérations artificielles), ne permettait pas un effet convaincant.

Ces premiers tests sur des sources musicales nous ont tout de même amenés à douter un peu plus quant aux limites inhérentes à la WFS. Nous nous sommes aussi demandés si le jugement que nous avançons n'était pas aussi du à notre plus grande critique vis à vis des contenus musicaux. En effet, en présupposant que notre culture sonore vienne en grande partie de la musique, il est probable que nous soyons un peu plus clairvoyants sur un certain nombre de problèmes, mis en avant par le mixage musical lui-même, codifié et récurrent sur de nombreuses productions : spectre large avec présence d'extrêmes aigus, latéralisation marquée de certaines sources, ... Autant d'effets que la WFS tend à distordre.

Ainsi, et ce de manière encore plus marquée que pour les ambiances, le mixeur devra opérer un choix entre la synthèse de front d'onde et la stéréophonie :

- Respecter le contenu fréquentiel l'amènera plutôt à opter pour la restitution discrète stéréophonique ;
- Proposer une image sonore frontale cohérente, indépendante de la position d'écoute, le conduira à la restitution WFS.

Protocole 2 : groupe « musique » placé dans l'espace WFS ; source discrète stéréophonique (surrounds actifs via panner 5.1) (détails en Annexe 1.7)

Ce protocole vise à mettre à contribution des canaux que nous avons jusqu'ici laissé de côté, les canaux dits de « surround », parfois appelés (et à tort) les « arrières ». Cette dernière terminologie est un leurre, en ce sens que les canaux ainsi nommés sont également diffusés par des enceintes latérales, souvent plus nombreuses que les enceintes arrières dans une salle de cinéma¹³. L'usage de ces canaux, également appelés Ls pour « Left surround » et Rs pour « Right surround », répond à un certain nombre de codes et de recettes, plus ou moins malléables en fonction des directions esthétiques requises par la narration. Cela dit, une chose peut être avancée et considérée comme une généralité : le réflexe d'attention, déjà très largement décrit dans la littérature (cf. 1.2) semble conditionner l'usage de ces canaux.

Les canaux surrounds, selon ce qui a été dit précédemment, seraient alors « limités » à 3 usages récurrents (notons bien que les usages cités ne prennent pas en compte les cas plus rares) :

- **Les évènements ponctuels** : Dans ce cas, les sources sonores font référence à un élément visuel, le plus souvent mis en mouvement, se dirigeant soit du cadre vers le hors-champ, soit dans le sens inverse. Ces effets trouvent souvent leur efficacité dans la furtivité de leurs stimuli visuels associés, en ce sens qu'il

¹³ Le terme anglophone « surround », de sa traduction basique « entourer », exprime bien le fait que ce type de haut-parleurs ne se limite pas à une implantation à l'arrière de la salle de cinéma.

serait difficile de ne pas faire intervenir le réflexe d'attention pour une source sonore trop lentement mise en mouvement de l'avant à l'arrière de la zone d'écoute (un tel effet pourrait exister, mais ne semble pas être récurrent dans les mixages déjà produits).

- **Les effets « d'ouverture »** : Cas visant à briser quelque peu la continuité sonore bâtie lors du montage son, les canaux surrounds servent alors à « ouvrir » le champ sonore, ce qui lui accorde potentiellement une largeur apparente plus importante. Cette technique est souvent utilisée pour mettre en place un contraste entre deux plans, le cas fréquent montrant successivement un plan plutôt serré puis un plan plus large. Notons que la sensation d'ouverture ainsi produite n'est pas forcément rattachée à un changement visuel, mais peut aussi faire l'objet d'un effet détaché de la succession « plan serré / plan large ».

- **L'enveloppement** : Il s'agit du critère principal mis en avant lorsque le son surround est évoqué, le terme « d'immersion » faisant lui aussi souvent son apparition. Généralement réservée aux ambiances et aux sources musicales, cette sensation est usuellement obtenue via l'usage de deux techniques : la spatialisation d'éléments sonores routés plus ou moins intensément dans les canaux Ls et Rs ; la spatialisation à l'arrière d'une réverbération (associée à des éléments sonores plutôt frontaux). Cette dernière technique est précisément celle que nous avons étudiée dans les protocoles rapportés dans la suite de ce document.

On se rend alors compte des limites de la spatialisation à l'aide de l'unique panner

5.1. Utiliser cet outil (et uniquement cet outil) peut difficilement amener à une sensation d'enveloppement globale, même si l'effet reste satisfaisant au sweet spot. La source, par effet de précedence, est localisée là où elle est diffusée en premier : si le spectateur est placé au fond de la salle, il y a de fortes probabilités pour que les haut-parleurs de surround soient perçus en premier. L'enveloppement s'en retrouve complètement faussé. En WFS, cette sensation semble être préservée quelque soit la position d'écoute dans la salle, par ce simple remplacement des pôles du groupe, par leur simple écartement. Bien évidemment, s'approcher brutalement des haut-parleurs amènera à percevoir la contribution de chacun d'entre eux à la synthèse du front d'onde. Cela dit, rares sont les occasions où le spectateur pourra autant s'approcher des transducteurs électro-acoustiques en salle de cinéma.



Fig. 2.17 : En restitution discrète, une ligne horizontale décrit la zone d'écoute où une bonne balance avant/arrière est perçue. En arrière de cette zone, les canaux perçus sont perçus plus intensément.

Protocole 3 : groupe « musique » placé dans l'espace WFS ; source discrète stéréophonique (surrounds actifs via réverbération)

L'idée pour ce troisième test mis en place est de faire intervenir une technique fréquemment utilisée : la source stéréophonique est placée dans les canaux frontaux (L & R), une réverbération stéréophonique associée est quant à elle placée dans les canaux surrounds afin de créer un enveloppement a priori efficace. L'enveloppement recherché est ajusté de telle sorte à avoir une sensation équivalente au sweet spot entre les deux systèmes, comme lors des tests précédents. L'effet de réverbération choisi rend une pièce de moyen volume, une « Small Hall », issue du plugin natif de Nuendo, *Reverence* (cf. Fig. 2.18 suivante). Comme évoqué précédemment, cette réverbération est uniquement routée à l'arrière du panner 5.1. Afin de s'écarter du réflexe d'attention, nous avons également filtré l'effet, de sorte d'atténuer les fréquences aigues¹⁴.

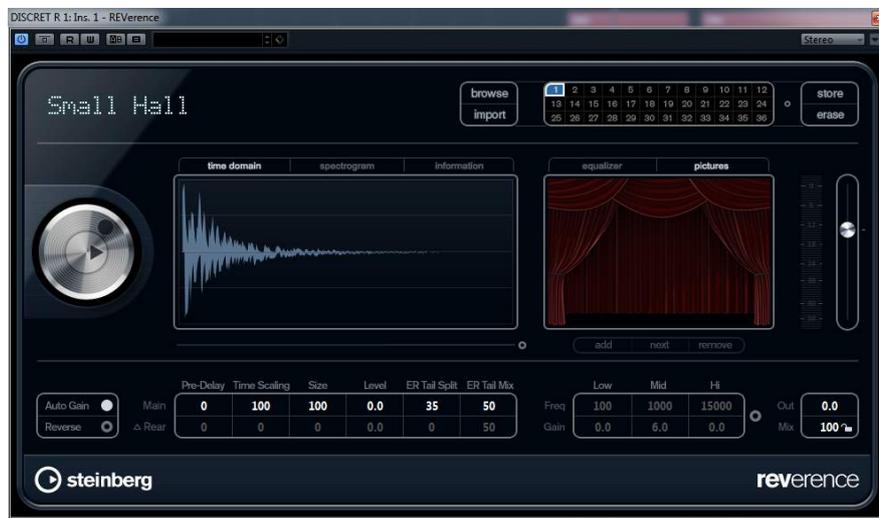


Fig. 2.18 : Fenêtre principale du plugin Reverence. Un preset d'une réverbération de type « Small Hall » est chargé.

¹⁴ Dans *Goodbye, Lenin !* par exemple (cf. Bibliographie), une écoute des canaux arrière sur la scène d'introduction peut nous faire percevoir la même musique qu'à l'avant, soumise à une réverbération. Cette technique est aujourd'hui très utilisée pour « adapter » une source musicale à un système 5.1.

Au sweet spot, la sensation d'enveloppement est convaincante, agréable. La perception des différents extraits musicaux nous semble élargie. La WFS semble toujours un peu plus enveloppante, bien qu'au sweet spot nous avons tendance à préférer le dispositif discret : le contenu fréquentiel de la source originale en ressort plus respecté. En se déplaçant vers l'arrière de la salle, l'enveloppement sur le système 5.1 persiste. Cela dit, du fait de l'éloignement des haut-parleurs frontaux, nous avons pu ressentir une translation du champ sonore vers l'avant, chose qui n'est pas perçue en WFS : grâce à la contribution des haut-parleurs latéraux, la scène sonore frontale semble être perçue de la même manière que pour une position d'écoute au sweet spot.

Ainsi, la technique de la réverbération nous semble très efficace pour le traitement de la musique off sur le système 5.1. Les canaux de surround sont utilisés pour ce qu'ils sont a priori : entourer, envelopper. Sur l'extrait *Time* notamment, l'enveloppement nous paraissait très intéressant du fait de son contenu pauvre en transitoires. En WFS, l'enveloppement nous semble beaucoup plus efficace, du fait de la conservation de « présence » sur les canaux frontaux : les haut-parleurs semblent être placés au même endroit, bien que pour une position d'écoute en fond de salle, la distance nous séparant d'eux augmente réellement.

Protocole 4 : groupe « musique » placé dans l'espace WFS ; source discrète stéréophonique (surrounds actifs via « upmix »)

Outre la technique de réverbération surround, les plugins « d'upmix » sont souvent utilisés par les mixeurs pour rendre multicanale une source stéréophonique. La technique d'upmix consiste à générer des signaux différents à partir du signal stéréophonique frontal, dans le but de nourrir les canaux de surround, ainsi que, potentiellement, le canal central et le canal Lfe. En d'autres termes, ce type de plugin agit de manière quasi-similaire à ce que nous avons présenté sur la réverbération dans les paragraphes précédents. D'ailleurs, si nous écoutons en solo les canaux surrounds générés par upmix, on peut percevoir un rendu très similaire à de la réverbération de petit volume. La principale démarcation de ces softwares réside dans leur simplicité d'interface, permettant d'obtenir des résultats perceptifs satisfaisants avec peu de paramétrage en amont.

Parmi les plugins d'upmix utilisés, nous avons choisi *Anymix* du développeur *Iosono*, ce dernier étant directement inclus dans la version de Nuendo (v 6.5). Nous avons choisi le preset « Movie Stem Basic Neutral » proposé par le plugin, en se basant sur le rapport frontal/surround a priori équilibré. Par suite, nous avons ajusté les paramètres du plugin toujours dans le but d'obtenir, au sweet spot, une largeur apparente équivalente par rapport à la source musicale synthétisée en WFS. Notons que l'un des paramètres de ce plugin, le « highcut », est fixé à 1,5 kHz, ce qui ne semble pas incohérent avec la valeur de filtre que nous avons paramétré lors du protocole précédent (cf. 2.2.3, protocole 3). Le concept d'upmix semble donc prendre en compte l'atténuation des aigus sur le signal envoyé en surround.



Fig. 2.19 : Anymix, l'un des plugins d'upmix utilisés par les mixeurs. Un switch « Enable » permet d'activer l'upmix, ou de le désactiver le cas échéant.

Au sweet spot, nous noterons tout d'abord que les deux systèmes de restitution produisent des résultats perceptifs très similaires, à la différence près qu'une certaine sensation de profondeur est audible en WFS pour les extraits *Time* et *Lacrimosa*. Si l'on s'éloigne dans la salle, en revanche, les différences commencent à se faire ressentir : les éléments percussifs des extraits *Bride of Deluxe* et *Hard Sun* ont tendance à exciter quelque peu l'effet de surround produit par Anymix. Pour des positions ainsi reculées, la WFS propose comme évoqué précédemment, un rendu tout à fait homogène et enveloppant.

Le plugin d'upmix dispose tout de même d'une certaine plus value par rapport aux techniques de réverbérations surrounds : le « plaquage » de l'image sonore frontale vers l'avant de la salle, constaté précédemment en 5.1 pour une position reculée dans la zone d'écoute, n'est plus réellement perceptible avec l'utilisation d'Anymix. Le plugin nous donne en effet accès à un paramétrage frontal, permettant de rendre l'image sonore plus fournie et plus homogène sur l'avant (notamment via le remplissage du canal central). Cet effet nous semble très intéressant pour les stimuli orchestraux, mais peine toujours à totalement nous convaincre pour les deux autres stimuli : la contribution trop marquée des canaux arrières sur les percussions tend à rendre les haut-parleurs un peu trop perceptibles (Exemple : coup de caisse claire sur *Hard Sun*).

Bien que le rendu d'Anymix soit le plus proche de ce que nous obtenons via synthèse de front d'onde, la WFS présente une autre différence perceptive, peut être plus complexe à percevoir dans une petite salle telle que le laboratoire de Sonic Emotion : le champ sonore semble « remplir » la salle. En tant qu'auditeurs, nous avons l'impression d'être baignés dans ce champ, même pour une position d'écoute lointaine à l'arrière de la zone d'écoute. On comprend alors l'intérêt potentiel de la WFS, qui prévoit une contribution des rampes latérales de haut-parleurs, en vue de la restitution d'une source bipolaire large. Via les techniques discrètes de réverbération ou d'upmix, une séparation entre l'image sonore frontale et le surround pouvait être perçue, ce qui ne permet pas, selon nous, un enveloppement aussi satisfaisant qu'en WFS.

Enfin, notons tout de même un effet indésirable amené par une telle configuration WFS : un tel élargissement des pôles de la source musicale nous amène à une distorsion de l'image frontale. En effet, pour des positions d'écoute fortement latéralisées, du fait de la contribution des haut-parleurs latéraux à la synthèse de la source, cette même source, qui était auparavant localisée au centre, est

maintenant localisée sur le côté le plus proche de là où nous nous trouvons. Par exemple, sur *Hard Sun*, pour une position d'écoute à l'extrême gauche de la salle, la voix du chanteur (originellement mixée au centre) est perçue à gauche de l'image sonore frontale (cf. figure Fig. 2.20 suivante). Notre trop grande latéralisation nous a alors amené au phénomène de précedence, impliquant une inversion de la contribution temporelle des haut-parleurs : les haut-parleurs latéraux étaient devenus les premiers à contribuer à la synthèse de la source, et donc logiquement perçus en premier. Notons que sur le dispositif discret, même avec l'usage d'Anymix, cette distorsion inhérente à la stéréophonie est préservée. Simplement, nous aurions aimé l'éviter sur le système WFS. Ce constat doit maintenant être évalué dans une plus grande salle, dans laquelle se rapprocher des rampes latérales de haut-parleurs ne sera pas aussi aisé.

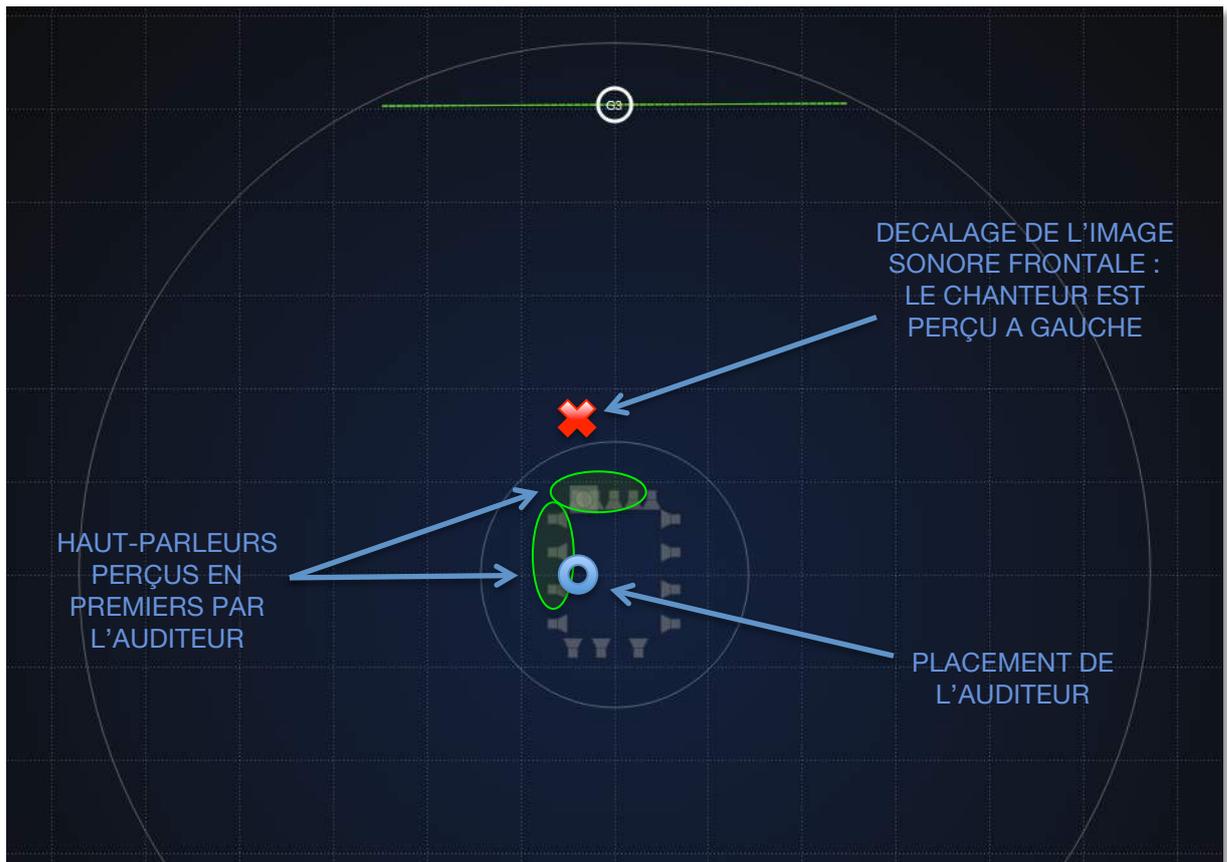


Fig. 2.20 : Illustration de l'effet de précedence malheureusement obtenu en WFS. Pour une position d'écoute très excentrée, la contribution de la rampe latérale gauche nous fait percevoir un décalage de l'image sonore frontale : le chanteur est perçu à gauche.

Analyse globale : la spatialisation de la musique « off »

En conclusion de cette série de tests sur la musique « off », nous pourrions noter la forte dépendance aux stimuli des effets obtenus par synthèse de front d'onde : le contenu fréquentiel et dynamique de certains extraits musicaux (*Bride of Deluxe* et *Hard Sun*) nous a amené à percevoir un certain nombre d'artefacts pouvant se faire plus discrets sur les stimuli orchestraux (*Time* et *Lacrimosa*). Ainsi, sur les stimuli contenant plus de transitoires, nous avons tendance à préférer le dispositif 5.1, l'altération spectrale amenée par la WFS s'exprimant tout de même de manière assez intense.

Dans l'usage des surrounds, la WFS propose un rendu satisfaisant, en ce sens qu'elle permet un enveloppement homogène, et ce quelque soit notre position dans la zone d'écoute. Le dispositif 5.1 (excepté lors de l'unique usage du panner 5.1) permet une spatialisation intéressante de par la mise en action des canaux Ls et Rs, que ce soit via une technique de réverbération, ou via l'usage d'un plugin d'upmix. Cela dit, là où des logiciels supplémentaires sont nécessaires en 5.1 pour disposer d'un effet convaincant, la WFS amène à un enveloppement indépendant de notre position d'écoute, et ce par simple élargissement du groupe bipolaire dans le Performer.

Notons qu'afin d'éviter les distorsions d'image sonore frontale précédemment constatées en WFS (cf. figure Fig. 2.20, p. 136), amenées par la très grande largeur du groupe bipolaire, une telle astuce pourrait être imaginée : lier l'apport des techniques abordées pour le dispositif discret, avec les groupes bipolaires du système WFS. A l'avant se trouverait le groupe bipolaire correspondant à la source musicale, d'une largeur plus restreinte de ce que nous avons proposé dans ce chapitre, tandis qu'à l'arrière serait placé un groupe bipolaire de réverbération, censé permettre une homogénéisation du champ sonore autour de l'auditeur, et

amener à un potentiel enveloppement. Une telle approche reste à être évaluée en conditions de mixage.

2.3. Premières conclusions sur l'utilisation de la WFS au cinéma

Loin d'établir les constats précédents comme étant des lois inébranlables, nous sommes tout à fait conscients que pour avancer un semblant de méthode réutilisable, il nous faudrait dans un premier temps soumettre une série conséquente de tests à un panel de sujets. Cette démarche ne nous offrirait pas nécessairement des éléments de réponse satisfaisants, et ne nous amènerait certainement pas à figer un ensemble de techniques, mouvantes, indéniablement, à chaque nouvelle création. Notons que les techniques actuelles associées aux systèmes standards de type 5.1, sont pour la plupart issues non pas d'études particulières dans le domaine, mais d'une transmission d'un certain savoir faire, acquis d'expérience en expérience, au cours de toute histoire du son au cinéma. Nous invitons par conséquent les utilisateurs et les professionnels à s'approprier l'outil à leur tour, afin d'effectuer les mêmes constats, d'en avancer d'autres, pour que s'édifie une liste de techniques potentielles, capables d'intégrer la WFS dans une chaîne de postproduction au cinéma.

Si cette première série d'expérimentations sur les matières sonores récurrentes au cinéma et leur comportement en WFS ne présente aucune valeur théorique ou scientifique, notre démarche empirique nous a tout de même conduit à émettre plusieurs hypothèses sur l'utilisation de la WFS au cinéma. Elle constitue une approche préliminaire, qui, nous le rappelons, fait pour le moment abstraction du rapport à l'image.

Nos premiers tests sur la voix nous ont permis d'entrevoir à la fois le potentiel et la difficulté du traitement de ce type de source dans un espace WFS. La modification du timbre, en particulier sur les hautes fréquences, est un problème rencontré pour plusieurs placements de source. Les positions à proximité du banc de haut-parleurs frontal semblent particulièrement sensibles à ce problème tout comme les positions éloignées qui donnent également à la voix une largeur importante, difficilement exploitable dans la logique d'un cinéma soucieux d'une certaine forme de réalisme. Les positions intermédiaires semblent d'avantage envisageables puisque plus respectueuses du timbre.

Nos tests sur la mise en perspective de cette voix par sa distance holophonique, associée à des effets de salle, nous ont permis d'évaluer l'apport de la forme du front d'onde sur la création de la perspective. Cette perspective, met en jeu de nouveaux paramètres désormais contrôlables par le mixeur. Il s'agit dès lors de prendre en compte la distance relative entre une source et son effet de réverbération associé, ainsi que la largeur de cet effet. Ces deux paramètres agissent sur l'impression de profondeur et la création d'un espace autour de la voix, derrière la rampe frontale de haut-parleurs. Il nous est difficile d'émettre des conclusions générales tant la variabilité des effets synthétisables est grande, et semble être fonction des types de voix et des réverbérations utilisées, ni même de quantifier l'apport de la structure de l'onde sur la perception de la distance (cf. [Corteel, 2004], p. 234). On constate cependant qu'une impression plus importante de profondeur est permise par la WFS. On remarque par ailleurs une fusion plus marquée (où la sensation "présence + réverbération" est beaucoup moins importante), une distribution de l'énergie spatiale plus homogène, ainsi qu'un respect de l'image frontale sur une plus grande zone d'écoute que sur un système discret (cf. [Berkhout et al., 1997], etc.). L'écoute « exo-centrée », avancée dans la partie 1.4.2, nous semble perceptivement correcte.

Cependant nos hypothèses à ce sujet ne peuvent dépasser le cadre de nos expérimentations limitées en temps et en moyen. D'autres questions naissent naturellement de ces premiers constats. La gestion d'une prise de son direct, composée de la voix et d'un effet de réverbération lui étant propre, laisserait certainement moins de marge de manœuvre qu'un placement holophonique de voix-off. Les problématiques d'équilibre et de traitement de la perche et des HF sont certainement à reconsidérer dans le cadre d'un espace WFS, ces deux microphonies mettant en jeu deux esthétiques de voix différentes. Il se pourrait par exemple que les effets de proximité rencontrés sur les HF demandent un traitement fréquentiel et dynamique d'autant plus particulier si l'on cherche à mettre une voix en profondeur.

Concernant le déplacement des sources, nous avons volontairement choisi de ne pas exploiter d'avantage cet aspect de la spatialisation. Si la WFS semble à même de proposer des déplacements plus cohérents et plus fluides, nous avons pu constater que des modifications importantes du timbre apparaissaient dès lors qu'une voix était déplacée dans l'espace WFS. Pour des raisons maintes fois évoquées dans la littérature (cf. par exemple [Baiblé, 1998]), le placement géométrique des voix en fonction de leur position à l'image se heurte à d'autres problématiques qui nous semblent dépasser le cadre du dispositif sonore, et qui tiennent d'avantage à la nature du son au cinéma. L'exemple du champ/contre-champ fait désormais figure de poncif lorsqu'il s'agit d'illustrer la difficulté de placer de manière cohérente des voix sur des corps, dans un espace découpé, évolutif, et reconstruit par le montage.

Le problème du timbre en fonction de la position de la source évoqué pour les voix s'est confirmé lors de nos tests sur les ambiances frontales. Il existe une zone critique en proximité de la ligne des haut-parleurs, où l'on perçoit une détérioration marquée dans le rendu des hautes fréquences. La perte de ces

informations sur les ambiances peut parfois orienter notre préférence vers le système discret qui, malgré ses défauts, propose un meilleur respect de la composition fréquentielle. D'autre part, le fonctionnement des systèmes WFS, amenant plusieurs enceintes à contribuer au front d'onde, permet un plus grand remplissage de la zone centrale, ce qui conduit parfois à la sensation d'une source moins large, en comparaison au système discret. Ainsi, sur le dispositif stéréophonique, pour une position d'écoute au sweet spot, certaines sources peuvent paraître plus immersives car de largeur apparente plus importante. Cependant, l'homogénéité frontale ainsi que la zone d'écoute élargie nous semblent être des avantages considérables pour la WFS, d'autant que plus la source est éloignée, plus la perception du système de restitution est atténuée, au profit d'une sensation d'espace sonore que l'on pourra qualifier, avec précaution, de réaliste.

Cette première approche du système nous amène à envisager une configuration de mixage particulière, tout en nous permettant de nous projeter dans une utilisation créative de la WFS. Les sensations associées aux voix semblent conduire à une utilisation prudente (ou du moins raisonnée) de ce type de source. Plusieurs problèmes ont été constatés, et d'autres, liés à l'interaction avec l'image, apparaîtront sans doute. Cependant cette utilisation prudente ne doit pas nous empêcher d'exploiter les potentialités perçues en ce qui concerne la création de perspective sur les voix au cinéma. La distance holophonique peut être considérée comme un paramètre complémentaire aux traitements déjà utilisés (rapport champ direct/champ réverbéré, traitements fréquentiels et dynamiques, niveau), mais devra probablement être utilisée avec prudence, puisque la mise en profondeur d'une source associée à un élément visuel induit des problèmes de distorsion angulaire pour les positions excentrées, comme vus dans [Melchior et al., 2003]. Pour autant, nous pensons que les informations de profondeur contenues dans l'image, alliée aux effets « d'aimantation audiovisuelle » (cf.

[Chion, 2003]), nous laisseront une marge de manœuvre suffisamment importante pour mettre en perspective les voix dans un mixage via la distance holophonique, sans créer un décrochage entre le champ visuel et le champ sonore. Ce raisonnement s'applique également à d'autres éléments ponctuels que les voix, matérialisés à l'image (sons "in"). De fait, une plus grande liberté sera probablement envisageable pour des éléments visuels ayant une présence moins affirmée à l'écran, ou complètement absents du cadre (sons "off").

Dans le plan frontal, la stabilité des ambiances, l'homogénéité, et le remplissage de tout l'espace nous semblent être un atout important en vue d'une création sonore cinématographique. Pour des ambiances dont le contenu est matérialisé à l'image, des précautions sont à prendre dans le cas d'un éloignement trop important de la source pouvant conduire également à des problèmes de distorsion angulaire. Pour des ambiances plus larges, le champ créatif ouvert par la WFS, notamment de par le rendu convaincant de la profondeur, nous semble important puisque l'étagement de ces matières permettrait une discrimination plus grande des différents plans de la scène sonore. La différenciation de ces plans semble en effet provoquer une sensation de relief, restituable de manière différente, probablement moins évidente, avec un système multicanal discret. La plupart de nos écoutes portant sur des ambiances seules, la question de la perception ou non de ce relief au sein d'un mélange de plusieurs sources reste entière.

Les constats effectués sur les avantages des systèmes stéréophoniques par rapport à la WFS ne nous semblent pas négligeables et nous conduisent à envisager une hybridation des dispositifs de restitution, et par conséquent une hybridation des méthodes de mixages. C'est sur les matériaux sonores musicaux que la nécessité de proposer un système hybride nous a semblé la plus forte. Il est très difficile de statuer de manière globale sur la valeur d'un système par rapport à l'autre, les rendus obtenus étant fortement liés au contenu des extraits que nous évaluons,

ainsi qu'aux techniques utilisées pour « élargir » la musique en 5.1. Sur des objets à l'origine stéréophoniques, la contribution des rampes latérales propose, en WFS, un bien meilleur enveloppement pour tous les auditeurs. Le 5.1 amène lui une stabilité fréquentielle qui n'est pas négligeable. Cependant, nous avons bien conscience que la seule évaluation d'objets musicaux stéréophoniques est extrêmement limitante pour le travail en 5.1, mais aussi et surtout en WFS. Travailler avec un objet dont la structure répond à des codes intrinsèquement liés au monde stéréophonique ne permet pas d'évaluer finement l'apport d'un système à synthèse de front d'onde. Par nature, ce dernier demande une autre approche, probablement plus segmentée, avec par exemple une organisation en *stems*, organisation souvent utilisée pour le mixage de contenu musical en 5.1.

Il s'agit dès lors de considérer la WFS comme un système complémentaire aux systèmes discrets, et de proposer une structuration des outils permettant cette utilisation composite, mélangeant des sources synthétisées en WFS, et d'autres reproduites sur le système discret standardisé. On désignera alors au cas par cas les éléments pour lesquels la synthèse de front d'onde nous semble envisageable et intéressante pour la narration, et ceux pour lesquels il sera préférable d'utiliser les canaux discrets.

En mettant ainsi en jeu de nouvelles problématiques autour de l'espace sonore, il est possible d'entrevoir que l'introduction d'une image demandera à penser autrement la scène audiovisuelle et à l'envisager sous le prisme de la WFS. Raphaël Mouterde, dans [Mouterde, 2004], proposait plusieurs hypothèses sur le placement des sources, par un raisonnement géométrique, qui éviterait tout conflit entre l'espace sonore et l'espace visuel. Nous aurons de ce fait l'occasion de vérifier ces hypothèses dans la partie 3.4.

Enfin, notre travail préliminaire nous a permis de constater que le mixage d'un film en WFS demandera sans doute un travail de réflexion mené en amont par le montage son. Si ce travail avait lieu, il irait dans le sens d'une décomposition et d'un découpage encore plus importants qu'à l'heure actuelle, en évitant au maximum l'utilisation d'ambiances globales ou d'éléments généraux comprenant plusieurs natures de son, et mettant déjà en jeu des sensations de relief et de perspective. La construction des sessions de montage actuelles dans l'optique d'un mixage 5.1, comprenant des sons de formats différents destinés à alimenter les 6 canaux, ne nous semble pas nécessairement optimisée pour la WFS. La synthèse de front d'onde définit une autre répartition des objets dans l'espace sonore : la prise de son, et donc le contenu sonore, s'affranchit du contexte de restitution [Berkhout et al., 1993]. Par exemple, nous avons constaté à plusieurs reprises que l'éloignement d'une source monophonique pouvait lui donner une largeur presque comparable à celle d'une source stéréophonique, tout en la répartissant de manière homogène et cohérente sur l'espace frontale.

Pour aller plus loin, il s'agirait de considérer autrement le montage son et surtout les matériaux sonore utilisés par le monteur. Dès lors, si nous envisageons la WFS comme un outil, il faudrait définir une autre chaîne de production, dans laquelle les acteurs de la bande son au cinéma disposeraient d'environnements techniques appropriés. Ce que nous entendons par ici, c'est que l'étape du montage son dans le cadre d'un film mixé en WFS peut difficilement se passer d'une salle équipée en conséquence. Cette étape du montage dépassant le contexte de notre sujet, une étude supplémentaire pourrait permettre d'aborder le type de matériaux sonores nécessaires, ou du moins optimisés, à la création sonore cinématographique en WFS.

3. Mise en pratique de la WFS dans le mixage cinéma

Cette troisième partie vise à confronter les principes de la WFS et ses outils, aux particularités d'un mixage cinéma. Elle sera pour nous l'occasion d'étendre nos expériences pratiques débutées en partie 2 (effectuées sur des matériaux sonores isolés) à des constats relatifs aux mélanges de sources, au mixage à proprement parler. Tout comme la restitution discrète qui, en mettant en jeu différents types de sources simultanées, est soumise à des problèmes fréquentiels (masquage) et spatiaux (cohérence entre les plans sonores), le mixage en WFS nous confrontera sans doute à des problématiques similaires. Nous évoquerons puis nous tenterons de résoudre ces problèmes via des méthodes de traitement qui seront les nôtres. A nouveau, cette partie décrira notre démarche, mais ne formulera en aucun cas des techniques génériques au mixage WFS. Il s'agit surtout de documenter les difficultés techniques et pratiques de la WFS et de proposer une première approche de solutions envisageables. Par ailleurs, cette partie sera surtout pour nous l'occasion de mettre la synthèse de front d'onde au devant de l'image cinématographique. Nous verrons en pratique, comment la relation entre la perspective sonore décrite dans la partie précédente et l'image projetée en deux dimensions s'établit.

Nous terminerons, dans un premier lieu, une description plus détaillée des outils propres à la WFS débutée en 2.1. Nous amènerons ensuite le lecteur à une réflexion relative aux caractéristiques spécifiques, et donc aux contraintes, de la postproduction sonore au cinéma. Nous tenterons ensuite de répondre à ces contraintes, en proposant une configuration de mixage qui sera la nôtre dans la suite de notre étude. Nous relaterons enfin nos expériences en WFS, à travers le mixage de six séquences sélectionnées selon les critères d'étude évoqués en partie 4.

3.1. Les outils développés par Sonic Emotion

La société Sonic Emotion, fondée au début des années 2000 à l'issue des études menées par Etienne Corteel en collaboration avec le projet CARROUSO, développe aujourd'hui le *Sonic Wave 1*, processeur de traitement WFS que nous avons présenté en 2.1.2. Le Wave 1 est au cœur de l'arborescence proposée par Sonic Emotion. C'est son contrôle, via des interfaces *software*, qui permettra la mise en « espace WFS » de sources sonores.

Deux logiciels distincts sont utiles à l'exploitation du Wave 1 :

- Le *Performer*, que nous avons déjà décrit en 2.1.1, utilisé pour placer et visualiser les objets sonores
- Le *Designer*, utilisé pour configurer le système, i.e. au positionnement des haut-parleurs, au routing de sortie, et à la calibration du système.

