

UNIVERSITÉ PARIS VIII

UFR Arts

Département de Musique

**La spatialisation des musiques électroniques
du point de vue de l'auditeur :
perception, immersion, concert spatialisé**

Amélie Nilles

Mémoire de Master 2

Sous la direction de Anne Sèdes

2017 – 2018

Je remercie, en premier lieu, ma directrice de recherche Anne Sèdes qui a toujours su me guider de par ses idées et conseils, et m'a montré une grande confiance dès le début de mes recherches de Master. Elle m'a également donné carte blanche pour l'organisation du concert *Immersion Sonore* à la Maison des Sciences de l'Homme Paris Nord.

Je remercie Eric Lyon qui a pris soin de répondre avec rapidité et précision à mes nombreuses questions concernant le portage d'une œuvre spatialisée sur différents systèmes.

Enfin, un grand merci à toutes les personnes qui m'ont aidée à mettre en place la première édition du concert *Immersion Sonore*, et notamment Arnault Damien, Frank Gillardeaux, Quentin Nivromont, Filippos Sakagian et João Svidzinski.

Table des matières

Introduction.....	5
I. La perception de l'espace sonore.....	9
I.1. La perception auditive de l'espace.....	10
I.1.1 Les indices de localisation d'une source sonore : ITD, IPD, ILD/IID.....	11
I.1.2 La localisation d'une source sonore (suite) : les indices spectraux, le corps de l'auditeur.....	12
I.1.3 En résumé : la localisation et la distance.....	15
I.2. Les facteurs visuels, mémoriels et sociologiques liés à la perception de l'espace.....	18
I.2.1 Le rôle de la vision dans la perception de l'espace : points de vues scientifiques et sociologiques.....	19
I.2.2 Le rôle de la mémoire dans la perception de l'espace.....	20
I.2.3 Le travail de « l'oreille spatiale ».....	22
I.3. Conclusions.....	25
I.3.1 Influence des connaissances en psycho-acoustique sur les techniques de spatialisation.....	25
I.3.2 Remarques sur différentes situations d'écoute de la musique électronique spatialisée : le concert et l'écoute au casque.....	29
II. La notion d'immersion dans la musique électronique spatialisée.....	32
II.1. Ce qui est à l'origine de l'immersion.....	34
II.1.1 La réverbération et le sentiment d'enveloppement.....	34
II.1.2 Les conditions techniques pour obtenir de l'immersion (systèmes de diffusion). ..	36
II.1.3 Les « attributs immersifs » et les techniques de spatialisation conduisant à l'immersion.....	40
II.2. Ce qui est conséquence de l'immersion.....	42
II.2.1 Immersion et « naturalness » : une expérience sonore naturelle.....	42
II.2.2 Immersion et ressenti par le corps de l'auditeur.....	45
II.2.3 L'espace sonore immersif comme moyen d'expression et de communication.....	47
II.3. Conclusion.....	50
III. Le concert spatialisé vécu comme une expérience par l'auditeur. Le cas pratique du concert <i>Immersion Sonore</i>.....	52
III.1. Le cas pratique du concert spatialisé <i>Immersion Sonore</i>	53
III.1.1 Présentation du concert : le déroulement, les aspects techniques.....	53

III.1.2 Ma participation au concert en tant qu'artiste : mes choix de spatialisation.....	56
III.1.3 Les retours du public : réponses au questionnaire et commentaires.....	61
III.2 Réflexions sur le cas général du concert de musiques électroniques spatialisées, suite à l'évènement <i>Immersion Sonore</i>	65
III.2.1 De la nécessité de pouvoir adapter son œuvre spatialisée à tout type de systèmes de diffusion et de salles.....	65
III.2.2 Repenser l'organisation globale du concert spatialisé : la place de l'artiste, la place de l'auditeur.....	70
III.2.3 Les concerts spatialisés participent à la formation de « l'oreille spatiale » et au développement de l'attention.....	76
III.3. Conclusion	78
Conclusion.....	81
ANNEXE A : Questionnaire <i>Immersion Sonore</i>	86
ANNEXE B : Programme <i>Immersion Sonore</i>	88
ANNEXE C : <i>Immersion Sonore</i> : modélisation du dispositif de haut-parleurs dans l'Auditorium et visualisation du placement des répondants au questionnaire. .	89
ANNEXE D : Outils de spatialisation utilisés pour le concert <i>Immersion Sonore</i>	90
ANNEXE E : Synoptique du dispositif technique du concert <i>Immersion Sonore</i>	92
ANNEXE F : Interview de Eric Lyon.....	93
Bibliographie.....	95

Introduction

« No viable approach to spatial audio can ignore how the listener perceives and thinks, nor can it limit its considerations simply to the physical acoustics of reproduction. We need to engage spatial audio in all of its intricacies. Any approach to spatial audio appropriate to our time ought to be based in knowledge of human perception and cognition. »¹ [Kendall, 2010]

« Even though humans perceive sound in 3D, the majority of electroacoustic works are still presented on two-dimensional (2D) loudspeaker configurations (Sazdov et al., 2007)². According to some, the creative possibilities of presenting electroacoustic music in three dimensions have not been adequately investigated (Sazdov et al., 2007 ; Normandeau, 2009)³. The need to investigate how sound moves and how music is perceived in an immersive environment is required in order to fully realize the creative compositional possibilities of 3D loudspeaker configurations (Sazdov et al., 2007; Normandeau, 2009). »⁴ [Lynch & Sazdov, 2011]

Alors que le travail de la dimension *espace*, dans les musiques électroniques, intéresse toujours plus de compositeurs, la notion d'espace sonore, à l'aube de 2020, semble toucher d'avantage de secteurs. Outre le cinéma qui emploie depuis de nombreuses années déjà le son spatialisé dans ses salles, et développe des moyens de reproduire le même type d'écoute dans

1 « Aucune approche viable en son spatial ne peut ignorer comment l'auditeur perçoit et pense, ni ne peut limiter ses considérations à l'acoustique physique de la reproduction. Nous devons travailler le son spatial dans toutes ses complexités. Toute approche du son spatial en accord avec notre époque doit être fondée sur la connaissance de la perception et de la cognition humaines »

2 L'auteur fait référence à : Sazdov, R., Paine, G., Stevens, K., 2007, *Perceptual Investigation into envelopment, spatial clarity, and engulfment in reproduced multi-channel audio*, Proceedings of AES 31st International Conference, London.

3 L'auteur fait référence à : [Normandeau, 2009].

4 « Même si les humains perçoivent le son en 3D, la majorité des œuvres électroacoustiques sont toujours présentées sur des configurations de haut-parleurs à deux dimensions (2D) (Sazdov et al., 2007). Selon certains, les possibilités créatives de présenter la musique électroacoustique en trois dimensions n'ont pas été suffisamment étudiées (Sazdov et al., 2007, Normandeau, 2009). La nécessité d'étudier comment le son se déplace et comment la musique est perçue dans un environnement immersif est impérative pour réaliser pleinement les possibilités créatives compositionnelles des configurations de haut-parleurs 3D (Sazdov et al., 2007; Normandeau, 2009). »

le cadre domestique, le secteur du jeu vidéo utilise également les systèmes 5.1 et le binaural pour l'écoute au casque. De nombreuses radios en ligne proposent des contenus audio en binaural, et l'avènement de la vidéo à 360° et de la réalité virtuelle pousse les ingénieurs du son à travailler la spatialisation. Ainsi, une majorité de personnes, adeptes de jeux vidéos ou de nouvelles technologies, a accès à du contenu audio spatial, sans nécessairement écouter de la musique spatialisée et, malgré elle, se sensibilise à ces nouvelles conditions d'écoute du son. En musique, les compositeurs travaillant avec l'espace semblent être attirés par l'écoute dite "naturelle", au plus proche de la "réalité". Autrement dit, on souhaite que la musique spatialisée sonne comme "naturelle", non pas dans son degré de précision, ni dans sa ressemblance à la réalité – au contraire : on est habitué à entendre des choses normalement impossibles spatialement lorsque l'on écoute de la musique spatialisée, car on l'accepte à des fins artistiques [Kendall, 2010] –, mais plutôt dans sa qualité d'écoute : essayer de ne pas entendre les haut-parleurs qui ne sont qu'un intermédiaire entre la source sonore et l'auditeur. Nous avons vu, dans notre mémoire de première année⁵, que certaines technologies de reproduction favorisent cette écoute dite "naturelle"⁶. Mais ce n'est pas suffisant : l'acoustique du lieu de diffusion, mais aussi la façon avec laquelle on utilise les sons, on compose l'espace, sont extrêmement importantes dans le résultat final et de cela va aussi dépendre la manière avec laquelle l'auditeur va recevoir l'oeuvre spatialisée. À ce propos, Denis Smalley [1991] appelle « espace superposé » l'espace ultime que l'auditeur perçoit, qui est la combinaison des espaces composés et de l'espace d'écoute. Ainsi, pour comprendre quelle sera la conséquence audible de tel geste ou telle technique de spatialisation, qu'elle soit considérée comme inhérente à l'oeuvre, comme un moyen de structurer l'oeuvre, ou encore comme une ornementation (cf. [Nilles, 2016]), il est essentiel, pour le compositeur, de pouvoir "contrôler" en amont le résultat sonore spatial et donc indispensable de se placer du point de vue de l'auditeur. Ce constat nous amène à notre première question : comment l'espace sonore est-il perçu ? D'après les citations de Kendall [2010], Lynch et Sazdov [2011] présentées au début de cette introduction, il semblerait que tout compositeur de musique électronique s'intéressant à la spatialisation du son et désireux d'aller plus loin soit tenu de connaître les bases de la perception humaine de l'espace sonore. Dans [Nilles, 2016] nous avons tenté de donner aux compositeurs quelques pistes créatives pour travailler l'espace sonore ou *avec* l'espace sonore en discutant de diverses techniques de spatialisation du son et de technologies de reproduction de l'espace. À présent, nous souhaitons nous intéresser aux façons selon lesquelles l'auditeur reçoit la musique électronique spatialisée, d'un point de vue théorique, mais aussi d'un point