Le protocole de communication OSC

Le Wave 1 peut être contrôlé par protocole OSC (*Open Sound Control*), un outil de communication communément utilisé et développé au départ pour les applications musicales. Les messages OSC sont basés sur le protocole UDP, et consistent en une chaîne de caractère associée à des arguments. Un "kit de messages" est spécifié pour permettre le contrôle de la position des sources sonores, de leur niveau, de leur égalisation, de leur routing...

En particulier, des messages de "groupes de sources" permettent à l'utilisateur de contrôler efficacement plusieurs sources comme une entité unique, en spécifiant la largeur du groupe ainsi que sa position. Cette technique de groupe est celle que nous avons déjà utilisée en 2.2.2 et 2.2.3, pour le traitement des ambiances stéréophoniques et des musiques. Un groupe peut de base être formaté en « bipole », mais également en 5 pôles suivant une organisation 5.0, etc.

L'utilisation d'un protocole simple tel que l'OSC permet de contrôler le processeur Wave 1 avec des programmes basés sur ce même protocole, en laissant la possibilité à l'utilisateur de développer des interfaces dédiés ou des patches. *TouchOSC* (du développeur Hexler) et *Lemur* (du développeur Liine) sont des exemples d'applications mobiles, alors que des applications de bureau "standalone" telles que *Max/MSP* ou *Pure Data* peuvent être utilisées.

Le protocole OSC permet par ailleurs de ne pas faire transiter de traitement audionumérique entre les machines. Les instructions, sous forme de texte, sont soumises au système d'exploitation du Wave 1, qui se charge seulement ensuite de réaliser les traitements propres à la synthèse de front d'onde, grâce aux *DSP* internes. Cette structure de communication simplifie grandement la tâche des machines, et permet notamment de minimiser la latence entre les interfaces, qui ne se chargent plus qu'à transmettre des messages textuels.

Le Designer : configuration et calibration du système

Le Designer présente deux modes :

- Une mode « offline » utilisé pour les placement des haut-parleurs et le routing de sortie ;
- Un mode « online » utilisé pour la calibration du système

Configuration « offline »

Dans le mode « offline », l'utilisateur crée une configuration de haut-parleurs en les plaçant dans un espace virtuel 3D. Les enceintes peuvent être créées individuellement, ou être générées par la création d'une ligne ou d'un cercle de haut-parleurs. Cette configuration basique étant créée, il est alors possible de réaliser des ajustements de position afin de s'adapter au mieux aux contraintes de l'environnement (portes, escaliers, autres équipements). Notons que selon Sonic Emotion, la synthèse de front d'onde via le Wave 1 ne nécessite pas une disposition régulière des haut-parleurs.

Les haut-parleurs peuvent être organisés en groupe, appelés "sous-systèmes". Chaque sous-système représente une section spécifique de l'installation. Par exemple, un sous-système pourrait représenter la façade (haut-parleurs situés derrière l'écran), un autre, les haut-parleurs latéraux, un troisième les haut-parleurs arrières, tandis qu'un dernier constituerait le *subwoofer* (cf. 3.3.1).

Chaque sous-système est ensuite associé à un algorithme de rendu spatial : *WFS main*, ou *WFS support*. Tous les sous-systèmes en mode *WFS main* seront groupés pour former le système de rendu principal, qui peut être soit totalement périphérique, soit limité à la façade. Les sous-systèmes en mode *WFS support* seront considérés comme des lignes complémentaires de haut-parleurs qui seront utilisées pour synthétiser des positions de source plus éloignées, en venant contribuer au front d'onde créé par le ou les sous-système(s) dit(s) "main".

Par suite, les haut-parleurs virtuels nouvellement placés sont assignés à des sorties physiques dans une matrice de routing. Pour les haut-parleurs "multi-voies", chaque entrée de ces derniers est affectée à une bande de fréquence particulière : basses, mediums, ou aigus.

Une fois la configuration réalisée, elle est chargée dans le processeur via la fonction *upload*, celle-ci reliant les informations aux sorties physiques. Des "filtres WFS" sont alors spécifiés dans le but de réduire le travail en temps réel du processeur. Enfin, en amont de l'*upload*, le Designer lance une procédure de vérification visant à valider la configuration et à indiquer à l'utilisateur les potentielles erreurs (câblage absent, liaison défectueuse, ...).

Calibration "online"

Une interface de test offre un outil pour tester les haut-parleurs individuellement (ou le routing effectué) en sélectionnant une entrée donnée du système et en la dirigeant directement vers une sortie/un haut-parleur donné. C'est précisément cette fonction que nous avons utilisée lors de notre calibration décrite en 3.3.3.

Le mode "online" du Wave Designer comporte un filtre paramétrique complet pour chaque haut-parleur, incluant filtres *highshelf/lowshelf* et *highpass/lowpass* (avec jusqu'à 48dB/oct), et 8 filtres de type *bell*. De plus, chaque égaliseur paramétrique peut être ajusté selon 3 modes :

- Individuellement par haut-parleur ;
- Semi-globalement par sous-système (ligne frontale de haut-parleurs, haut-parleurs de surround, ...) ;
- Globalement pour tous les haut-parleurs.

Cette égalisation peut être effectuée manuellement, ou en chargeant des presets (fournis par les fabricants d'enceinte, ou enregistrés préalablement, par exemple).

Notons que la WFS par Sonic Emotion requiert une puissance acoustique équivalente pour tous les haut-parleurs d'un même sous-système, que celui-ci soit associé à un algorithme de rendu de type *main*, ou de type *support*.

Des ajustements de niveau peuvent être effectués en mesurant le niveau sonore pour chaque enceinte, en plaçant le dispositif de mesure à 1 mètre. Cela dit, la plupart des problèmes devraient être atténués en utilisant des haut-parleurs et des amplificateurs de même puissance dans chacun des sous-systèmes.

Une gestion des haut-parleurs multi-voies est rendue possible et permet le contrôle de la fréquence de coupure. Des fonctions de niveau, de delay, d'inversion de phase, et de compression basique sont également mises à disposition pour chaque sortie physique du système. Ces réglages peuvent être sauvegardés avec les presets d'égalisation évoqués ci-dessus.

Des outils additionnels permettent un paramétrage de niveau et de délai pour chaque sous-système, en permettant par exemple une balance entre la ligne principale WFS ("main") et la ligne de support ("support").

La configuration complète (positionnement des haut-parleurs, routing, égalisation, alignement du niveau & délai, ...) peut être sauvegardée comme un preset global et être rappelé plus tard, sur une différente machine par exemple.

Gestion de l'automation

L'*automation*, ou « automatisation », est un procédé permettant de sauvegarder l'évolution temporelle d'un paramètre (ou de plusieurs paramètres) sur une échelle temporelle, communément appelée *timeline*. La plupart des

logiciels audio, "séquenceurs", incluent cette fonctionnalité depuis plusieurs années, offrant ainsi la possibilité à l'utilisateur d'écrire les valeurs d'un paramètre associé à une piste audio, puis de les lire a posteriori : le paramètre concerné suit alors les variations préalablement écrites. C'est également ce procédé qui est à la base des mixages audionumériques réalisés aujourd'hui dans l'industrie cinématographique.

A l'heure actuelle, le Performer n'inclut pas de gestion de l'automation à proprement parler. En effet, il faudrait pour cela que le processeur Wave 1 embarque une interface de synchronisation, de type MTC ou LTC, tels que le font les séquenceurs audio à l'heure actuelle¹. Pour palier à cela, une solution de type plugin développée par l'IRCAM est mise à disposition des ingénieurs du son et offre un lien simple entre le Performer, le Wave 1, et l'automation de n'importe quel séquenceur audio². Ce plugin, du nom de *Tosca*, est représenté par l'interface suivante :

¹ MTC (Midi Timecode) ou LTC (Longitudinal Time Code) sont des protocoles différents de timecode, permettant tous deux la synchronisation de différentes interfaces (audio ou vidéo), via l'envoi de courts messages numériques à intervalle régulier.

² Tosca est disponible aux formats AAX (Avid Pro Tools), VST (Nuendo, Pyramix, ...), ou AU (Logic Pro).



Fig. 3.1 : Interface du plugin Tosca. « ID #7 » fait référence à la source n°7 dans le Performer.

Le fonctionnement du plugin est basé sur le protocole OSC décrit précédemment. Un paramètre du Performer est associé à une adresse OSC contenant des arguments (= des valeurs). Dans notre cas, c'est la position de la source que nous cherchons à automatiser : elle est représentée dans le plan polaire par un azimut et un rayon.

```
<parameter index="1" name="Set/SrcPos/A" min="-180" max="180"
scaling="linear"/>
```

Sur ce message OSC décrivant la position de la source (SrcPos), « SrcPos/A/1 » est l'adresse OSC représentant l'azimut (A) de la source, pour la première instance (1) du plugin Tosca. Cela signifie qu'avec cette adresse, on sera capable de contrôler la source n°1 affichée dans le Performer. Le plugin Tosca se chargera, quant à lui,

de transcrire les paramètres OSC lui arrivant, en valeurs d'automation comprises par le séquenceur utilisé. Par ailleurs, « 180 » ou « -180 » sont des exemples d'argument, i.e. d'azimut (en degré), pour cette adresse OSC.

Notons qu'un fichier texte au format .xml devra être joint avec Tosca, ou être sauvegardé à tout autre emplacement qui permettra de le recharger facilement depuis le plugin (bouton "Load"). Le contenu de ce fichier (fait d'une unique ou de plusieurs adresses) décrit simplement les paramètres du Wave 1 qui seront rendus accessibles dans le séquenceur audio. Lors de nos expériences, nous appellerons ce fichier "Adresses".

Une configuration du plugin est nécessaire avant l'utilisation de celui-ci en automation :

- "input port" est le port d'entrée du plugin. Lorsque la position d'une source est modifiée dans le Performer, c'est par ce port qu'arrivent les adresses OSC modifiées. Elles sont ensuite transcrites en données d'automation par le Tosca. Par défaut, on pourra utiliser le port 5555. Si celui-ci est occupé, on tâchera de le modifier dans le Performer, menu « Network »³.
- "output port" est le port de sortie du plugin. C'est l'interface entre le séquenceur et le processeur Wave 1. Lorsqu'une valeur d'automation est lue dans le séquenceur, elle est transcrite en paramètre OSC par Tosca, puis envoyé au Wave 1 via liaison ethernet. Le numéro de ce port est fixé à 4000 : c'est le port fixe du Wave 1.
- "output IP" correspond à l'adresse IP du Wave 1.

³ Rappelons qu'un port réseau ne peut être ouvert qu'une seule fois.

- "enable" permet d'activer la connexion lorsque la configuration est terminée
- Enfin, lors du chargement du plugin sur plusieurs pistes, il faudra parfois recharger le fichier .xml "Adresses" décrit précédemment.

Physiquement parlant, c'est par liaison ethernet que transitent les informations entre le processeur Wave 1, le Performer, et l'ordinateur hébergeant le séquenceur audio (et donc le Tosca). Sonic Emotion a pour le moment opté pour une structure réseau triangulaire :

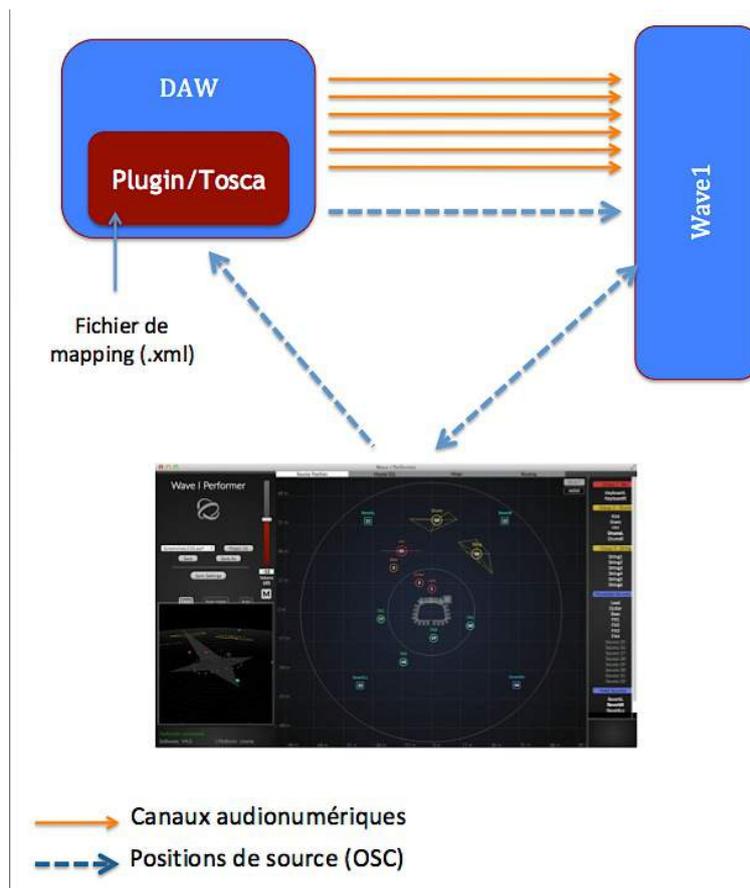


Fig. 3.2 : Intégration réseau des interfaces pour l'exploitation du Wave 1.

Dans notre configuration, le plugin ToscaA communique en local avec le Performer par la carte réseau interne. Notons qu'il est tout à fait envisageable d'avoir le séquenceur (muni du ToscaA) sur une machine, et le Performer sur une autre.

Pour résumer, avec ToscaA inséré sur une piste du séquenceur, les paramètres du Performer deviennent automatisables depuis ce même séquenceur, le plugin jouant le rôle d'interface ethernet entre les deux machines. Graphiquement, on pourra accéder aux courbes décrivant l'évolution de ces paramètres dans la majorité des séquenceurs audio. Enfin, rappelons que dans notre cas, seuls les paramètres d'azimut et de rayon seront nécessaires pour rendre automatisable la position d'une source dans le Performer.

3.2. Particularités techniques de la post production cinématographique

Envisager l'ensemble de la question de la WFS au cinéma impose de s'attacher d'une part à l'étude de son fonctionnement dans une approche créative et perceptive et, d'autre part, à son adaptation dans un environnement technique et industriel parfois complexe, soumis à certaines normes et certains critères de mise en œuvre qui répondent eux-mêmes à une logique d'efficacité productive. Dans notre démarche, nous nous sommes questionnés sur la compatibilité des différents outils de la WFS avec le cinéma, sachant que les dispositifs pour le moment dédiés à la WFS sont plutôt orientés vers le live et la sonorisation, domaines qui répondent à d'autres problématiques de mise en œuvre. Les problèmes soulevés dans cette sous-partie trouveront une forme de réponse concrète dans la partie 3.3 présentant la configuration de mixage que nous proposons. Ces réponses prennent en compte les impératifs techniques liés au mixage car il nous semble important de maintenir, dans cet environnement de travail, une partie des outils d'expression dores et déjà à disposition du mixeur.

3.2.1. Une configuration hybride et modulable

Un des premiers aspects de la postproduction audiovisuelle concerne le nombre élevé de pistes en provenance du montage son. Ces pistes contiennent des matériaux sonores, d'origines et de fonctions différentes. Des ambiances aux effets, en passant par les directs, la musique, le bruitage, les post-synchro, les pistes de VI (version internationale), le mixeur gère un panel considérable de sources aux formats différents qui demandent des traitements particuliers et spécifiques. Ces différents éléments sont généralement organisés par stems, que le mixeur constitue en routant ses tranches de console vers des bus de sortie.

Les processeurs WFS sont aujourd'hui limités en nombre d'objets synthétisables (32 ou 64). Il ne semble donc pas envisageable de reproduire la totalité des éléments à mixer en WFS. Face à la complexité grandissante des montages son, même 64 canaux/objets ne suffiraient pas. De plus, nous avons constaté que certaines sources sonores se comportaient mieux sur une diffusion discrète, en terme de timbre et de largeur notamment. Il est donc important de laisser au mixeur la possibilité de pouvoir choisir sur quel système il est préférable de reproduire telle ou telle source, ce choix n'étant pas uniquement guidé par un impératif technique car il constitue également une décision esthétique.

Plutôt que de mettre en place deux systèmes différents cohabitant au sein d'une même installation, il nous paraît plus judicieux de proposer un seul système hybride, qui autorise à la fois une diffusion 5.1 et WFS sur la même configuration d'enceintes. Il faut donc mettre en place une solution de routing qui permette d'accéder aux envois vers des canaux discrets ou vers des objets WFS, et que ce

choix ne soit en rien définitif. Les pistes du montage son arrivant séparément sur des tranches de la console, il est tout à fait envisageable que le mixeur décide d'envoyer vers le 5.1 un élément du montage placé sur une piste et que plus tard, pour un autre élément placé sur cette même piste, la WFS soit préférée.

On voit donc ici se dégager un premier impératif matériel, celui de l'utilisation de la console comme outil de routing⁴. Cet outil est de toute manière un élément qu'il faut impérativement considérer, puisqu'il est véritablement le point névralgique d'un auditorium de mixage et qu'il est l'instrument de travail du mixeur, son principal moyen d'expression. Cette expression se caractérise en partie par l'équilibre ou la balance effectués mais aussi par la fluctuation de cet équilibre via la modification dynamique des différents paramètres utilisables au sein d'une console (niveau, correction fréquentielle, dynamique, envoi vers des réverbérations, panoramiques). Un mixage se construit par de multiples *passes*, en modifiant le rapport des sons entre eux, en proposant une évolution de ces rapports, des textures, des positions des sources dans l'espace. Nous abordons ici la question de l'*automation* et de la modification, au cours du temps, des paramètres évoqués, mais aussi de l'enregistrement et de la relecture de ces paramètres. Cette question ne se pose quasiment pas sur des configurations dédiées au spectacle vivant. Elle est cependant centrale au cinéma, et également d'actualité en mixage musique. L'utilisation de la console permet de régler la question pour l'automation des égaliseurs, filtres, compresseurs, envois auxiliaires. Cependant, concernant la spatialisation en WFS des différents objets, le problème persiste, puisque le positionnement des sources ne répond plus aux lois de panoramiques implantés dans les consoles.

⁴ Ce routing évolutif nous semble incontournable dans le cas d'un mixage s'appuyant sur un montage son pensé en 5.1. Si l'édition des sons était envisagée pour la WFS, avec par exemple des pistes dédiées et directement routées vers des canaux objet, cet impératif serait à reconsidérer.

A l'heure actuelle, les consoles utilisées en mixage cinéma ne prennent pas en compte de protocole d'échange avec les processeurs WFS, protocole qui pourrait centraliser cette opération d'écriture directement dans la console. Cette spatialisation a donc lieu en aval, via des logiciels dédiés et développés pour la WFS. La position de ces objets doit pourtant pouvoir évoluer progressivement au cours d'une séquence, ou être placée à un endroit différent plus tard dans le film. Les outils de la WFS, à l'heure actuelle, n'intègrent pas de référence temporelle (*timecode*), ce qui élimine d'emblée toute possibilité d'écriture d'automation au sein de ces outils. Il faut donc trouver un support d'écriture pour ces modifications. Nous pensons qu'il est nécessaire d'envisager cette écriture au sein d'un logiciel familier de la postproduction cinéma, et ce pour une raison essentielle. Les conditions de travail et de production imposent aujourd'hui des délais drastiques à tout les corps du métier. Il est donc de plus en plus courant de débiter un mixage alors que le montage image n'a toujours pas de caractère définitif. Dans le cas d'un changement de ce dernier, il faut alors procéder à une reconfiguration du montage son, étape qui consiste à adapter la version précédente à la nouvelle par déplacements, coupes, suppressions, ajouts d'éléments. Pour le mixage, il faut dans ce cas replacer au bon endroit tous les paramètres déjà écrits dans l'automation de la console. Ce travail demande une certaine rigueur et peut parfois être fastidieux. Aujourd'hui plusieurs logiciels ont automatisé l'opération et agissent directement sur les *timelines* des sessions rendant la tâche moins pénible. L'écriture de l'automation de spatialisation des objets WFS doit donc être réalisée sur un support permettant cette étape de reconfiguration, qu'elle soit manuelle ou automatique.

3.2.2. Un système électro-acoustique adapté

Cet environnement technique va de paire avec une configuration électro-acoustique qui a également certaines particularités. La première d'entre elles concerne la coloration des enceintes de cinéma utilisées derrière l'écran. Ces enceintes frontales utilisent un haut-parleur classique pour la reproduction des fréquences graves et une chambre de compression pour les aigus. La fréquence de raccordement entre ces deux modes de transduction se situe généralement autour de 1kHz, quoique que cette valeur puisse varier de manière importante d'un modèle d'enceinte à un autre. La zone d'écoute étant très large au cinéma, les chambres de compression permettent d'améliorer le rendement du moteur et la reproduction des fréquences élevées dans toute la salle. L'utilisation de cette technique n'est pas neutre, puisqu'elle est source de distorsions harmoniques, qui donnent entre autres cette couleur particulière au « son cinéma ». De plus, la présence du pavillon induit un phénomène de directivité de la diffusion sonore qui dépend de ses caractéristiques propres. Ces enceintes de cinéma ont également la capacité de descendre facilement dans le bas du spectre et fournissent des graves généreux et précis. Les enceintes de sonorisation descendent généralement moins bas et ce manque de grave est compensé par du *bass management* en envoyant également le signal dans un ou plusieurs *subwoofers*. L'utilisation du caisson de grave au cinéma est tout autre, il constitue un canal à part entière (le Lfe, pour *Low Frequency Effects*), alimenté par des signaux dédiés (même s'il reste possible d'envoyer dans le Lfe un pourcentage du signal alimentant un ou plusieurs canaux du 5.0) [Dolby, 2000]. En d'autres termes, ce canal est d'avantage considéré comme un effet de grave voir d'infra-grave - ajouté à des enceintes qui en contiennent déjà beaucoup – plutôt qu'un renfort de basse pour palier à un manque du système principal.

On voit donc ici que le cinéma à tout intérêt à envisager un dispositif d'enceintes disposant de bonnes performances dans le bas du spectre, cette zone fréquentielle étant fortement travaillée au mixage. Au delà de l'apport « spectaculaire » sur certains effets, cette zone permet de donner du corps et du poids aux sources sonores qui le demandent, narrativement parlant, et nous considérons qu'elle fait partie des moyens d'expression d'un mixeur dans son accompagnement et son approche du film. Si nous apportons ces précisions techniques, c'est que la question du type de haut-parleurs pour la WFS au cinéma se pose. Pour les raisons évoquées ci-dessus, nous pensons que le bas du spectre est une zone trop importante pour être renvoyée littéralement et techniquement dans un *subwoofer* via *bass management*. Concernant la coloration des pavillons on peut, dans un premier temps, considérer que cette couleur fait partie intégrante du médium cinéma : on irait donc plutôt vers l'utilisation de ce type haut-parleur (en plus grand nombre) pour la WFS. Cependant plusieurs points de vue pourraient s'opposer à cette vision « conservatrice ». Tout d'abord, nous l'avons dit, la directivité donnée par les pavillons, ne semble pas tout à fait convenir avec les principes physiques de la WFS et les préconisations des spécialistes, qui encouragent plutôt l'utilisation d'enceintes avec une directivité douce et progressive (en théorie omnidirectionnelle, cf. 1.4). De plus, en supposant que la WFS permette d'établir une autre habitude d'écoute dans le cinéma, le maintien de cette couleur ne semble pas si primordiale dans le cadre de notre étude. Enfin, le coût de ce type de haut-parleurs est considérable et leur taille ne se prête pas, à première vue, à une implémentation aisée. Cela peut en effet rendre prohibitive l'installation de la WFS dans les salles obscures.

3.3. Proposition de configuration pour un mixage cinéma

L'installation que nous proposons ici tente d'apporter une réponse pratique aux conditions techniques établies en 3.2. C'est un premier pas vers ce que pourrait être une configuration de mixage cinéma en WFS, prenant en compte diverses contraintes matérielles et budgétaires.

Cette configuration à été mise en place pendant deux semaines à l'auditorium de L'ENS Louis-Lumière. Elle nous a permis, dans un premier temps, de concrétiser nos idées concernant la mise en pratique technique du système. Nous avons pu, dans un second temps, mixer et constituer un corpus en vue des tests perceptifs. Au cours de ce travail, nous avons pu d'une part vérifier la validité de notre configuration ainsi mise à l'épreuve de la pratique, puis évaluer la pertinence de nos observations détaillées en partie 2, dès lors qu'une composante image s'ajoute au dispositif.

Synoptique de l'installation (page suivante)

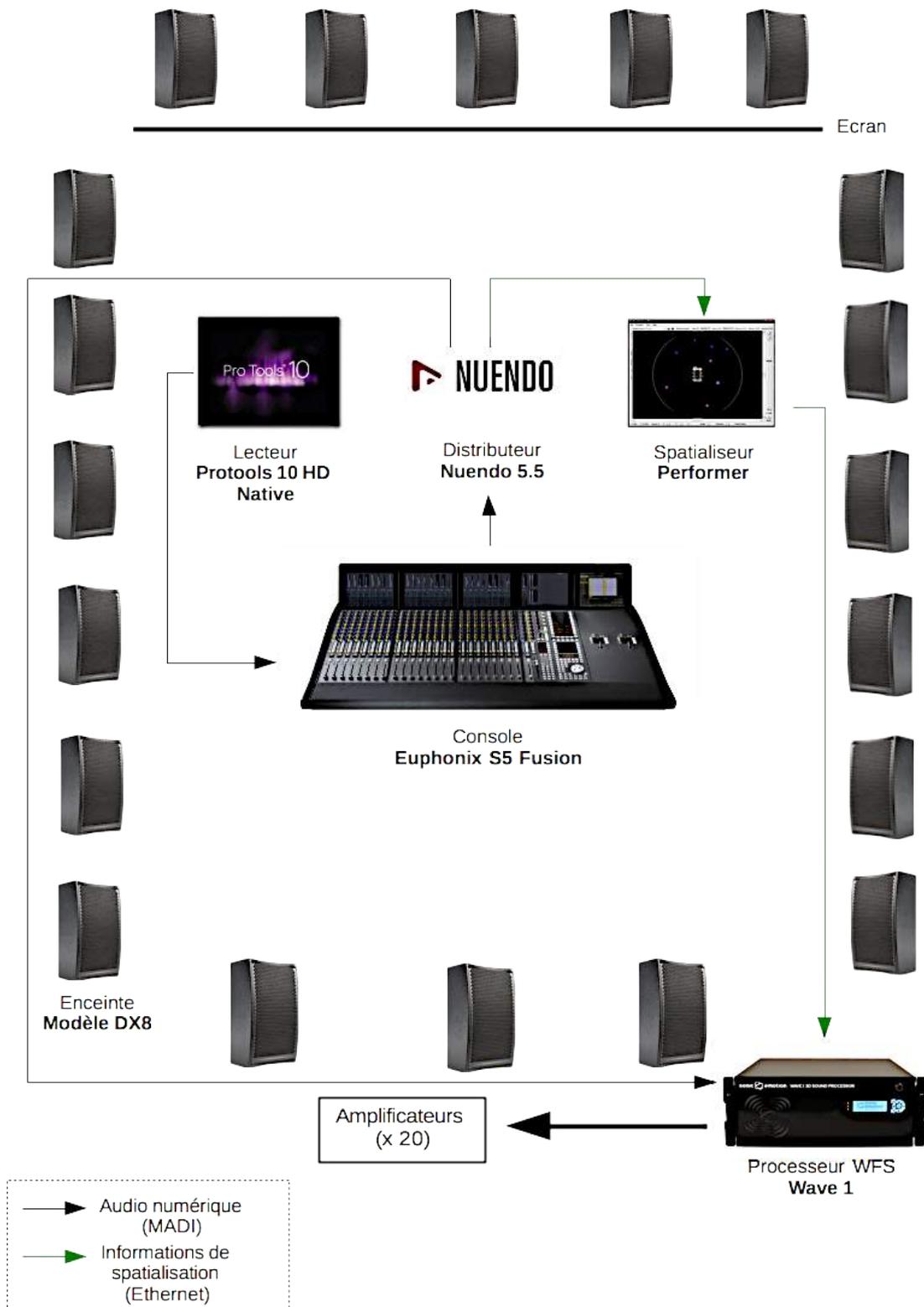


Fig. 3.3 : Synoptique de l'installation faite dans l'auditorium de l'ENS Louis-Lumière.

3.3.1. Choix et disposition des enceintes

Nous disposons, pour notre partie pratique, de vingt enceintes DX8 de la marque APG. Ces enceintes de sonorisation nous semblent bien adaptées à la WFS puisque leur conception et leur technologie coaxiale permettent une directivité et une diffusion plus douce sur l'ensemble du spectre (cf. Annexe 3.1). Cependant, le lecteur pourra trouver ce choix contradictoire avec certains éléments évoqués en 3.2.2 sur la nécessité de maintenir un dispositif électroacoustique spectralement proche de celui en place dans les cinémas. Ce choix s'est fait principalement pour deux raisons pratiques qui renvoient à d'autres remarques effectuées à ce sujet, toujours en 3.2.2. La première tient à l'opportunité que nous avons de disposer de ce matériel sur une telle durée. Il est plus difficile financièrement parlant, de louer une quantité identique d'enceintes conçues pour le cinéma. La seconde nous confronte à la corpulence de ces enceintes et à leurs conditions de mise en place : pas de fixation sur pied, masse importante, réglage du cross-over à prévoir.

Ce choix d'enceinte étant effectué, nous devons penser leur disposition dans l'espace. Une première approche consiste à envisager une répartition régulière, par un espacement identique entre chaque centre acoustique des haut-parleurs, dans l'ensemble de la pièce. Dans notre cas, au vu du nombre d'enceintes disponibles et de la surface à couvrir, ce raisonnement aurait conduit à un espacement trop important entre les haut-parleurs. De plus, Baiblé nous invite à remarquer une « *inégaie répartition technique des secteurs auditifs par rapport à l'écran* » [Baiblé, 1998] dans les systèmes multicanaux actuels, indiquant que les quatre zones (avant, arrière, gauche et droite) ne proposent pas un enveloppement homogène du spectateur et sont surtout incohérentes avec le rapport que nous entretenons à

l'image. La « *sectorisation d'une salle de cinéma* » [Bailblé, 1998] appelle à une certaine polarisation vers l'avant (quoique se soit déjà un peu le cas avec le LCR « classique ») et donc à un renfort de la zone frontale. Partant de cette remarque et de nos expériences, nous décidons de densifier le maillage de haut-parleurs derrière l'écran tout en restreignant le nombre d'enceintes arrières (celle réellement situées le long du mur du fond) afin de disposer d'un nombre suffisant de DX8 pour proposer un écartement raisonnable sur les cotés. Favoriser le nombre d'enceintes à l'avant permet au processeur de préciser la forme des fronts d'ondes et ainsi d'apporter une plus grande finesse de localisation dans cette zone, lieu de focalisation de l'attention du spectateur. Nous choisissons donc de moins favoriser la couverture de la zone arrière, qui nécessite selon nous, une moins grande précision.

Ces choix nous conduisent à la disposition suivante, dans un auditorium de dix mètres de long sur sept mètres de large :

- cinq enceintes derrière l'écran ;
- six enceintes sur chaque coté ;
- trois enceintes à l'arrière.

Cette configuration donne un espacement de 110cm entre les cinq enceintes frontales, de 160cm entre les six enceintes latérales, et de 180cm entre les trois enceintes arrières. Nous faisons le choix de placer les enceintes frontales uniquement derrière l'écran de quatre mètres de base, et de ne pas les faire dépasser de part et d'autre. L'hybridation du système envisagée en 3.2.1 nous impose de disposer de sources de diffusion placées de manière identique au 5.1, en LCR. Ainsi, nous préférons densifier notre maillage frontal derrière l'écran, sans couvrir toute la longueur du mur. Acoustiquement, le placement d'enceintes proches d'un angle conduit à plusieurs réflexions rapprochées pouvant induire du

filtrage en peigne, ceci perturbant le fonctionnement du système. Ainsi, pour ne pas trop « exciter » les angles de la salle, et pour compenser le manque de couverture dans cette zone, les premières enceintes latérales gauche et droite sont tournées à 45° vers la zone d'écoute, assurant une meilleure homogénéité et une continuité entre la façade et les cotés.

Les enceintes sont disposées sur des pieds, le centre acoustique est placé à un mètre quarante de hauteur, soit légèrement plus haut que la hauteur des oreilles d'un spectateur assis. Pour compenser le manque d'absorption acoustique à l'avant de l'auditorium (bois lisse et peint), ce qui pouvait causer des phénomènes de *flutter echo*, conduisant potentiellement à des associations d'ondes mal maîtrisées, nous avons décidé d'ajouter des panneaux absorbants à notre dispositif. Outre le fait que nous ayons perceptivement résolu le problème du *flutter echo*, nous pensons que le comportement en hautes fréquences de notre installation est dès lors plus adapté à ce que préconise la théorie de la WFS, i.e. une salle de diffusion la plus « mat » possible (cf. 1.4.2).

Un seul des deux *subwoofers* Meyer est utilisé pour le canal Lfe, mais aussi pour un renfort de grave du système principal, les modèles DX8 ne permettant pas selon nous de descendre assez bas en fréquence pour du contenu cinématographique. Ce *subwoofer* est placé derrière l'écran, dans le mur THX accueillant le système permanent de l'école.

L'ensemble des paramètres de cette configuration est reporté et sauvegardé sous forme de preset dans le *Designer*. A ce stade, nous n'avons utilisé que le mode « offline » du *Designer* (cf. 3.1). Nous créons alors quatre sous-systèmes. Le premier correspond aux enceintes frontales, le second aux latérales, le troisième aux arrières, tandis que le quatrième est dédié au *subwoofer*. Cette configuration est chargée dans le processeur Wave 1, via la fonction *upload*, qui permet au

processeur de « prendre conscience » du nombre et de la position des haut-parleurs dont il dispose.

Cette disposition nous place dans des conditions de travail qui, sans être optimales (du fait d'un écartement inter-enceintes non négligeable), sont tout à fait acceptables. En effet, selon Etienne Corteel, nous nous situons dans des espacements suffisamment bien gérés par le Wave 1 pour construire des fronts d'onde corrects.



Fig. 3.4 : Vue de l'auditorium. L'écart entre les enceintes latérales est de 160 cm.



Fig. 3.5 : Vue de l'auditorium. Entre les enceintes latérales sont placés des panneaux absorbants.

3.3.2. Organisation de session et routing

Nous utilisons une session contenant tous les montages son des séquences que nous allons mixer, lue dans un Pro Tools HD Native. La session contient également deux pistes auxiliaires pour les effets de réverbération. Nous utilisons deux instances du plugin *Altiverb 7*, l'un mono vers stéréo, l'autre mono vers quadriphonique. Les pistes de cette session sont routées individuellement sur les 64 canaux disponibles en sortie de l'interface HD MADI vers les entrées de l'Euphonix System 5. Nous récupérons donc chaque piste de la session Pro Tools sur autant de *strips* dans la console. Les différentes sessions de montage ayant été réorganisées pour proposer un nombre raisonnable de pistes, l'ensemble de ces dernières tient sur un seul *layout* de console, composé de deux couches accessibles via la fonction *swap*. Cette réorganisation nous permet de pouvoir mixer l'ensemble des sons uniquement avec l'automatisation de la console, et nous évite ainsi d'avoir recours à un pré-routing interne dans Pro Tools, et de passer par le protocole de télécommande *Eucon*.

Comme nous l'avions expliqué en 3.2.1, il nous semblait indispensable de disposer d'un système hybride proposant à la fois un routing de nos *strips* vers le 5.1 et un autre vers les objets WFS. On pourra y voir ici une analogie assez forte avec le fonctionnement de l'Atmos qui propose des *Beds* multicanaux « classiques » et des pistes objets. Pour cela, nous utilisons les *Mix Busses* de la console⁵. L'envoi vers ces bus est automatisable, ce qui répond au besoin de pouvoir changer nos assignations au cours du temps, et ce pour chaque *strip*. Nous créons alors un premier *Mix Bus* 5.0 qui correspond, comme son nom l'indique, à l'envoi vers la partie multicanale classique de notre système. Par défaut, toutes les *strips* de

⁵ En configuration multicanal classique, ces bus sont assimilés à des stems.

notre projet sont routées vers ce bus. A ce premier bus 5.0 se rajoute six bus mono que nous nommerons OBJ de 1 à 6, puis neuf bus stéréo que nous nommerons OBJ de 8 à 16. Ces bus correspondent aux envois vers des objets WFS : les sons routés vers ces bus sont donc destinés à être synthétisés par front d'onde.

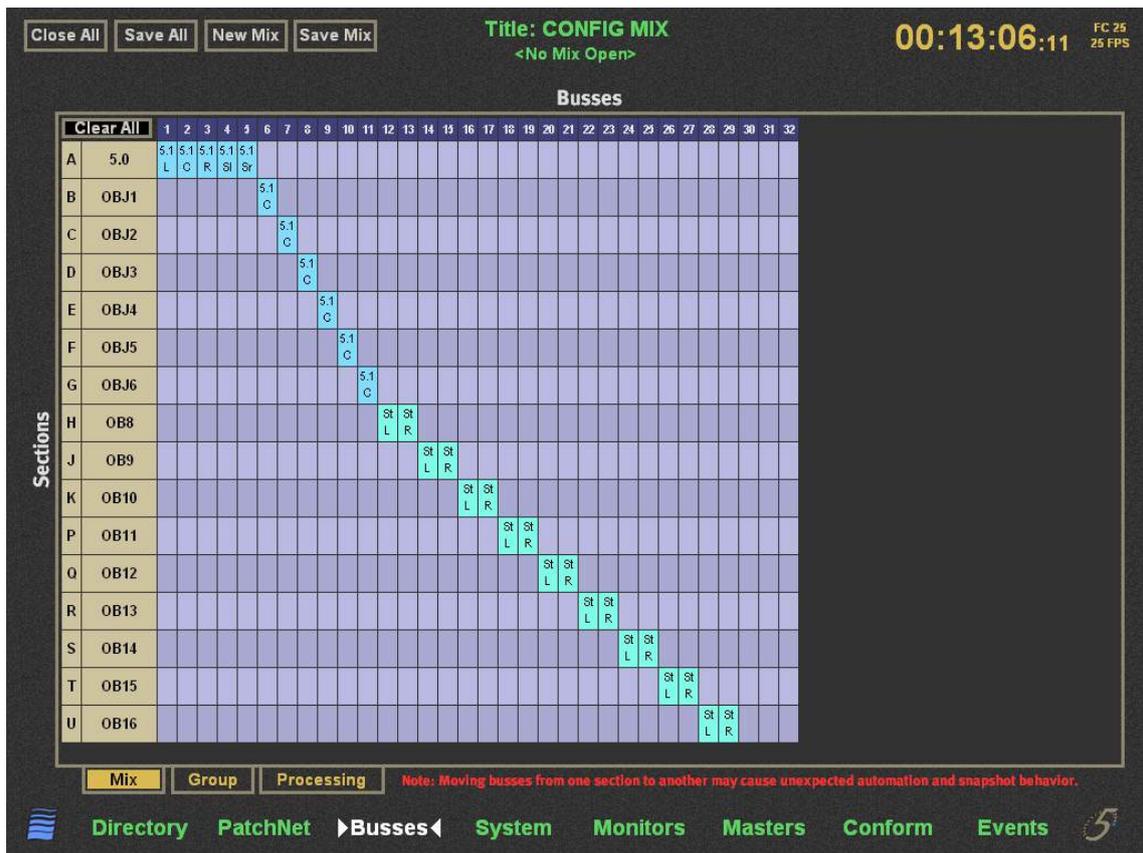


Fig. 3.6 : Configuration des Mix Busses dans la console.

Le système de l'Euphonix propose bien 32 canaux disponibles pour les *Mix Busses*, ce qui correspond exactement au nombre d'objets pris en charge par le Wave I. Seulement, il n'est possible de créer que 16 *Mix Busses*. Ayant besoin de plusieurs bus de petit format, cette limite ne nous permet de router que 29 canaux, regroupés sous nos 16 *Mix Busses* à disposition (1 bus 5.0, 6 bus mono, 9 bus stéréo). Pour contourner le problème et étendre nos possibilités, nous avons

utilisé deux bus auxiliaires, qui, si leur niveau d'envoi est fixé à 0, peuvent être utilisés de la même manière que les *Mix Busses*. Le premier bus auxiliaire correspond à l'envoi vers le Lfe, le second à l'envoi vers un dernier objet mono (OBJ7). Notons que deux autres bus auxiliaires sont également utilisés pour un envoi vers nos réverbérations dans Pro Tools.



Fig. 3.7 : Configuration des Aux Busses dans la console.

Chaque bus (*Mix Bus* ou *Aux*) est ensuite routé en MAD1, canal par canal, des sorties de la console vers les entrées d'une carte RME reliée à Nuendo. Le patch d'entrée de Nuendo reprend donc la matrice et le format de sortie des 16 *Mix Busses* et des deux *Aux Busses* de l'Euphonix. On récupère ainsi notre bus 5.0 (MADI 1-5), le bus Lfe (MADI 6), nos sept bus objet mono (MADI 7-13) et nos neuf

bus stéréo (MADI 14-31). Par souci de clarté, l'ordre des canaux sera conservé tout le long de la chaîne. Ces entrées sont assignées à des pistes dans la *timeline* de Nuendo, dont le nombre et le format recopient celui du patch d'entrée : une piste 5.0, une piste mono LFE, sept « pistes objet » mono et neuf « pistes objet » stéréo. Sur les pistes dédiées aux objets WFS, on insère une instance du plug-in *Tosca* par canal (donc deux plug-ins pour les pistes stéréo). Le fonctionnement de ce plug-in est détaillé en 3.1. Une fois l'autorisation d'écriture et de lecture activée dans Nuendo, les deux paramètres utiles au placement des objets dans l'espace WFS (distance et azimuth) seront inscrits dans le séquenceur sous la forme de lignes d'automations. Chaque « piste objet » de Nuendo correspond donc à un objet (ou un groupe pour les pistes stéréo) dans le *Performer*. Finalement, nous pouvons résumer cette configuration en disant qu'une « piste objet » dans Nuendo est un support déporté à l'écriture de l'évolution de position d'un objet dans le *Performer*, la correspondance entre un objet (du *Performer*) et une piste (de Nuendo) étant permise par le renseignement de l'ID dans chaque instance du *Tosca*.

Mise en pratique de la WFS dans le mixage cinéma

Nom de Bus	HP	Périphérique Audio	Port Périphérique
5.0 In	5.0	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 1
- Droite			HDSP MADI 3
- Centre			HDSP MADI 2
- Surround Gauche			HDSP MADI 4
- Surround Droit			HDSP MADI 5
LFE In	Mono	ASIO Hammerfall DSP	
- Mono			HDSP MADI 6
Mono In	Mono	ASIO Hammerfall DSP	
- Mono			HDSP MADI 10
Mono In 2	Mono	ASIO Hammerfall DSP	
- Mono			HDSP MADI 11
Mono In 3	Mono	ASIO Hammerfall DSP	
- Mono			HDSP MADI 12
Mono In 4	Mono	ASIO Hammerfall DSP	
- Mono			HDSP MADI 13
Mono In 5	Mono	ASIO Hammerfall DSP	
- Mono			HDSP MADI 14
Mono In 6	Mono	ASIO Hammerfall DSP	
- Mono			HDSP MADI 15
Mono In 7	Mono	ASIO Hammerfall DSP	
- Mono			HDSP MADI 16
Stereo In	Stéréo	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 17
- Droite			HDSP MADI 18
Stereo In 2	Stéréo	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 19
- Droite			HDSP MADI 20
Stereo In 3	Stéréo	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 21
- Droite			HDSP MADI 22
Stereo In 4	Stéréo	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 23
- Droite			HDSP MADI 24
Stereo In 5	Stéréo	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 25
- Droite			HDSP MADI 26
Stereo In 6	Stéréo	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 27
- Droite			HDSP MADI 28
Stereo In 7	Stéréo	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 29
- Droite			HDSP MADI 30
Stereo In 8	Stéréo	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 31
- Droite			HDSP MADI 32
Stereo In 9	Stéréo	ASIO Hammerfall DSP	
- Gauche			HDSP MADI 33
- Droite			HDSP MADI 34

Fig. 3.8 : Patch d'entrée de Nuendo

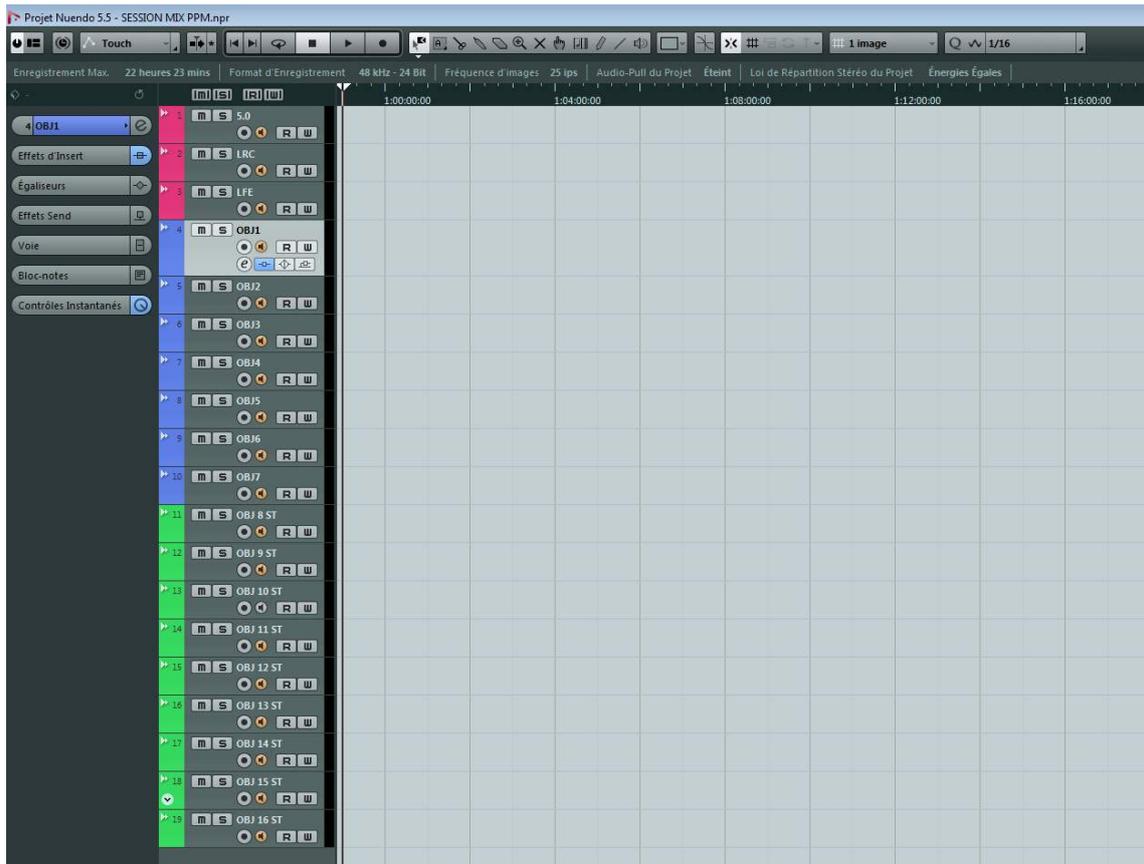


Fig. 3.9 : Timeline du séquenceur Nuendo

Toutes les pistes de Nuendo sont ensuite routées vers les sorties MAD1 de la RME. Le lecteur constatera que le patch de sortie est identique au patch d'entrée et que l'ordre de canaux dans la trame MAD1 est conservé. En terme de routing, Nuendo est ici complètement transparent. Sa place sur la chaîne du signal n'est due qu'à la nécessité d'un support d'écriture à l'automatisation de placement des objets. Parallèlement à la liaison MAD1, un câble RJ45 relie la carte réseau de l'ordinateur sur lequel sont installés Nuendo et le Designer/Performer, à la carte réseau du Wave I. Notons qu'il aurait été tout à fait envisageable d'installer ces deux logiciels (Nuendo et Performer) sur deux machines différentes, mais que toujours dans un souci de clarté et d'ergonomie, nous avons préféré n'en utiliser qu'une : cette machine est notre interface de spatialisation. La liaison ethernet permet

d'envoyer au processeur toutes les informations de position des objets décidées par l'utilisateur via l'interface du Performer, ou relues par les automatisations du Tosca dans Nuendo⁶. Le Performer propose une matrice d'entrée qui permet d'assigner les canaux d'entrée du processeur aux 32 objets disponibles. Pour notre configuration, nous avons conservé un routing en diagonale, c'est-à-dire que le canal 1 sortant du Nuendo est considéré comme l'objet 1 par le Wave 1, et ainsi de suite.

Le Wave 1 reçoit donc des signaux audionumériques par liaison MADI, et des informations de position concernant chaque objet, par liaison ethernet : nous sommes bien placés dans un contexte de mixage *orienté objet*. C'est en intégrant ces deux types de données que le processeur calcule le signal qu'il doit fournir à chaque enceinte du dispositif pour reconstruire des fronts d'ondes tenant compte de ces informations de positionnement⁷.

Dans notre routing, les six premiers canaux MADI entrant dans le Wave 1 correspondent à notre bus 5.0 (MADI 1-5, de format L-C-R-LS-RS) et à notre envoi LFE (MADI 6). Ces six canaux, désormais considérés comme des objets par le processeur, ne doivent pourtant pas être synthétisés en WFS mais bien être envoyés respectivement sur des points de diffusion précis, en d'autres termes, sur une seule enceinte par canal. Comme nous l'avons expliqué en 2.1.1, le Performer permet de « débrayer » la synthèse WFS d'un objet dans un ou plusieurs sous-systèmes, et d'assigner cet objet à une ou plusieurs enceinte de ce(s) sous système(s). Prenons un exemple concret :

⁶ C'est également via cette liaison que le processeur reçoit les infos de configuration du designer.

⁷ Cf. : encadré iv p. 120.

Le canal 1 de notre trame MAD1 correspond à la gauche (L) de notre 5.0. Cet objet doit donc être envoyé uniquement dans l'enceinte de gauche, du sous-système 1 (frontal). Dans l'onglet routing du Performer (cf. Fig. 3.10), il suffit de se placer dans l'encadré correspondant à l'objet concerné (ici objet 1) et de décocher les trois autres sous-systèmes (latéral, arrière, sub). Peu importe sa position, l'objet ne sera donc plus reproduit par ces sous-systèmes. Cocher l'option *manual* pour le sous-système frontal désactive la WFS pour cet objet dans ce sous-système. La fonction *assign* permet ensuite de définir sur quelle(s) enceinte(s) du sous-système l'utilisateur choisit de diffuser cette en discret, dans notre cas l'enceinte de gauche. Le procédé est identique pour le quatre autres canaux du 5.0 et du Lfe, qui lui n'est envoyé que dans le sous-système 4 (celui du *subwoofer*). Pour les canaux Ls et Rs, nous décidons d'alimenter avec ces signaux les deux dernières enceintes latérales, ainsi que les premières enceintes de la rampe arrière. L'enceinte de la rampe arrière n'est pas alimentée.

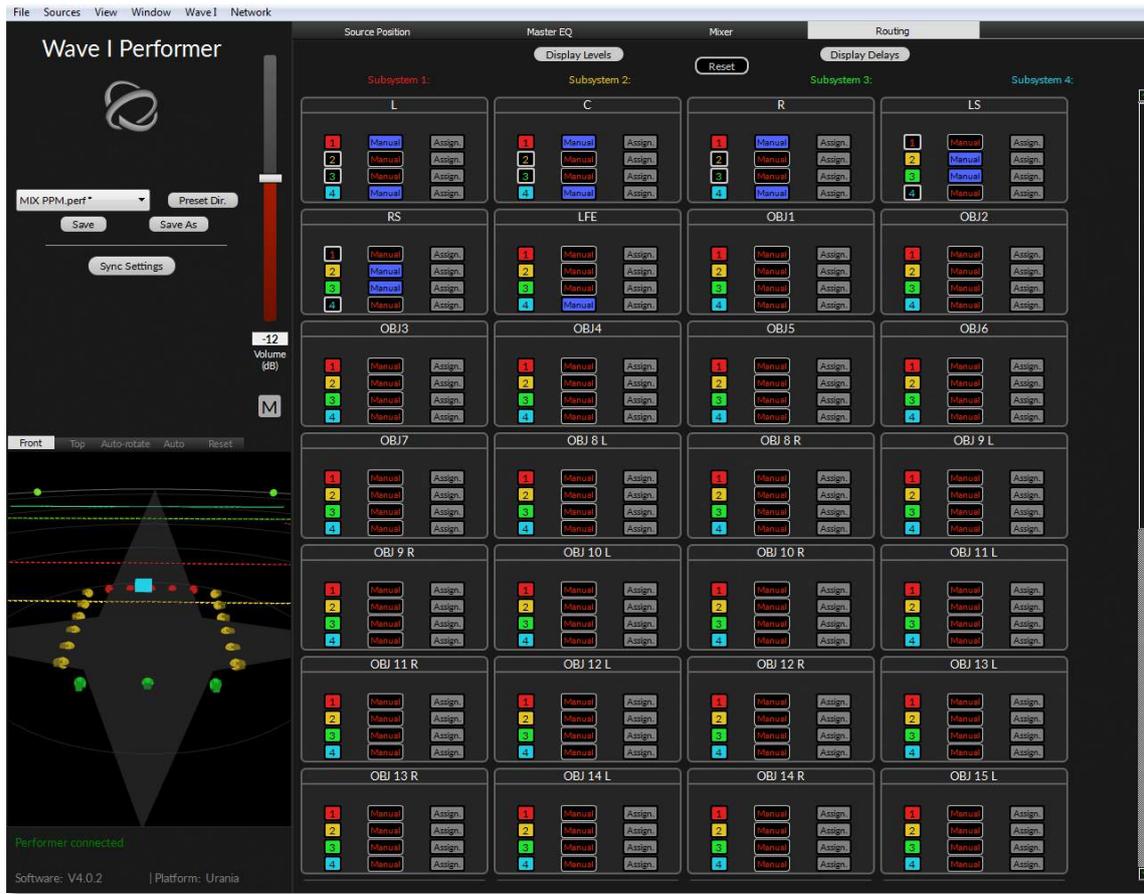


Fig. 3.10 : Routing dans le Performer, utile à notre session de mixage

Les signaux d'alimentation calculés par le Wave 1 sortent du processeur via liaison MAD1 et attaquent un convertisseur numérique/analogique. Ce convertisseur utilise directement le signal MAD1 pour extraire un signal d'horloge nécessaire à cette étape de conversion. Le processeur fournit alors vingt et un signaux destinés à alimenter les vingt enceintes du système, ainsi que la *subwoofer*. Ces signaux sont récupérés en analogique en sortie du convertisseur, via des connectiques SUB-D 25. Ils attaquent ensuite dix amplificateurs stéréo auxquels sont reliées les enceintes, le caisson de grave étant actif.

3.3.3. Calibration

Un des principaux défauts de notre installation tient à l'utilisation de plusieurs modèles d'amplificateurs, chacun ayant ses caractéristiques propres, notamment en terme de puissance. L'étape de calibration du système est donc primordiale pour garantir son bon fonctionnement. Son principe et son objectif en WFS restent les mêmes qu'en multicanal, mais sa mise en œuvre est légèrement différente. Le lecteur trouvera en Annexe 3.2 les spécifications techniques d'un des amplificateurs utilisés.

Au préalable de cette calibration, nous avons chargé dans le *Designer* des presets d'enceintes, que Sonic Emotion et APG ont élaborés ensemble. Ces presets permettent de corriger spectralement la non-linéarité des DX8 en vue d'une utilisation dans un dispositif WFS. Deux presets sont fournis, en fonction de l'utilisation ou non d'un caisson de grave en *bass management*. Le preset 1, défini pour une installation sans *bass management*, propose notamment une bosse de 10 dB à 77Hz ($Q = 1,48$). Le preset 2, défini pour un système renforcé par un *subwoofer*, propose une réponse coupée à 115Hz à -24dB/octave. Après avoir chargé le preset 2 sur toutes les enceintes et comparé avec une bruit rose envoyé à la fois sur la DX8 centrale *bass managé*, et l'enceinte de centre du système Meyer installé à l'école, nous ressentions toujours un manque dans le grave. Le preset 1 proposait lui, une coloration bien plus proche des Meyer dans cette zone fréquentielle. Nous avons donc choisi ce preset pour l'ensemble des enceintes de notre installation.



Fig. 3.11 : Preset de l'enceinte DX8, avec bass management.



Fig. 3.12 : Preset de l'enceinte DX8, sans bass management.

Un paramètre primordial à prendre en compte dans cette étape de calibration est la présence d'un système 5.1 au sein de notre installation WFS. Nous devons donc mettre en place une sorte de double alignement, l'un tenant compte de la norme SMPTE RP200 définie pour le 5.1, et le second assurant un niveau identique, enceinte par enceinte, pour la WFS. Nous décidons de calibrer la WFS par rapport au 5.1 et non l'inverse. Ce raisonnement nous semble plus propice à nous amener naturellement vers des niveaux cohérents pour une diffusion cinéma, le 5.1 imposant alors une forme de gabarit de niveau pour le reste du système. Pour des

raisons expliquées dans la suite du texte, nous ne calibrerons dans un premier temps que l'enceinte de centre associée, par *bass management*, au *subwoofer* (les raisons de ce choix sont expliquées plus loin). La calibration de ce canal central nous servira ensuite de référence, pour tout le système, et nous finirons finalement par l'ajustement du 5.1. L'ensemble des ajustements de niveau s'effectue directement dans le *Designer* que l'on bascule en mode « online », ce qui nous donne accès aux réglages de niveau de chaque haut-parleur indépendamment, via l'utilisation d'un curseur de gain qui permet un réglage de +12dB à - ∞.

Un bruit rose mesuré à 0 dB VU (niveau électrique) est utilisé pour les mesures. Ce bruit est généré dans la console, via la fonction *OSC*, puis est routé dans le bus 5.0, au centre, afin de n'alimenter que l'enceinte centrale et le *subwoofer* utilisé pour le *bass management*. Nous rappelons au lecteur que les canaux de notre 5.0 sont assimilés à des objets discrets, n'alimentant respectivement qu'une seule enceinte, renforcée dans le bas du spectre par le caisson de grave.

Le niveau fourni par le couple enceinte/sub alimenté par le bruit rose est alors mesurée au sonomètre, pondération C, en mesure lente. On ajuste ensuite le master fader (niveau de l'ensemble du système) du *Performer* de telle sorte à obtenir un niveau de pression de 85 dB_{SPL} (C) pour une mesure effectuée au *sweet spot*. En agissant sur le master fader nous réglons grossièrement un niveau d'ensemble pour le système, ce qui nous permettra d'effectuer des réglages plus fins lorsque nous travaillerons enceinte par enceinte. Ce master fader est alors réglé à - 19,2 dB. Le niveau de *subwoofer* est ajusté individuellement au cours de cette mesure dans le *Designer* pour obtenir un niveau de grave perceptivement cohérent avec l'enceinte centrale du système Meyer de l'école. Cet ajustement est effectué par une écoute comparative entre les deux systèmes. Notons qu'une

retouche sur le niveau du *subwoofer* demande obligatoirement d'effectuer de nouveau une mesure sur le couple enceinte/sub.

Cette première étape de calibration effectuée, nous mesurons ensuite le niveau fourni par cette enceinte centrale, à une distance d'un mètre, pour le même bruit rose, toujours avec le même réglage de sonomètre, mais sans l'action du *subwoofer*. Nous obtenons un niveau de 95 dB_{SPL} (C). La seconde étape consiste à reporter ce niveau sur l'ensemble des enceintes. Un piège dans lequel il ne faut évidemment pas tomber consisterait à router notre signal de test vers un objet et de placer cet objet devant l'enceinte mesurée. On ne mesure alors pas le niveau de l'enceinte, mais bien le niveau de cet objet synthétisé en WFS, et donc potentiellement étalé sur plusieurs haut-parleurs. Pour effectuer une mesure enceinte par enceinte rigoureuse, le mode « online » du *Designer* propose une fonction *Test*. Ce mode permet à l'utilisateur de choisir un signal d'entrée parmi les entrées audionumériques du Wave 1, et de sélectionner l'enceinte sur laquelle on souhaite envoyer ce signal. Lorsque ce mode *Test* est activé, il prend la main sur le routing du *Performer*. Nous pouvons donc utiliser notre bruit rose, routé dans le centre de notre bus 5.0 (Input MADI 2 pour le Wave 1) pour calibrer l'ensemble de notre installation.

Les enceintes sont réglées une à une via leur curseur de gain dans le *Designer* pour obtenir un niveau, à un mètre de leur centre acoustique, de 95 dB_{SPL} (C).

Le niveau des enceintes réglé, nous devons ensuite ajuster le niveau de notre système 5.0. Si nous n'avons pas effectué ce réglage en premier comme nous l'avions d'abord envisagé, c'est que les canaux surround alimentent plusieurs enceintes (trois pour Ls, trois pour Rs). La norme de calibration 5.1 demande un niveau de 82 dB_{SPL} (C) pour chaque canal surround. Il nous semble compliqué voire impossible de parvenir à ce niveau en réglant le gain de chaque haut-parleurs,

tout en maintenant l'alignement de 95 dB_{SPL} par enceinte nouvellement effectué, qui doit être impérativement conservé pour un bon fonctionnement du système WFS. Ce niveau ne pouvant plus être touché, nous pouvons contourner ce problème en réglant le niveau, non pas des enceintes, mais des deux canaux arrières. Les canaux du 5.0 étant assimilés à des objets discrets, il nous est possible d'appliquer un *trim* à ces objets pour chaque sous-système dans le *Performer*. Ainsi nous pouvons obtenir un niveau de 82 dB_{SPL} pour chaque canal surround sans toucher au niveau des enceintes dans le *Designer*. Pour éviter les effets de précédences, les canaux arrières sont décalés de 2,4 ms.

La dernière étape de notre calibration consiste à corriger la couleur spectrale globale du système. La coloration des enceintes de sonorisation favorise la partie haute du spectre, ce qui rend notre système agressif. De plus nous n'avons pas évoqué en 3.2.2 une autre particularité des auditoriums de mixage : la courbe ISO X. Cette courbe part du constat suivant : un son diffusé dans un grand espace, à travers un système de diffusion ayant une courbe de réponse plate, paraît plus brillant à cause des réverbérations de la pièce. Définie par la norme ISO 2969 et ANSI/SMPTE 202M, cette courbe permet de limiter l'agressivité originale de la reproduction sonore dans une grande salle par l'application d'un filtre de -3dB/octave à partir de 2kHz. Il ne s'agit pas pour nous de reproduire cette courbe à l'identique. Nous souhaitons avant tout corriger la coloration de notre système pour des conditions de travail agréables, et ainsi se rapprocher du rendu d'un auditorium de mixage cinéma. C'est par une écoute comparative d'un bruit rose diffusé sur le système Meyer de l'école (soumis à la courbe ISO X) et sur notre installation que nous effectuons la correction spectrale, en utilisant le *Master EQ* du *Performer*, s'appliquant à toutes les enceintes. Les réglages de ce *Master EQ* sont les suivants :

- *Low Shelf*, -2,9 dB à 10234 Hz

- *Notch*, -1,4 dB $Q=0,68$ à 5024 Hz
- *Notch*, -1,6 dB, $Q=0,68$ à 306 Hz

La correction dans le bas médium à 306 Hz, nous permet de corriger une résonance trop prononcée du couplage enceinte/pièce, pouvant aussi être due à l'utilisation du preset 2 pour la courbe de réponse des enceintes. Enfin, cette correction impose d'ajouter un *Make up gain* de 3,6 dB via le *Master EQ* pour retrouver un niveau de 85 dB_{SPL} au *sweet spot* pour les canaux frontaux du 5.0.

3.4. Expérience de mixage

Cette configuration établie, nous avons pu commencer à travailler sur du contenu et confronter notre installation à la réalité et à l'exigence pratique d'un mixage. Tout au long de nos séances, nous avons eu pour objectif commun de réaliser des mixages qui nous semblaient adaptés à l'image proposée. Ces choix ont été les nôtres, et ne forment en aucun cas une approche généralisable du mixage en WFS. Cela dit, le lecteur pourra ici trouver des éléments dans l'intérêt de s'approprier lui-même les outils, et expérimenter à son tour les effets de la WFS associés à une image 2D. C'est également dans cette partie que nous avancerons un avis quant à la validité de la technologie WFS confrontée au support visuel cinématographique.

Une description des séquences mixées est mise à disposition du lecteur en 4.1.3, p. 215. Comme nous l'évoquons plus tard, ces séquences ont été choisies dans le but de mettre en avant certains aspects de la spatialisation sonore, dans le but de comparer le rendu du système 5.1 avec celui du système WFS. Nous souhaitons souligner que même si nous avons eu le sentiment d'optimiser nos choix de mixage en WFS, les mixages 5.1 peuvent avoir été mieux réussis, du fait de notre plus grande expérience avec ce type de dispositif. Par ailleurs, bien que cette caractéristique soit une borne de notre sujet, nous avons à disposition des montages son pensés pour le 5.1. Il est tout à fait probable qu'avec un contenu différent et une organisation différents nous aurions obtenus des résultats différents. La question de penser le montage son « en objet » est une étude intéressante à laquelle nous n'avons volontairement pas prêté attention.

3.4.1. Ergonomie, méthode et espace(s) de travail en WFS

Dans cette courte partie, nous présentons au lecteur notre méthode de travail qui nous a suivi pendant une semaine de mixage. Au fur et à mesure des séquences, des défauts sont apparus, des manques se sont fait ressentir, et une autre manière d'appréhender les différentes étapes d'un mixage s'est construite.

Centralisation des outils nécessaires au mixage

C'est d'abord dans la multiplication des machines que nous avons vu un premier défaut à notre installation. En déportant les outils de spatialisation et de son écriture en dehors de la console, nous les avons séparé et espacé physiquement, créant des pôles de travail spécifiques. Il est pourtant d'appréhender le mixage comme une succession d'étapes indépendantes les unes des autres. La spatialisation ne peut s'envisager sans un retour sur les niveaux, le filtrage, et inversement. Dans le cas d'un unique opérateur, cela impose des allers et retours qui peuvent rendre le mixage plus long et fatigant. Cela dit, ces déplacements peuvent permettre de contrôler l'image sonore à un autre point qu'au *sweet spot*. L'ajout d'outils supplémentaires impose également une bonne maîtrise et connaissance de chacun d'entre eux, ce qui demande une concentration et des compétences supplémentaires. Plusieurs solutions pourraient être envisagées afin de simplifier certaines opérations.

D'une manière générale, ces solutions consistent à regrouper les différents outils au sein d'une même unité de contrôle. Une référence temporelle et des fonctions d'écriture de l'automation pourraient être implémentés dans le logiciel de

spatialisation WFS, ce qui permettrait de s'affranchir d'un outil spécifique et dédié à cette tâche, comme nous l'avons fait avec *Nuendo*. Cela limiterait en effet le nombre de manipulations nécessaires pour placer un objet dans l'espace sonore.

Un échange d'informations entre la console et le Performer pourrait permettre un gain de temps sur un certain nombre de tâches et rendre le travail plus intuitif. Par exemple, à l'heure actuelle, lorsque un objet est renommé dans le Performer, l'information n'est pas reportée dans la console. Une telle fonction s'avèrerait très pratique. On pourrait également imaginer un dialogue direct entre les deux machines pour l'affichage des objets : dès lors qu'une source sonore serait routée dans un objet (via l'automation de routing de la console), cet objet serait immédiatement affiché et sélectionné sur le Performer. Seraient alors affichées sur l'interface uniquement les objets dont l'utilisateur a besoin. Cela dit, le *Performer* permet déjà de masquer les objets inutiles à un moment du mix. Notons qu'une collaboration entre Sonic Emotion et le constructeur de consoles *Studer* est déjà en cours.

S'il semble compliqué d'envisager l'implémentation complète d'un *panner* WFS au sein de toute les marques de consoles, on pourrait très bien imaginer des interfaces tangibles, sur des tablettes tactiles par exemple, qui permettraient de s'affranchir de la souris pour contrôler les objets. Cette interface pourrait être placée à côté du mixeur pour limiter ses déplacements. Soulignons tout de même que le travail dans le *Performer* est très agréable, intuitif, et que l'interface semble vraiment pouvoir s'adapter à beaucoup de situations. L'espace est très clairement lisible en 2D, comme nous l'avons utilisé, mais également en 3D, si l'on envisage des haut-parleurs zénithaux. Par ailleurs, les premiers retours utilisateurs sur les outils *Atmos* sont beaucoup moins élogieux, pointant des problèmes de lisibilité, de représentation de l'espace, et des manques de modularité.

Compatibilité des outils et limite de notre configuration

Nous avons vu au en 3.3.2 que le nombre de sorties physiques de la *System 5* de l'école était limité à 32, ce qui correspond au nombre de canaux d'entrée du *Wave 1*. Cela dit, le nombre de *Mix Busses* (ou *Stems*) virtuels reste limité à 16, ce qui pourrait réellement devenir problématique dès lors que le *Wave 1* sera en mesure d'accueillir 64 voire 128 objets sonores. En effet, dans l'état actuel des choses, en considérant que l'envoi vers des objets WFS ne peut se faire que par un routing évolutif des bus, nous ne pouvons envisager que 32 objets mono (ou 16 groupes bipolaires), ou d'une manière hybride, 8 objets mono et 8 groupes bipolaires. Dans notre cas, nos sessions ne mettant en jeu qu'un nombre raisonnable de pistes, cette limitation du nombre d'objet ne nous est pas vraiment apparue comme une contrainte. Mais au regard des capacités de l'*Atmos*, de l'*Auro 3D*, et en envisageant un travail sur des séquences plus complexes et fournies, il faudrait reconsidérer les capacités de routing.

Cette méthode choisie pour définir le chemin de sortie de chaque source n'est pas non plus sans causer un certain nombre de difficultés. En effet, l'automatisation de routing de la *System 5* ne fonctionnait pas comme nous l'entendions : l'enchaînement « *Preview / Punch Preview* », manipulation très souvent utilisée pour l'écriture d'automations globales (sur toute une séquence par exemple), ne nous amenait pas au changement de routing attendu. De fait, nous avons procédé à une écriture directe du routing en « *Fill Start Region / Fill End Region* », méthode pas forcément très pratique pour comparer le rendu des différents chemins choisis (en 5.1 et en WFS, par exemple).

L'idéal serait d'implémenter, dans la console, une fonction de routing dédiée à la WFS. De la même manière que la fonction de routing classique apparaisse sur la console (bouton *Mix* sur chacun des tranches), un autre switch permettrait

d'accéder aux objets WFS. Une page de la console (un *Layout*) pourrait alors rassembler les différents canaux visant à être envoyés dans le Wave 1. Ce nouvel étage pourrait alors permettre un certain nombre d'ajustements, de compensations, avant la synthèse de front d'onde : modification du spectre, du niveau...