⁵ Nous faisons référence à [Nilles, 2016].

⁶ La Wave Field Synthesis notamment.

de vue plus pratique, en considérant la situation de concert spatialisé. Notre première partie sera ainsi consacrée à la perception de l'espace sonore. Nous verrons tout d'abord quelques bases de psycho-acoustique, puis nous nous intéresserons aux autres facteurs dont dépend la perception de l'espace, c'est-à-dire la vision et la mémoire. En effet, toujours dans l'optique de ne pas perdre de vue la musique électronique et les concerts spatialisés, nous nous demanderons en quoi la vision influence la perception de l'espace sonore et s'il est possible pour l'auditeur de travailler son « oreille spatiale » au fil du temps. L'ensemble de ces notions, remis dans le contexte de la musique spatialisée en partie I.3 – notamment en se posant la question de la validité des techniques de spatialisation usuelles sur le plan perceptif – , servira de base théorique pour le reste du présent mémoire.

D'autre part, le sondage réalisé par Peters [2010] sur un ensemble de compositeurs de musiques électroniques travaillant avec la dimension *espace* met en évidence l'importance du paramètre « immersion ». Payri [2011] vient confirmer ce propos en déclarant que le nouvel enjeu de la musique spatialisée est l'immersion et non le mouvement du son. En outre nous avons constaté, dans [Nilles, 2016], que la notion d'immersion est la conséquence de nombreuses techniques de spatialisation du son, et qu'elle revient très fréquemment lorsque l'on parle de spatialisation car elle semble intéresser d'innombrables compositeurs issus de tout type de musiques électroniques. Ainsi, l'envie de créer un espace immersif pourrait constituer un point commun entre l'ensemble de ces musiques... Néanmoins, nous venons de le voir, il est essentiel de toujours se positionner du point de vue de l'auditeur lorsque l'on parle d'espaces sonores. Qu'en est-il de la perception de l'immersion en musique électronique spatialisée ? Quelles sont ses conséquences sur l'auditeur du point de vue de l'audition, des émotions, de l'attention ? La notion d'immersion sera alors au centre de nos considérations dans une seconde partie. Nous tenterons de la définir le plus largement possible en nous intéressant, dans un premier temps, à la façon dont on peut créer un espace immersif et, dans un deuxième temps, aux conséquences perceptives de ce type d'espace. Lynch et Sazdov [2011]⁷ proposent les dispositifs 3D pour créer un espace immersif et "naturel". Nous nous poserons donc la question de la validité des systèmes stéréo, 2D et 3D dans la reproduction d'un espace sonore immersif. De même, travailler l'immersion est-il un moyen d'obtenir une écoute naturelle ?

Dans la troisième partie de ce présent document, nous souhaitons proposer un regard plus pratique en ce qui concerne la perception de l'auditeur de musique électronique spatialisée, en considérant la situation du concert spatialisé, à caractère immersif, et

⁷ Voir la citation en début d'introduction du présent document.

notamment le cas particulier du concert *Immersion Sonore* qui s'est tenu à la MSH Paris-Nord, le 7 Décembre 2017. Un questionnaire a été distribué permettant de recueillir l'avis des auditeurs ; ces réponses vont permettre d'orienter le propos et de nous poser certaines questions en rapport avec l'influence de la vision lors d'un concert spatialisé, le sentiment d'immersion, ou encore la place de l'auditeur par rapport au dispositif de haut-parleurs. Enfin, nous avons fait l'hypothèse, en organisant le concert *Immersion Sonore*, qu'il était possible de rassembler différents publics et divers genres de musiques électroniques au sein d'un même concert, par le biais de l'écoute de l'espace – cette volonté provient essentiellement du fait que l'auteur de ce présent mémoire écoute et compose de la musique à mi-chemin entre plusieurs genres (jazz, chanson et musique électroacoustique). En effet, le milieu des concerts de musiques spatialisées, en France, nous semble encore assez restreint même s'il tend, ces dernières années, à s'ouvrir : il s'agit souvent d'un public d'habitues qui écoute généralement le même type de musique. Nous chercherons à vérifier notre hypothèse en nous demandant si la notion d'immersion participe à la création d'une expérience sonore et spatiale et peut ainsi stimuler l'attention de l'auditeur. Dans ce cas, est-ce que la création d'espaces sonores, en musique, a la capacité de modifier l'écoute du temps ? Cette déclaration de Denis Smalley, dans [Austin, 2000], pourrait être un premier élément de réponse : « The spatial experience of electroacoustic music is one of the particular aspects it has to offer that no other musical art has to offer in such variety or with such vividness... so, that makes it unique. »⁸

Ainsi, les parties I et II vont poser les bases théoriques de notre questionnement, que nous confronterons en partie III au cas pratique du concert de musiques électroniques spatialisées. L'auditeur sera toujours placé au centre de la réflexion.

8 « L'expérience spatiale de la musique électroacoustique est l'un des aspects particuliers qu'elle a à offrir, qu'aucun autre art musical ne peut proposer avec tant de diversité et de vivacité... donc c'est ce qui la rend unique. »

I. La perception de l'espace sonore

« However, it is argued that before effective spatialisation approaches can be formulated, it is first necessary to understand how spatial audio is perceived. »⁹ [Lynch, 2014]

Dans cette partie, nous allons présenter les bases de la perception auditive humaine de l'espace et mettre ainsi en lumière les divers facteurs qui l'influencent, à plus ou moins grande échelle. Ces notions touchent principalement à la psycho-acoustique, mais d'autres facteurs entrent en jeu comme la vision et la mémoire, abordées d'un point de vue plus sociologique et naturaliste. Nous nous interrogerons sur la possibilité ou non pour l'auditeur de travailler son "oreille spatiale", dans le but de lier cela à l'écoute régulière de musiques spatialisées. L'objectif est de fournir des outils au compositeur travaillant l'espace sonore en partant d'un angle différent par rapport à celui employé dans notre précédent écrit¹⁰ : celui non négligeable de l'auditeur. De nombreuses expériences ont été mises en œuvre par des neuroscientifiques ou des musicologues au cours de ces cinquante dernières années. Nous allons nous appuyer de certaines d'entre elles pour expliciter notre propos. En revanche, nous gardons toujours en tête que ces connaissances vont nous servir pour la musique électronique spatialisée ; elles ont donc été sélectionnées en fonction et ne sont, par conséquent, pas toujours extrêmement pointues. Nous invitons le lecteur souhaitant un contenu plus approfondi concernant la perception auditive spatiale à se référer au livre très complet de Jens Blauert : *Spatial Hearing: The Psychophysics of Human Sound Localization* [1997].