Méthodes de travail

Ces sessions de mixage ont été l'occasion pour nous de comprendre un peu plus le fonctionnement de la WFS avec une image. N'ayant pas vraiment d'expérience sur ce format et étant conscients des conflits spatiaux que celui-ci peut générer dans un contexte audiovisuel, nous avons très souvent vérifié nos placements d'objets et cela en plusieurs endroits de la salle. C'est une démarche qui nous semble importante, puisqu'elle vise à vérifier que notre spatialisation ne crée pas d'aberrations de localisation pour certains spectateurs. Nous pensons que ce travail restera nécessaire dans le cadre d'une diffusion WFS au cinéma, mais il est certain qu'avec plus d'expérience il serait possible d'anticiper les problèmes.

Au fur et à mesure que nous avançons dans le mixage de nos séquences, nous avons trouvé une méthode pour placer les sources tout en permettant au mixeur de concentrer d'avantage son travail autour de la console, lui évitant ainsi trop de déplacements vers le logiciel de spatialisation. Pour cela nous avons défini pour chaque séquence plusieurs positions d'objet mono et stéréo « par défaut ». Ces positions sont écrites dans *Nuendo* pour chaque séquence d'un film par exemple. Suivant les séquences traitées, on cherchera à proposer un maillage d'objets plus ou moins resserrés dans l'espace WFS, en fonction de la profondeur et du relief à exploiter. Cette solution permet au mixeur de router les sons en fonction du placement des objets/groupes dans le performer. On peut dès lors imaginer un positionnement des objets dans l'ordre croissant, du plus proche au plus lointain.

Le mixeur a alors la possibilité de tester plusieurs positions très rapidement (en *Preview*) et d'évaluer leur pertinence. Cette méthode fonctionne particulièrement bien pour les ambiances et certains effets. Dès lors que les positions prédéfinies ne fonctionnent plus pour un son, on devra router vers des objets « libres » que l'on placera en conséquence. Dans notre cas, envisager trois positions d'objets mono et stéréo à l'avant, deux à l'arrière, nous laisse suffisamment d'objets libres pour les sons demandant un placement particulier ou un déplacement dynamique. Si nous n'avons pas pu emmener cette façon de travailler jusqu'au bout et bien en évaluer les limites, cela nous semble être une configuration de travail plus confortable.

3.4.2. Possibilités de spatialisation en WFS

Comme nous l'avons déjà évoqué ailleurs dans ce document, nous envisageons le mixage comme une activité faite d'allers et retours entre la création d'un effet et la perception de celui-ci. Cet aspect propre au travail du son, et plus largement à n'importe quel objet esthétique, nous permet d'envisager une analyse découpée, entre description de l'effet et méthodes nécessaires à sa création. Sur ces séances de mixage dans l'auditorium, nous sommes passés par plusieurs étapes, plus ou moins récurrentes, faites d'une multitude et d'une diversité d'effets créés. Par la suite, nous nous attarderons uniquement sur les quelques effets principaux constatés, dont ceux qui ont pu mettre en évidence des différences entre le dispositif 5.1 et le dispositif WFS.

Globalement, pour chaque séquence, nous avons essayé de respecter la géométrie de l'image. Non pas que cette approche naturaliste ait été une condition inébranlable pour nous, mais elle semble être grandement induite par la présence de l'image, à la fois guide de la narration, et guide de l'espace investi par les protagonistes. Par exemple, dans *La Menace d'une Rose*, le premier plan sur la Tamise nous a amené à éloigner toutes les ambiances de type « fond d'air urbain » et à rapprocher les quelques effets de clapotis du fleuve, dans le but de créer du relief. Notons bien que par « éloigner », nous entendons introduire de la distance holophonique, i.e. éloigner les objets dans le *Performer*. Nous espérions ainsi modifier la structure de l'onde restituée. En effet, une source éloignée est censée se rapprocher d'une diffusion en onde plane (cf. 1.4.3).

L'un des premiers effets perçus ne fut pas en lien avec le relief dans l'espace sonore. En effet, ce fut la stabilité de l'image sonore frontale, indépendante de notre position d'écoute dans la salle, qui nous est apparue de prime abord. Ce que nous avons constaté sur les séries de tests effectués chez Sonic Emotion se vérifiait aussi dans l'auditorium de l'école. Là où en 5.1 l'image sonore se décale pour des positions excentrées, une image large et cohérente est récupérée en WFS. Notons pour exemple l'ambiance d'orage dans la séquence *Gabriela*. En WFS, et ce même pour une position excentrée, la diffusion de l'orage reste enveloppante, large. C'est une caractéristique que nous ne parvenions pas à obtenir en 5.1. Ce groupe bipolaire avait simplement été éloigné dans le *Performer* de sorte à mettre à contribution toute la façade ainsi qu'une partie des haut-parleurs latéraux. C'est aussi sur cette ambiance que les basses fréquences des rampes latérales nous ont été très utiles et se sont avérées, selon nous, très efficaces.

En WFS, la mise à contribution d'un plus grand nombre de haut-parleurs a également eu pour résultat une plus grande densité sonore perçue. Selon nous,

sur un mixage de plusieurs ambiances, cela rendait l'image sonore plus homogène et plus enveloppante. A l'inverse, en 5.1, nous pouvions parfois percevoir les sons « placés » dans les enceintes. Ce fut par exemple le cas dans la séquence *Mundos Immundos*, où les vagues latéralisées en 5.1 rendaient les haut-parleurs plus perceptibles, le système moins transparent. En WFS, la latéralisation associée à une distance holophonique assez importante (plusieurs mètres virtuels) mettaient toutes deux à contribution un plus grand nombre de haut-parleurs, ce qui selon nous avait tendance à adoucir la perception des vagues, à les rendre peut-être plus réalistes.

Le relief obtenu nous a dans un premier temps semblé perceptible pour un certain nombre de séquences. Cependant, pour la plupart des éléments visuels « in », ce relief nous a fait ressentir une certaine « artificialité ». A l'image des vagues frappant les rochers dans la séquence *Mundos Immundos*, la distance holophonique apportée à des objets visibles dans le champ apportait, selon nous, une certaine décorrélation entre la source sonore ponctuelle et le reste du mixage. De plus, nous verrons dans la partie suivante (cf. 3.4.3) que le relief est difficilement perçu pour les éléments sonores non-ponctuels, telles que des ambiances de type « fond d'air », des rumeurs... C'est sur les stimuli brefs hors-champs que le relief apporté par la WFS nous a semblé le plus probant. Dans la séquence *Couper le voile*, la femme criant depuis se balcon n'apparaît pas dans le cadre, et la distance holophonique induite éloignait effectivement cette voix en lui offrant une diffusion élargie. Cette distance virtuelle perçue contrastait de manière efficace avec le reste de l'action au premier plan, à savoir l'homme parlant depuis le seuil de son immeuble. Nous avons retrouvé ce type d'effets sonores ponctuels dans *La Menace d'une Rose*, avec les bruits métalliques provenant de la zone industrielle hors-champs. Cela dit, ce type d'éléments sonores s'est fait assez rare dans l'ensemble de nos séquences.

Pour finir, la large étendue spectrale sur les rampes latérales et arrières nous ont permis un certain nombre d'effets intéressants que le dispositif 5.1 était incapable de reproduire aussi efficacement. La reproduction des basses fréquences par les canaux surrounds n'est pas une caractéristique propre à la WFS, puisque Dolby prévoit une telle possibilité avec l'*Atmos*. Cependant, l'introduction de distance holophonique à l'arrière, associée aux basses fréquences, nous a permis de créer des effets d'enveloppement où les canaux arrières étaient beaucoup moins perçus, et ce même pour des positions d'écoute reculées dans la zone d'écoute. Ce fut le cas pour l'ambiance d'orage dans *Gabriela*, mais également pour les ambiances maritimes dans la même séquence, lorsque le personnage principal rejoint le bord de mer. Nous sommes par exemple parvenus à spatialiser des effets de vent à l'arrière, sans que ces derniers soient gênants même pour un auditeur placé à l'arrière de la salle. Nous nous sentions plus enveloppés, sans pour autant être décrochés de l'écran. En 5.1, comme nous l'avions remarqué chez Sonic Emotion (cf. 2.2.3) un tel résultat était possible au *sweet spot*, mais impossible à préserver pour toute la zone d'écoute : pour des positions d'écoute reculées, le réflexe d'attention mêlé à l'effet de précedence (entre la façade et la rampe arrière) faisaient à nouveau leur apparition.

Nous avons également pu définir cet enveloppement dans une plus grande largeur globale du mixage, comme si celui-ci dépassait légèrement l'écran. Cette largeur était aussi amenée par les enceintes situées proches des bords de l'écran, et mises à contribution par des ambiances éloignées dans le *Performer*. C'est une caractéristique intéressante qui pourrait être envisagée sur d'autres systèmes de diffusion, même discrets.

3.4.3. Plusieurs effets discordants

Bien que des apports assez évidents aient émergé de notre expérience en WFS, plusieurs effets nous ont paru gênants en condition de mixage, et même discordants avec la projection d'une image 2D. Le plus évident d'entre eux, et peut être le plus gênant, touchait à la mise en mouvement des objets sonores. En plus de la difficulté rencontrée à mettre en mouvement les sources de manière efficace, nous avons au final obtenu des effets qui ne nous satisfaisaient pas. Dans *Petite Bête*, nous avons tenté de suivre la sortie du champ de la mère avec une latéralisation de sa voix. En plaçant l'objet WFS à une distance plutôt cohérente avec ce que nous avions à l'image (environ 5 mètres virtuels pour la voix de la mère), nous ne parvenions pas à opérer une latéralisation fluide sans qu'un phénomène de « saute » entre les enceintes ne soit perçu. Cela dit, nous avons peut être effectué un déplacement de l'objet trop près du banc de haut-parleurs. En procédant ainsi, on concentre l'énergie du son sur une seule enceinte, les haut-parleurs adjacents recevant alors un léger signal retardé pour créer un front d'onde sphérique. En mettant l'objet en mouvement, on déplace cette concentration d'énergie sur le haut-parleur voisin (et ainsi de suite), d'où cette perception de mouvement saccadé, amplifié par les différences de rendu fréquentiel entre les haut-parleurs. L'exemple le plus marquant fut le mouvement des voitures dans *La Menace d'une Rose*. Ce déplacement lent, caractérisé par un passage de véhicule du centre à la droite de l'image, nous faisait percevoir d'avantage d'aberrations spectrales (en plus des « sautes » de haut-parleur en haut-parleur) du fait du registre fréquentiel très large de la voiture. Nous avons donc, pour ces cas précis et pour d'autres mouvements de source un peu plus anecdotiques, préféré les effets réalisés avec le dispositif 5.1. Le système WFS ne nous a pas semblé assez transparent.

Par ailleurs, les problèmes spectraux, évoqués pour ces déplacements d'objets mais aussi précédemment dans notre étude (cf. partie 2), se sont manifestés pour un certain nombre de sources sonores. Les plus touchées, de manière assez évidente, furent les ambiances de large spectre, comme nous l'avions constaté en 2.2.2. Dans *Mundos Immundos*, le travail habituellement complexe des matériaux de type « vagues » ou « vents » a été rendu d'autant plus difficile en WFS de par l'introduction d'artefacts spectraux. En 5.1, ces deux types de matière jouaient déjà l'une contre l'autre, et avaient tendance à s'associer en une masse globale qui pouvait sembler cohérente, mais de laquelle il était parfois difficile d'en extraire les différentes composantes. Les problèmes spectraux induits par la synthèse de front d'onde, en plus de nous faire parfois percevoir un certain phénomène de « phasing », ne rendaient pas particulièrement aisé le travail de ces matières. Ce fut également le cas pour le passage d'avion dans la séquence *Louis le Manchot*, où un effet de « filtrage en peigne » pouvait être perçu. Notons tout de même que cet effet ne nous a pas dérangé dans un premier temps, puisqu'il pouvait être assimilé à certains phénomènes fréquentiels que nous avons l'habitude de percevoir dans le ciel. Le problème résidait dans le fait que nous ne voulions pas obtenir cet effet : nous le subissions plutôt que de le provoquer. En conséquence, le système a une nouvelle fois montré qu'il n'était pas transparent.

La perspective sonore en WFS n'a pas non plus été sans induire plusieurs problèmes perceptifs. En pratique, nous avons obtenu des effets marqués assez facilement, comme cela fut le cas pour la voix de la femme dans la séquence *Couper le voile*. Pour un placement de l'objet à une distance de 10 mètres dans le *Performer*, nous percevions une perspective assez efficace. Or, le phénomène de parallaxe décrit en 1.4.3 n'a pas tardé à se manifester. Après avoir bougé dans la zone d'écoute, nous avons commencé à percevoir un certain décalage entre la source sonore et son stimulus visuel, qui lui était fixé à l'écran 2D. Une distorsion angulaire, comme celle décrite dans [Melchior et al., 2003] et que nous avons

évoqué en 1.4.4, a été perçue, ce qui a modifié notre regard global sur le dispositif WFS. Pour d'autres sources, cette distorsion pouvait même nous faire percevoir la source en dehors de l'écran. En effet, certaines vagues latérales de *Mundos Immundos* étaient perçues hors cadre pour une position d'écoute à gauche dans la salle. La distorsion angulaire est illustrée en figure Fig. 1.28, p. 74.

Comme nous l'avons indiqué en 1.4.4, nous pensions que « l'effet ventriloque », également mis en avant dans [Corteel, 2004], allait nous permettre de préserver une certaine « aimantation audiovisuelle ». Cependant, en WFS, l'introduction d'une nouvelle dimension géométrique - en profondeur - dans l'espace sonore semble nous faire apprécier ces phénomènes différemment. En apportant une nouvelle information de spatialisation à la source grâce à la distance holophonique, nous précisons d'avantage sa position en l'éloignant virtuellement. Il semblerait que la forme donnée au front d'onde nous rende plus exigeant et plus critique quant aux discordances entre le champ visuel et le champ sonore. Nous nous attendons à une certaine précision que nous n'obtenons pas. Remarquons, à l'inverse, que le dispositif 5.1 peut parfois jouir de « l'effet ventriloque », même si son action est souvent limitée au centre de l'écran. En effet, le terme de ventriloque provient de la diffusion voix au cinéma qui sont, la plupart du temps, diffusées dans le canal central.

De plus, le relief évoqué en 3.4.2, qui nous est d'abord apparu comme un apport intéressant, nous a rapidement fait ressentir un certain « décrochage » de certains éléments sonores par rapport à l'image, comme si ces sources ne faisaient plus partie du mixage global. Cette profondeur et ce relief en WFS seraient-ils actifs au prix de moins « d'aimantation audiovisuelle » ? Nous nous demandons si ce phénomène perçu n'est pas également dû à nos habitudes d'écoute sur le dispositif 5.1. Au vu de ces constats, on pourrait alors penser que l'introduction de distance holophonique est à réserver pour les éléments hors-champs. Cependant, le travail

minutieux effectué sur ces objets peut s'avérer quelque peu inefficace face à la prépondérance de l'image cinématographique.

En effet, la présence du cadre en deux dimensions a considérablement atténué les phénomènes de relief et de profondeur que nous avons par exemple pu constater en partie 2. Ce « tassement » des plans sonores par l'image était tel que nous avons même eu des difficultés à reconnaître le mixage 5.1 du mixage WFS pour la séquence *Mundos Immundos*, par exemple. Cette perte de relief induite par l'image concerne en particulier tous les « étages » de plans sonores que nous avons construit avec les ambiances, et mais aussi avec les réverbérations. Dans *Louis le Manchot*, la distance holophonique apportée sur les ambiances d'insecte est bien trop fine pour que nous percevions une différence avec le système 5.1. Dans *Couper le voile*, les réverbérations extérieures (qui avaient notamment été utilisées pour les claquements de portières) ne forment pas en soi un espace plus profond, si ce n'est que nous pouvions les percevoir un peu plus larges qu'en 5.1. Ces constats brisent de nombreuses attentes que nous avons vis à vis de la WFS. Les apports, parfois considérables, que nous avons constatés en son seul (sans image) ne semblent pas avoir résistés au dispositif cinématographique. Certains effets restent cependant perceptibles, comme la voix à l'étage dans *Couper le voile*, ou encore la voix de la mère dans *Petite Bête*, mais une autre question se pose sur la pertinence de ces effets de spatialisation.

Préciser la localisation des sources, et donc par conséquent, préciser la scène sonore, semble rendre plus fréquentes les manifestations de notre réflexe d'attention. A-t-on envie d'entendre *là* ce que nous voyons *ici*? Cette question nous semble être au centre d'un sujet plus global, également valable pour les dispositifs plus « classiques » comme le 5.1. Comme l'indique également Claude Baiblé dans [Baiblé, 2013], « l'écoute [est] alors attirée et dispersée en de multiples azimuts porteurs de sens, alors que le regard [a tendance à rester]

centré vers l'écran ». Cependant, et nous l'avons déjà évoqué, l'introduction de distance holophonique en WFS semble nous rendre plus exigeants, en ce sens que le moindre défaut de localisation devient d'avantage perceptible. La profondeur devient un paramètre de plus dans notre localisation de la source. En percevant et en identifiant l'effet de spatialisation, le spectateur se retrouve potentiellement détourné de la narration puisque son œil tente de suivre les stimuli sonores qui lui sont proposés. Ceci nous semble être en contradiction avec la volonté même d'impliquer toujours plus le spectateur dans la narration. En 5.1, bien que l'image sonore frontale ne soit plus correcte pour des positions d'écoute excentrées, nous avons pu percevoir une plus grande « stabilité », une plus grande « accroche » du son à l'image. En effet, ce dispositif sonore ne nous amenait pas à percevoir tant de détails, tant de relief : le champ des possibles dans l'espace sonore était réduit. Un nombre bien plus réduit de positions de sources étaient exploitables (que ce soit latéralement, ou en profondeur), ce qui empêchait concentrer notre attention sur une éventuelle cohérence de spatialisation. En bref, ces constats iraient dans le sens d'un cinéma « mono-centré », ou du moins, une expression majoritairement centrale de son contenu sonore. Rien de surprenant en effet, lorsque l'on sait aujourd'hui qu'en 5.1, la plupart de l'énergie d'un mixage se retrouve concentrée dans le canal central

3.4.4. Conclusions pratiques sur l'utilisation de la WFS en mixage cinéma

Cette partie nous a permis d'émettre un certain nombre de constats sur l'usage de la WFS, à la fois en condition de mixage (i.e. mélange de sources), mais aussi en présence d'une image. Ces deux caractéristiques, propres au dispositif cinématographique, avaient en effet été laissées de côté dans la partie 2 de notre étude. Ainsi, un certain nombre d'hypothèses avancées en 1.4.4 et en 2.3 ont pu être validées, même si ces validations ne tiennent que de notre propre capacité de jugement, et non pas d'une quelconque objectivisation. Précisons au lecteur que nous tenterons d'élargir ces constats dans la partie suivante, en mettant en place un test perceptif qui, nous l'espérons, pourra mettre en évidence certaines tendances sur la perception de contenu WFS lié à une image. L'un des buts de cette troisième partie a donc été de voir si les différents constats et les différentes hypothèses émis en partie 2 pouvaient résister à la diffusion simultanée de plusieurs sources sonore, et à la projection d'une image en deux dimensions sur un écran.

Nous avons vu par exemple que l'écoute exo-centrée était une posture tout à fait envisageable en WFS. L'image sonore frontale indépendante de la position d'écoute, que nous avons déjà constatée lors de nos tests chez Sonic Emotion, est une caractéristique que nous avons retrouvée dans l'auditorium de l'école. D'autre part, l'impression d'homogénéité de l'image sonore, chose qui nous était apparue de manière assez évidente pour la diffusion des ambiances dans la partie 2, s'est à nouveau faite ressentir, bien qu'en 5.1 (du fait du mélange de plusieurs sources) une certaine « densité » puisse également être perçue. Enfin, un enveloppement

efficace et intéressant a pu être constaté, et ce même pour la diffusion assez intense d'objets sonores à l'arrière, tels que des ambiances d'orage, des ambiances maritimes, ou des fonds d'air urbains. Pour des positions assez reculées dans la zone d'écoute, nous continuions à entendre une image frontale dense et cohérente, tout en nous sentant enveloppés par l'ambiance de surround. Ce constat n'a effectivement pas été appliqué au dispositif 5.1, pour lequel nous percevions un peu trop facilement la diffusion des canaux arrières.

En contrepartie, la WFS a pour nous été source d'un certain nombre de problèmes perceptifs, d'abord intrinsèquement liés à des artefacts sonores, puis ensuite liés à des effets plus ou moins capables de s'adapter à la projection audiovisuelle. La mise en mouvement des objets sonores n'est, selon nous, pas assez transparente pour réaliser un effet suffisamment fluide. L'exemple des déplacements de véhicule a en effet été décrit en 3.4.3 comme étant plus convaincant en 5.1 qu'en WFS, du fait des artefacts spectraux qu'induit la synthèse de front d'onde. Par ailleurs, l'étagement des plans sonores que nous avons essayé de créer en WFS n'a semble-t-il pas résisté à la présence d'une image fixe. C'est effectivement notre posture analytique de mixeur, détachée pendant un temps de la narration, qui nous a permis d'émettre ce constat. Nous avançons donc que les volontés de relief et de profondeur exprimées par le mixage sont assez fortement « tassées » par la présence de l'image. Enfin, les distorsions angulaires ont été effectivement dépendantes de la position d'écoute, ce qui va, selon nous, à l'encontre d'un spectacle audiovisuel identique pour tous les spectateurs.

Remettons cela dit le contexte de notre étude face aux limitations technologiques qui lui sont propres. Le système dont nous disposions, bien qu'étant d'une qualité suffisante pour effectuer nos expérimentations, était probablement jonché de défauts principalement dus à l'absence de calibration spectrale. Notre démarche n'a effectivement pas pris en compte de corrections fréquentielles (cf. 3.3.3). De

plus, les enceintes de petite taille nous ont obligé à effectuer un certain nombre de calibrations dans le bas du spectre qui étaient, par manque de temps et de moyen, réalisées « à l'oreille », sans matériel de mesure particulier. Après discussion avec Etienne Corteel sur le sujet, cette façon de faire nous a tout de même semblé être la bonne, et il ne serait pas pertinent de ramener tous les problèmes perceptifs rencontrés à ces quelques contraintes propres à notre installation. Nous nous demandons par ailleurs si toutes les observations effectuées dans cet auditorium seraient également valides sur un autre dispositif WFS. D'autres tests seraient à mener, mais le manque de temps et de moyens déjà évoqués ont limité nos investigations. La question de la transportabilité évoquée en 1.4.4 n'a donc pas été traitée, mais mériterait une étude à part entière. La difficulté de « raccorder » les installations entre elles, déjà bien connue en 5.1, ne serait qu'accrue en WFS du fait du plus grand nombre de haut-parleurs. Les techniques d'égalisation multicanale avancées dans [Corteel, 2004] seraient à considérer dans le contexte particulier de la diffusion sonore au cinéma.

Parallèlement à cela, nous pensons que la conception des outils s'est tout à fait adaptée aux exigences de la postproduction sonore. Quelques optimisations restent à prévoir (centralisation des outils sur une même machine, interfaces tangibles à imaginer), mais ce dispositif pratique, développé en 3.3.2 puis en 3.4.1, nous a semblé tout à fait exploitable. Cette configuration nous a permis d'investiguer puis de présenter une méthode de travail en WFS qui, nous l'espérons, pourra être une nouvelle source d'inspiration aux études futures.

Globalement, à l'issue de ces sessions de mixage, les limitations de la technologie WFS nous semblent d'avantage perceptives que technologiques. Selon nous, les effets « pertinents » de la synthèse de front d'onde sont assez minces en contexte de diffusion cinématographique. La distance holophonique pourrait être prise pour exemple : si son utilisation n'est pas adaptée à la plupart des sons « in », et est

plutôt réservée aux sources sonores hors-champs (généralement moins importantes pour la narration), est-ce que cette distance virtuelle, propre à la WFS, comporte un intérêt dans la diffusion sonore cinématographique ? De même, la cohérence de l'image sonore pour des positions d'écoute « exo-centrées » ne semble pas être un effet suffisamment évident pour aller dans le sens d'une diffusion WFS au cinéma. En effet, le caractère « mono-centré » du mixage, caractérisé par l'omniprésence de l'action au centre de l'image (comédiens associés à des voix, des bruitages, des effets, ...) a tendance à naturellement ramener notre attention vers le milieu de l'écran. En 5.1 déjà, cette prépondérance du canal central (vis à vis des ambiances plus faibles) semble atténuer les effets de précedence propres à la stéréophonie, décrits en 1.2. Ainsi, l'intérêt du système WFS à développer une image sonore cohérente quelque soit la position d'écoute du spectateur ne nous semble pas nécessaire dans la majorité des cas proposés par le cinéma. Par ailleurs, les ambiances stéréophoniques de forte intensité, placées dans un mixage dépourvu de quelconque élément sonore au centre, pourraient former un contre-exemple où la WFS aurait un intérêt significatif pour élargir la zone d'écoute. Cela dit, ces cas nous semblent isolés et peu fréquents dans la plupart des formes narratives développées aujourd'hui en salle obscure.

La validité perceptive de la distance holophonique, ainsi que la cohérence entre le champ sonore et le champ visuel semblent donc être deux questions centrales qui surgissent de notre étude pratique. Nous verrons dans la partie suivante que ces questions pourront faire l'objet de critères d'étude à part entière. Par ailleurs, en 1.4.4, nous introduisons nos hypothèses en nous demandant si les apports de la WFS allaient être suffisants pour lui imaginer une intégration dans la postproduction sonore au cinéma. Bien que notre expérience de mixage nous fournisse déjà un nombre assez important de constats, nous pensons que de proposer un test perceptif à un panel de sujets pourra nous aider à prendre du recul sur les différents problèmes que nous avons rencontrés. Cela pourra

également nous permettre de recueillir d'autres avis perceptifs, élargissant ainsi notre vision du sujet.

4. Evaluation

perceptive de la WFS

au cinéma

Cette dernière partie permet de soumettre nos constats et nos hypothèses avancées dans les parties précédentes au regard et à l'écoute d'autres personnes. La mise en place d'un système WFS dédié au cinéma et son utilisation nous a permis de constituer un corpus de séquences mixées à la fois en 5.1 et en WFS que nous avons utilisées pour nos tests. Nous proposerons tout d'abord une contextualisation des études perceptives autour de la WFS, ce qui nous permettra de justifier nos choix concernant l'aspect formel des tests et de dégager les critères d'études envisagés. Les parties précédentes nous apportent suffisamment d'éléments de réflexion pour formuler deux hypothèses que nous chercherons à valider ou non au cours de cette étude. Nous apporterons également des précisions sur les six séquences du corpus.

L'analyse des tests fera l'objet d'une approche scientifique qui donnera lieu à une interprétation des résultats développée dans la partie finale, interprétation mise au regard de notre expérience de mixage.

4.1. Propos liminaires à la mise en place des tests

4.1.1. Contexte d'études

Dans notre approche de la WFS, la mise en place de tests perceptifs ne nous est pas d'abord apparu comme fondamentale puisque notre étude était centrée sur le travail du mixeur et son utilisation du système. En effet, cette démarche bornée à une vision technico-professionnelle pouvait, de notre premier point de vue, se passer d'une approche perceptive du sujet si elle se soumettait aux regards des concernés, c'est-à-dire des mixeurs. Il nous semblait pourtant que cette question de la perception était au cœur de notre problématique, qu'elle était à la fois nécessaire à la caractérisation d'un dispositif sonore mais, qu'avant tout, elle guidait de manière intrinsèque le travail du mixeur dont la finalité a vocation d'être « perçue » par un public. Ainsi nous ne pouvions nous contenter d'étudier ce système uniquement dans ses pratiques technologiques. De plus, notre travail serait incomplet si les hypothèses que nous avançons n'étaient pas soumises à une forme d'évaluation. Car, si c'est dans son fonctionnement que ce système nous est apparu, c'est bien vers le regard porté à l'objet achevé que nous entendions amener notre réflexion.

Il s'agissait alors pour nous d'évaluer l'apport de la WFS et de vérifier si la construction d'un espace sonore différent via l'utilisation d'un système qui, par nature, propose une autre approche d'illusion perceptive, influence notre perception d'un objet audiovisuel. Nous avons donc choisi d'envisager nos tests sous une forme comparative et ce pour deux raisons. Premièrement, l'évaluation

d'un système d'écoute nous semble bien plus aisée lorsque nous disposons d'éléments de comparaison connus. Deuxièmement, il s'agissait bien d'évaluer la WFS dans un cadre audiovisuel, domaine dans lequel le 5.1 est largement implanté, maîtrisé et semble, malgré tout, bien dimensionné pour répondre aux problématiques perceptives du son à l'image. Son usage codifié (qui résout sans doute une partie des problèmes perceptifs qui lui incombent, cf. 1.2) conduit également à une habitude d'écoute mais aussi à un langage propre à une écriture sonore particulière et commune. Il nous semblait alors intéressant d'envisager comment la WFS, sans prétendre redéfinir la création sonore au cinéma, pouvait étendre la palette d'effets actuels et comment cet apport était perçu, en ayant bien conscience qu'il ne fallait à aucun moment envisager une nouvelle machine à écrire le son et tenter de dégager des règles scénographiques figées, nous évitant ainsi toute théorisation issue d'expériences pratiques. Si notre démarche est inductive, elle se limite pour nous à l'étude des phénomènes de perception. Nous ne pouvons en effet prétendre à une recherche langagière qui implique à la fois un processus d'écriture et de lecture.

Il était nécessaire pour nous d'établir un corpus qui sans prétendre être exhaustif, permettait d'aborder divers aspects de la spatialisation au cinéma et mettait en jeu différentes matières sonores, différents espaces et différents genres. La sélection des séquences dépendait des critères d'études que nous devions pouvoir travailler sur l'ensemble du corpus. Le choix de ces critères était donc une condition préalable à l'établissement définitif de ce corpus, d'autant que les sujets participants aux tests devaient pouvoir juger de ceux-ci sur chaque séquence.

Un impératif technique exigeait également d'avoir à notre disposition les éléments sonores séparés. L'accès aux sessions de montage son permettait de pouvoir réellement mixer nos séquences en WFS et ne pas se contenter d'une

simple redistribution par front d'onde d'un mixage 5.1 figé (*upmix*), ce qui nous semblait brider considérablement l'évaluation du système.

Un test comparatif demande une certaine rigueur pour comparer efficacement ce qui peut l'être. Nous envisagions dans un premier temps d'utiliser les mixages 5.1 déjà existant et de ne repartir du montage son que pour le mixage en WFS. Cette approche nous permettait de gagner plus de temps sur le mixage en WFS, mais présentait un—inconvenient majeur : nous avions nécessité de proposer deux mixages cohérents l'un par rapport à l'autre au risque de biaiser les tests. Il nous fallait donc d'être en mesure d'adapter et de retravailler le 5.1 en fonction de ce que nous faisons ou de ce que nous voulions faire en WFS. La solution inverse (qui consistait à utiliser un mix 5.1 déjà fait) nous semblait limiter l'exploitation du potentiel du système. Cette solution n'aurait pas été honnête, dans le sens où nous n'aurions pu qu'accentuer les différences en mettant en valeur le rendu d'un système par rapport à un autre. De plus, nous voulions, pour certaines séquences, nous affranchir de la musique qui pouvait prendre le pas sur certaines subtilités sonores, et nous concentrer d'avantage sur d'autres matières. Il a donc été décidé d'effectuer deux mixages, l'un en 5.1, l'autre en WFS, avec à chaque fois la volonté d'exploiter toutes les possibilités des systèmes pour arriver à un résultat convaincant et comparable. Le lecteur trouvera des précisions sur notre travail de mixage dans 3.4.

4.1.2. Critères d'études

Les conclusions des parties précédentes nous permettent d'émettre deux hypothèses que nous chercherons à valider ou à rejeter.

Hypothèse 1 : La WFS propose de nouvelles possibilités de perspective sonore dans le cadre audiovisuel.

Il s'agit alors d'évaluer si la distance holophonique déjà constatée reste perçue et « résiste » au dispositif cinématographique proposant une image à deux dimensions.

Hypothèse 2 : Cette perspective sonore ne crée pas de conflits de perception audiovisuelle.

Dès lors que cette perspective reste perçue, est-elle compatible avec le contrat sous-jacent de la diffusion cinématographique ?

Ces hypothèses posées, il s'agit ensuite de définir plusieurs critères d'études proposant un angle d'approche permettant de valider ou non nos postulats.

Le panel des critères envisageables dans une étude portant sur la WFS au cinéma est large et la littérature à ce sujet fournit plusieurs éléments distincts. Certains critères semblent liés à une étude précise d'un phénomène propre à la WFS. Il s'agit, par exemple, de critères visant à déterminer les incohérences de position entre l'image et le son, cherchant à quantifier l'angle maximal admissible pour la distorsion de perspective [Melchior et al., 2003]. Ces études, d'ordre quasi physiologique ou structurelle, visent à étudier un aspect très spécifique et détaché, à première vue, d'un travail créatif (quoiqu'on puisse considérer qu'en

définissant des limites perceptives au couplage d'un système image/son, elles définissent également des limites esthétiques et des codes d'utilisation). Ces études mettent aussi en jeu des critères qui s'éloignent de notre approche. En effet, nombreuses d'entre elles utilisent des stimuli déconnectés du contexte au sein duquel nous prétendons nous placer, i.e. le son au cinéma. Elles présentent cependant un intérêt pour notre étude, puisqu'elles pointent dans leur ensemble, le problème général de la cohérence (ou des cohérences) entre un espace visuel défini par deux dimensions et un espace sonore qui semble désormais pouvoir s'exprimer en trois¹ dimensions.

Cette question de la cohérence est au cœur de notre sujet qui s'intéresse à la compatibilité perceptive d'un système WFS avec le spectacle audiovisuel. Elle nous permet à la fois d'étudier la « validité spatiale » d'un mixage en WFS, et peut mettre à jour les limites établies par le genre d'étude précédemment évoqué, mais dans le cadre de stimuli liés à un contexte cinématographique. Nous nous plaçons ici dans un cadre sémantique et, dans des situations cinématographiques, il est quasiment certain que les seuils déterminés en laboratoire soient moins pertinents, en partie à cause de l'effet de sens en jeu dans notre contexte d'étude. Par ce critère de cohérence, nous entendons ici la fusion entre le champ visuel et le champ sonore, entre l'action vue et entendue, entre les sensations visuelles et sonores.

De cette notion de cohérence découle un critère plus spécifique mais qui questionne presque autant les interactions entre image et son et leurs espaces de représentation : celui de la distance. L'ajout d'une sensation de profondeur à une

¹ Cette notion de « son 3D » est sujette à diverses interprétations (cf. 1.3). Notre système ne contenait pas d'enceintes zénithales. Nous pouvons considérer que la WFS ajoute une troisième dimension en rendant possible le placement virtuel en profondeur des objets sonores.

image établie sur un plan demande à être interrogé, d'une part sur sa pertinence (mais dans ce cas, le critère de cohérence semble plus approprié), mais également sur sa perception ou non. En d'autres termes, il s'agit de vérifier si cette profondeur permise par la WFS (et constaté en partie 2) « résiste » lorsqu'elle est associée à un dispositif cinématographique. Par dispositif nous entendons à la fois, le cadre de projection d'un objet audiovisuel (l'écran), mais également le contexte de production sonore que constitue le mixage, associant et mélangeant divers matériaux sonores. « *Même si le sonore est analysable, démontable en ses constituants élémentaires, c'est en masses, en ensembles, en compréhension dirait le mathématicien moderne, qu'il sera joué* » avance Daniel Deshays². Nos tests et nos constats effectués chez Sonic Emotion portaient uniquement sur des sources sonores isolées et nous n'avons aucune expérience en ce qui concerne le mélange de plusieurs sources sur plusieurs plans.