9 « Cependant, on soutient qu'avant de formuler des approches de spatialisation claires, il est d'abord nécessaire de comprendre comment l'espace sonore est perçu. »

10 [Nilles, 2016]

I.1. La perception auditive de l'espace

En psycho-acoustique, pour déterminer l'emplacement d'une source sonore par rapport à la tête, l'espace est défini par trois plans : le plan horizontal (correspondant à 0° d'élévation), le plan médian ou l'élévation (correspondant à 0° d'azimuth) et le plan frontal (une source dont les coordonnées correspondant à 0° d'azimuth et 0° d'élévation se tient directement face à la tête) [Moore, 1989] [Blauert, 1997] (cf. Figure1 ci-dessous). La localisation d'un son dans cet espace par l'auditeur dépend pour autant de propriétés temporelles, d'amplitudes et spectrales. Des modifications de ces caractéristiques, induites par des mouvements de tête ou de la source sonore elle-même, influencent la perception de la direction de ce son. Il faut noter que les oreilles reçoivent plein d'informations qu'elles ne peuvent distinguer, elles sont principalement des récepteurs. Contrairement aux yeux, on ne peut pas, à proprement parler, "fermer nos oreilles". C'est le cerveau qui va analyser et traduire ces informations en données spatiales.

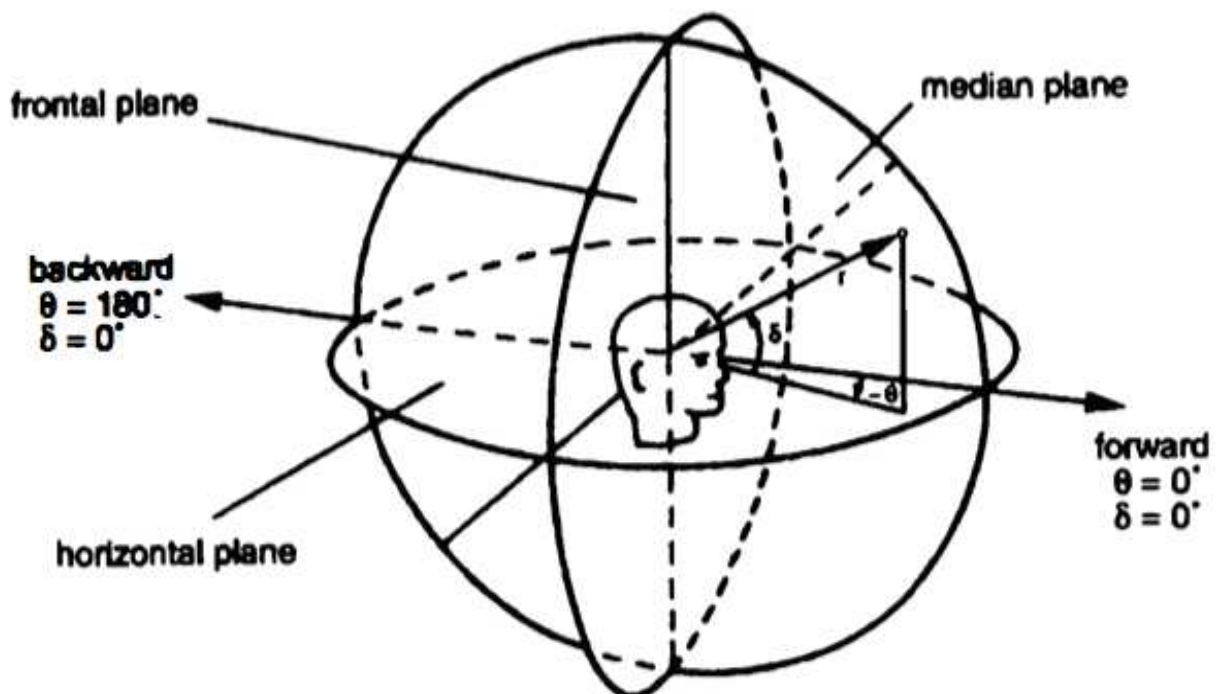


Figure1 : Système de coordonnées utilisé pour déterminer la position des sources sonores dans l'espace, par rapport à la tête de l'auditeur. θ représente l'azimuth, δ représente l'élévation et r la distance. (source : [Moore, 1989])

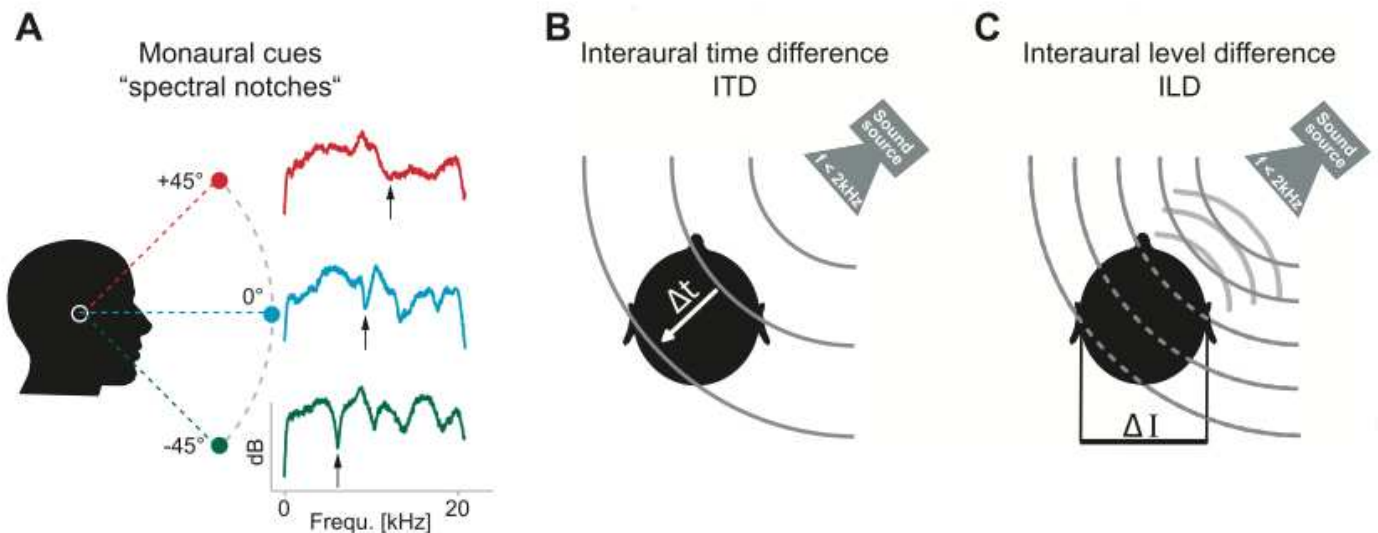
I.1.1 Les indices de localisation d'une source sonore : ITD, IPD, ILD/IID

« The majority of spatial perception is dependent on the listener having two ears, although certain monaural cues have been shown to exist – in other words it is mainly the differences in signals received by the two ears that matter. »¹¹ [Rumsey, 2001]

Plusieurs critères entrent en jeu lorsque l'on considère la localisation d'une source sonore dans l'espace, dépendants chacun du fait que l'on possède une oreille de chaque côté de la tête :

- l'indice interaural de temps ITD (Interaural Time Difference) : il s'agit d'une valeur de temps qui traduit la différence de marche que subit l'onde sonore pour atteindre l'oreille opposée. L'ITD permet de localiser la source sonore en direction de l'oreille qui reçoit le signal le plus tôt. Le calcul de cet indice est surtout utile avec les fréquences basses, inférieures à 800Hz [Ekeberg Henriksen, 2002]. En effet, au-dessus de 1600Hz, la tête de l'auditeur devient une « barrière acoustique » pour les fréquences de courte longueur d'onde [Kendall, 1995] [Blauert, 1997].
- l'indice interaural de phase IPD (Interaural Phase Difference) est en fait un ITD. Il s'agit du calcul des « différences temporelles de la structure fine » [Périaux & al., 2015]. L'IPD est pertinente principalement pour les fréquences basses (inférieures à 1000Hz) et surtout pour des signaux sinusoïdaux [Moore, 1989] [Rumsey, 2001] .
- l'indice interaural d'intensité ILD (Interaural Level Difference) ou IID (Interaural Intensity Difference) : il s'agit de la différence de niveau qui traduit la diffraction et l'absorption que subit une onde sonore, dans les hautes fréquences (supérieures à 1000Hz) [Dean, 2014]. En effet, « aux basses fréquences, l'ILD est très faible car la tête diffracte peu l'onde incidente. » [Périaux & al., 2015]. Rumsey [2001] précise que le calcul de l'ILD est effectif pour des sources proches de l'auditeur : lorsque les sources sonores sont loin, la distance parcourue par chaque signal est tellement grande que la différence d'amplitude reçue par chaque oreille est négligeable et ne peut donc aider à la localisation.

¹¹ « La perception de l'espace est majoritairement dépendante du fait que l'auditeur a deux oreilles, bien que certaines propriétés monaurales ont été démontrées – en d'autres termes, c'est principalement les différences entre les signaux reçus par les deux oreilles qui sont importantes à considérer. »



*Figure 2 : Mécanismes de la localisation sonore chez les êtres humains
(source : [Grothe & al., 2010])¹²*

En résumé, les sons qui n'arrivent pas directement de l'avant ou de l'arrière de la tête parviennent toujours plus tôt à l'une des deux oreilles ; c'est cela qui crée ce que l'on appelle une ITD, et ceci est valable principalement pour les fréquences basses. D'autre part, il y a un effet de masquage qui se produit envers l'oreille la plus éloignée de la source, pour les longueurs d'onde inférieures ou égales au diamètre de la tête (fréquences hautes), ce qui a pour conséquence de créer une différence de niveau (ILD) entre les deux oreilles de l'auditeur [Grothe & al., 2010]. Il faut noter que « l'ITD et l'ILD sont des critères dont la valeur est constante pour des positions de la source et de la tête fixes. [...] pour un même angle d'incidence, la différence de marche est exactement la même que le son vienne de l'avant ou de l'arrière. » [Périaux & al., 2015]. Ce sont alors de petits mouvements de tête, entre autres, qui vont aider l'auditeur à résoudre ces indéterminations. Nous allons en parler plus longuement dans ce qui suit.

I.1.2 La localisation d'une source sonore (suite) : les indices spectraux, le corps de l'auditeur

Nous l'avons évoqué précédemment, du fait de sa taille importante, la tête de l'auditeur

¹² Une erreur s'est glissée dans la Figure 2 : en C, il faut lire « $f > 2\text{kHz}$ ».

agit comme une barrière acoustique pour les sons de hautes fréquences. D'autre part, des réflexions et résonances causées par la forme du pavillon de l'oreille externe, mais aussi par les épaules et l'ensemble du corps en général, modifient le spectre de la source sonore. Blauert [1997] a mis en évidence que, pour un son constitué d'un spectre globalement uniforme, des régions de ce spectre sont soit renforcées, soit atténuées, lors de positions particulières de la source sur le plan horizontal ou médian. « The sum of all these effects is a unique head-related transfer function or HRTF for every source position and angle of incidence, including different elevations and front-back positions. »¹³ [Rumsey, 2001] La fonction HRTF relève d'une association complexe de pics de fréquences et de creux qui varient systématiquement avec la direction du son par rapport à la tête et qui est unique pour chaque position dans l'espace. Il s'agit d'une fonction « à trois variables : la fréquence ou le temps, la localisation de la source sonore en coordonnées sphériques par exemple (azimut, angle de radiation et d'élévation) et la morphologie de l'individu. » [Périaux & al., 2015]. Il paraît alors difficile pour l'auditeur de séparer les caractéristiques spectrales propres aux sources de celles ajoutées par les HRTFs. Pourtant, il semble y arriver en comparant les HRTFs monauraux (dépendant du signal reçu par chaque oreille) à des patterns déjà connus et enregistrés par le cerveau (cf. partie I.2.2). Cela fonctionne d'autant plus avec des sources en mouvement ; en effet, si le spectre d'une même source semble se modifier au cours du temps, c'est que cela est dû à son changement de localisation et donc aux HRTFs [Rumsey, 2001].

Si la tête de l'auditeur reste immobile, l'ITD n'est pas suffisante pour définir la position de la source dans l'espace. Nous avons ouvert cette discussion en fin de partie I.1.1 en donnant l'exemple d'un son positionné à l'avant ou à l'arrière de la tête. De façon plus générale, on parle de « cône de confusion » pour désigner l'ensemble des positions possibles d'une source sonore dans l'espace, qui rendront le même indice d'ITD (voir Figure3, ci-après) [Moore, 1989]. Ce sont alors les mouvements de tête de l'auditeur qui l'aideront à localiser le son – à condition que celui-ci soit de longue durée ou bien répété dans le temps. Les indices spectraux reçus par chacune des deux oreilles vont être modifiés et vont ainsi venir compléter les informations nécessaires pour déterminer la provenance du son. Pour toutes ces raisons, on peut noter qu'être positionné face au son facilite sa localisation. De plus, si la rotation de la tête n'est accompagnée d'aucun changement dans la position de l'image sonore, on peut déduire que le son est localisé soit au-dessus, soit au-dessous de l'auditeur.

13 « La somme de tous ces effets est une unique fonction de transfert relative à la tête, ou HRTF, pour chaque position de la source et angle d'incidence, incluant l'élévation et les positions avant-arrières. »

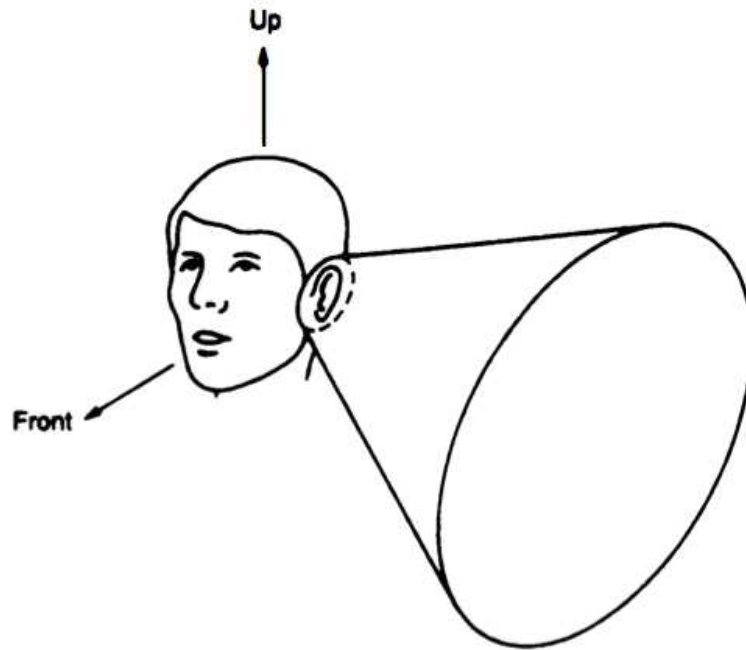


Figure3 : « Un cône de confusion pour une tête sphérique et un ITD donné. Toutes les sources sonores à la surface du cône vont produire ce même ITD. » [Moore, 1989] (source : [Moore, 1989])

Les fonctions HRTFs dépendent de la morphologie de l'individu, les mouvements de la tête sont essentiels dans le processus de localisation d'une source sonore ; nous pouvons également remarquer que le pavillon de l'oreille joue non seulement un rôle très important dans la détermination des fonctions HRTFs mais aussi pour un autre phénomène que l'on appelle la « résonance concha » [Moore, 1989] [Rumsey, 2001]. Il s'agit d'une résonance créée par la cavité principale du centre du pavillon externe ; elle agirait dans le sentiment d'externalisation du son. En d'autres termes, sans cette morphologie particulière, on entendrait tout à l'intérieur de notre tête. Il ne faut bien évidemment pas négliger l'influence de la réverbération des sons dans l'espace.

D'une manière générale, c'est tout le corps de l'auditeur qui entre en jeu dans la perception de l'espace sonore, en tant qu' "objet" réflecteur d'ondes sonores, mais aussi en tant que système nerveux. Celle-ci n'est pas seulement auditive, il s'agit en fait d'une expérience vécue par le corps tout entier (« a full body experience » [Smalley, 1991] [Schmele, 2011]). L'auditeur se situe dans l'espace ; son corps constitue le point de référence autour duquel il va déterminer la position de chaque source sonore, par le biais de ses deux oreilles, mais aussi de ses mouvements de tête, s'il se déplace, etc. Ce sont les différents axes du corps qui vont donner un contexte aux événements spatiaux qui apparaissent dans la vie quotidienne

[Ekeberg Henriksen, 2002] [Kendall, 2010].

I.1.3 En résumé : la localisation et la distance

Nous venons d'apprendre que, pour localiser un son, nous nous basons sur les indices de temps (ITD, IPD) pour les basses fréquences (environ inférieures à 2000Hz), d'intensité (ILD ou IID) pour les hautes fréquences (environ supérieures à 2000Hz) et sur les indices spectraux reçus par notre appareil auditif. L'ITD et l'ILD sont plutôt utiles pour la localisation droite–gauche, tandis que les fonctions HRTFs nous permettent de déterminer si le son provient de l'avant, de l'arrière, du dessus ou du dessous [Ekeberg Henriksen, 2002]. « D'autre part, il semblerait que la localisation latérale est génétique, tandis que l'élévation est un mécanisme qui s'acquiert par la vision (calibration par la vision) » [Périaux & al., 2015]. Nous reviendrons sur ce dernier point en partie I.2.1. Il est important également de considérer l'effet de précedence (ou « effet de Haas » ou encore « loi du premier front d'onde » [Blauert, 1997]) qui est une conséquence directe de l'interprétation des ITD et ILD : si un même son est diffusé, au même moment, en différents points de l'espace, c'est en direction du signal sonore qui atteindra en premier l'une des deux oreilles (et donc le plus proche) que l'on attribuera l'emplacement de la source¹⁴ (cf. partie I.3.2).