Pour évaluer cette notion de distance, nous avons choisi de regrouper deux termes liés à la perception de l'espace : la profondeur et le relief. Nous définissons ici la profondeur comme l'impression de distance entre l'auditeur et le dernier plan sonore, et le relief comme la possibilité de distinguer plusieurs plans sonores dans l'espace. S'il est possible de concevoir, à titre de contre exemple, des scènes sonores sans relief mais avec beaucoup de profondeur (et inversement), ces critères sont en pratique relativement liés.

Par ailleurs, il nous semblait intéressant de soumettre à nos sujets un critère qui semble en apparence plus subjectif et moins analytique : celui de la préférence. Il n'y a pas de lien direct, a priori, entre la perception d'une scène sonore cohérente et/ou profonde, et la préférence du spectateur. De par ce critère de préférence, nous ne cherchons en aucun cas une manière de statuer une forme de validité

² Deshays Daniel, *Pour une écriture du son*, Klincksieck, Paris, 2006, p.41.

narrative ou émotionnelle (notion finalement assez étrange et difficilement analysable), mais plutôt d'obtenir des informations sur ce qui pouvait échapper aux deux premiers critères.

D'autres critères ont été envisagés dans notre étude, en particulier celui de l'immersion ou de l'enveloppement [Moulin, 2011]. Si ce terme est aujourd'hui récurrent dans la création sonore au cinéma, nous pensons que son usage reste malheureusement conceptuellement lié à l'utilisation des surrounds et des niveaux forts visant à donner la sensation aux spectateurs d'être plongés dans la scène, dans une forme d'exaltation sensoriel. Nous nous interrogeons donc sur la validité de cette notion d'immersion ainsi conceptualisée et nous nous demandons si l'immersion au cinéma ne passe pas aussi par la création d'une forme de perspective dans l'image amenant le spectateur dans l'univers du film. Sans être redondant, le critère de cohérence paraît également pouvoir englober (du moins partiellement) cette notion d'immersion. En effet, l'immersion étant souvent quantifiée par la balance son frontal/son surround, on peut supposer qu'une plus grande cohérence à l'image sera obtenue si le spectateur n'est pas trop attiré par des éléments sonores diffusés derrière lui. Ce critère de cohérence regroupe également d'autres notions telles que la dissociation spatiale (sensation que le son et l'image ne sont pas au même endroit) et le critère d'adhésion (impression que la scène est réaliste en terme de rapport image/son). D'autres critères tels que le degré de spatialisation [Moulin, 2011] et l'impression de distance [Noguès et al., 2003] auraient pu également se substituer à celui de la profondeur et du relief, mais nous avons fait le choix de ces deux dernières notions qui nous semblent moins obscures et plus imagées que les précédentes.

Il nous semblait également intéressant d'apporter une autre composante à notre recherche concernant la zone d'écoute étendue. Dans le cadre du spectacle vivant, les compétences de la WFS à ce sujet ne sont plus à prouver. Cependant, la

question se repose entièrement dans le cadre d'une diffusion associant un système à synthèse de fronts d'ondes à un contexte cinématographique. Il nous semblait logique d'introduire alors une variable intra-sujet, en instaurant deux positions d'écoute : une première centrée et une autre excentrée. La zone d'écoute élargie ne fait pas partie de nos critères d'études à proprement parlé. Cette variable permettra d'analyser les résultats obtenus sur nos critères d'études (profondeur-relief, cohérence, préférence) selon des points d'écoutes différents mais ne permettra pas de statuer sur l'élargissement de la zone d'écoute éventuel dans notre contexte.

Ces critères d'études proposent une approche non-exhaustive des différents effets de perceptions inhérents à la WFS associé à une diffusion cinématographique. En sois, l'étude perceptive que nous menons ne permet pas d'aborder l'ensemble des effets constatés lors de nos travaux préalable. Elle constitue une extension de ce que nous avons déjà constaté et permettra de valider ou non certaine tendances.

4.1.3. Constitution et description du corpus

Le choix des critères d'étude nous a amené à la constitution d'un corpus de six séquences, extraites de six films différents. Ces séquences présentaient dans leur réalisation un cadre de travail propice à l'exploration de plusieurs aspects de la spatialisation au cinéma, et nous permettaient d'appréhender l'apport d'un système WFS au regard de nos trois critères d'étude et de nos hypothèses de départ. Ces séquences sont extraites de films sur lesquels nous avons travaillé personnellement. Nous en proposons ici au lecteur un bref descriptif narratif et technique, afin que la suite de notre rédaction ne paraisse pas trop en dehors de notre démarche.

1) *Mundos Immundos*, 2014, 2'10, Documentaire

Trois jeunes chiliens marchent au bord d'une falaise, dans un décor aride, où se dressent de vieilles maisons en bois sans doute abandonnées. La séquence est composée de deux plans fixes dont nous avons réduit la durée pour le test. Dans le premier plan, les trois personnages entrent un à un, gauche cadre, marchent sur un petit chemin de terre bordé d'une végétation sèche et s'éloignent vers une des maisons. On aperçoit une mer forte et agitée au loin, sur la droite du cadre. Dans le second plan, ils gravissent une colline rocheuse en haut de laquelle se dresse une autre maison. La mer est fortement agitée et très présente au cadre, avec de grosses vagues frappant les rochers.

Pour cette séquence, nous disposons de plusieurs ambiances de vent et de mer avec des textures différentes. Nous avons également plusieurs éléments

ponctuels renforçant les vagues les plus importantes, quelques insectes et oiseaux. Les deux plans décrivant un espace immense au milieu duquel évoluent des corps (dont la présence devient de plus en plus dérisoire), cette séquence nous semblait être l'occasion en particulier d'explorer la construction d'espace, de profondeur et de relief, ainsi que la précision de la localisation permise par la WFS.



Fig. 4.1 : Mundos Immundos, de Charlotte BAYER-BROC, photogramme.

2) *Petite Bête*, 2014, 1'05, Fiction

Deux enfants jouent à un jeu vidéo sur une grande tablette tactile. La mère apporte le goûter qu'elle pose sur la table et demande à un des enfants d'enlever son manteau. La séquence est composée de trois plans fixes. Le premier plan est un insert sur les doigts des enfants jouant sur la tablette. Le second est un plan rapproché d'un des garçons vexé d'avoir perdu la partie. Le troisième plan, plus

large, nous permet de voir l'ensemble du salon où rentre la mère par la droite du cadre, prend le manteau de l'enfant, et quitte le champ par la droite également.

Pour cette séquence nous disposons du montage des directs, composé d'une perche et de plusieurs HFs, ainsi que de plusieurs ambiances permettant de créer l'espace du salon. La scène, plutôt simple, pouvait nous permettre d'utiliser la WFS sur la voix de la mère afin de la placer en profondeur sur le plan large, et d'évaluer la pertinence des fronts d'onde sur la distribution énergétique des effets de réverbération. L'extrait proposant deux valeurs de plan (gros plan sur l'enfant, plan large de l'ensemble du salon avec la mère), nous pouvions exploiter l'hybridation du système en laissant les voix des plans serrés dans l'enceinte centrale, et celles du plan large en WFS.



Fig. 4.2 : Petite Bête, de Margaux REMAURY, photogramme.

3) La Menace d'une Rose, 2014, 1'45, Fiction

A Londres, dans une zone industrielle, un homme monte dans sa voiture pour rejoindre un autre personnage situé à quelques mètres de lui. Les deux hommes échangent un bref dialogue avant de monter dans une autre voiture et de partir. Le premier plan est un plan d'ensemble d'un port le long de la Tamise, avec des bâtiments et une activité urbaine au loin. Les deux derniers plans nous montrent la voiture s'engouffrant dans la City puis passant dans une zone pavillonnaire.

Le montage son nous donnait accès aux sons directs, ainsi qu'à plusieurs ambiances urbaines variées, composées entre autres d'éléments ponctuels (bruits métalliques, sirène de bateau, ...). Ces éléments ponctuels étaient directement liés au plan ou renvoyaient au hors champ. Cette séquence nous donnait la possibilité d'utiliser la WFS sur les déplacements de source (ici la voiture), sur le placement d'objets sonores liés au hors champ, sur les effets d'espace via l'utilisation de réverbérations extérieures, et sur l'échelonnement des plans sonores par le positionnement des ambiances.



Fig. 4.3 : *La Menace d'une Rose*, de Thomas LHERMITTE, photogramme.

4) *Couper le voile*, 2014, 1'30 Fiction

Dans une rue calme, la nuit, un homme termine sa cigarette sur le seuil de son immeuble lorsqu'il entend sa femme crier par la fenêtre, appelant son chien qui s'est visiblement enfuit. Au même moment arrive une voiture, dont le conducteur sort précipitamment pour empoigner le passager, et le sortir violemment de son véhicule.

Nous disposions du montage des directs ainsi que de plusieurs éléments sonores captés lors du tournage. Les ambiances, composées de rumeurs de villes, permettaient la construction d'un tapis sonore que nous pouvions placer en profondeur, pour créer du relief entre l'action et l'environnement sonore lointain. Cela nous semblait d'autant plus pertinent que la séquence met en scène l'irruption de plusieurs événements dans le calme d'une ville endormie. Le placement des cris de la femme en hors champ et l'arrivée de la voiture pouvaient se préciser par usage de la WFS. Les éléments dynamiques tel que les cris du conducteur, le claquement de la portière, et le redémarrage de la voiture, nous semblaient pouvoir être associés à une réverbération de rue que nous pouvions placer en WFS pour apporter une cohérence à l'espace.



Fig. 4.4 : *Couper le voile*, de Maximilien WILLMOCH, photogramme.

5) *Louis Le Manchot*, 2014, 1'25, Fiction

Pendant la seconde guerre mondiale, un groupe de résistant attend l'arrivée nocturne d'un avion venu déposer des colis. Les hommes assis au coin du feu se lèvent et s'agitent lorsqu'ils entendent l'avion arriver. L'engin passe deux fois. Des colis tombent au loin et les hommes courent dans leur direction.

La séquence mettait en jeu les sons directs déjà emprunts d'une rumeur d'insectes nocturnes. Ces directs étaient complétés par du bruitage, particulièrement utile sur les chutes de colis. A cela s'ajoutait des ambiances de champ la nuit, composées de chants de grillons et d'une rumeur légère, ainsi que des pistes d'effets pour les passages de l'avion. L'extrait nous semblait intéressant à envisager en WFS pour plusieurs raisons. La mise en espace de la réverbération extérieure associée aux voix, et la mise en relief des ambiances, permettaient selon nous de sculpter l'espace et de placer ces hommes au milieu de nulle part,

dans leur solitude. Le montage son découpé au plan permettait de jouer différentes largeurs d'ambiances en fonction des échelles de plan. Le plan serré sur le visage du personnage principal pouvait, par exemple, proposer moins d'espace hors champ, et ainsi se recentrer sur le jeu du comédien. Enfin, nous comptons sur cette séquence pour utiliser la WFS sur les passages d'avions. Si l'effet tient au spectaculaire, il nous semblait intéressant de donner de la distance à cet avion, distance qui par défaut pouvait s'assimiler à de la hauteur. Les regards des personnages suivant les déplacements de l'engin dans le ciel (sans que jamais il ne soit montré), l'éventuelle fluidité de déplacement de l'objet pouvait apporter un autre degré de réalisme à la séquence.



Fig. 4.5 : Louis Le Manchot, de Camille JAULENT, photogramme

6) Gabriella, 2014, 3'50, Fiction

Sur une île méditerranéenne, une jeune femme fugue de chez elle. Elle traverse une forêt étrange et lugubre avant de rejoindre le bord de mer, prison

naturelle et infranchissable. Plusieurs plans de coupe nous dévoilent l'espace montagneux et sauvage de l'histoire. Une voix off, hispanophone, nous narre le parcours intérieur de ce personnage condamné sur cette île. De retour chez elle, son père l'enferme entre les murs de sa chambre : nouvelle prison plus étroite. La fin de la séquence se tourne vers l'onirisme et alterne entre des plans de la jeune fille au visage fiévreux, les yeux rivés vers l'extérieur éclairés par le soubresaut lumineux d'éclairs dans le ciel, des plans de son père jouant seul aux cartes, des plans du village, et des plans de papillons agglutinés sur une ampoule en stop motion.

Le montage son proposait une voix-off sur une piste dédiée, des directs renforcés par du bruitage, et un montage d'ambiances et d'effets lié aux différents environnements que propose la séquence. Beaucoup de ces éléments étaient liées à la nature et certains d'entre eux tendaient à rendre l'univers intrigant et fantastique. Des aboiements de chiens constituaient le leitmotiv de cette séquence et pouvaient symboliser à la fois la peur, l'enferment et l'autorité du père. Cette prise de son était l'œuvre de l'ingénieur du son sur le plateau.

Cette séquence, la plus longue de notre corpus, était pour nous l'occasion d'explorer la tenue d'un mixage en WFS sur une durée plus importante, ainsi que le passage entre différents univers sonores. La narration ayant le temps de s'installer, nous mettions à l'épreuve les subtilités de la WFS face à l'emprise du récit. La première partie de l'extrait nous permettait de synthétiser globalement l'ensemble des techniques mises en œuvre précédemment. La dernière séquence, de par son détachement au réel, nous permettait d'envisager un placement de sources plus libre, et d'évaluer le potentiel créatif ouvert par la WFS lorsqu'il ne s'agit plus de « coller » à l'image de manière réaliste, mais au contraire de s'en détacher et d'en proposer un autre rapport.

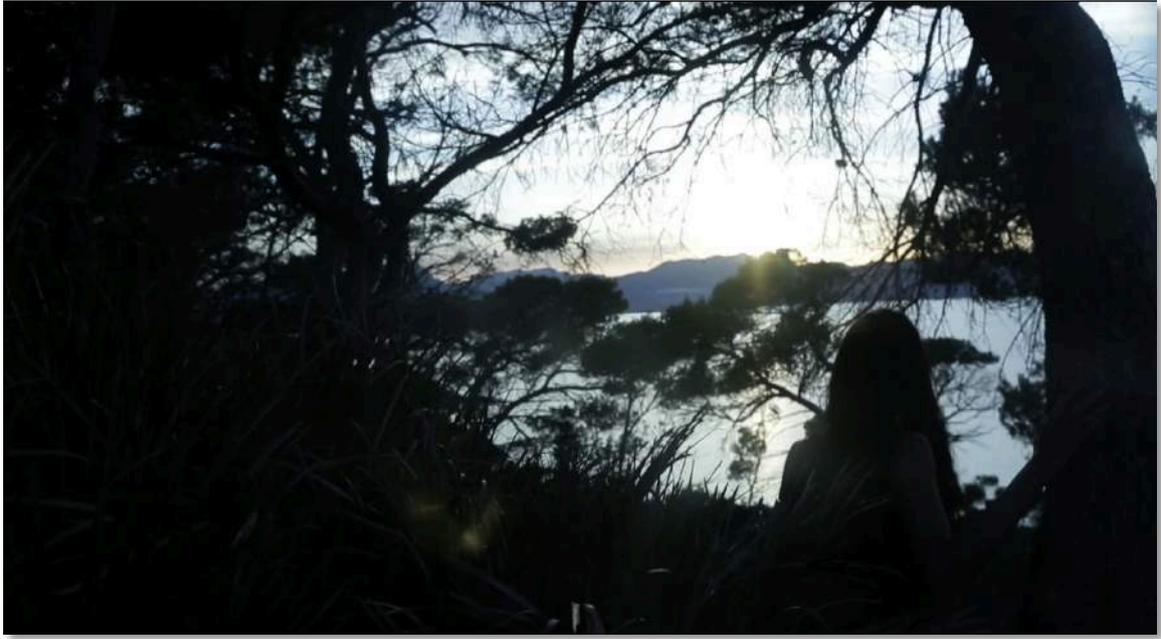


Fig. 4.6 : Gabriela, de Anna LEYSENS, photogramme.

4.2. Protocole, analyse et discussion

4.2.1. Protocole de tests

Les tests se sont déroulés dans l'auditorium de mixage de Louis Lumière. La configuration d'écoute et de projection est identique à celle décrite dans 3.3.

Chaque session de test mettait en place deux séries de projection. Dans chaque série les six séquences et leurs deux mixages associés ont été projetés. Chaque séquence était donc projetée deux fois de suite, une fois en WFS, une fois en 5.1. L'ordre des séquences au sein d'une série était aléatoire, ainsi que l'ordre de diffusion des mixages pour une même séquence. Pour la deuxième série de projection, l'ordre des mixages était inversé par rapport à la première série. Cette organisation nous permettait de minimiser l'effet d'ordre et l'influence des réminiscences de la première écoute sur la seconde. Le lecteur trouvera en annexe l'ordre des séquences pour chaque session de test. Un exemple est donné ici à titre indicatif :

SESSION I

SERIE 1

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	4/ COUPER LE VOILE	5.1	WFS
2	3/ LA MENACE D'UNE ROSE	WFS	5.1
3	5/ LOUIS LE MANCHOT	5.1	WFS
4	6/ GABRIELA	5.1	WFS
5	2/ PETITE BETE	WFS	5.1
6	1/ MUNDOS	WFS	5.1

SERIE 2

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	2/ PETITE BETE	5.1	WFS
2	1/ MUNDOS	5.1	WFS
3	4/ COUPER LE VOILE	WFS	5.1
4	3/ LA MENACE D'UNE ROSE	5.1	WFS
5	6/ GABRIELA	WFS	5.1
6	5/ LOUIS LE MANCHOT	WFS	5.1

Le sujet devait évaluer relativement la première et la deuxième projection de chaque séquence (c'est-a-dire l'une par rapport à l'autre) selon trois critères que nous rappelons ici :

- Profondeur et relief
- Cohérence
- Préférence

Afin que les sujets évaluent bien les mêmes notions, une note descriptive définissant chaque critère est lue avant le début des tests. A la fin de chaque séquence, le sujet dispose alors d'une minute pour évaluer les trois critères et ajouter des remarques éventuelles. Chacun de ces critères est noté sur une échelle non graduée, proposant uniquement trois repères : deux pour les extrémités de l'échelle, un pour le centre. Le sujet est invité à indiquer sur cette échelle l'extrait qui lui a semblé le plus représentatif du critère étudié. Par exemple, pour le critère de profondeur et de relief, si le sujet place une croix à gauche c'est que l'extrait 1 lui a semblé proposer plus de profondeur et de relief que l'extrait deux. Plus sa croix est située à gauche, plus la différence entre les deux extraits lui a semblée importante. Une consigne en préambule du questionnaire indiquait la manière de remplir les échelles.

Deux positions d'écoute ont été définies, l'une centrée (position A), l'autre excentrée (position B). Chaque sujet se voyait attribuer une place pour la durée totale du test.

Limites a priori du protocole

Effet de fatigue :

Le test, de part sa longueur (environ 55 minutes pour 30 d'extraits), son aspect répétitif, et l'attention qu'il demande, engendre une fatigue pouvant générer d'autres effets par répercussion :

- La longueur ressentie entraîne un phénomène de lassitude, l'attention étant bien moins certaine à la fin, et il sera sans doute de plus en plus difficile de répondre à la question liée au critère de préférence.
- Il y a un phénomène d'habituation aux différents extraits au cours du test. Ainsi, une fois la connaissance des extraits acquise, on se focalisera de plus en plus sur l'aspect analytique (cohérence et profondeur) en oubliant la préférence. Le jugement de ce critère risque donc d'évoluer, en particulier entre les deux séries.

Difficulté de jugement :

Les sujets doivent combiner à la fois une écoute analytique et une écoute plus passive. Les critères de profondeur et de relief demande une écoute précise de certains éléments spatialisés. Le critère de cohérence demande lui d'analyser à la fois la séquence sur des points précis mais fait aussi écho à un sentiment général moins analytique.

Contrastes et disparités entre les deux mixages :

Les sujets devaient comparer le rendu de deux mixages, diffusés respectivement sur deux systèmes différents. Lors de nos sessions de mix, il nous est rapidement apparu que le rendu des sources sonores était différent entre les deux systèmes. Le rapport qu'impose la WFS à la distance, à la spatialisation, mais aussi au timbre,

est a priori différent du 5.1. C'est évidemment cette différence que nous cherchons à évaluer et il était entendu que les deux mixages seraient contrastés. Dans nos mixages, nous avons donc cherché à exploiter ces deux approches de spatialisation. Néanmoins, un certain nombre de choix (largeur apparente, qualité spectrale, balance) devaient demeurer communs aux deux systèmes, choix qu'il fallait reporter d'un mixage à l'autre. Pour des raisons de temps et des questions techniques, ce report n'a pas pu être fait de manière aussi précise que nous l'aurions souhaité, créant ainsi des disparités d'intention esthétique (mais non intentionnelle d'un point de vue technique). Avec un peu plus de temps, il aurait été possible d'homogénéiser d'avantage nos mixages pour évaluer des différences perceptives uniquement liées à nos critères d'étude.

Ordre de projection :

Le choix d'inverser systématiquement l'ordre des mixages entre les deux séries induit également un facteur de reconnaissance chez le sujet. Un sujet qui identifie ce procédé peut être tenté de reporter ses notes de la première série sur la deuxième. Conscient du problème, nous avons malgré tout fait ce choix qui nous permettait d'obtenir des résultats affranchis de l'effet d'ordre.

4.2.2. Analyse statistique des résultats

Vingt personnes ont participé à ce test. Toutes étaient sensibilisées au domaine du son. Le protocole mis en place nous permet de mener une analyse sur trois facteurs selon trois variables indépendantes (critères de profondeur-relief, cohérence à l'image, préférence). Pour nous placer dans l'hypothèse d'une distribution normale et effectuer ainsi une analyse de la variance (ANOVA), les résultats obtenus sont normalisés.

La variable **séquence** (S) et la variable **répétition** (R) sont définies comme des variables inter-sujets. Cela signifie qu'elles sont partagées par tous les sujets. La variable S nous permet d'analyser l'influence des différentes séquences proposées sur nos trois variables indépendantes. La variable R nous informe de l'influence de la répétition du test (deux fois) sur ces mêmes variables. La variable **groupe** (G) est une variable intra-sujet : elle n'est pas partagée par tous les sujets. En effet, certains d'entre eux étaient positionnés au centre, tandis que d'autres étaient excentrés. Cette variable G quantifie l'influence de la position des sujets sur les réponses données, toujours selon nos trois variables indépendantes, i.e. nos trois critères d'étude. Nous serons donc en mesure de connaître, à chaque fois, l'influence de S, R, et G sur l'évaluation de nos 3 critères de départ.

Le calcul de la valeur du seuil de significativité (p) pour chaque facteur (S, R ou G) nous permet de savoir si ces facteurs ont une influence significative sur chacune des variables indépendantes (profondeur-relief, cohérence à l'image, préférence). On considère que si la valeur de p est inférieure à 0,05, elle est significative pour le facteur concerné, c'est-à-dire qu'elle ne tient pas au hasard. De la même manière si ce facteur s'avère significatif pour la variable indépendante étudiée ($p \leq$

0,05 pour la préférence par exemple), on calcule une valeur de p pour chaque variable de ce facteur (la séquence, la répétition, et le groupe). On vérifie alors si les résultats obtenus sont significatifs pour toutes les variables de ce facteur. Par exemple, si le facteur S (séquence) s'avère significatif pour un critère d'étude, on calcule un p pour chaque séquence. Cependant, notons bien que l'ANOVA ne nous permet d'évaluer qu'une seule chose : les différences perçues par nos sujets sont-elles significatives et ne tiennent-elles pas au hasard ? Si des différences sont perçues de manière significative, nous ne pouvons pas évaluer le niveau de cette différence. Par exemple si pour le critère de cohérence à l'image, les sujets du groupe B ont préféré de manière significative la WFS, on ne pourra affirmer s'il s'agit d'une préférence clairement marquée, ou d'une tendance légère.

Profondeur et relief :

facteurs	p
R	0,437
S	0,03
G	0,361

Pour la variable indépendante profondeur et le relief, seul le facteur S nous donne une valeur significative ($p = 0,03$). Nous obtenons alors des valeurs significatives pour quatre de nos six séquences soumises au sujet.

	p	Mean Difference
S1	0,883	-0,055
S2	0	1,735
S3	0,392	0,3425
S4	0,021	0,945
S5	0,04	0,8375
S6	0,004	1,01

Sur les séquences 1 et 3, il semble que les sujets n'aient donc pas perçu de différence significative de profondeur et de relief entre la diffusion en 5.1 et la diffusion en WFS. On obtient cependant des différences significatives pour les séquences 2, 4, 5 et 6. Pour ces quatre séquences, cette différence est significativement supérieure à 0 (*Mean Difference*) ce qui indique que les sujets ont ressenti d'avantage de profondeur et de relief sur le format WFS que sur le format 5.1³.

La valeur de p pour le facteur groupe (G) est de 0,361. On peut donc en déduire que la position des sujets n'a pas influencé leur perception de la profondeur et du relief et qu'ils ont donc globalement répondu de la même manière qu'ils soient situés en position A ou B.

La valeur de p pour le facteur répétition (R) est de 0,437. La répétition (R) du test n'a, semble-t-il, pas influencé nos sujets, qui ont globalement répondu de manière cohérente au critère de profondeur et de relief sur les deux séries de projections.

³ Notre échelle allant de -5 à 5 et étant centrée sur 0, nous avons adopté la stratégie suivante dans le report des notes : la WFS était notée en positif, alors que le 5.1 correspondait à une note en négatif. Ainsi, une moyenne positive indique un choix pour la WFS, une moyenne négative un choix pour le 5.1.

Concrètement cela signifie que l'ordre de projection des séquences, ainsi que l'ordre de projection du 5.1 et de la WFS au sein de chacune de ces séquences, n'ont globalement pas influencé les réponses de nos sujets sur leur perception de la profondeur et du relief : ils ont eu la même stratégie de réponse. Globalement, si un sujet a perçu, lors de la première série, que la séquence 1 proposait plus de profondeur et de relief sur la première projection -qui correspondait alors au format WFS-, son choix s'est également porté sur ce format lors de la série 2, peu importe l'ordre de projection des formats, et l'ordre de diffusion des séquences.

Cohérence à l'image :

facteurs	p
R	0,595
S	0,009
G	0,098

De la même manière que pour le critère de profondeur et de relief, nous n'obtenons pas de valeurs significative de p pour les variables de répétition et de groupe. Cela signifie que les réponses des sujets pour la cohérence ne dépendent pas de leur position et qu'ils ont adopté la même stratégie de réponse pour les deux séries.

La valeur de p pour S est de 0,009. Cela signifie que comme pour le critère de cohérence, les sujets ont perçu des différences entre les deux formats en fonction des séquences.

	p	Mean Difference
S1	0,166	0,5125
S2	0,588	0,225
S3	0,002	-1,42
S4	0,21	0,525
S5	0,611	-0,2175
S6	0,129	-0,52

En regardant en détail, il n'y a finalement que sur la séquence 3 où nous observons une influence du système sur la perception de la cohérence à l'image. Généralement, les sujets ont trouvé que le mixage 5.1 était plus cohérent avec l'image pour cette séquence (*mean difference* = -1,42).

Préférence :

facteurs	p
R	0,805
S	0,078
G	0,024

On observe uniquement une influence du groupe pour le critère de préférence. La séquence et la répétition n'ont apparemment pas d'influence significative sur la préférence de la WFS ou du 5.1.

	p	Mean Difference
G1	0,001	-0,83667
G2	0,755	0,07667

Il n'y a que pour le groupe 1 qu'on observe une préférence significative d'un système par rapport à l'autre. Globalement, les sujets placés au centre ont préféré le système 5.1.

4.2.3. Interprétation et discussion

L'analyse statistique des résultats nous permet de dégager doré et déjà plusieurs informations concernant les différences perçues entre un système 5.1 et un système WFS. Cependant il nous semble important d'analyser ces résultats sous le prisme de notre travail et des particularités des séquences que nous avons mixées. Il s'agit d'apporter au lecteur des informations complémentaires permettant une interprétation pratique des résultats obtenus, et de développer la discussion débutée en 3.4.4 sur les particularités du mixage en WFS d'un point de vue perceptif. Nous disposons également de nombreuses remarques écrites fournies par les sujets.

Influence du facteur de répétition

Concernant tout d'abord le facteur de répétition (R), il semble que ce facteur n'ait aucune influence sur les réponses données par nos sujets et ce pour les trois critères qu'ils avaient à évaluer. Nos doutes concernant la trop longue durée du test, qui aurait pu générer une fatigue d'écoute et influencer les résultats, semblent levés, bien que cette longueur fasse fréquemment l'objet de plusieurs commentaires chez les personnes ayant participé au test. Certaines d'entre elles ont parfois signalé qu'elles avaient été distraites ou avaient été moins concentrées pour une séquence particulière. C'est un facteur d'erreur qu'il faut prendre en compte dans la valeur donnée à nos résultats.

Profondeur et relief

Nous obtenons des résultats très intéressants sur la perception de la profondeur et du relief. A de nombreuses reprises, les sujets ont signalé une plus grande clarté de l'extrait, une écoute agréable et ont évoqué des sensations de relief supplémentaires en WFS. Très souvent, c'est sur des éléments ponctuels que leur attention se focalise, le travail fait sur les ambiances plus diffuses étant moins souvent mentionné. Sur *Gabriela*, la plus longue séquence du corpus, la perspective semble être toujours perceptible, bien que la narration puisse s'installer et que l'attention du spectateur soit moins focalisée sur le son et son espace associé. En termes de profondeur et de relief, il n'y a que sur les séquences 1 et 3 (*Mundos Immundos* et *La Menace d'une Rose*) que nos sujets n'ont globalement pas perçu de différence entre le 5.1 et la WFS. C'est un constat que nous avons déjà personnellement fait lors de nos sessions de mixage et nous nous attendions à ce genre de résultats.

La séquence 1 était en effet une séquence délicate à mixer, tant en 5.1 qu'en WFS (cf. 3.4). Par nature, mettre en place de la profondeur et du relief dans ce genre de séquence demande un travail fréquentiel délicat que nous n'avons pas eu le temps de mener de manière idéale. De plus la séquence proposait des plans où la mer est à la fois extrêmement présente au cadre mais aussi très éloignée en terme de distance apparente. Il fallait donc réussir à accompagner sa présence visuelle, sa puissance, sans pouvoir s'appuyer sur des éléments sonore aqueux trop riches dans le haut du spectre, au risque de ressentir une trop forte proximité qui ne correspondrait pas à l'image. Dans ce cas de figure -un cadre visuel avec beaucoup de profondeur de champ mais des matériaux sonores complexe- la forme du front d'onde ne semble pas suffisante pour localiser les sons et percevoir d'avantage les différents placements d'objet dans la profondeur. Notre oreille semble dans cette séquence, plus sensible au timbre pour extraire des éléments de spatialisation dans un environnement fréquentiellement complexe. En ce sens, lors du mixage

en 5.1, nous obtenions des résultats de spatialisation très satisfaisants et la WFS ne nous a pas permis d'amener un niveau supplémentaire de profondeur et de relief suffisamment perceptible en présence de l'image.

Le timbre de ces ambiances larges étant affecté par la mise en espace WFS, cette séquence était sans doute une de celle qui présentait le moins d'éléments spatialisés par front d'ondes de tout notre corpus. Plutôt que de vouloir exploiter de manière déraisonnable le système, nous avons préféré rediriger certains sons vers le 5.1 et mettre en place un mixage cohérent et d'une qualité aussi satisfaisante que possible. Pour ces raisons, les résultats obtenus sur la profondeur et le relief nous paraissent cohérents avec les sensations éprouvées et les difficultés rencontrées lors du mixage.

La séquence 3 (*La Menace d'une Rose*) présentait le même genre de problématique que la séquence 1, à un niveau moindre. La spatialisation en WFS des ambiances urbaines, assez diffuses, ne semble pas avoir été perçue. Cette séquence présentait également des éléments ponctuels hors-champs dont la synthèse par front d'onde n'a pas apporté de sensation de relief et de profondeur supplémentaire par rapport au 5.1. Cela nous surprend d'avantage, car la profondeur et le relief que nous avons constaté lors du mixage n'ont pas été perçus de manière significative par nos sujets. Cependant, les résultats obtenus pour cette séquence 3 doivent être considérés avec prudence pour des raisons que nous évoquerons dans les paragraphes suivants.

Sur les quatre autres séquences, les sujets ont perçu plus de profondeur et de relief sur le format WFS. C'est en effet un aspect de la spatialisation qui semble bien ressortir de nos tests. La perception d'une perspective ne semble pas dépendre de la position d'écoute puisque le facteur G n'est pas significatif sur ce critère d'étude. C'est un résultat qui tend également à confirmer l'augmentation de

la largeur de la zone d'écoute en WFS, les sujets ayant globalement perçu plus de profondeur et de relief sur ce système, peu importe leur position. Cependant cette perception de la perspective semble être accrue pour certains matériaux sonores.

La voix et l'espace

Les remarques des sujets sur les feuillets de réponse nous apportent des éléments complémentaires de réflexion autour de la perspective de certains éléments distincts. La séquence 2 (*Petite bête*) fait l'objet de remarques récurrentes concernant la voix de la mère. Cette séquence était la plus épurée de notre corpus, avec des directs, des ambiances légères et quelques bruitages simples. Beaucoup de commentaires font état d'une sensation d'espace plus importante autour de la voix de la mère, d'une sensation plus réaliste du lieu et de son acoustique. Certains évoquent cette sensation réaliste tout en avouant avoir été perturbés par une localisation précisée de la source. Il ne s'agit pas pour nous de statuer sur la validité de notre mixage, et les éléments dont nous disposons ne nous permettent pas d'aller plus loin dans notre conclusion. Cependant il semble que le cadre de cette séquence 2 soit d'avantage propice à la perception d'une forme de perspective, qu'elle soit valide ou non. C'est un constat que nous avons effectivement fait lors du mixage. Dans cette scène sonore simple, épurée, comportant des matières bien distinctes, la perspective se dégage plus facilement. En ne spatialisant en WFS qu'un seul élément, mis au regard des autres matières envoyées sur les canaux discrets, il est plus facile de sentir et de mettre en avant une distance apparente sur cet objet. De plus, la voix, élément sémantique dans un cinéma « verbo-centré », est l'objet d'attention principale du spectateur, ce qui explique sans doute les nombreuses remarques la concernant.

On retrouve des commentaires similaires sur les séquences mettant en jeu des éléments vocaux spatialisés en WFS. C'est le cas de la séquence 4 (*Couper le voile*), où une voix féminine hors-champs a été synthétisée par front d'onde. Les commentaires soulignent une sensation accrue d'espace et de profondeur pour la WFS, ce qui est cohérent avec les résultats obtenus (cf. 4.2.2). A l'inverse de la séquence 2 (*Petite bête*), où la synthèse de front d'ondes semblait, de manière générale, être bien reçue, les remarques sur la séquence 4 (*Couper le voile*) sont plus nuancées. Certains sujets ont ressenti un défaut de localisation, signalant que la voix est bien perçue en profondeur mais pas à l'endroit où elle devrait se situer logiquement, c'est-à-dire en hauteur. Nous rappelons que notre système ne permettait pas une spatialisation zénithale des objets. Ce constat fait écho à une sensation que nous avons ressentie en mixant nos séquences et que nous avons déjà évoquée : en précisant la distance holophonique d'un objet par la forme des fronts d'onde, on attend une localisation géométrique précise de cet objet.