Nous répertorions ci-dessous ce qu'il est important de retenir en ce qui concerne la localisation d'un son sur le plan horizontal :

- les mouvements de la tête permettent à l'auditeur de résoudre certains conflits, notamment dus au « cône de confusion » ;
- la vision joue un rôle non négligeable que nous tâcherons de développer en partie I.2.1 ;
- les sources sonores mobiles apportent plus d'informations que les sources fixes. L'auditeur peut ainsi mesurer les changements entre les différentes informations perçues et résoudre certaines incertitudes ;

14 « la source perçue en premier est identifiée comme l'unique provenance. » [Périaux & al., 2015].

- l'effet de précedence va favoriser la détermination de la localisation ; c'est grâce à lui que les réflexions du son dans l'espace ne viennent pas brouiller la perception de l'emplacement de la source : « The precedence effect is of fundamental importance for acoustics. It is because of this that the auditory event is perceived in the direction of the original sound source and the reflected sound does not disturb this orientation. This does not mean, however, that the reflected sound would be completely masked. »^{15 16} (voir partie I.3.2).

Nous pouvons ajouter que la précision avec laquelle nous percevons l'espace n'est pas la même selon l'endroit d'où provient la source sonore. Blauert [1997] a nommé « localisation blur »¹⁷ cette marge d'erreur, exprimée en degrés. Elle est évaluée à 1° à l'avant de l'auditeur [Blauert, 1997]. Il s'agit de la perception spatiale la plus accrue, à la fois dans le plan horizontal et dans le plan vertical (4° de marge d'erreur). Sur les côtés, plus la source sonore provient des extrémités plus la valeur du « flou de localisation » augmente jusqu'à atteindre 10°. À l'arrière, elle s'élève à 6°, la plus grande indétermination étant au-dessus de la tête avec une marge d'erreur de 22° [Kendall, 1995] [Blauert, 1997]. Bien entendu, ce « flou de localisation » varie en fonction des caractéristiques temporelles et spectrales du son, mais aussi en fonction de l'expérience acquise par l'auditeur au cours de sa vie (cf. partie I.2.2). Par exemple, la localisation des sons de basses fréquences sera moins précise que celle des sons dont l'énergie spectrale est plutôt contenue dans les médiums et hautes fréquences [Ekeberg Henriksen, 2002].

Arrêtons-nous quelques instants sur le cas particulier de la localisation verticale, puisque l'élévation intéresse tôt ou tard tout musicien souhaitant travailler avec le son en trois dimensions. On sait déjà que notre perception auditive de l'espace est plus sensible sur l'axe horizontal que sur l'axe vertical, et plus robuste lorsque nous sommes dans une situation présentant un important bruit de fond [Zwiers & al., 2001]. La localisation verticale des sons dépend majoritairement de variations subtiles des indices spectraux, qui elles-mêmes découlent de notre morphologie (oreille externe et corps) [Payri, 2011]. Par exemple, pour un son donné, on observe un filtrage *notch* (bande de fréquences très étroite) de son spectre du

15 Jens BLAUERT, *Hearing Of Music In Three Spatial Dimensions*, Mis-en-ligne le 30 Mars 2015 sur https://www.researchgate.net/profile/Jens_Blauert [consulté le 12/10/2017]

16 « L'effet de précedence est d'une importance fondamentale en acoustique. C'est grâce à cela que l'évènement audio est perçu dans la direction de la source sonore d'origine et que le son réfléchi ne perturbe pas cette orientation. En revanche, cela ne veut pas dire que le son réfléchi sera complètement masqué. »

17 « flou de localisation »

grave à l'aigu, selon sa localisation sur l'axe vertical, du bas vers le haut (voir la Figure2 p.12) [Grothe & al., 2010]. On peut déjà en déduire que chaque individu ne perçoit pas les sons avec le même degré de précision sur le plan vertical. De même, un sinus qui présente un spectre d'étendue extrêmement réduite sera parfaitement localisé en terme d'azimut, mais absolument pas en terme d'élévation [Middlebrooks, 1992] [Zwiers & al., 2001]. En effet, les neuroscientifiques Dobрева, O'Neill et Paige [2011] ayant réalisé des expérimentations afin de mettre en évidence l'influence de l'âge du sujet sur la perception verticale des sons, ont appuyé les résultats de recherches antérieures¹⁸ en confirmant que le spectre sonore doit être le plus étendu possible au-dessus de 3000Hz pour que la localisation d'une source sur l'axe vertical soit la plus précise possible. Ainsi, les personnes plus âgées témoignant d'une baisse de l'audition dans les hautes fréquences (supérieures à 10kHz) localisent moins bien les sources élevées que les personnes plus jeunes. S'il est alors essentiel que les sources sonores soient constituées de fréquences hautes pour être correctement localisées verticalement, le contraire n'est pas vrai : un spectre sonore uniquement compris entre 10kHz et 20kHz ne sera pas suffisant [Dobрева & al., 2011].

La perception de la distance peut être absolue ou relative. Elle fonctionne beaucoup mieux lorsque nous percevons différentes sources en même temps et que nous pouvons les comparer. Voici quelques remarques notables quant à son "fonctionnement" :

- L'auditeur peut se baser sur la différence de volume perçu : une source éloignée sera de niveau plus faible qu'une source plus proche. Ceci est d'autant plus efficace avec des sons familiers ; sinon, la perception des niveaux est exagérée. Pour des sons très proches de la tête, l'ILD est également utile pour déterminer la distance de chacun [Moore, 1989] [Périaux & al., 2015]. On peut noter que, dans le cas d'une source mobile qui s'éloigne de plus en plus, la valeur de l'ILD se rapproche de 0 tandis que la distance augmente.
- Avec le chemin parcouru pour atteindre l'oreille externe, on constate une modification du spectre sonore en fonction de la distance, due à l'absorption de l'air, du vent, au taux d'humidité ou encore à cause de divers obstacles rencontrés. Ce sont les

¹⁸ Voir : Carlile S, Leong P, Hyams S., 1997, *The nature and distribution of errors in sound localization by human listeners*. Hear Res 114: 179–196 ; Carlile S, Delaney S, Corderoy A., 1999, *The localisation of spectrally restricted sounds by human listeners*, Hear Res 128: 175–189 et King, R.B., Oldfield, S.R., 1997, *The impact of signal bandwidth on auditory localization: implications for the design of three-dimensional audio displays*. Hum Factors 39: 287–295.

fréquences aiguës qui subissent la plus forte atténuation [Rumsey, 2001].

- L'auditeur peut également se référer au rapport *champ direct* sur *champ réverbéré*, en particulier dans le cas d'un son non familier. En effet, un son lointain est plus réverbérant qu'un son proche. Surtout, on observe moins de différence de temps entre la source directe et l'arrivée des premières réflexions. Aussi, le spectre du son réfléchi est généralement différent du spectre du son direct et cela peut également constituer un critère pour évaluer la distance [Zahorik, 2002].

« The multiplicity of cues to sound location provides a certain redundancy in the system, so that even under very difficult conditions (reverberant rooms or brief sounds) we are still capable of quite accurate localization. »¹⁹ [Moore, 1989] Et comme nous avons pu déjà l'évoquer, la vue mais aussi l'expérience acquise par chaque individu au cours de sa vie sont autant de facteurs qui contribuent à notre perception de l'espace.

I.2. Les facteurs visuels, mémoriels et sociologiques liés à la perception de l'espace

Outre le fait de connaître quelques bases de psycho-acoustique, il nous paraît important, lorsque l'on s'intéresse à la musique spatialisée et notamment à la situation de concert, de comprendre quel est le rôle de la vue dans la perception de l'espace. Plus généralement : quels sont les autres facteurs qui entrent en jeu dans la perception d'un espace sonore ? L'auditeur peut-il apprécier ou du moins accepter un espace sonore, sans image ? Peut-il puiser dans sa mémoire pour l'aider à localiser ? Peut-il se créer ses propres images ? Nous nous demandons également si un auditeur non expérimenté, à force d'expériences au cours de sa vie, mais aussi en les provoquant en assistant à des concerts de musique spatialisée par exemple, a la capacité ou non d'affiner sa perception de l'espace par le son.