Des remarques rejoignent nos impressions concernant la voix-off dans la séquence 6 (*Gabriela*). C'est un matériau que nous avons déjà abordé lors de nos tests chez Sonic Emotion (cf. 2.2.1). Lorsque la voix-off est placée dans l'espace WFS, sans être trop « profonde » mais suffisamment éloignée pour mettre à contribution plusieurs haut-parleurs frontaux, on perd la sensation de « point de diffusion » localisé dans l'enceinte centrale. Le système s'efface, la voix perd sa matérialité et apparaît dans un espace, qui n'est pas celui de l'image à proprement parler. Le caractère omniscient nous semble renforcé. C'est un raisonnement très empirique et qui n'engage que nous, mais nous trouvons cette possibilité très intéressante dans la construction d'une bande son. Il semble que cet effet ait été perçu, plusieurs remarques pointant une différence entre le 5.1 (où la voix était diffusée dans l'enceinte centrale) et la WFS, mais notre conclusion sur ce procédé s'arrête ici. En dehors de notre avis personnel, les résultats que nous obtenons ne nous permettent pas d'évaluer l'apport esthétique de ce procédé.

L'hypothèse exposée en 4.1.2, concernant la perception de la profondeur et du relief malgré la présence d'une image à deux dimensions, semble donc validée. Nous ne pouvons évidemment pas décrire exhaustivement tous les contextes qui constituent un cadre propice au renforcement de cette sensation. Nous avons cependant pu constater que dans des environnements sonores complexes, sur des matériaux larges, la perception d'une forme de perspective semble plus délicate. C'est sur des événements singuliers, qui focalisent l'attention du spectateur, que la WFS trouve un cadre d'expression et de mise en application efficace.

La question de la cohérence

Les sujets n'ont globalement pas perçu de différences entre la WFS et le 5.1 concernant ce critère et ce, quelle que soit leur position. La théorie nous avait amené à envisager qu'une plus grande cohérence aurait pu être perçue sur le système WFS pour les sujets excentrés. Cela n'est pas le cas ou, du moins, ce n'est pas ce qu'avancent nos résultats.

Attachons nous d'abord à décrire les résultats obtenus pour la séquence 3 (*La Menace d'une Rose*), seul stimulus où nous obtenons une valeur significative concernant la cohérence. Dans cette séquence, deux passages de voiture ont été spatialisés. Il s'agissait de déplacer le son en fonction de la position du véhicule dans l'image. Pour le 5.1, nous avons utilisé une méthode classique en s'aidant du *panner* de la console. En WFS, nous avons écrit une automation en déplaçant les objets concernés dans le *Performer*. Nous avons eu beaucoup de difficultés à effectuer un déplacement cohérent (cf 3.4.3). Notre démarche globale étant de fournir des mixages dont nous étions satisfaits, nous aurions pu dans ce cas nous retourner vers un panoramique classique sur le système discret. Nous avons

pourtant fait le choix de laisser cette « erreur » dans le mixage final pour évaluer son impact réel sur nos sujets. Nous savions qu'au *sweet spot*, à l'endroit où nous mixions, le mouvement saccadé en WFS serait perçu, et nous nous doutions qu'il aurait aussi un impact sur le groupe excentré. Il semblait intéressant de quantifier cet impact. Le résultat est sans appel, puisque c'est sur cette unique séquence que nos sujets ont trouvé une véritable différence de cohérence à l'image entre les deux systèmes, et ont ainsi préféré le 5.1.

Cette séquence a donc un aspect particulier qu'il nous semblait important de préciser. D'ailleurs, plusieurs remarques concernent des déplacements irréguliers d'objets, mettant en avant des problèmes « de phase », à interpréter ici comme étant des problèmes fréquentiels. Très souvent, ces remarques concernent des déplacements d'objets avec un spectre large. Pour la séquence 6 par exemple, le déplacement du personnage dans une forêt, au milieu de branches et de feuilles, a posé problème à plusieurs de nos sujets. Une tendance intéressante tient également au fait que les déplacements d'objets, qu'ils soient effectués en 5.1 ou en WFS, sont souvent critiqués par nos sujets.

Cela étant dit, il ne semble pas que la WFS, par un autre degré d'illusion perceptive qu'elle proposerait, puisse être considérée comme un système apportant une plus grande cohérence entre le champ visuel et le champ sonore. Cependant, si elle n'en apporte pas plus, elle ne semble pas pour autant incompatible avec la représentation cinématographique : c'est en tout cas ce que nos tests affirment. Nos séances de mixage nous ont permis de comprendre assez rapidement les tenants et les aboutissants de la distorsion angulaire inhérente à la WFS dans un contexte audiovisuel. Cet effet, détaillé dans 1.4.4, s'est rapidement fait ressentir pour des sources placées en profondeur et/ou très latéralisées. Cependant notre point de vue sur la question est moins catégorique que celui de Raphaël Mouterde [Mouterde, 2004], qui avait établi un certain

nombre de consignes et de règles à ne pas transgresser pour éviter ces conflits, en limitant l'espace possible des objets sonores à la proximité de l'écran. A l'usage, comme il l'avait également prévu, la marge de manœuvre semble plus large. Il est cependant certain que ce phénomène de distorsion angulaire restera une problématique constante dans le mixage d'un objet audiovisuel en WFS, et que l'aimantation audiovisuelle ou « l'effet ventriloque » trouvent rapidement leurs limites. Conscients du problème, nous avons très souvent écouté nos mixages à d'autres endroits qu'au sweet spot en ayant un regard critique sur notre travail de spatialisation afin d'y déceler d'éventuels aberrations de localisation. Bien que nous ayons essayé de minimiser ce phénomène, il nous a semblé délicat de l'éviter complètement. Nous pensions que ce problème ressortirait d'avantage dans les résultats chiffrés de nos tests, tant certaines distorsions nous semblait à nous, mixeurs, évidentes. Cela dit, des remarques fréquentes ont été effectuées sur des soucis de localisation pour certains objets WFS, allant parfois jusqu'à une source sonore perçue en dehors de l'écran pour des positions d'écoute excentrées. Les vagues latérales dans *Mundos Immundos*, ainsi que le passage d'avion dans *Louis le Manchot*, constituent les exemples les plus flagrants de ce type d'aberration.

Il est également très intéressant de constater la non-existence de lien entre le critère de profondeur et de relief, et celui de la cohérence. Une logique purement géométrique voudrait que si l'on ajoute de la profondeur sur des éléments sonores associés à des éléments visuels (certes projetés sur un support plat, mais mis en perspective par des rapports de taille et de mise au point contenus dans l'image), alors la sensation de cohérence serait plus forte. Il n'en est rien puisque nos sujets ont bien perçu plus de profondeur et de relief en WFS, mais cela n'implique visiblement pas de sensation de cohérence accrue.

La préférence

Notre dernier critère d'étude fournit un résultat extrêmement intéressant. Il semble en effet que les sujets placés au *sweet spot* aient préféré les mixages en 5.1. A cet endroit, ils sont idéalement placés pour que l'illusion perceptive du 5.1 fonctionne. Il semblerait donc que l'illusion de la WFS ne soit pas aussi convaincante en ce point que celle d'un système discret. La théorie, là aussi, nous avait amené à envisager une préférence au moins pour les sujets excentrés. Il n'en est rien, ou du moins, nos résultats ne tendent pas à valider cette hypothèse. Lors de nos sessions de mixage, nous nous sommes souvent rendus compte que le système 5.1, malgré ses défauts liés au *sweet spot*, semblait très bien calibré et adapté à la diffusion audiovisuelle, notamment en termes de stabilité des sources et de respect du timbre. De plus en recoupant les résultats obtenus sur les différents critères, nous voyons que la préférence n'est aucunement liée à la perception de la profondeur ou du relief, ni à la cohérence du champ visuel et du champ sonore. Si les sujets ont bien perçu, sur la majorité des séquences, une perspective plus importante en WFS, cela n'a pas d'influence sur leurs goûts et leurs préférences. L'enjeu de la spatialisation au cinéma n'est peut être finalement pas là. La cohérence à l'image ne semble pas non plus rentrer en compte dans la préférence d'un système par rapport à un autre, puisque les sujets n'ont pas perçu de différences fondamentales entre les deux formats. Il s'agit dès lors de questionner ce que les sujets ont entendu par préférence. Si nous avons pourtant lu une note descriptive avant le début de chaque session, il est, pour des raisons évidentes, très difficile de savoir ce qui détermine la préférence d'une personne pour un format ou un autre. L'habitude d'écoute est probablement un élément déterminant. Le 5.1, par l'esthétique codifiée qu'il introduit, nous habitue à des placements de sources, des sensations d'espace, que nous avons tenté d'interroger autrement lors de mixages en WFS. C'est peut-être ce repère presque culturel qui a conduit nos sujets à préférer le 5.1. Cela dit, ce que nous avançons ici ne reste qu'une hypothèse.

CONCLUSION

Ce mémoire visait à évaluer, sous plusieurs aspects, l'inscription des systèmes à synthèse de front d'onde dans un contexte de postproduction cinématographique.

En proposant une autre approximation du champ acoustique réel, un autre degré d'illusion perceptive, la synthèse de front d'onde présentait une approche différente de la spatialisation du son. L'implantation actuelle d'installations commerciales conséquentes nous amène à envisager que le nombre de haut-parleurs dans une salle de cinéma pourrait ne plus être considéré comme une limite, et c'est en cela qu'il nous semble légitime d'envisager la WFS dans ce contexte.

L'étude théorique que nous proposons dans ce mémoire a posé les bases pour la compréhension des principes de la synthèse de front d'onde. Ces bases ont alors constitué un socle de réflexion et nous ont permis d'émettre, par anticipation, plusieurs hypothèses sur l'utilisation de la WFS au cinéma.

Ces premières hypothèses posées, nous avons effectué une première prise en main du système WFS et de ses outils associés. Nous avons proposé une approche comparative. Elle fait état de nos expériences visant avant tout à évaluer le comportement des matériaux d'une bande son cinématographique sur un système WFS, sans image. Les travaux que nous avons menés sur la voix, les ambiances et la musique nous ont permis d'entrevoir le potentiel des « fronts d'onde » dans la perception d'un autre espace, et ce sur une zone d'écoute étendue. Les premières limites sont également apparues mettant à jour des problèmes spectraux sur des

matériaux fréquemment complexes. Nous avons pu constater que la segmentation des éléments sonores au cinéma demanderait à être repensée si l'on voulait s'affranchir d'une partie de ces problèmes. Au terme de cette étude préliminaire, nous avons également pu esquisser les premiers traits d'une configuration plus complexe et dédiée au cinéma qui doit, selon nous, proposer une hybridation entre un système discret « classique » (type 5.1) et un système WFS.

La suite de notre travail nous a logiquement conduit à envisager la WFS au cinéma de manière concrète. Il nous semblait important de décrire, dans un premier temps, ce cadre de travail tel qu'il est conçu à l'heure actuelle. Les problématiques techniques que nous avons dégagées de cette étude nous ont permis de concevoir une installation permettant le maintien d'un espace de travail et de méthodes cohérentes avec la WFS. Cette approche technique nous a ensuite amené à mettre en place un système dans l'auditorium de l'ENS Louis-Lumière. Nous avons tenté d'adapter la WFS au cinéma et non l'inverse, dans un environnement à vocation professionnelle. La configuration que nous proposons tente de répondre à la fois à un critère de maintien des conditions et des outils propres au mixage cinéma, tout en apportant des modifications qui permettent d'intégrer les notions d'objets sonores et les principes de la WFS sur un système hybride.

Nous avons ensuite mis à l'épreuve de la pratique cette configuration lors de sessions de mixage de six séquences différentes. Nous avons détaillé les constats effectués lors de ces travaux, qu'ils soient techniques ou perceptifs, et qui faisaient dès lors état de la présence d'une image. Si les outils dont nous disposions n'étaient pas, au départ, pensés pour le cinéma, ils se sont révélés tout à fait adaptés pour envisager une intégration de ceux-ci dans un auditorium, moyennant quelques ajustements. Pour autant, le potentiel créatif perçu dans nos

CONCLUSION

premières expérimentations s'est vu considérablement réduit par la présence de l'image.

Le mixage cinéma en WFS est un travail exigeant. Il impose une autre approche de la matière sonore et pose de nouvelles problématiques perceptives mettant en jeu plusieurs conflits autour de l'espace audiovisuel. Son apport nous semble limité à des contextes audiovisuels particuliers, sur des éléments sonores distincts. D'autres approches pourraient être envisagées à condition de « dé-construire » autrement le montage son, bien qu'elles ne solutionneraient qu'une partie des problèmes que nous avons rencontrés.

Les tests perceptifs mis en place par la suite s'appuyaient sur les séquences que nous avons remixées en 5.1 et mixées en WFS. Ces tests étaient pour nous un moyen de compléter une approche qui, jusqu'ici, nous était personnelle. Ils proposaient aux sujets une comparaison des deux systèmes selon trois critères. Dans notre contexte, les résultats obtenus confirment que la profondeur et le relief permis par la WFS est un phénomène qui « résiste » à la projection d'une image à deux dimensions. Cependant la perception de cette perspective nouvelle ne semble pas conduire à la sensation d'un espace sonore plus cohérent avec l'image, ni même entraîner une adhésion affective de nos sujets. Il reste à questionner ce qui tient, dans nos résultats, de nos habitudes d'écoutes (voire notre conditionnement), sans doute déroutées par les effets instaurés par la WFS. Une analyse de certains phénomènes en jeu dans les séquences, sous le prisme de notre travail de mixage, ont proposé une autre lecture des résultats obtenus permettant d'en affiner leur interprétation.

Au vu des constats établis par la pratique et des résultats obtenus lors des tests perceptifs, l'intérêt de la WFS au cinéma nous semble limité. Techniquement, le système nous semble tout à fait adaptable aux particularités de la postproduction

CONCLUSION

cinématographique. La limite à son utilisation ne se trouve pas dans la conception des outils de la WFS. Si nous avons décelé des imperfections dans le couplage d'un système WFS avec les matériaux sonores du cinéma, il ne nous semble pas non plus que ce soit dans ce couplage que résident les principales contraintes de la synthèse de front d'onde au cinéma.

C'est bien dans la nature intrinsèque du dispositif cinématographique que se trouvent les limites d'utilisation de la WFS au cinéma. Si ce système permet de créer une autre forme de perspective sonore, il se heurte à la puissance de l'image. Nous envisageons la WFS comme une solution aux problèmes perceptifs liés au 5.1, mais d'autres difficultés sont apparues. Les effets perceptibles de la WFS au cinéma sont limités et demandent un travail de positionnement précis afin de garantir une sensation d'espace cohérente pour l'ensemble du public, et respecter ainsi le « contrat du spectacle cinématographique ». La force des phénomènes entrant en jeu dans la perception d'un objet audiovisuel semblent instaurer un cadre, limitant la spatialisation sonore dans le cinéma actuel, et qui semble dépasser la technologie de diffusion du son, quel qu'en soit son principe.

Enfin, il y a fort à parier que le cinéma contemporain, de par sa conception de l'espace, son découpage, son rapport ambiguë au réel, constitue également un cadre qui limite l'investigation sonore. La WFS ne trouverait-elle pas un terrain d'expression plus vaste dans un autre cinéma, à l'instar du cinéma expérimental dont une des préoccupations est d'interroger en profondeur les matériaux, les outils et les techniques qui le façonne ?

BIBLIOGRAPHIE

[André et al.] **ANDRE Cédric, EMBRECHTS Jean-Jacques, G. VERLY Jacques**, *Adding 3D Sound to 3D Cinema : Identification and evaluation of different reproduction techniques*

[Altman, 1995] **ALTMAN Rick**, *Sound Theorie, Sound Practice*, Cineaste 21 n° 1-2, 1995

[Baiblé, 1990] **BAIBLE Claude**, *L'Audiophile*, n°51, Janvier 1990

[Baiblé, 1998] **BAIBLE Claude**, *L'image frontale, le son spatial*, dans *Cinéma et Dernières Technologies*, sous la direction de F.Beau, P.Dubois, G.Leblanc, Ina, Paris, 1998

[Baiblé, 2013] **BAIBLE Claude**, « Atmosphère, atmosphère, est-ce que j'ai une gueule d'atmosphère ? », dans *La Lettre du CST*, n°144, p. 22, 2013

[Barani, 2009] **BARANI Raphaël**, *Discrimination spatiale dans une scène sonore complexe diffusée par un système à synthèse de front d'onde*, Mémoire de fin d'étude, ENS Louis-Lumière, 2009

[Beck] **BECK Jay**, *A quiet revolution: changes in American film sound practices*, University of Iowa, 2003

[Becker, 2014] **BECKER Jeff**, *Les tendances de mixage cinéma*, Mediakwest magazine n°7, p. 24, 2014

[Berkhout, 1988] **BERKHOUT A.J.**, *A holographic approach to acoustic control*, Journal of the Audio Engineering Society, vol. 36, 1988

[Berkhout et al., 1993] **BERKHOUT A. J., DE VRIES D., VOGEL P.**, *Acoustic control by wave field synthesis*, Journal of the Acoustical Society of America, vol. 93, 1993

[Berkhout et al., 1997] **BERKHOUT A. J., DE VRIES D., SONKE J. J.**, *Array technology for acoustic wave field analysis in enclosures*, Journal of the Acoustical Society of America, vol. 102, 1997

[Besse, 2013] **BESSE Alain**, « Dossier : Spatialisation du son », dans *La Lettre du CST*, n°144, mars 2013

[Blauert, 1999] **BLAUERT Jens**, *Spatial hearing : The psychophysics of human sound localization*, MIT Press, 1999

[Boone & de Bruijn, 2003] **BOONE Marinus M.**, de **BRUIJN Werner P.J.**, *Improving speech intelligibility in teleconferencing by using Wave Field Synthesis*, convention paper 5800, 114th AES Convention, Amsterdam, 2003

[Bresson, 1975] **BRESSON Robert**, *Note sur le cinématographe*, Gallimard, Paris, 1975

[Brix et al., 2001] **BRIX Sandra**, **SPORER Thomas**, **PLOGSTIES Jan**, **CARROUSO** – *An European Approach to 3D-Audio*, convention paper 5314, 110th AES Convention, Amsterdam, 2001

[Canonville, 2006] **CANONVILLE Christian**, *Dimension sonore au cinéma : variations imaginatives*, Médiamorphoses n°18, p. 46-51, 2006

[Caulkins, 2007] **CAULKINS Terence**, *Caractérisation et contrôle du rayonnement d'un système de Wave Field Synthesis pour la situation de concert*, Thèse de Doctorat : traitement du signal, Université Paris VI – Pierre et Marie Curie, 2007

[Chion, 1991] **CHION Michel**, *L'Audiovision*, Nathan, Paris, 1991

[Chion, 1998] **CHION Michel**, *Traité d'Acoulogie*, Nathan, Paris, 1998

[Chion, 2003] **CHION Michel**, *Un art sonore le cinéma*, Cahiers du Cinéma, coll. Essais, Paris, 2003

[Corteel & Nicol, 2003] **CORTEEL Etienne**, **NICOL Rozenn**, *Listening room compensation for Wave Field Synthesis. What can be done ?*, AES 23rd International Conference, Copenhagen, 2003

[Corteel, 2004] **CORTEEL Etienne**, *Caractérisation et extension de la Wave Field Synthesis en conditions réelles*, Thèse de Doctorat : traitement du signal, Université Paris VI – Pierre et Marie Curie, 2004

[Corteel 2, 2004] **CORTEEL Etienne**, *Création et manipulation de scènes sonores pour la Wave Field Synthesis*, Cahier Louis-Lumière n°2, Automne 2004

[Corteel et al., 2014] **CORTEEL Etienne, GLAETTLI Peter, FOULON Raphael, FRAULY Reda, HAHN Ingo, HEINIGER Roger, PELLEGRINI Renato S.**, *3D speaker management systems – Mixer integration concepts*, 28th Tonmeisterstagung – VDT International Convention, 2014

[Deshays, 2006] **DESHAYS Daniel**, *Pour une écriture du son*, Klincksieck, Paris, 2006

[Dolby, 2000] **Dolby Laboratories Inc.**, *What is the LFE channel ?*, 2000, révisé en 2011

[Dolby, 2012] **Dolby Laboratories Inc.**, *Dolby Atmos : Un son de prochaine génération pour le cinéma*, Document de prestation technique, 2012, disponible sur www.dolby.com

[Dolby 2, 2012] **Dolby Laboratories Inc.**, *Dolby Atmos : Recommandation techniques pour le cinéma*, Document de prestation technique, 2012, disponible sur www.dolby.com

[Dolby, 2013] **Dolby Laboratories Inc.**, *Authoring for Dolby Atmos Cinema Sound Manual*, Issue1, Software v1.0, 2013, disponible sur www.dolby.com

[Dupas, 2007] **DUPAS Aymeric**, *Multicanal au cinéma : un canal zénithal ?*, Mémoire de fin d'études, ENS Louis-Lumière, 2007

[Hugonnet, 2000] **HUGONNET Christian, WALDER Pierre**, *Théorie et pratique de la prise de son stéréophonique*, Editions Eyrolles, 2000

[Husson, 2011] **HUSSON Thibault**, *Sonorisation de concert par synthèse de front d'onde*, Mémoire de fin d'études, ENS Louis-Lumière, 2011

[Journal of the Society of Motion Picture Engineers, 1941], *Journal of the Society of Motion Picture Engineers*, Aout 1941

[Kerins, 2011] **KERINS Mark**, *Beyond Dolby (Stereo) cinema in the digital sound age*, Indiana University Press , Blommington, 2011

[Melchior et al., 2003] **MELCHIOR Frank, BRIX Sandra, SPORER Thomas, RODER Thomas, BEATE Klehs**, *Wave Field Synthesis in combination with 2D video projection*, AES 24th International Conference on Multichannel Audio, Banff, 2003

[Millot & Pelé, 2006] **MILLOT Laurent, PELE Gérard**, « La sensation d'espace dans le spectacle audiovisuel », dans *Une architecture du son*, Cahiers de Champs Visuels n° 1 & 2, L'Harmattan, Paris, 2005

[Moulin, 2011], **MOULIN Samuel, NICOL Rozenn, GROS Laetitia, MAMASSION Pascal**, *Influence of the audio rendering on 3D audiovisual experience*, Lannion, 2011

[Mouterde, 2004] **MOUTERDE Raphaël**, *Etude perceptive en vue de l'utilisation d'un système WFS au cinéma*, Mémoire de fin d'études, ENS Louis-Lumière, 2004

[Murch, 1992] **MURCH Walter**, *In a blink of an eye, a perspective on film editing*, James Press, 1992

[Nicol, 1999] **NICOL Rozen**, *Restitution sonore spatialisée sur une zone étendue : Application à la téléprésence*, Thèse de Doctorat, Université du Maine, Le Mans, 1999

[Noguès et al., 2003] **NOGUES Mathieu, CORTEEL Etienne, WARUSFEL Oliver**, *Monitoring distance effect with wave field synthesis*, proc. of the 6th Int. Conference on Digital Audio Effects, London, 2003

[Pelé, 2012] **PELE Gérard**, *Etudes sur la perception auditive*, L'Harmattan, Paris, 2012

[Perrott, 1993] **Perrott David R.**, *Auditory and visual localization : Two modalities, one world*, AES 12th International Conference, 1993

[Rébillat, 2008] **REBILLAT Marc**, *Cohérence spatiale des informations visuelles et auditives dans un environnement de réalité virtuelle*, Mémoire de Master : mécanique et ingénierie des systèmes, Université Paris VI – Pierre et Marie Curie, 2008

[Richard, 2002], **RICHARD Elen**, « A dictatorial approach to bass management in surround music mixing », dans *Surround Professional*, October 2002

[Rumsey, 2001] **RUMSEY Francis**, *Spatial Audio*, Focal Press, Oxford, 2001

[Sporer, 2004] **SPORER Thomas**, *Wave Field Synthesis – Generation and reproduction of natural sound environments*, proc. of the 7th Int. Conference on Digital Audio Effects, Naples, 2004

BIBLIOGRAPHIE

[Sporer & Klehs, 2004] **SPORER Thomas, KLEHS Beate**, *Wave Field Synthesis in the real world : part 2 – in the movie theatre*, convention paper 6055, 116th AES Convention, Berlin, 2004

[Start, 1997] **START E. W.**, *Direct sound enhancement by Wave Field Synthesis*, Thèse de Doctorat, Université de Delft, Delft, Pays-Bas, 1997

[Theile et al., 2003] **THEILE Günther, WITTEK Helmit, REISINGER Markus**, *Potential Wave Field Synthesis applications in the multichannel stereophonic world*, AES 24th International Conference : Multichannel audio, the new reality, 2003

[Van Baelen et al.] **VAN BAELEN Wilfried, BERT Tom, CLAYPOOL Brian, SINNAEVE Tim**, *Auro-3D – A new dimension in cinema sound*, disponible sur www.auro-3d.com

SITOGRAPHIE

[Aldred] **ALDRED John**, *Fantasia !*, disponible sur <http://www.filmsound.org/film-sound-history>, consulté en mai 2015

[Edison] *Projet EDISON 3D*, Catégorie : Contenus et interactions numériques, disponible sur <http://www.agence-nationale-recherche.fr/?Projet=ANR-13-CORD-0008>, consulté en mai 2015

[Holophony], *Holophony.net*, disponible sur www.holophony.net, consulté en mai 2015

[Thom] **THOM Randy**, *Are Movies Getting Too Loud ?* , disponible sur <http://filmsound.org/randythom/loud-movies.htm>, consulté en mai 2015

FILMOGRAPHIE

Couper le Voile, réalisé par Maximilien Willmoch, 22 min, 2014

En Solitaire, réalisé par Christophe Offenstein, 97min, sorti en France le 25 août 2013

Gabriela, réalisé par Anna Leysens, 13 min, 2013

Good Bye, Lenin !, réalisé par Wolfgang Becker, 118 min, sorti en France le 10 septembre 2003

BIBLIOGRAPHIE

Gravity, réalisé par Alfonso Cuarón, 91min, sorti en France le 23 octobre 2013

Into the Wild, réalisé par Sean Penn, 147 min, sorti en France le 9 janvier 2008

La Menace d'une Rose, réalisé par Thomas Lhermitte, 21 min, 2013

Le Hobbit : Un Voyage Inattendu, réalisé par Peter Jackson, 169 min, sorti en France le 12 décembre 2012

Louis le Manchot, réalisé par Camille Jaulent, 23 min, 2015

Mundos Immundos, réalisé par Charlotte Bayer-Broc, 66min, 2015

Petite Bête, réalisé par Margaux Remaury, 12 min, 2014

Rebelle, titre original *Brave*, réalisé par Mark Andrews et Brenda Chapman, 95 min, sorti en France le 1 août 2012

Requiem for a Dream, réalisé par Darren Aronofsky, 110 min, sorti en France le 21 mars 2001

Skyfall, réalisé par Sam Mendes, 143 min, sorti en France le 26 octobre 2012

ANNEXE 1 : Détails des protocoles mis en place en partie 2

ANNEXE 1.1 : Traitement de la voix, mise en espace – Protocole 1 : voix-off seule - Détails

Détails du protocole 1 : voix-off seule

→ *Premier placement : objet WFS au niveau de l'enceinte centrale*



Sur ce premier placement, nous notons tout d'abord une légère différence spectrale entre la source WFS et la source discrète. La source WFS nous semble légèrement moins définie, comme si une partie de l'extrême aigu avait été atténuée. Spatialement, la source WFS semble se résumer à un point centré, légèrement plus étendu que la source discrète.

→ Deuxième placement : objet WFS un peu plus en retrait, au niveau du premier cercle du Performer

Ce deuxième placement nous amène à percevoir un élargissement spatial de la voix-off en WFS. En effet, celle-ci ne semble plus se résumer à un point, mais forme une zone plus large autour de l'enceinte centrale. Cet effet nous semble rester cohérent quelque soit notre position d'écoute dans la salle : la voix-off reste placée au centre (bien qu'étant apparemment plus large), et ce même pour des positions d'écoute excentrées. La comparaison avec la source discrète est maintenant flagrante : le passage entre l'une et l'autre nous permet d'apprécier le caractère étendue de la source WFS, alors que la source discrète semble désormais beaucoup plus ponctuelle. D'un point de vue spectral, l'atténuation du registre aigu observé lors du premier placement nous semble moins perceptible pour cette deuxième position de source WFS.

→ Troisième placement : objet WFS beaucoup plus en retrait, au plus proche du second cercle du Performer

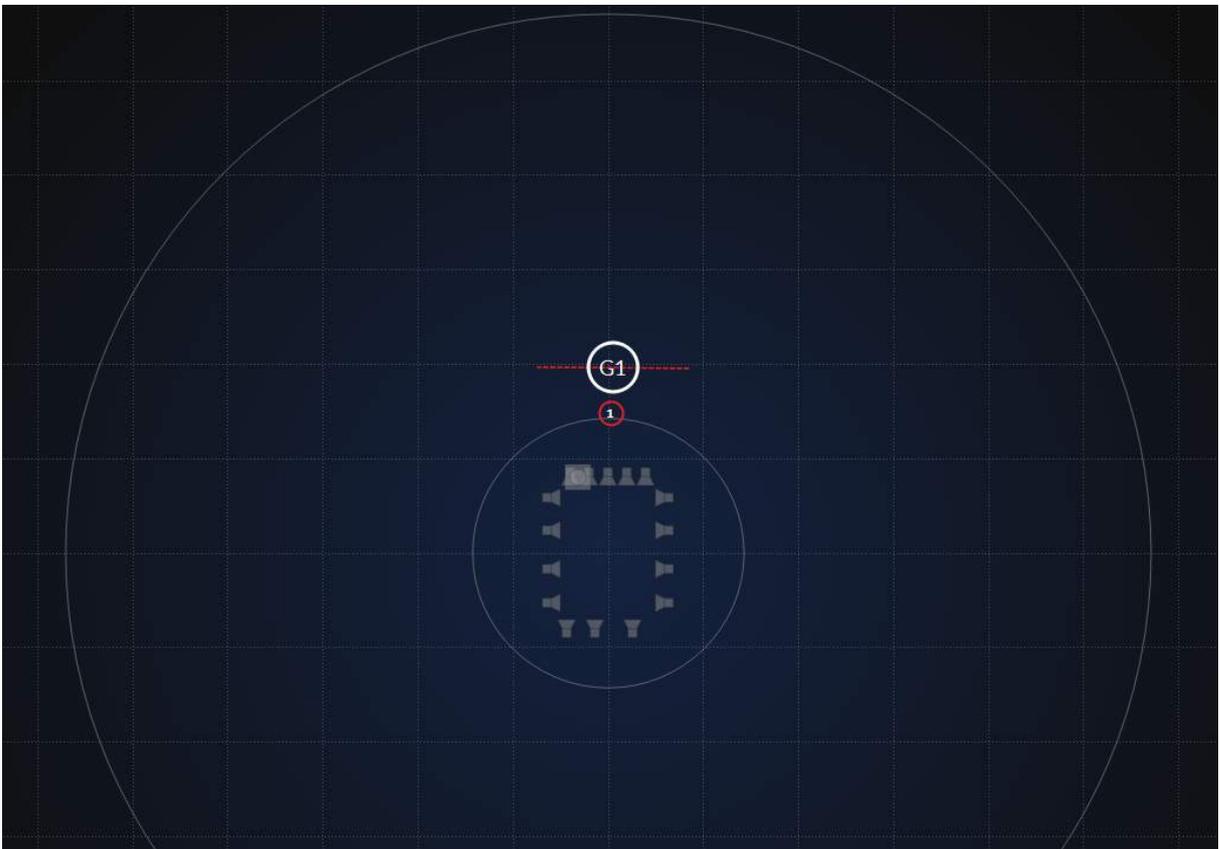


Cette dernière position nous fait entendre une voix-off très large, moins localisable. En se rapprochant d'une diffusion en ondes planes, la source ne présente plus d'origine ponctuelle mais semble provenir d'un espace situé en face du point d'écoute. En effet, d'un point de vue technologique, chaque enceinte est amenée à contribuer à la synthèse de la source, ce qui amène à renforcer cette sensation de largeur et de non-appartenance à un stimulus ponctuel. Notons également une modification spectrale plus prononcée, avec à nouveau une atténuation des aigus, plus audible que précédemment.

ANNEXE 1.2 : Traitement de la voix, mise en espace – Protocole 2 : voix-off associée à une réverbération - Détails

Détails du protocole 2 : voix-off associée à une réverbération

→ *Première configuration : objet WFS à hauteur du premier cercle du Performer, groupe de réverbération (stéréo) derrière la source et relativement proche de celle-ci*



Cette première position de source nous semblait être intéressante dans un objectif de placement de celle-ci dans un espace -en l'occurrence une salle de petit volume-. La réverbération matérialisée, par un groupe de deux sources, a par défaut été placée en retrait de la source, à proximité de cette dernière.

L'ajout d'un effet de réverbération en WFS, nous amène à questionner la position de cet effet relativement à la position de la source. Pour ce premier placement d'objet WFS, plusieurs positions de réverbération ont été testées :

Une première manipulation nous a permis de constater que plus on éloigne le groupe de réverbération de la source dans l'axe vertical du Performer, plus ces deux éléments tendent à se dissocier. Dans le but de créer un espace cohérent entre une voix et son lieu de diffusion, ce qui représente nombre des cas lors d'un mixage en vue d'une création cinématographique, il nous semblait alors plus approprié de rapprocher le groupe de réverbération de sa source. Etienne Corteel nous a également conseillé cette organisation après avoir constaté que lors du mixage de projets majoritairement musicaux, un placement des réverbérations juste derrière la source apportait une fusion intéressante entre le "champ direct" et le "champ réverbéré". A l'inverse, dans un objectif de décorrélation, et en vue d'un effet s'éloignant des propriétés physiques de diffusion d'un stimulus sonore dans un espace, ce "placement de réverbération" pourrait s'avérer être intéressant, dans un contexte de création musicale par exemple.

La gestion de la réverbération en WFS demande également de s'interroger sur la distribution latérale de celle-ci. La fonction de groupe, récemment implantée dans le Performer, permet entre autres de modifier la largeur de celui-ci. La largeur d'un groupe correspond simplement à l'écartement de ses bornes, et donc, dans le cas d'un groupe stéréo, à l'écartement entre le canal gauche et le canal droit.

Le test suivant consistait donc à faire varier la largeur du groupe afin d'évaluer son impact sur la mise en espace de la voix. Perceptivement, nous obtenions une fusion comparable, seule la sensation de largeur étant apparemment différente. En effet, nous avons opté pour une première largeur de groupe « réverbération » équivalente à la distance entre les deux haut-parleurs extrêmes de la rampe frontale. Ainsi, dans ce cas, seuls les haut-parleurs frontaux participaient à la synthèse de la réverbération. Dans une nouvelle configuration, où la largeur du groupe « réverbération » dépasse les bornes de la zone de fenêtrage frontal, les haut-parleurs des rampes latérales contribuent à la création de l'effet. Ceci n'est pas sans créer une sensation d'espace large, certes immersive, mais rendant d'avantage audible la présence des haut-parleurs.