¹⁹ « La multiplicité des repères à la localisation des sons fournit une abondance d'informations redondantes, si bien que même sous des conditions difficiles (espaces réverbérants ou sons courts) nous sommes toujours capable d'une bonne précision dans la localisation. »

I.2.1 Le rôle de la vision dans la perception de l'espace : points de vues scientifiques et sociologiques

« Events with visible sources are more plausible than ones where nothing can be seen. »²⁰
[Kendall, 2010]

Les indices sonores spatiaux que nous venons d'expliciter en partie I.1 et, plus généralement, l'ensemble des caractéristiques sonores nous aidant à la perception de l'espace, sont souvent dépendants du contexte dans lequel nous sommes et peuvent être influencés par des informations apportées par d'autres sens, notamment la vision. En effet, l'orientation visuelle agit activement sur notre perception de l'espace, et « la plus forte représentation spatiale implique l'intégration des informations provenant des différents sens » [Périaux & al., 2015].

Sans la vision, de nombreuses études ont montré que la perception de la distance dans un environnement non réverbérant ne peut pas se faire uniquement avec la perception auditive – à moins, bien entendu, que l'on ait un point de référence d'une distance connue auquel on peut comparer différentes sources sonores [Rumsey, 2001]. Ceci n'est qu'un exemple ; plus généralement, des scientifiques ont observé que « la vision et l'audition traitent l'information spatiale dans la même région du cerveau, et agissent ensemble dans l'élaboration de cartographies spatiales (Auerbach et Sperling 1974²¹; Hollander 1994²²) »²³ [Ekeberg Henriksen, 2002] L'auteur cite également Warren²⁴ qui a découvert que la localisation de sources sonores invisibles est beaucoup plus effective avec les yeux ouverts qu'avec les yeux fermés. De même, on localiserait de façon plus précise en pleine lumière que dans le noir. Les neuroscientifiques Zwiers, Opstal et Cruysberg [2001] qui ont réalisé des tests sur des personnes aveugles depuis la naissance (ou depuis la petite enfance), ont d'ailleurs démontré que nous avons besoin de la vue pour percevoir l'espace de façon optimale. Les non-voyants de naissance sont en effet désavantagés lorsqu'il s'agit de localiser un son se trouvant sur l'axe vertical (angle d'incidence situé entre $- 50^\circ$ et $+ 50^\circ$) dans la zone frontale uniquement, correspondant à la surface recouverte par la vue. En revanche, ceci est surtout avéré lorsqu'ils

20 « Les événements constitués de sources visuelles sont plus convaincants que ceux où il n'y a rien à voir. »

21 L'auteur fait référence à : Auerbach, C., et Sperling, P., 1974, *A Common Auditory-Visual Space : Evidence for its Reality*, In « Perception and Psychophysics ».

22 L'auteur fait référence à : Hollander, A., 1994, *An Exploration of Auditory Shape Perception*, Master's thesis. The University of Washington, Seattle.

23 Texte d'origine : « vision and audition process spatial information in the same area of the brain, and support each other in the process of spatial mapping (Auerbach and Sperling 1974 ; Hollander 1994). »

24 Warren, D. H., 1970, *Intermodality Interactions in Spatial Localisation*, In « Cognitive Psychology ».

sont confrontés à une situation où le rapport signal/bruit est élevé. De plus, ils développent une « hyper-compensation » auditive au niveau des différences d'intensité (ILD / IID) qui leur permet de définir l'azimut de façon plus précise qu'une personne voyante.

Mettons à présent en évidence l'influence de la vision dans la perception de l'espace d'un point de vue plus sociologique. Contrairement aux sons que l'on peut percevoir à 360°, le champ de la vision est limité, au mieux, à l'hémisphère frontal. C'est donc uniquement par le biais de notre audition que l'on perçoit l'espace à l'arrière de la tête. Dans un cas d'écoute au casque en binaural par exemple, la plupart des auditeurs ont tendance à placer les sons à l'arrière de la tête ; ils ont du mal à percevoir une scène frontale puisqu'il y a absence de stimuli visuel [Rumsey, 2001]. À l'inverse, si l'on visualise sur un écran, face à nous, la source d'un son qui est en fait diffusé derrière nous, on imagine que le son vient de l'avant et non de l'arrière. La vision accapare donc l'attention sur l'audition et va jusqu'à nous induire en erreur. Un dernier exemple de l'influence de la vue sur notre façon de percevoir l'espace est à nouveau mis en évidence dans une situation présentant uniquement des caractéristiques sonores : lorsque l'on diffuse le son d'un avion à un groupe de personnes, avec un angle d'incidence compris entre 0 et -50° (avec des haut-parleurs qui seraient situés au sol ou bien au casque en binaural), la plupart des gens vont tout de même lever la tête car ils ont l'habitude de l'entendre voler au-dessus d'eux, bien que les indices spectraux ne correspondent pas à ce cas de figure ! Il s'agit ici d'un son familier et cela nous renvoie à l'importance de la vue d'une part, dans la perception de l'espace, mais aussi aux connaissances emmagasinées peu à peu avec les expériences quotidiennes, qui développent en nous des réflexes sur lesquels on se repose par la suite.

1.2.2 Le rôle de la mémoire dans la perception de l'espace

« [...] certain time and frequency relationships between signal components [...] tend to lead the brain to link them to form a common object. Where that object has been previously recognized and labelled, memory enables one to identify it. »²⁵ [Rumsey, 2001]

D'après [McAdams & Bigand, 1994] et [Rumsey, 2001], le cerveau humain analyse

25 « [...] certaines relations temps-fréquence entre les composants d'un signal [...] amènent le cerveau à les lier afin de former un objet commun. Si cet objet a été précédemment reconnu et nommé, la mémoire nous permet de l'identifier. »

les scènes sonores en formant des objets : il classifie les sons en fonction de leur morphologie et de leurs propriétés spectrales²⁶. L'auditeur peut alors identifier à nouveau ces objets grâce à ce qu'il a déjà vécu et enregistré mentalement, il puise dans son expérience. La notion de sons familiers est alors très importante quand il s'agit de percevoir une distance. En effet, celle-ci ne se fait pas avec le même degré de précision selon si le son nous est déjà connu ou non. Tout d'abord, lorsque la source sonore nous est familière, ce n'est plus tant le niveau sonore de cette dernière qui importe pour évaluer sa distance, mais son spectre car nous comprenons rapidement s'il a subi ou non des modifications, et en quoi cela nous indique s'il est loin ou proche [Ekeberg Henriksen, 2002]. Au contraire, pour un son qui nous est tout à fait inconnu, la marge d'erreur dans l'évaluation de sa distance peut s'élever jusqu'à 20% [Moore, 1989]. Zahorik [2002] explique que lorsque ce son est proche de l'auditeur, sa distance est sur-évaluée, et inversement s'il est loin, sa distance est sous-évaluée. De même, la localisation d'un son est bien meilleure lorsque celui-ci nous est familier : toute personne est capable de localiser une voix humaine, par exemple, de façon précise. Ceci est d'autant plus vrai si cette voix appartient à quelqu'un que l'on connaît [Blauert, 1997].

La mémoire joue donc inévitablement un rôle dans la perception de l'espace. Le neuropsychologue Richard Gregory reconnaît d'ailleurs que la perception en général s'appuie à 90% sur la mémoire, contre seulement 10% sur les sensations (l'ensemble des stimuli externes, provenant de l'environnement, et des stimuli internes, produits par le cerveau) [Gregory, 1998]. Si l'on considère un concert de musique électronique et notamment de musique dite "acousmatique", où l'audience est souvent plongée dans le noir, à l'abri de stimuli visuels extérieurs, ce que va percevoir l'auditeur en terme d'espace ne va pas dépendre uniquement de l'acoustique du lieu et des caractéristiques techniques du système de diffusion, mais aussi et surtout de ses représentations mentales [Kendall, 2010]. Schmele, dans son mémoire [2011], appuie ces propos : « When an audience member listens to sounds in an acousmatic context, he usually imagines what these sounds could relate to in the real world, perhaps having a visual or even a haptic perception. »²⁷. La perception de l'espace est donc très personnelle et dépend des expériences de chacun, et notamment de l'environnement dans lequel on évolue, de notre culture [Ekeberg Henriksen, 2002] [Blessner & Salter, 2009]. Des personnes confrontées à un même espace sonore n'auront pas la même écoute selon si elles vivent habituellement en ville ou à la campagne, en Alaska ou à New-York, etc. Enfin, nous

26 Voir aussi : Kendall, Gary S., 1991, *Visualization by Ear: Auditory Imagery for Scientific Visualization and Virtual Reality*, In « Computer Music Journal », 15:4. Cambridge, Massachusetts: MIT Press.

27 « Quand un membre de l'audience écoute des sons dans un contexte acousmatique, il imagine généralement à quoi ces sons peuvent correspondre dans le monde réel, en développant peut-être une perception visuelle ou même tactile. »

pourrions oser affirmer que la perception de l'espace est entièrement subjective : à la fois le calcul des indices sonores spatiaux définis en partie I.1 et la recherche de références gardées en mémoire sont influencés par nos sentiments, nos émotions, notre état d'esprit sur le moment. Et réciproquement, la perception de l'espace peut jouer sur nos émotions. C'est tout à fait ce sur quoi repose la notion de « Aural Architect » (que l'on pourrait traduire par « architecte auditif » ou encore « architecte du sonore ») exposée par Blesser dans son livre *Spaces speak... are you listening ?* [2009] qu'il définit comme un créateur d'espaces sonores qui ne se base pas sur des caractéristiques physiques tel que le ferait un « architecte acoustique » (un acousticien). Un « Aural Architect » peut consciemment créer et jouer sur les émotions, influencer les sociétés... et inversement, le contexte social peut affecter directement tout « architecte du sonore » : « Those who built cathedrals and those who designed virtual electroacoustic spaces were seldom aware of how their social context influenced their spatial creations. »²⁸ [Blesser & Salter, 2009]. On peut mettre ici en relation l'expérience réalisée et décrite par Rumsey [2001], ayant pour but de déterminer quelles propriétés spatiales sont les plus aptes à provoquer des émotions positives chez l'auditeur. Il résume : « In this experiment, it seems that high level of envelopment and room impression created by surround sound, rather than accurate imaging of sources, were the descriptive features most closely related to positive emotional responses. »²⁹ La perception de certaines propriétés spatiales liées à la notion d'immersion serait alors capable de susciter en nous des émotions positives. La façon avec laquelle on expérimente l'espace sonore est bien en lien avec notre propre vécu, et donc notre mémoire.

I.2.3 Le travail de « l'oreille spatiale »

« I argue that most people learn from environmental sound around them [...] »³⁰ [Dean, 2014]

La perception n'est pas passive, elle n'est pas seulement une réponse quasi automatique à des stimuli externes, visuels et auditifs ; elle est également active car, pour être optimale, elle demande à l'auditeur d'être en mouvement, de bouger la tête, mais aussi de puiser dans sa

28 « Ceux qui ont construit des cathédrales et ceux qui ont créé des espaces électroacoustiques virtuels étaient rarement conscients de l'influence de leur contexte social sur leurs créations spatiales.. »

29 « Dans cette expérience, il semble qu'un fort degré d'enveloppement et d'effet de salle créés par un son à caractère "immersif", plutôt que par des sources présentant une image sonore très précise, sont les critères les plus aptes à susciter des réponses émotionnelles positives. »

30 « Je soutiens que la plupart des gens apprennent à partir de l'environnement sonore qui les entoure »

mémoire, d'acquérir de l'expérience afin de pouvoir s'en servir par la suite. Il semble donc logique de pouvoir affiner notre perception de l'espace de façon consciente, en travaillant ce que nous appelons notre « oreille spatiale », qui correspond à tout ce qui permet à l'auditeur de percevoir l'espace par le biais du son.

Nous avons vu en partie I.1.2 que les filtres HRTFs dépendent de notre morphologie, et notamment de celle de nos pavillons d'oreilles externes qui sont propres à chacun de nous, autant que peut l'être une empreinte digitale. Lorsque l'on attribue un HRTF différent du sien à une personne donnée – via l'écoute binaurale par exemple –, celle-ci a du mal, au début, à s'orienter ; elle est comme déséquilibrée. Puis, petit à petit, elle parvient à s'adapter jusqu'à percevoir l'espace avec la même acuité que si elle possédait ses propres HRTFs [Rumsey, 2001]. L'être humain peut donc s'adapter afin de percevoir au mieux l'espace qui l'entoure : il est par conséquent capable de *travailler* son « oreille spatiale ». On peut également citer l'expérience rapportée chez [Grothe & al., 2010] qui consiste à insérer une sorte de prothèse dans chaque oreille de façon à ce que « le passage du son dans le tympan soit permise, mais que la forme du pavillon soit altérée ; ainsi la performance des auditeurs est dégradée de sorte qu'ils ont du mal à distinguer la localisation des sons sur le plan vertical (bien que la discrimination gauche-droite ne soit pas altérée) »³¹. En effet, la perception gauche-droite va se reposer, dans ce cas, uniquement sur le calcul des ITD et ILD. L'expérience s'étend sur plusieurs semaines et les résultats montrent que l'erreur de localisation verticale est réparée au fil du temps : il y a donc bien un apprentissage qui se fait. Et lorsque ces prothèses sont retirées, il ne reste aucune séquelle : les auditeurs comprennent aussitôt l'espace de la même façon qu'auparavant. « La sensibilité aux indices spectraux est [donc] représentative d'un comportement d'apprentissage »³² [Grothe & al., 2010]. Ceci démontre bien l'importance de la mémoire dans le processus de perception de l'espace.

Comme énoncé en partie I.2.1, la vision nous aide à développer notre perception de l'espace dans la zone frontale. Le fait d'avoir un stimulus auditif accompagné d'un stimulus visuel permet de garder en mémoire une information de façon plus précise que si nous n'avions eu qu'un stimulus audio. Ainsi, les non-voyants de naissance localisent avec moins d'acuité que les personnes voyantes les sources sonores frontales. En revanche, ils développent une « hyper-compensation » pour le reste, ce qui montre que même sans la vue,

31 Texte original : « [...] that permit the passage of sound to the tympanic membrane, but alter the shape of the concha, human listeners' performance is degraded to the extent that they have trouble distinguishing the location of sounds in the vertical plane (although left-right discriminations appear unaltered). »

32 Texte original : « sensitivity to spectral cues represents a learned behavior. »

ils apprennent énormément au fur et à mesure de leurs expériences. Au contraire, c'est certainement le fait que les voyants se reposent davantage sur la vision que sur l'audition que leur perception spatiale hors du champ visuel est déficiente par rapport à celle des personnes aveugles. Nous pouvons alors supposer que développer notre écoute de l'espace, sans aucun stimulus visuel, par le biais de la musique électronique spatialisée, peut en effet aider à affiner notre perception spatiale globale. À ce propos, Rumsey [2001] fait référence à Komiyama³³ qui apporte des arguments démontrant qu'il existe bien une différence de précision dans la perception auditive de l'espace entre des sujets experts et non experts (les sujets experts étant des personnes ayant l'habitude de travailler avec ou d'écouter de la musique spatialisée). L'expérience présentée consiste à déterminer à partir de quelle différence d'azimuts entre le stimulus visuel et la source sonore correspondante les sujets interrogés détectent une distorsion de directionnalité, c'est-à-dire que l'image et le son ne semblent pas provenir du même endroit. Les novices seront perturbés par une distorsion de directionnalité à partir de 20° d'écart entre l'image et le son, contre seulement 11° pour les experts. Nous avons donc une preuve que nous pouvons rendre plus sensible notre perception auditive de l'espace à force d'écoutes multiples, comme un entraînement. En revanche, cela ne signifie absolument pas qu'il faut être musicien : en effet, la formation musicale ne favorise aucunement la perception auditive de l'espace. Payri [2011] a réalisé une expérience sur des sujets musiciens et non musiciens afin d'évaluer leur capacité à reconnaître des patterns spatiaux, en employant divers sons (des sons simples, ou présentant des variations temporelles d'intensité, une distribution spectrale différente, etc.). L'écoute se déroulait à la fois via des haut-parleurs mais aussi au casque. Il en a conclu que les musiciens ne sont pas meilleurs que les non-musiciens dans la reconnaissance de patterns spatiaux. Ces derniers ont acquis une expérience grâce à « l'audition active » qu'ils ont développée au fil du temps, n'ayant jamais suivi de « formation musicale ». Ils sont alors plus réceptifs et présentent moins *d'a priori* dus à l'éducation.

Ainsi l'écoute « active » et décomplexée de musique électronique spatialisée, en plus de l'expérience acquise par chacun en évoluant dans l'environnement, pourrait bien, au fur et à mesure, participer au développement de « l'oreille spatiale ». Il existe diverses situations d'écoute possibles de musiques électroniques spatialisées auxquelles chacun peut être confronté, de façon consciente ou à son insu, et qui mériteraient de se généraliser : l'écoute domestique en stéréophonie ou 5.1, au casque en binaural grâce, notamment, à l'émergence de sites-web ou podcasts permettant d'écouter en binaural³⁴, les concerts ou séances d'écoute, les

33 Komiyama, S., 1989, *Subjective evaluation of angular displacement between picture and sound directions for HDTV systems*, J.Audio Eng. Soc., 37, 4, pp.210-214.