Nous avons donc choisi, pour poursuivre les tests, une position du groupe « réverbération » proche de la source et une largeur ne mettant à contribution que

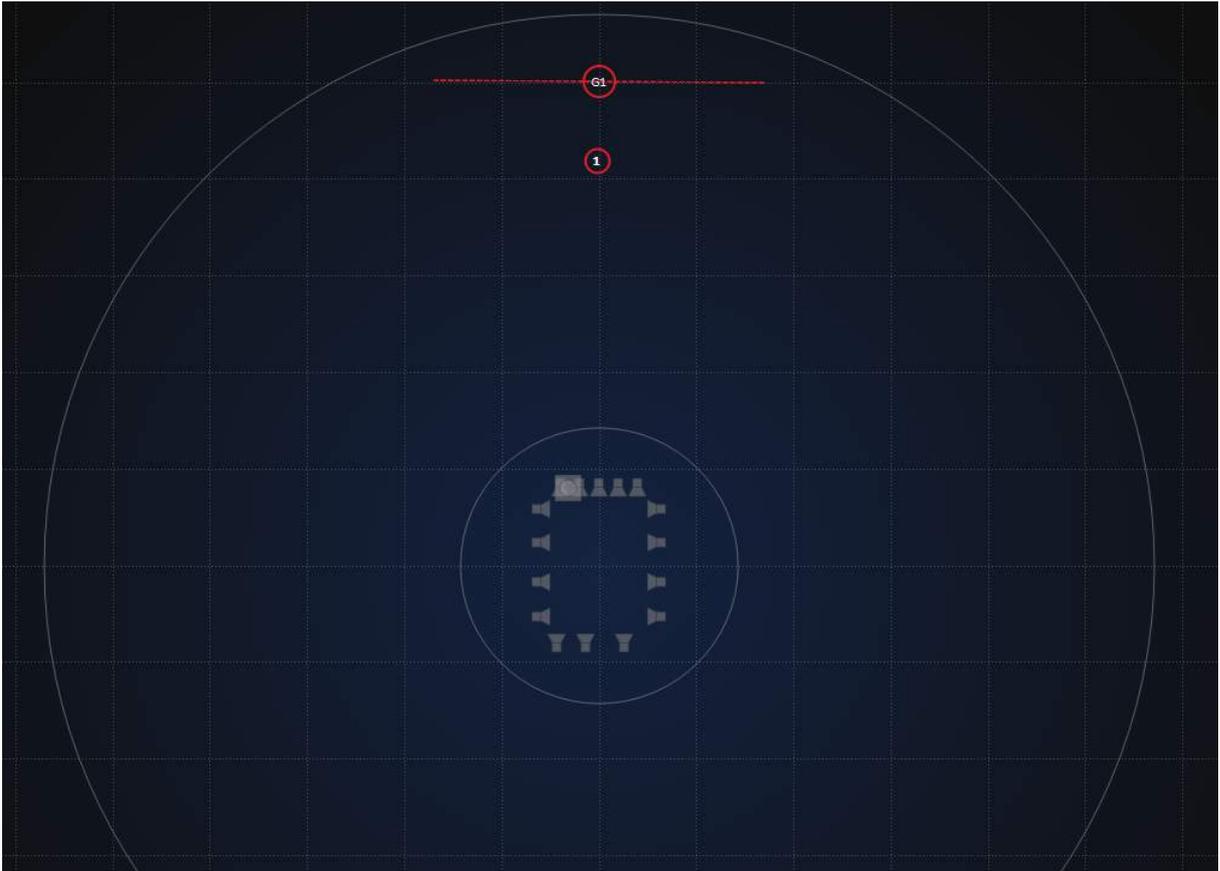
les haut-parleurs de façade. Nous avons par suite tenté d'approcher cet effet pour la source discrète et sa réverbération associée. La voix était alors directement envoyée dans l'enceinte centrale ; la réverbération était quant à elle routée en stéréo, utilisant les deux haut-parleurs placés aux positions extrêmes.

Lors du passage entre le dispositif discret et le dispositif WFS, la différence principale concernait la fusion entre le son direct et le son réverbéré. Comme évoqué précédemment, la synthèse de front d'onde nous a amené à percevoir une certaine fusion entre le signal original et sa réverbération associée. Sur le dispositif discret, ce dernier utilisant le canal central et les deux canaux latéraux, nous parvenons à entendre la source au centre, séparée de la réverbération latérale qui lui est associée. Notons à nouveau que sur un tel système (un 5.1 standard, par exemple), cet effet aurait pu être évité (ou du moins atténué) par l'utilisation de la "divergence". En redistribuant une portion de la réverbération stéréo dans le canal central, astuce souvent utilisée, la fusion entre la source directe et le champ réfléchi peut être accrue.

On constate ici qu'en termes de méthode, l'utilisation de la synthèse de front d'onde pourrait présenter réel un apport dans l'usage des effets de réverbération. Là où une étape de "distribution" du signal est nécessaire sur un système à restitution discrète -par usage de la divergence-, cette distribution est directement réalisée sur une système WFS, avec qui plus est, un signal différent envoyé dans chacune des enceintes.

Pour un point d'écoute central, i.e. au « sweet spot », l'effet obtenu en restitution discrète ou en WFS pourrait être comparable. Le résultat perçu en un point d'écoute excentré en est tout autre : sur le système discret, la réverbération semble uniquement provenir du canal gauche (pour une position d'écoute à gauche dans la salle), alors qu'en WFS, une perception des deux canaux semble être conservée (la réverbération paraît large).

→ Deuxième configuration : objet WFS reculé au niveau du deuxième cercle du Performer, groupe de réverbération derrière la source et relativement proche de celle-ci, largeur du groupe "réverbération" reconsidérée



En plaçant la source et sa réverbération à une position plus reculée, le premier constat touchait à la largeur apparente du groupe de réverbération : nous ne retrouvions pas l'effet de salle souhaité, mêlant à priori premières réflexions, et champ diffus. Avec une largeur de réverbération identique à celle de la première configuration testée, le champ diffus n'apparaissait pas en tant que tel, et semblait être masqué par le son direct, provoquant ainsi un certain "gonflement" de la voix dans certaines parties du spectre. C'est en écartant les bornes du groupe de réverbération que nous sommes parvenus à retrouver un effet cohérent, à savoir une source placée de manière assez lointaine dans un espace de petit volume.

Afin d'accentuer cet effet et de le rendre plus évident, nous avons décidé de filtrer les fréquences extrêmes de la voix, comme pour simuler une atténuation de certaines fréquences lors de l'éloignement d'une source (cf. [Rumsey, 2001], p. 35). Pour rendre comparables les deux systèmes de restitution, nous avons appliqué les mêmes paramètres de filtrage à la source WFS et à la source discrète, ce qui nous a semblé satisfaisant car perceptivement similaire.

Notre objectif étant d'évaluer une position plus reculée de source, il fallait rapprocher celle-ci d'une dispersion lointaine, ceci se traduisant également par une prépondérance du champ réverbéré sur le champ direct (cf. [Rumsey, 2001], p. 35). Nous avons alors opté pour diminuer le niveau du son direct, et augmenter quelque peu le niveau de la réverbération.

Pour un comportement spectral très proche, une largeur apparente semblable, un rapport direct/réverbéré équivalent, les deux systèmes présentaient une différence perceptive importante au niveau de la sensation de perspective. En WFS, nous percevons un espace, créé au-delà de la ligne frontale de haut-parleurs. La source semble être placée au fond de l'espace créé, et répondre à des lois géométriques la mettant en perspective par rapport à notre position d'écoute -qu'elle soit centrale ou excentrée-. En restitution discrète, la sensation de profondeur semble être créée, mais se limiter à l'éloignement d'une "ligne sonore", comme si l'espace situé en amont de cette ligne lointaine n'était pas perçu.

Par ailleurs, le placement lointain de la source et de sa réverbération synthétisée en ondes planes propose une image sonore cohérente quelque soit notre position d'écoute dans la salle. A contrario, en diffusion discrète, et pour une position d'écoute latéralisée, nous ne percevons qu'un canal de réverbération sur les deux, celui se situant le plus proche de nous (le canal gauche pour un point d'écoute "avant-gauche", par exemple). Cette distorsion de localisation de la réverbération nous a semblé plus importante que précédemment : comme déjà évoqué, nous avons en effet augmenté le niveau de la réverbération par rapport à celui du son direct, ce qui la rendait nécessairement plus audible.

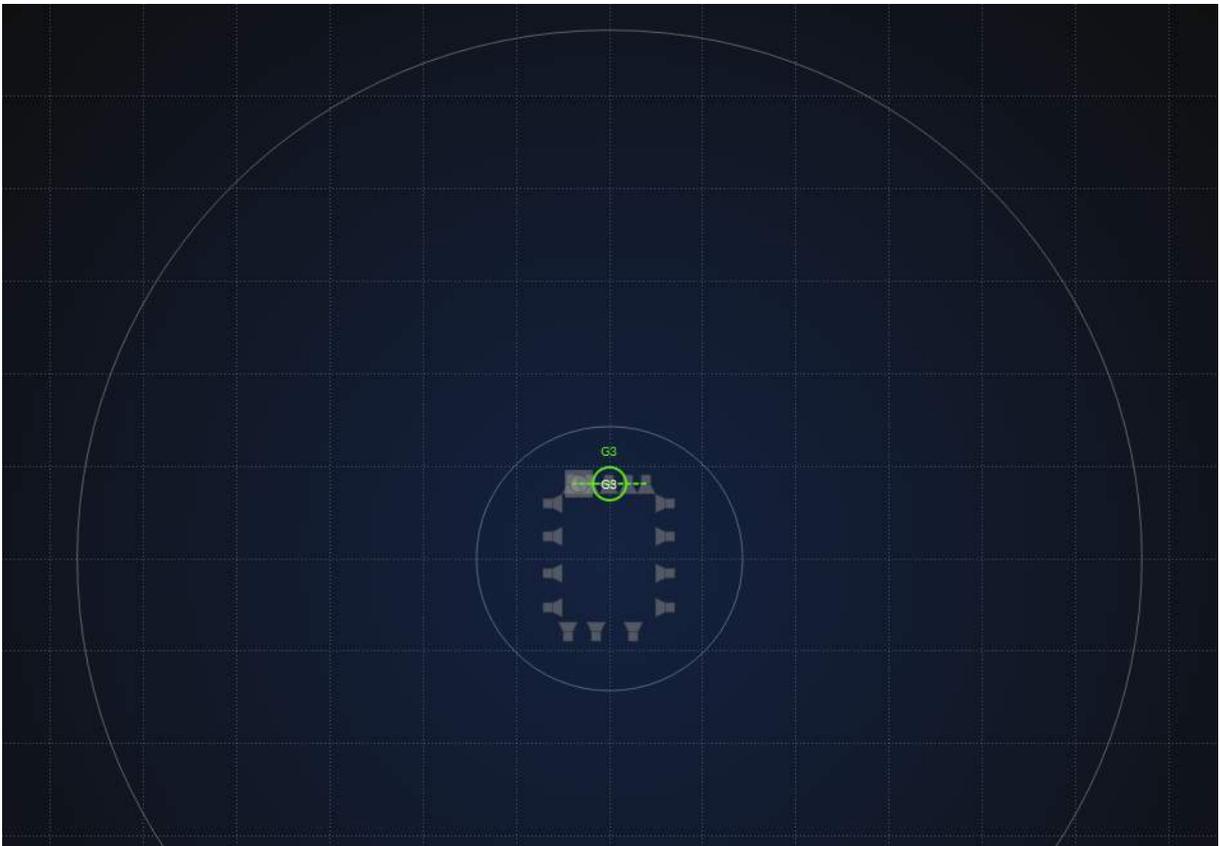
Enfin, au vue de ces deux configurations testées (source proche/réverbération associée, source lointaine/réverbération associée) nous pouvons remarquer une certaine corrélation entre l'éloignement de la source sur l'axe vertical du Performer, et la largeur apparente de la réverbération stéréophonique. Il semblerait en effet que plus la source est éloignée dans "l'espace WFS", plus la réverbération aurait nécessité d'être large, afin de respecter la perspective et la profondeur construites. Remarquons tout de même que ce constat a été fait pour une seule source, pour un certain type de réverbération -de petit volume-, pour deux positions de source, et pour un nombre limité de positions d'écoute. Il ne

peut constituer en aucun cas une généralité pour tout autre type de source, tout autre type de réverbération, ou toute autre position du couple « source ; réverbération ».

ANNEXE 1.3 : Cas des ambiances frontales –
Protocole 1 – Ambiance stéréophonique de contenu
fréquentiel large - Détails

Détails du protocole 1 : ambiance stéréophonique de contenu fréquentiel large

→ **Position 1 : groupe bipolaire "ambiance" au niveau des enceintes (un pôle dans chaque enceinte)**



Le premier élément perçu concerne la modification spectrale due à la synthèse WFS. Alors que pour les voix, la perte d'aigu se caractérisait par une perte d'informations transitoires, dans le cas des ambiances, elle se manifeste par la perte quasi-totale d'une composante aiguë de l'ambiance : le vent dans les feuilles. En effet, en diffusion discrète, l'ambiance nous paraît être plus "définie", plus "aérée", du fait de la préservation de cette composante venteuse.

Une sensation de resserrement est venue accompagner la perte d'aigus décrite ci-dessus. L'ambiance paraît plus compacte, et même moins immersive. De plus, ce resserrement aurait pu nous amener à percevoir une préservation de l'image stéréophonique quelque soit notre position d'écoute dans la salle. En réalité, et assez paradoxalement avec ce que la WFS est censée apporter, nous obtenons une image déséquilibrée pour des positions d'écoute latéralisées : à gauche, le pôle gauche de l'ambiance est quasi-exclusivement perçu ; à droite, il en est de même pour le pôle droit. Notons tout de même que cette distorsion d'image sonore, pour des positions d'écoute latéralisées, s'amplifie dès lors que l'on passe en restitution discrète.

En résumé, pour ce premier placement d'ambiance, et pour ce type d'ambiance, la synthèse de front d'onde ne nous semble pas être le moyen de restitution le plus adapté. Bien que de tels constats soient amenés à évoluer si le contenu de l'ambiance en venait à changer, nous pensons que de choisir ce positionnement WFS "par défaut" peut amener à mettre de côté un peu trop rapidement ce que la stéréophonie discrète peut encore apporter.

→ Position 2 : groupe bipolaire "ambiance" au niveau du premier cercle du Performer ; largeur du groupe reconsidérée



Pour cette deuxième position de l'ambiance, nous nous sommes servis de ce que nous avons constaté lors des tests précédents. Ainsi, nous avons quelque peu élargi la distance entre les deux pôles du groupe, de telle sorte à ce que ces derniers soient placés aux bords de la zone de fenêtrage frontal (cf 2.1.2).

Tout d'abord, on constate que pour cette position, la WFS respecte d'avantage le registre aigu de l'ambiance. Une modification du spectre reste perceptible, mais nous semble bien plus légère que pour la position précédente : on pourra d'avantage parler de "coloration" s'approchant de l'éloignement naturel d'une source.

Etrangement, en termes de largeur apparente, la perception de l'ambiance diffusée en WFS varie beaucoup selon la position d'écoute dans la salle. Pour une position d'écoute centrale, l'ambiance nous apparaît plus homogène, l'image semble mieux remplie. Ceci pourrait toujours s'expliquer de par la contribution des haut-parleurs centraux, participant eux aussi à la synthèse de l'ambiance bipolaire. De fait, une certaine impression de réalisme peut se faire ressentir : la distribution des canaux "gauche" et "droit" de l'ambiance sur l'ensemble de la

façade frontale tend à nous faire oublier le dispositif de restitution. Pour des positions d'écoute latéralisées, un phénomène de tassement devient perceptible. L'ambiance semble resserrée, les évènements ponctuels permettant de ressentir une certaine largeur se retrouvent regroupés au centre de l'image.

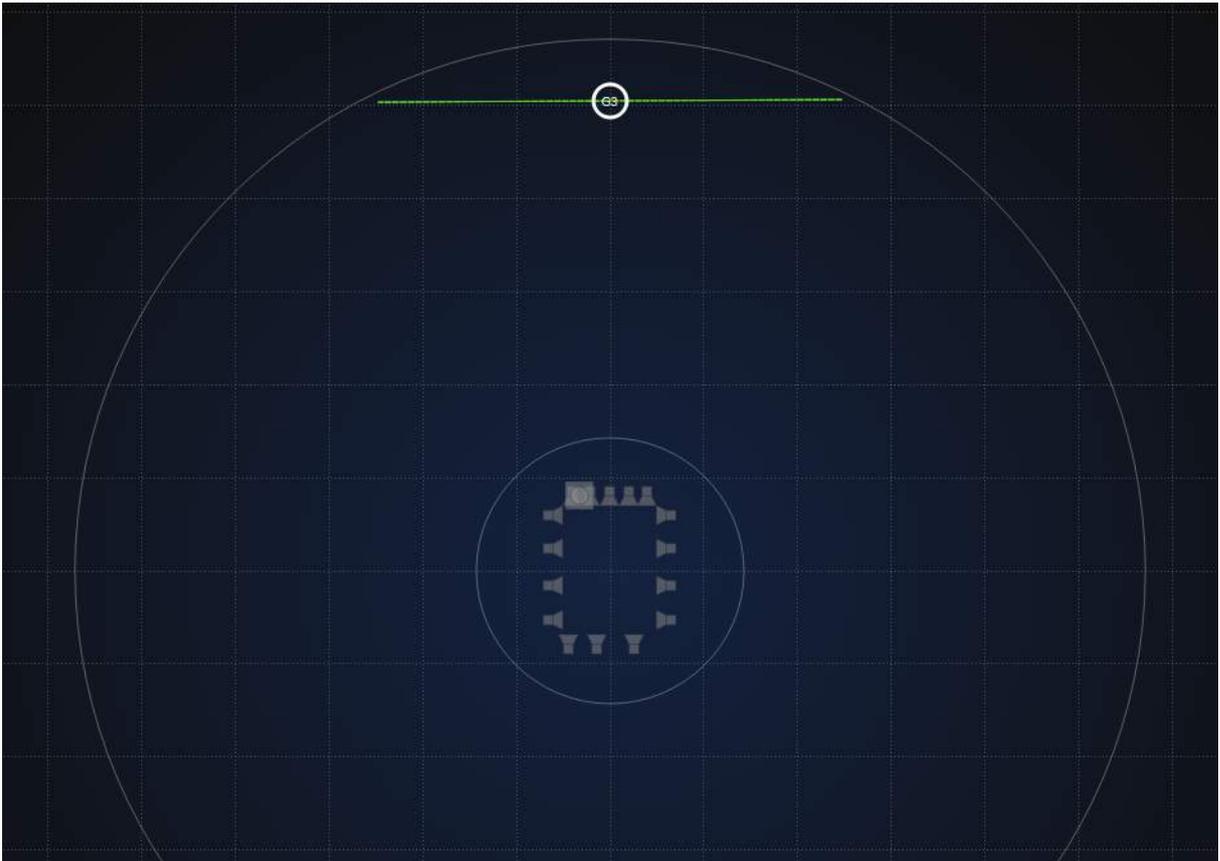
Ceci dit, cette impression de tassement est à mettre en regard de la cohérence de l'image stéréophonique. En effet, en WFS, la sensation d'une image sonore large est préservée, quelque soit notre position d'écoute dans la salle. Pour des positions latérales extrêmes, une largeur apparente peut continuer à être attribuée à l'ambiance, pouvant aller de l'enceinte gauche à l'enceinte inter-droite, pour un placement de l'auditeur à l'extrême gauche de la salle.



Suite au constat effectué sur le remplissage du centre de l'image, du à la contribution des enceintes centrales, une question relative au masquage s'est faite ressentir. Lors d'un mixage (i.e. lors d'un mélange de plusieurs sources), qu'advient-il des éléments habituellement placés au centre (bruitages, effets, et surtout voix), si ce même espace est déjà occupé par la synthèse WFS d'autres types de source (ambiances, musiques, ...) ? A priori, la WFS permettrait le

démasquage des sources et une meilleure intelligibilité dans le rendu d'une scène sonore complexe [Barani, 2009]. Ce postulat reste à vérifier dans le cas d'un mixage lié à une image.

→ **Position 3 : groupe bipolaire "ambiance" au niveau du second cercle du Performer ; largeur du groupe reconsidérée**



Comme pour la position précédente, nous avons dans un premier temps réadapté la largeur du groupe à sa position dans le Performer. Notons tout de même qu'un tel placement, censé éviter la contribution des rampes latérales de haut-parleur, conduit à une faible contribution des premières enceintes situées sur les murs latéraux. Cette conséquence était inévitable si nous voulions percevoir une largeur apparente équivalente, en stéréophonie discrète et en WFS, pour un placement d'écoute au sweet spot.

En comparaison avec la position 2, cette position de groupe semble offrir un meilleur respect du timbre. Une légère atténuation de l'aigu peut être perçue, sans que cela ne modifie le rendu global de l'ambiance de manière trop importante. De

plus, l'effet de tassement perçu pour la position précédente semble avoir totalement disparu, ce qui est plutôt de bonne augure pour un placement aussi éloigné et aussi large de groupe.

L'impression de réalisme semble être le critère le plus marquant à constater pour cette position 3 de l'ambiance. Cette sensation se caractérise majoritairement par "l'oubli" du système de restitution. Alors qu'en stéréophonie discrète, la contribution des canaux "Left" et "Right" est perceptible en tant que telle (i.e. "une paire d'enceintes diffusant un contenu sonore"), la contribution complexe des haut-parleurs en WFS semble rendre le système de restitution plus transparent. De plus, cette impression de réalisme est accrue par l'homogénéité du rendu de l'image sonore, quelque soit notre position d'écoute dans la salle.

ANNEXE 1.4 : Cas des ambiances frontales – Protocole 2 – Ambiance stéréophonique de contenu fréquentiel resserré - Détails

Détails du protocole 2 : ambiance stéréophonique de contenu fréquentiel resserré

→ Position 1 : groupe bipolaire "ambiance" au niveau des enceintes

Pour cette première position de groupe, on note très peu de différence entre la restitution stéréophonique et la restitution WFS : les deux ambiances nous paraissent de fait très semblables. Spectralement, aucune différence audible n'est à relever. En termes de largeur spatiale, l'ambiance WFS nous paraît très légèrement plus resserrée. Notons que ce tassement commence à nous faire percevoir une création de perspective.

Ainsi, contrairement à ce que nous avons remarqué pour les sources de large spectre, un tel placement d'ambiance, pour des sources à timbre plus resserré, ne nous semble pas inintéressant. Du fait qu'aucun artefact audible ne soit relevé, cette position de source conduit à des résultats perceptifs équivalents entre la stéréophonie et la WFS. Cela dit, l'avantage d'opter pour la WFS permet d'envisager des placements dynamiques au cours d'une séquence. Sans forcément parler de mouvement de source, on pourra aisément modifier le placement (et/ou la largeur) du groupe suivant les différents plan de la séquence : tandis qu'un plan serré nous amènerait naturellement à une telle position de groupe (au niveau des enceintes), un plan large nous conduirait certainement à élargir le groupe, et/ou l'éloigner. Ces modifications pourraient être effectués via l'automation incluse dans le séquenceur utilisé.

→ Position 2 : groupe bipolaire "ambiance" au niveau du premier cercle du Performer ; largeur réajustée

Une bonne impression d'espace se dégageait déjà de la restitution stéréophonique. Mais lors du passage en WFS, les différences se sont immédiatement faites remarquer.

Nous percevons tout d'abord une importante spatialité : malgré une certaine "monoïsation" de l'ambiance (maintenant bien connue pour cette position 2), un espace profond se crée au delà des haut-parleurs frontaux. De fait, ce qui est perdu en largeur semble être gagné en sensation de perspective. Par exemple, l'oiseau ponctuel (à la base placé plus loin dans le décor naturel, car contenant un champ réverbéré plus important) nous paraît être placé bien loin derrière la rampe frontale de haut-parleurs.

Une légère "décoloration" du registre aigu est à noter pour la source synthétisée en WFS. Remarquons que ce filtrage léger est plutôt cohérent avec un placement lointain de source [Rumsey, 2001]. Nous nous demandons également si cette très légère perte d'aigu ne participe pas au fait que nous percevions la source aussi loin.

Associée à la profondeur, une sensation de relief est mise en avant par le système WFS. Deux plans nous paraissent distincts, l'un contenant les chants d'oiseaux groupés (plus proche), l'autre contenant l'oiseau lointain. En WFS, nous obtenons cette sensation de manière immédiate. En stéréophonie, la perception d'un éventuel relief se fait plus progressivement. C'est au bout d'une certaine durée d'écoute que l'effet nous semble satisfaisant, comme si une accoutumance était nécessaire à la perception des différents plans. Bien entendu, si nous repassons en WFS, la différence de ressenti est évidente, et tend à nous faire dire que le rendu du relief est une plus value de la synthèse de front d'onde.

D'un point de vue immersif, la stéréophonie pourrait presque prétendre être plus efficace que la WFS, car de largeur apparente plus importante (pour une écoute au sweet spot). En WFS, cette "perte d'enveloppement", provoquée par le tassement de l'ambiance, pourrait probablement être rattrapée avec la mise à contribution des enceintes arrières. L'usage de la réverbération en surround est l'une des techniques privilégiées de certains mixeurs pour accroître le caractère immersif d'une ambiance, d'une source. Nous développerons en particulier ce type de techniques dans la partie suivante (cf. 2.2.3).

→ Position 3 : groupe bipolaire "ambiance" au niveau du second cercle du Performer ; largeur réajustée

Les effets de la restitution WFS considérés comme gênant, car dénaturant le contenu original de l'ambiance stéréophonique (atténuation des fréquences aiguës, resserrement spatial de l'ambiance), tendent à disparaître pour cette position 3.

D'une part, au sweet spot, les 2 systèmes de restitution offrent une ambiance à largeur apparente équivalente. Notons que dès qu'une position d'écoute latérale est adoptée, le système stéréophonique ne permet pas de rendre une telle impression de largeur, contrairement à la WFS qui tend à préserver cette caractéristique. Ainsi, l'ambiance nous paraît maintenant tant enveloppante en WFS qu'en stéréophonie pour une position d'écoute au sweet spot, et plus enveloppante en WFS pour une position d'écoute latéralisée. D'autre part, la légère décoloration de l'aigu se fait effectivement oublier pour ce 3ème positionnement de groupe.

En termes de sensation de profondeur, la position 3 semble être bien plus convaincante que les 2 précédentes. Un effet "holophonique" se fait même ressentir, et ce en particulier pour l'oiseau situé au fond de la scène sonore. On constate par conséquent que même si la source sonore dispose déjà d'informations spatiales natives (réverbérations et premières réflexions issues de la prise de son), le positionnement de cette source dans l'espace WFS est un apport quant à la perspective que le système est capable de créer.

ANNEXE 1.5 : Cas des ambiances frontales – Protocole 3 – Ambiance monophonique - Détails

Détails du protocole 3 : ambiance monophonique

→ **Position 1 : objet "ambiance" monopolaire placé au niveau de l'enceinte centrale**

La ponctualité de l'ambiance monophonique diffusée via le canal central ne nous permettait pas de comparer correctement cette dernière avec la source restituée en WFS. En effet, une trop grande différence de largeur apparente était perceptible. Pour compenser l'aspect resserré de l'ambiance diffusée en discret, nous avons décidé d'utiliser le paramètre de divergence. Nous l'avons ajusté de telle sorte à obtenir, au sweet spot, une largeur perçue équivalente. C'est avec la divergence réglée à 100% que nous obtenions un résultat satisfaisant, c'est à dire avec autant de signal sur les canaux L et R, que sur le canal central.

Par suite, en restitution discrète, le problème inhérent à la divergence s'est rapidement fait ressentir. Pour des positions d'écoute latéralisées, les canaux latéraux vers lesquels nous nous rapprochons ont tendance à se faire un peu trop entendre. Ne souhaitant à la base qu'élargir l'ambiance monophonique de foule, nous obtenons de fait un décalage de la source fantôme : en position d'écoute latérale gauche par exemple, un nouveau centre virtuel est créé entre l'enceinte gauche et l'enceinte inter-gauche.

En WFS, l'effet obtenu en est tout autre. Nous mouvoir dans la zone d'écoute nous fait toujours percevoir l'ambiance au même endroit, en tant que source positionnée juste derrière l'enceinte centrale. Bien qu'ayant un rendu un peu large pour un placement aussi proche de la rampe de haut-parleurs, l'objet sonore est toujours localisé au centre de l'écran.

→ Position 2 : objet "ambiance" monopolaire placé au niveau du premier cercle du Performer

Pour cette deuxième position de source, nous préservons une largeur apparente équivalente entre la restitution discrète et la restitution WFS. De même, la légère atténuation du registre aigu est toujours perceptible, caractérisée par la perte de "définition" sur les exclamations et autres hurlements de voix. Notons que ce léger filtrage nous semble à nouveau cohérent avec l'éloignement holophonique de la source. Au niveau des remarques récurrentes, on pourra aussi constater que la localisation de l'ambiance reste cohérente quelque soit notre placement dans la zone d'écoute.

Avec cette ambiance de foule très fournie, la sensation d'espace créé derrière la rampe frontale de haut-parleurs nous semble un peu moins évidente que pour les ambiances testées précédemment. Bien qu'en absolu cette impression de perspective soit moins évidente, une bascule rapide sur le système discret nous amène à percevoir une ambiance "plaquée" sur la ligne de haut-parleurs, comme si les différents plans potentiellement perceptibles avaient été fortement rapprochés, en plus de nous faire percevoir le système de restitution en tant que tel. Ainsi, on pourra à nouveau parler de réalisme quant à la diffusion de cette ambiance monophonique via synthèse de front d'onde.

→ Position 3 : objet "ambiance" monopolaire placé au niveau du second cercle du Performer

Pour cette troisième position de source, l'hypothèse de départ reste inchangée : au sweet spot, nous disposons d'une largeur apparente équivalente pour les deux systèmes. Toujours au sweet spot, un gain dans le registre bas medium est à constater en diffusion WFS. En effet, lors du passage sur le système discret, l'ambiance nous paraît moins "gonflante", moins "résonnante". Ce constat, au vu de la sensation de profondeur qu'amène à nouveau la WFS, semble quelque peu paradoxal : les effets découlant de la distance holophonique (profondeur, perspective), semblent être en conflit avec le gain sur ces fréquences, ces dernières conduisant habituellement à une certaine sensation de présence.

Outre la sensation d'espace créé derrière la rampe de haut-parleurs, nous obtenons à nouveau une cohérence de localisation de la source, indépendante de notre position dans la zone d'écoute. Pour des positions excentrées, la source paraît moins ponctuelle, plus large, et placée loin derrière la ligne frontale de haut-parleurs. Cela dit, la source reste localisable dans une direction donnée, chose que nous n'obtenons toujours pas (sans trop de surprise) en restitution discrète.

ANNEXE 1.6 : Spatialisation de la musique off – Protocole 1 - Détails

Détails du protocole 1 : groupe « musique » placé dans l'espace WFS ; source discrète stéréophonique (surrounds inactifs)

Les 4 extraits musicaux sont diffusée à la fois en WFS (via le groupe bipolaire créé), puis en stéréophonie (via les haut-parleurs gauche et droit de la rampe frontale). Nous avons testé deux positions différentes pour le groupe "musique" : l'une au niveau du premier cercle du Performer (position 2 sur les tests précédents) ; l'autre au niveau du second cercle du Performer (position 3 sur les tests précédents).

→ Position 1 : groupe bipolaire "musique" au niveau des haut-parleurs

Notons que nous avons rapidement abandonné la position 1 (i.e. pôles au niveau des haut-parleurs frontaux). En effet, après un premier essai à cette position, nous avons perçu une spatialité équivalente entre la restitution WFS et la restitution discrète. De plus, les problèmes spectraux relatifs à cette position étaient d'avantage audibles avec les extraits musicaux qu'avec les ambiances. Pour les ambiances, premiers éléments stéréophoniques que nous avons testé, il pouvait être nécessaire d'évaluer cette position pour tous les extraits. Pour les musiques, nous avons fait le choix de nous concentrer sur les positions 2 et 3, produisant à notre sens, des résultats plus intéressants.

→ Position 2 : groupe bipolaire "musique" proche du premier cercle du Performer

Quelques remarques globales peuvent être effectuées, valables pour les 4 extraits sélectionnés. Premier constat auquel nous sommes maintenant habitués : la musique, d'une largeur stéréophonique donnée, se ressert dès lors qu'on bascule sur le système WFS. Cet aspect est d'autant plus audible sur les musiques. En effet, une source musicale étant dotée de nombreux éléments rythmiques et

"percussifs", la densité en transitoires y est plus importante. Ainsi, les transitoires liées à l'écoute stéréophonique et nous faisant entendre des sources aux extrêmes de l'image sonore, se retrouvent tassées vers le centre de l'image en restitution WFS. De plus, du fait du caractère cyclique des extraits sélectionnés (i.e. structure musicale composée de mesures qui se répètent), les mêmes artefacts sont perçus plusieurs fois, et donc plus facilement notables.

La perte d'aigus est également un effet propre à la WFS pour cette position de groupe. Aussi, cet artefact est plus audible que sur les ambiances, ceci étant à nouveau causé par la grande quantité en transitoires des extraits musicaux. De fait, cet artefact est plus audible sur les extraits "Bride of Deluxe" et "Hard Sun", ces derniers contenant des éléments de percussion (électronique sur "Bride of Deluxe" ; acoustique sur "Hard Sun"). Notons que sur l'extrait "Time", l'atténuation des fréquences aigues se manifeste par la perte des frottements de corde, tout à fait perceptibles en restitution stéréophonique

En termes de sensations d'espace (profondeur, relief, perspective), la WFS ne semble pas apporter des modifications significatives à notre perception des extraits musicaux, contrairement à ce que nous avons pu noter sur les ambiances pour la même position de groupe. En revanche, en ce qui concerne l'image frontale, on préserve une perception équivalente de cette image quelque soit notre position dans la salle. Par exemple, pour une position d'écoute à l'extrême gauche, l'image sonore s'étale de l'enceinte gauche à l'enceinte inter-droite, ce qui est proportionnellement équivalent à la largeur apparente perçue pour une écoute au sweet spot. En stéréophonie, des positions d'écoute extrêmes nous font percevoir d'importantes distorsions de localisation. Celles-ci nous semblent d'avantage audibles sur les deux derniers extraits : leur contenance rythmique (percussion rythmiques ou acoustiques) les rendent particulièrement sensibles à ces déformations d'image frontale. Notons également que ces distorsions d'image nous semblaient moins perceptibles sur les ambiances.

→ Position 3 : groupe bipolaire "musique" proche du second cercle du Performer

Pour cette troisième position, nous avons, comme lors des tests réalisés sur les ambiances, ajusté la largeur du groupe sur la fenêtre frontale décrite au paragraphe X. Globalement, pour cette position lointaine de groupe, on pourra noter ces quelques remarques valables pour tous les extraits :

- Une atténuation du registre aigu se fait ressentir. Là où pour les ambiances cette atténuation était moins audible sur ce troisième positionnement, les extraits musicaux amènent à une perte d'aigus quasiment aussi perceptible que pour la position 2. Cet artefact nous a amené à rééquilibrer le spectre des sources WFS pour la suite des tests, afin de se pencher plus particulièrement sur les effets d'espace permis par la synthèse de front d'onde.
- Une image sonore large est préservée, indépendante de notre placement d'auditeur dans la salle. La source musicale nous paraît dense. Une certaine homogénéité, permise par la contribution d'un plus grand nombre de haut-parleurs, peut être perçue. Lors du passage en stéréophonie, les extraits musicaux nous paraissent tous plus "restreints" à la restitution de 2 canaux, ces mêmes canaux se faisant d'avantage ressentir. En restitution WFS, la perception du dispositif en est réduite, l'écoute nous paraît plus naturelle, et ne semble pas provenir des enceintes placées à quelques mètres devant nous.

Par suite, nous avons effectués quelques constats, variables suivant les extraits. Les deux premiers extraits peuvent être qualifiés "d'orchestaux". On pourra noter une prise de son globale (ou du moins en partie), impliquant un effet de salle plutôt réaliste : toutes les sources semblent en effet provenir d'une même espace. Sur le Requiem, le chœur, quelque peu mis en avant au mixage, semble être placé devant l'orchestre, ce dernier accompagnant les principales mélodies. Il en est de même pour Time, où la réverbération des cordes, semble être de même nature que celle réservée aux cuivres et aux autres sections de l'orchestre. Du fait de cette organisation "simple" de l'espace (ensemble de sources compact + réverbération globale), la restitution via synthèse de front d'onde nous semble très intéressante. La profondeur est rendue : nous percevons les sources, ainsi que l'espace créé autour de celles-ci. Pour ces deux extraits, et pour des positions d'écoute latéralisées, nous préférons la restitution WFS.

Les deux extraits suivants répondent à des critères esthétiques plus actuels. En effet, les sources semblent être issues d'un enregistrement de proximité, en témoignent les éléments de percussion, que ce soit dans *Bride of Deluxe* ou *Hard Sun*. Tandis que le premier extrait d'entre eux nous semble plus proche des codes de la musique électronique (synthétiseurs, boîte à rythme, nombreux effets non-réalistes, ...), le second tend vers la restitution d'un groupe folk, composé d'une instrumentation cohérente, faite de guitares acoustiques et de percussions acoustiques, entre autres. Bien que le second extrait puisse être considéré comme étant plus "authentique" que le premier, il répond très probablement aux techniques actuelles de prise de son et de mixage : forte compression dynamique, placements différents des sources dans l'espace, effet "présence + réverbération" valable pour plusieurs sources, ... Par conséquent, les effets d'espace amenés par la WFS nous semblent moins perceptibles que pour les deux extraits "orchestraux". La profondeur et la sensation d'espace créée derrière les enceintes ne se manifestent plus que pour quelques éléments de la partition : quelques "nappes" de synthétiseur pour *Bride of Deluxe*, et quelques réverbérations longues pour *Hard Sun*. Notons que pour une position d'écoute au sweet spot, la stéréophonie nous amène à une bonne sensation d'espace, caractérisée par la perception des différents plans spatiaux. Cela dit, dès qu'une position d'écoute latéralisée est adoptée, la WFS est le système de restitution que nous préférons.

ANNEXE 1.7 : Spatialisation de la musique off – Protocole 2 – Détails

Détails du protocole 2 : groupe « musique » placé dans l'espace WFS ; source discrète stéréophonique (surrounds actifs via panner 5.1)

Nous avons comparé le rendu de nos extraits musicaux, d'une part diffusés via le système discret après différentes méthodes de mixage mettant à contribution les canaux surrounds, et d'autre part diffusés via le système WFS en tant qu'objets bipolaires. Le protocole 2 fait par conséquent état d'une technique primaire quant à l'usage des canaux surrounds dans le but de créer un enveloppement : le « panning » de la source stéréophonique vers l'arrière. Les bornes ont été reculées d'environ 30% sur le panner 5.1, ce qui nous offrait un dosage « avant/arrière » satisfaisant au sweet spot, autrement dit un « bon » enveloppement. En WFS, la position 3 est conservée pour le groupe bipolaire, à la différence près d'un élargissement de ses bornes, toujours dans le but d'obtenir une largeur apparente équivalente au sweet spot, entre la source WFS et la source 5.1.

Notons tout d'abord que quelques difficultés ont été rencontrées pour régler la largeur du groupe. Une certaine sensibilité de l'outil s'est manifestée, nous obligeant à prendre de minutieuses précautions dans l'élargissement des bornes du groupe, faute d'une trop grande contribution des canaux latéraux à la synthèse de la source bipolaire. En d'autres termes, un seul scroll de souris pouvait nous amener à une perception de la largeur totalement différente. Notons que ce léger « casse-tête » ressemble fortement à celui d'un panner 5.1 standard, qui demande lui aussi une grande minutie lorsqu'il s'agit de « reculer » des sources : les canaux surrounds se font eux-aussi rapidement ressentir, la limite étant très fine entre leur perception et leur « transparence ».

Sur ces nouveaux tests, nos positions d'écoute ont assez fortement été modifiées par le contenu à évaluer. Nos déplacements dans la zone d'écoute se sont plutôt dirigés vers l'arrière de la salle, dans le but de comparer la balance avant/arrière

et la sensation d'enveloppement pour différents placements d'auditeur. En 5.1, le sweet spot nous semble être la seule position d'écoute intéressante en termes de dosage façade/surround. Notons tout de même que ce « bon » dosage reste valable sur toute une ligne tracée horizontalement entre les côtés de la salle et le sweet spot. En WFS, plus on s'en approche, plus les enceintes latérales deviennent perceptibles en tant que telles, et s'éloignent de leur rôle de « source secondaire » préconisé par la théorie de la synthèse de front d'onde : une position d'écoute trop rapprochée des haut-parleurs nous amène effectivement à les percevoir en tant que « sources primaires », ce qui entrave totalement l'effet permis par le processeur WFS.

Par suite, nous nous sommes déplacés vers le fond de la salle. En 5.1, comme évoqué précédemment, les haut-parleurs de surround sont rapidement perceptibles. L'effet d'enveloppement voulu peut être considéré comme biaisé : on obtient une diffusion « multisource » de l'extrait musical, ce qui n'est en aucun cas signe d'un enveloppement, mais amène plutôt à percevoir plusieurs fois la source, provenant de différents canaux. En WFS, malgré notre recul dans la zone d'écoute, le rendu obtenu reste homogène : même pour une position d'écoute tout au fond de la petite salle, la sensation d'enveloppement est préservée. On constate alors que les haut-parleurs latéraux participent à la synthèse du groupe : ils diffusent, très faiblement certes, mais semblent participer à l'effet créé.

ANNEXE 2 : Documents relatifs aux tests

ANNEXE 2.1 : Extrait du questionnaire de tests

Tests perceptifs : l'utilisation de la WFS dans la création sonore cinématographique

Date :

Nom/Prénom :

Profession/Activité :

Etes-vous sensible au domaine du son?

Consigne

Six séquences différentes vont vous être présentées. Chaque séquence sera projetée deux fois de suite. Vous devrez comparer les deux projections selon les critères proposés.

Le test sera répété une deuxième fois.

Comment remplir les échelles de notation ?

Pour chaque séquence, vous devez évaluer relativement l'extrait 1 et l'extrait 2 selon les critères proposés.

Ex:

Pour le critère de profondeur et de relief, placer une croix au milieu de l'échelle indique que vous n'avez pas perçu de différence concernant la profondeur et le relief entre les deux extraits. Placer une croix à gauche indique que l'extrait 1 vous a semblé proposer plus de profondeur et de relief. Notez que plus votre croix est à gauche, plus la différence entre les deux extraits vous a semblé importante.

Afin d'évaluer un critère supplémentaire, cochez la case correspondant à votre position pendant la session.

J'ai fait le test en position:

A B

La position A correspond à la position centrée dans l'auditorium.

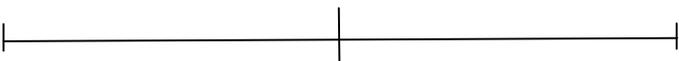
La position B correspond à la position excentrée dans l'auditorium.

Bon test à tous!

Série 1

Séquence 1

- Quel extrait vous a semblé proposer le plus de profondeur et de relief?



L'extrait 1 propose plus de relief et de profondeur que l'extrait 2

L'extrait 2 propose plus de relief et de profondeur que l'extrait 1

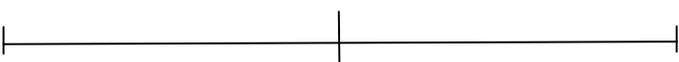
- Quel extrait vous a semblé le plus cohérent avec l'image?



L'extrait 1 est plus cohérent avec l'image que l'extrait 2

L'extrait 2 est plus cohérent avec l'image que l'extrait 1

- Quel extrait avez-vous préféré?



Extrait 1

Extrait 2

Remarque:

.....
.....

Séquence 2

- Quel extrait vous a semblé proposer le plus de profondeur et de relief?



L'extrait 1 propose plus de relief et de profondeur que l'extrait 2

L'extrait 2 propose plus de relief et de profondeur que l'extrait 1

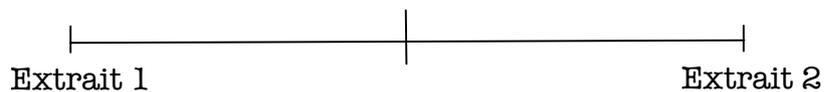
- Quel extrait vous a semblé le plus cohérent avec l'image?



L'extrait 1 est plus cohérent
avec l'image que l'extrait 2

L'extrait 2 est plus cohérent
avec l'image que l'extrait 1

- Quel extrait avez-vous préféré?



Remarque:

.....
.....

ANNEXE 2.2 : Ordre des séquences (tiré au sort)
pour chaque séance de test

SESSION I

SERIE 1

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	4/ COUPER LE VOILE	5.1	WFS
2	3/ LLMDR	WFS	5.1
3	5/LLM	5.1	WFS
4	6/GABRIELLA	5.1	WFS
5	2/PETITE BETE	WFS	5.1
6	1/MUNDOS	WFS	5.1

SERIE 2

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	2/PETITE BETE	5.1	WFS
2	1/MUNDOS	5.1	WFS
3	4/COUPER LE VOILE	WFS	5.1
4	3/LMDR	5.1	WFS
5	6/GABRIELLA	WFS	5.1
6	5/LLM	WFS	5.1

SESSION II

SERIE 1

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	4/COUPER LE VOILE	WFS	5.1
2	6/GABRIELLA	WFS	5.1
3	5/LLM	5.1	WFS
4	2/PETITE BETE	5.1	WFS
5	1/MUNDOS	WFS	5.1
6	3/LMDR	5.1	WFS

SERIE 2

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	2/PETITE BETE	WFS	5.1
2	4/COUPER LE VOILE	5.1	WFS
3	5/LLM	WFS	5.1
4	6/GABRIELLA	5.1	WFS
5	3/LMDR	WFS	5.1
6	1/MUNDOS	5.1	WFS

SESSION III

SERIE1

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	6/GABRIELLA	WFS	5.1
2	5/LLM	5.1	WFS
3	3/LMDR	5.1	WFS
4	2/PETITE BETE	5.1	WFS
5	1/MUNDOS	WFS	5.1
6	4/COUPER LE VOILE	WFS	5.1

SERIE 2

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	3/LMDR	WFS	5.1
2	4/COUPER LE VOILE	5.1	WFS
3	1/MUNDOS	5.1	WFS
4	6/GABRIELLA	5.1	WFS
5	2/PETITE BETE	WFS	5.1
6	5/LLM	WFS	5.1

SESSION IV

SERIE1

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	6/GABRIELLA	WFS	5.1
2	3/ LMDR	5.1	WFS
3	2/PETITE BETE	5.1	WFS
4	1/MUNDOS	5.1	WFS
5	5/LLM	WFS	5.1
6	4/COUPER LE VOILE	WFS	5.1

SERIE2

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	2/PETITE BETE	WFS	5.1
2	1/MUNDOS	WFS	5.1
3	5/LLM	5.1	WFS
4	3/LMDR	WFS	5.1
5	6/GABRIELLA	5.1	WFS
6	4/COUPER LE VOILE	5.1	WFS

SESSION V

SERIE1

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	1/MUNDOS	WFS	5.1
2	6/GABRIELLA	5.1	WFS
3	3/LMDR	WFS	5.1
4	2/PETITE BETE	5.1	WFS
5	4/COUPER LE VOILE	5.1	WFS
6	5/LLM	WFS	5.1

SERIE2

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	3/LMDR	5.1	WFS
2	2/PETITE BETE	WFS	5.1
3	1/MUNDOS	5.1	WFS
4	5/LLM	5.1	WFS
5	4/COUPER LE VOILE	WFS	5.1
6	6/GABRIELLA	WFS	5.1

SESSION VI

SERIE 1

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	1/MUNDOS	5.1	WFS
2	4/CLV	WFS	5.1
3	5/LLM	5.1	WFS
4	2/PETITE BETE	WFS	5.1
5	6/GABRIELLA	WFS	5.1
6	3/LMDR	5.1	WFS

SERIE 2

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	3/LMDR	WFS	5.1
2	6/GABRIELLA	5.1	WFS
3	4/CLV	5.1	WFS
4	1/MUNDOS	WFS	5.1
5	2/PETITE BETE	5.1	WFS
6	5/LLM	WFS	5.1

SESSION VII

SERIE1

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	6/GABRIELLA	5.1	WFS
2	5/LLM	WFS	5.1
3	1/MUNDOS	WFS	5.1
4	3/LMDR	5.1	WFS
5	2/PETITE BETE	WFS	5.1
6	4/COUPER LE VOILE	5.1	WFS

SERIE 2

ORDRE	SEQUENCE	1ere Projection	2eme Projection
1	1/MUNDOS	5.1	WFS
2	3/LMDR	WFS	5.1
3	2/PETITE BETE	5.1	WFS
4	4/COUPER LE VOILE	WFS	5.1
5	5/LLM	5.1	WFS
6	6/GABRIELLA	WFS	5.1

Cohérence à l'image

	Position (1 : centré ; 2 : excentré)	1 : Mundos	2 : Bete	3 : LMDR	4 : CLV	5 : LLM	6 : Gabriela	1 : Mundos	2 : Bete	3 : LMDR	4 : CLV	5 : LLM	6 : Gabriela
Christian Gil A	1	-2	0,8	2,7	2,6	3,9	4,6	3,6	0,4	0	-2,7	1,4	-0,5
Arthur S2 A	1	-1,2	0	-4,4	3,6	0,9	3,4	2,4	3,9	-2,9	3,3	1,5	-1,4
Charlotte A	1	1	-2,9	-1,3	3,4	0,5	-3,5	-5	5	-5	0	-5	0
Clément Fourment A	1	0,4	-4,2	-3	-3,2	-3,3	-1,5	0,7	-2,4	-1,1	1	0,4	-3,5
Hugo Deguillard A	1	0	-3	0	-0,7	-2,5	-1,6	0,5	-0,4	-1	-0,4	-1,2	0
François Salmon A	1	1,2	1,2	2,4	-3,7	-2,5	-3,2	1,3	1,5	1,1	-2,5	3,4	-1,6
Twargny A	1	4,2	0,9	-3,8	-2	2,9	-1,4	3,7	2,6	-3,8	1,4	2,5	-3,2
Tijou A	1	-3	2,4	-3,8	-4	-1,3	-1,2	5	3,4	-3,8	5	-3,6	-3,2
Antoine S2 A	1	-4,5	-2,8	-4,5	0,7	2,9	-1,7	3,3	-3	-2,1	-2,7	-3,6	-3
Hugo S1 A	2	0	-5	-5	0	-5	5	1	-3,1	-3,8	3,9	-3,4	-2,1
Pierre Chailloleau B	2	4,2	5	0	5	-5	0	0	5	-5	3	-2,8	0
Baptiste S1 B	2	-1	-3,9	-2	-1,3	2,5	0,2	1,4	-0,5	-2,8	2	-3,1	-1,2
Julien S1 B	2	-1	1,5	2	0,6	2,5	-2,8	0	0	-3,5	0,4	1,6	-2
Tarik B	2	-0,7	0	4	0	2,2	1,2	-1,1	1,2	-1,2	-0,6	-1,8	-1,5
Adrien Llave B	2	-2,2	-1,1	-1	0	0	-1,7	1,4	0	-4,2	0	0,6	1
Lucien S2 B	2	0	0,8	2,5	-1,7	4,4	1,5	0,6	0,2	3	2,6	3,5	1,6
Lucas S1 B	2	0	0	-5	-5	0,5	1	0	2	-1	5	-1,8	-2
Jonas B	2	2,8	1,4	3,2	2	0	-0,6	0	3,5	1,5	4,1	1,9	0
Nico B	2	3,2	-1,4	-2,4	0	-2,4	3,5	3	-2,9	-3,5	0	-1,3	0
Léa B	2	-2,7	2	1,2	1,9	0	0,6	0	0,9	0,5	0	0,9	0
facteurs			P				P	Mean					
R			0,595				0,166	0,5125					
S			0,009				0,588	0,225					
G			0,098				0,002	-1,42					
							0,21	0,525					
							0,611	-0,2175					
							0,129	-0,52					

ANNEXE 3 : Documentations techniques

ANNEXE 3.1 : Enceintes DX8 – Spécifications techniques

Série Dispersion

Enceinte compacte polyvalente

Deux voies coaxiales, charge bass-reflex
Un 20 cm à moteur ventilé (bobine 45 mm)
Un moteur HF 1" (bobine 45 mm)

Caractéristiques

Polyvalence façade/retour
Clarté sonore
Rendement acoustique élevé

Applications

Façade compacte
Retour de scène de proximité
Sonorisation de complément
Couverture répartie

Spécifications

200 W AES
97 dB SPL @ 1W à 1m
8 Ohms
105° conique
422 x 236 x 260 mm

2 protections thermiques
Fonctionne en mono-amplification
Subwoofer optionnel

Le modèle APG DX8 est une enceinte compacte deux voies, large bande. La section grave/médium est confiée à un haut-parleur de 20 cm à grande élongation chargé en bass-reflex et monté en coaxial avec la section un moteur à compression 1" pour la section aiguë.

Comme les autres enceintes de la série APG Dispersion, la DX8 est équipée d'un ensemble de haut-parleurs coaxiaux ce qui lui confère un champ acoustique cohérent sur une ouverture acoustique conique de 105°.

La DX8 se caractérise par son rapport taille/puissance, sa clarté de restitution sonore et son ergonomie particulièrement complète et sophistiquée : l'ébénisterie comporte un pan coupé à 30° pour une exploitation au sol en retour de scène et intègre deux poignées encastrées, une embase pour pied standard de 36 m et un réceptacle pour l'accessoire d'accroche amovible QRFS30. Des inserts métalliques permettent également le montage sur l'étrier ETDX8 en installation fixe.

Les processeurs dédiés SPX8 et LPDX8 permettent d'optimiser le fonctionnement de la DX8 en mode large bande et le filtrage en combinaison avec un caisson de basses.

Pour les renforts de basses APG préconise l'emploi des "subwoofers" suivants : SB115 ou TB115S.

DX8



Enceinte DX8 présentées en façade et en retour de scène

La DX8 est une enceinte satellite offrant une totale polyvalence entre les applications de retour de scène et de façade avec ou sans renfort de grave. La DX8 est destinée à la sonorisation des petites et moyennes surfaces ainsi qu'à la couverture répartie de grands espaces. Sa compacité et sa clarté sonore en font le produit idéal pour les applications où la discrétion doit être conjuguée à une précision de restitution sonore élevée.

La DX8 est utilisable suivant plusieurs configurations : elle peut être posée, montée sur pied, fixée à la verticale ou à l'horizontale grâce à son accessoire d'accroche rapide QRFS30. Pour l'installation en fixe, est prévu l'étrier optionnel ETDX8. La DX8 peut être équipée, d'un transformateur pour les applications de sonorisation distribuée sur lesquelles un câblage en ligne 70V/100V est requis (options T50 ou T100).

En mode large bande sans processeur, la DX8 est réservée aux applications vocales type conférence tandis qu'avec processeur, elle permet la diffusion musicale de haute précision. Elle constitue un retour de scène particulièrement discret et efficace en théâtre, sur de petites scènes et plateaux TV. Combinée avec des caissons de basses APG, elle permet de former des systèmes de façade compacts et portables dotés d'une importante capacité dynamique et haute qualité sonore.

Les processeurs LPDX8 et SPX8 ainsi que les processeurs numériques APG, permettent de gérer un système stéréo + subwoofer, respectivement avec ou sans protection active des haut-parleurs.

APG



Caractéristiques	Voie basse	Voie haute
Bande passante avec processeur (± 3 dB)	60 Hz à 2 kHz	2 kHz à 19 kHz
Bande passante sans processeur (± 3 dB)	80 Hz à 2 kHz	2 kHz à 19 kHz
Efficacité 1W à 1m, $f > 120$ Hz		97 dB SPL (1)
Niveau maxi continu à 1m, $f > 100$ Hz		118 dB SPL
Niveau crête à 1m, $f > 100$ Hz		124 dB SPL
Angles de couverture		105° conique (2)
Impédance nominale		8 ohms

Composants

Transducteurs	1 x 20 cm à moteur ventilé	1 moteur à compression 1"
Diamètre de la bobine	45 mm	26 mm
Type de charge	Bass reflex	Pavillon coaxial

Puissances

Amplification recommandée (3)	200 à 400 W
Instantanée	600 W
AES (3)	200 W

Construction et caractéristiques physiques

Ebénisterie	Multipils de bouleau marine feuilluré collé
Filtres passifs	Selfs à air, condensateurs à film plastique, résistances bobinées
Protactions contre les surcharges	Thermiques à réarmement automatique
Finition	Revêtement aquarethane noir chargé à haute résistance
Face avant	Grille en acier perforé, mousse acoustique de 5 mm noire
Connecteurs	2 SPEAKON NLM4P (4)
Poignée	1 poignée encastrée
Montage sur pied	Support femelle pour pied de 36 mm
Dimensions (H.L.P)	422 x 236 x 260 mm
Masse nette unitaire	12 kg
Masse brute, la paire emballée	28 kg

Options et Accessoires

QAFS30	Système d'accroche rapide amovible (30kg max)
QFS2	2 x cornières métalliques intégrée pour accroche rapide
ETDX8 (5)	Etrier de fixation d'installation fixe pour enceintes DX8
FCDX8	Flight Case pour 2 enceintes DX8
Options de peinture et tropicalisation (6)	Nous Consulter

Traitement du signal

Le SPX8 assure les fonctions de filtrage, égalisation, protection et la distribution de signal pour un système de façade stéréo en DX8 éventuellement combinées avec des "subwoofers" APG.

Le LPDX8 comporte en plus les fonctions de protection active des haut-parleurs par simulation des paramètres limitatifs : déplacement de la membrane, montée en température de la bobine, clip de l'amplificateur. Il permet d'optimiser l'alignement en phase des "subwoofers" par rapport aux enceintes satellites et comporte un clavier de sélection permettant d'adapter le type de "subwoofer" APG utilisé.

Les processeurs numériques assurent les fonctions de filtrage, égalisation, protection et la distribution de signal pour tous types de systèmes. Les "presets" usine sont disponibles auprès d'APG France.

APG FRANCE - 19 bis rue des écoles - Site Valnor ZI Haute - 95500 Le Thillay
Tel. : +33 (0)1 30 18 92 70 - Fax : +33 (0)1 30 18 92 71 - e.mail : contact@apg.tm.fr - www.apg.tm.fr

(1) L'efficacité de 97 dB s'entend après atténuation ; avant atténuation, l'efficacité est de 100 dB. Cette atténuation permet non seulement de linéariser le système mais aussi d'augmenter la puissance admissible dans la section considérée.

(2) La dispersion est contrôlée à partir de 600 Hz, les angles donnés sont peu significatifs en dessous de cette fréquence.

(3) L'amplification recommandée pour exploiter les caractéristiques de dynamique, de qualité sonore et de fiabilité des produits APG doit correspondre au minimum à la puissance admissible AES. En cas d'utilisation à faible puissance (proximité, diffusion répartie etc.), une amplification inférieure est acceptable, sans toutefois être inférieure à la moitié de la puissance AES.

La norme de puissance AES correspond à une utilisation pendant 2 heures, en bruit rose filtré (facteur de crête de 6 dB) sur une décade dans la bande utile.

(4) Les SPEAKON sont câblés point chaud en 1+, point froid en 1-.

(5) L'option ETDX8 comprend un étrier de fixation simple réservé à l'installation en fixe de l'enceinte.

(6) La DX8 peut être fournie dans différents coloris sur la base d'une référence PANTONE ou RAL. Pour l'exploitation en milieu tropical ou humide, 2 niveaux de protection et traitement imperméabilisant peuvent être appliqués.

5 ANS GARANTIE Une garantie totale de cinq ans couvre les pièces et les transducteurs contre tout vice de fabrication dans les conditions normales d'utilisation des produits.

Diffusion : Novembre 2013

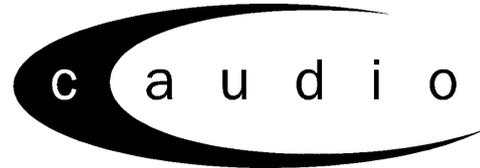
APG mène une politique de recherche et de développement destinée à l'amélioration de ses produits. Pour cette raison, de nouveaux matériaux, méthodes de fabrication et changements de principe peuvent être introduits sans avertissement préalable. De ce fait, un produit APG peut différer sous certains aspects de sa description publiée, toutefois, sauf indication contraire, ses caractéristiques seront supérieures ou égales à celles publiées.

APG

ANNEXE 3.2 : Amplificateurs utilisés pour les rampes latérales – Spécifications techniques

TECHNICAL SPECIFICATIONS

GB Series



Power Ratings

GB202	
Into 8 Ohms	145W
Into 4 Ohms	200W
Into 2 Ohms	250W
GB402	
Into 8 Ohms	260W
Into 4 Ohms	400W
Into 2 Ohms	570W
GB602	
Into 8 Ohms	370W
Into 4 Ohms	600W
Into 2 Ohms	840W

Benefits of the GB Series

- Simple reliable design
- Heavy duty PSU incorporating oversized Toroidal transformer
- Clip Limiter and Subsonic Filters
- Efficient Forced Air Cooling System
- All Aspect Protection Circuitry
- Electronically Balanced inputs
- Extended Capability via AMPSAP Protocol

Cost-effective high performance power amplifiers

The GB Series breaks new ground in affordable power. Primarily intended as a competitively priced, reliable and powerful work horse for the musical instrument market, the GB Series has quickly established itself as a must have amongst touring musicians. Such is the quality and practicality of the range that it has also found favour amongst a wide range of users including discos, clubs, bars, venues, theatre installers and cinema sound specifiers.

The implementation of the AMPSAP protocol means that sophisticated internally fitted system controllers are available to drive active loudspeaker systems.

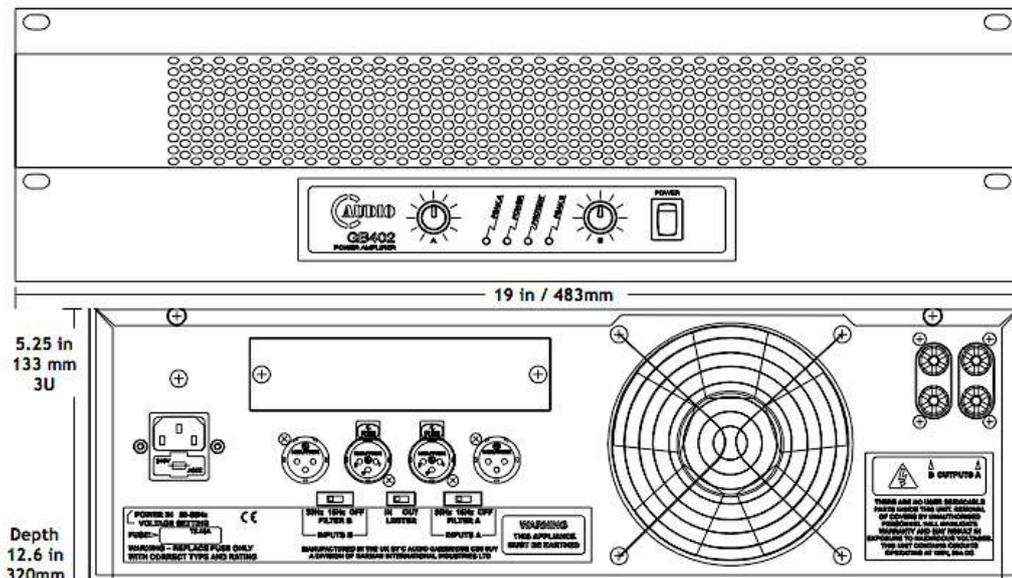
The GB Series may be bridged externally for higher power mono operation.

Technical Specifications

Power Ratings	GB202	GB402	GB602
<small>Measured per channel, both channels driven at 1kHz to no more than 0.1% THD+N</small>			
8 Ohms	145Wrms	260Wrms	370Wrms
4 Ohms	200Wrms	400Wrms	600Wrms
2 Ohms	250Wrms	570Wrms	840Wrms
<small>Note: 2 Ohm capability is thermally limited</small>			
Bridged Mono			
8 Ohms	400Wrms	800Wrms	1200Wrms
<small>Measured per channel, both channels driven at 1kHz to no more than 1% THD+N</small>			
8 Ohms	155Wrms	275Wrms	400Wrms
4 Ohms	225Wrms	445Wrms	670Wrms
2 Ohms	260Wrms	610Wrms	880Wrms
<small>Note: 2 Ohm capability is thermally limited</small>			
Sensitivity <small>(for full rated power at 4 Ohms)</small>	0.725Vrms	1.025Vrms	1.26Vrms
Frequency Resp.	10Hz-40kHz, +0/-1dB		
Damping Factor	>200 (ref. 8 Ohms)		
Hum and Noise	-100dB (ref. rated 4 Ohm power, 20-20kHz)		
Channel Separation	60dB at 1kHz		
Slew Rate	17V per microsecond		
Rise Time	5 microseconds		
Input Impedance	20 kOhm balanced		
Input Connectors	Female XLR & Male parallel out (pin2 Hot)		
Subsonic filters	12dB/oct, 30Hz/15Hz/OFF		
Clip Limiter	Linear Optocoupler		
Output Connectors	Tamper Proof Binding posts		
Output Circuit Type	Bipolar Class AB		
Power Requirement	100-120 or 220-240Vac±20% 50/60Hz		
Weight - Net	13Kg 30lbs	14Kg 31lbs	15Kg 33lbs
Shipping - Gross	17Kg 37lb	18Kg 38lbs	19Kg 42lbs
Dimensions	483mm x 320mm x 133mm (3U) 19in x 12.6in x 5.25in (3 rack units)		

Trade Descriptions Act: C Audio have a policy of continued product improvement and accordingly reserve the right to change features and specifications without prior notice.

January 1999



AMPSAP Options

The GB's Advanced Multi-tasking Primary Signal Access Port means that option cards may be fitted to implement basic crossover functions.

The AMPSAP card for the GB Series provides a stereo 2-way crossover, featuring a sub-sonic high pass filter on the LF outputs and variable constant directivity equalisation on the HF outputs.

The four outputs are provided on XLR connectors on the rear panel of the option card.

The amplifier shall have two channels, each capable of producing an output of 200/400/600 Watts into a 4 Ohm load with both channels driven. Each input shall be electronically balanced and have a CMRR of greater than 50dB at 1kHz, and effective filtration against RF and DC.

Full rated output with a 4 Ohm load shall be achieved by an input signal not exceeding 1.3Vrms per channel.

Each channel shall have a +0/-1dB frequency response from 10Hz to 40kHz at 1W and full rated power into a nominal 4 Ohm and shall exhibit harmonic distortion not exceeding 0.01% at 1kHz. Hum and noise shall be at least 100dB below full rated output when measured over a 20Hz to 20kHz bandwidth with

50 Ohm input termination. Channel separation shall be in excess of 60dB at 1kHz. The amplifier shall be stable into any load configuration with any combination of open or grounded input connection, and shall protect itself and its loudspeaker loads against mismatched, short or open circuit loads, or any failure which might otherwise cause a DC offset voltage to appear at its output.

Muting circuits shall automatically disconnect the output loads during power-up and power-down and a self-resetting thermal sensing system shall be incorporated to protect the power transistors against over-temperature operation. A two-speed forced air cooling fan shall be incorporated, venting warm air out of

the front panel. Each amplifier channel shall have a rotary level control accessible from the front panel and carry an output signal clip indicator which accurately indicates clipping irrespective of output loading or mains supply voltage. LEDs shall be provided to indicate the status of AC power and the amplifier's own protection systems. Provision shall be made for fitting signal processing modules within the amplifier, compatible with C Audio's proprietary AMPSAP protocol. Audio input shall be via mirrored XLR connectors and outputs via tamper proof binding posts.

The amplifier shall be a model GB 202/402/602 manufactured by C Audio, Potters Bar, England.



C Audio

A Division of Harman International Industries Ltd,
Summit Road, Potters Bar, Herts EN6 3JB England
Tel +44 (0)1707 643643 Fax +44 (0)1707 660755
<http://www.c-audio.com>

BSS Audio USA

Harman Pro North America Inc
1449, Donelson Pike, Nashville TN37217
Tel (615) 360 0277 Fax (615) 360 0480

ANNEXE 3.3 : Processeur Wave 1 – Spécifications techniques

Technical Specification: Sonic Wave I

Description	Wave field synthesis rendering unit for processing of up to 24 input and up to 64 output channels.	
Audio Input	Digital MADI (coaxial or optical) or ADAT Internal playback engine (HD option) UltraLight ADAT only	24 24 8
Audio Outputs	Digital MADI (coaxial or optical) or ADAT UltraLight ADAT only	32/64 16
Audio Processing	Sample rate and bit depth	48 kHz 24 Bit
Storage	SSD (HD Version Only)	60 GB
Connectivity	Ethernet / LAN (RJ45)	Yes
Latency	Processing-Time: End-to-End	< 7ms
Power Supply	Electric Consumption	480W
Required AC Mains	230 Volts / 50 Hz with PFC 115 Volts / 60 Hz with PFC	5.0A 8.0A
AC Mains Connector	Integrated EMI Filter	IEC
Dimensions	19" Rack Case (Units) 430 x 480 x 180 mm (W x D x H)	4 U
Safety Standard	CE, FCC	



Sonic Emotion Pro Systems
Eichweg 6
8154 Oberglatt (Zürich) Switzerland

Tel: +41 44 850 08 38
Email: pro@sonicemotion.com
www.sonicemotion.com



ANNEXE 3.4 : Convertisseur *MADI to Analog* – Spécifications techniques

ANDIAMO 2


[Description](#)
[Technical data](#)
[Download](#)


MADI Ports (I/O)*: 2 x SC-Socket multi-mode
or
1 x SC-Socket multi-mode & 1 x coaxial BNC, 75 Ω

Analog (I/O): 8 x DSUB-25, balanced (TASCAM pinout)
Analog level (I/O): +6 dBu / +15 dBu (digitally switchable) or
+9 dBu / +18 dBu (digitally switchable) or
+15 dBu / +24 dBu (digitally switchable)

A/D

SNR: -115,5 dB RMS (20 Hz - 20 kHz) / -118 dB(A)
THD @ -1dBFS: -113 dB
Frequency Response: -0,15 dB (10 Hz) / -0,15 dB (20 kHz)

D/A

SNR: -113,5 dB RMS (20 Hz - 20 kHz) / -116,5 dB(A)
THD @ -1dBFS: -100 dB
Frequency Response: -0,5 dB (10 Hz) / -0,15 dB (20 kHz)

Global

Word clock: coaxial BNC (75 Ω termination switchable)
USB: USB 2.0 for firmware updates and remote control
Sample rates: 44.1, 48, 88.2, 96 kHz (+/- 12,5%)
MADI formats: 48k / 96k Frame, 56 / 64 channel, S/MUX
Power supply: 2 x 84 V to 264 V AC / 47 Hz to 63 Hz / safety class 1, phaseredundant
Dimensions: Width 19" (483 mm)
Height 1 HU (44,5 mm)
Depth 10" (254 mm)
Weight: about 3 kg