34 L'Hyperradio par Radio France : <http://hyperradio.radiofrance.fr> et BBC Radio 3 :

installations sonores, les salles de cinéma, ...³⁵ Ces situations d'écoute sont de plus en plus nombreuses mais sont encore vécues comme des expériences singulières pour les novices. Pourtant, elles pourraient permettre à tout un chacun de s'habituer et de s'entraîner à l'écoute de l'espace, en faisant même partie du quotidien pour certains. Quelques actions sont lancées pour sensibiliser les gens, dès le plus jeune âge. Le GRM (Groupe de Recherches Musicales) a proposé par exemple un atelier d'éveil le 8 Février 2018 à Radio France afin de sensibiliser les enfants à l'écoute de la musique électroacoustique spatialisée.

I.3. Conclusions

Dans cette partie I, nous avons tenté de poser les bases du fonctionnement de la perception auditive spatiale, impliquant à la fois l'appareil auditif, le corps de l'auditeur, la vision, la mémoire, l'apprentissage par l'expérience, ... afin de donner au compositeur de musique électronique des pistes pour travailler la spatialisation du son. Il nous semble important de chercher à comprendre comment l'auditeur perçoit l'espace sonore puisque c'est à lui que s'adresse la musique électronique spatialisée ainsi que la situation de concert. Une fois ces "rudiments" de psycho-acoustique rassemblés, nous pouvons en tirer certaines conclusions quant au travail de l'espace dans la musique, mais aussi en ce qui concerne le concert spatialisé et les dispositifs de diffusion. Ces observations nous serviront de point de référence dans la suite du présent mémoire (parties II et III).

I.3.1 Influence des connaissances en psycho-acoustique sur les techniques de spatialisation

La perception auditive de l'espace dépend, dans un premier temps, d'une combinaison d'indices de temps (et phase), d'intensité et spectraux. Il semble donc essentiel, pour le compositeur de musique électronique spatialisée, de connaître quelques bases de psycho-

<https://www.bbc.co.uk/radio3>

³⁵ Le lecteur intéressé peut se référer à la partie I de [Nilles, 2016].

acoustique pour comprendre comment choisir un son en fonction de ce qu'il souhaite donner à entendre, ou comment modifier le spectre de ce son à son avantage. Ainsi, pour simuler l'éloignement ou le rapprochement d'une source sonore par exemple, il est important de savoir comment fonctionne la perception auditive de la distance. Nous avons vu, en partie I.1.3, que les fréquences hautes sont plus vite atténuées que les fréquences basses avec la distance, dû aux frottements et à l'absorption de l'air [Ekeberg Henriksen, 2002]. Ceci donne des pistes au compositeur qui souhaite faire entendre une source sonore au loin, ou encore qui s'éloigne ou se rapproche. À nouveau, il faut considérer l'importance du contenu spectral si l'on veut que l'auditeur puisse bien localiser un son. En effet, ce sont surtout les hautes fréquences, dues à leur taille de longueur d'onde, qui entrent en interaction avec le pavillon de l'oreille externe et qui permettent de générer les indices spectraux indispensables à la localisation [Rumsey, 2001]. Ainsi, le spectre de ce son doit absolument posséder des fréquences au-dessus de 6kHz [Moore, 1989].

L'élévation (et la dé-élévation) est souvent utilisée en musique électronique spatialisée, notamment pour des installations sonores, lorsque l'on a la possibilité de disposer d'un système 3D. Pour percevoir un son en hauteur, nous l'avons vu précédemment (partie I.1.3), il est indispensable que le spectre de ce son soit très étendu et riche en hautes fréquences au-dessus de 3kHz. En revanche, pour percevoir un son comme provenant du bas, il est préférable que le spectre du son soit étroit et concentré dans les fréquences graves, en-dessous de 2kHz [Kendall, 2010]. Dans les deux cas, la localisation verticale étant moins précise pour l'oreille humaine, on entend rarement le son provenir exactement du haut-parleur du haut ni de celui du bas ; on perçoit plutôt un entre-deux (entre le haut-parleur et nous-même). Kendall [2010] ajoute que, si le son présente un spectre très étroit (bande) et constitué de fréquences très hautes, on pourrait percevoir le son comme provenant du plafond. Dans tous les cas, un son situé sur l'axe vertical est généralement perçu comme proche de l'auditeur.

À la fois les facteurs culturels et le contenu fréquentiel des sons influencent notre perception de l'espace [Ekeberg Henriksen, 2002]. Les sons correspondant à une note grave ou aiguë sont souvent perçus – tout du moins imaginés – réciproquement comme provenant du bas ou du haut, à cause de la notation occidentale sur partition qui représente les fréquences basses en-dessous des fréquences hautes. Surtout, les HRTFs et leur dépendance au spectre d'origine du son pour communiquer des informations précises de localisation mettent en évidence que des variations dans le spectre sonore induiront des changements de localisation. Ainsi, si l'on modifie le spectre d'une source sonore immobile, celle-ci peut sembler s'élever ou au contraire s'abaisser, provenir de l'avant ou bien de l'arrière [Ekeberg

Henriksen, 2002]. Pour cela, il faut ajouter des composantes fréquentielles à certains endroits du spectre (créer des pics de fréquences). Si l'énergie du spectre sonore est contenue (ou centrée) :

- entre 250 et 600Hz, entre 3000 et 6000Hz, ou encore autour de 16kHz, le son semble être originaire de l'avant ;
- autour de 1000Hz ou bien de 12kHz, le son semble être localisé à l'arrière ;
- autour de 8000Hz, le son semble provenir du dessus de la tête [Blauert, 1997] [Rumsey, 2001] [Ekeberg Henriksen, 2002].

Blauert³⁶ ajoute que les sources sonores présentant un spectre riche entre 2000 et 7000Hz sont perçues comme proches de l'auditeur, tandis que celles dont le spectre est centré autour de 1000Hz paraissent diffuses. C'est certainement en partie sur ces informations que s'appuie la constitution d'un Acousmonium. Annette Vande Gorne décrit la répartition et le rôle de chacun des groupes de haut-parleurs : « La couleur, la réponse spectrale des transducteurs (du grave au suraigu), joue aussi un rôle. On les divise en cinq familles : les graves, appelés « contrebasses » (10 à 400Hz) en stéréo très large qui couvre tout le champ de la salle, ou au centre s'il n'y a qu'un « subwoofer » ; les médium « creux » (250 à 1000Hz) appelés à remplir l'espace ; les médium « clairs » (400 à 3000Hz) responsables du maximum d'audibilité du message, et particulièrement de la voix humaine ; les médiums « brillants » (3000 à 8000Hz) qui renforcent la présence, la vie microscopique des êtres sonores ; enfin, un groupe démultiplié de haut-parleurs suraigus (8000 à 16000 Hz) qui en précise les bords d'attaque. On se servira de petits haut-parleurs aigus pour préciser les contours en présence dans le public. On utilisera les trompettes suraiguës pour mieux localiser un ensemble placé loin. Les graves sont diffus. Des haut-parleurs de piètre qualité (médium creux) joueront sans problème un rôle de masse diffuse. » [Vande Gorne, 2002]. Ces informations peuvent être appliquées au spectre d'un son pour tenter d'obtenir des rendus spatiaux similaires.

Pour aller plus loin, si l'on fait varier les pics de fréquences (filtres *notch*) d'un spectre sonore, on aura le sentiment de percevoir le son « "bouger" sur un arc autour de la tête de l'auditeur compris entre l'avant et l'arrière. »³⁷ [Ekeberg Henriksen, 2002]. Bien sûr, cela dépendra du type de son. Le mouvement, en effet, « de la source ou bien de l'auditeur, engendre la variation dans le temps des informations fournies par les ITDs, les ILDs et les HRTFs » [Ekeberg Henriksen, 2002]. Un autre indice important dans la perception du mouvement est ce que l'on appelle « l'effet Doppler » : il correspond à un changement de

36 Jens BLAUERT, *Hearing Of Music In Three Spatial Dimensions*, Mis-en-ligne le 30 Mars 2015 sur https://www.researchgate.net/profile/Jens_Blauert [consulté le 12/10/2017]

37 Texte original : « [...] "move" in an arc above the listener's head between the front and behind. »

hauteur lorsque la distance entre le son et l'auditeur varie. Il donne ainsi des indications sur la vitesse avec laquelle le son s'éloigne ou se rapproche. Il est très facile pour le compositeur de mettre cela en application ; de nombreux logiciels et plugins proposent de simuler l'effet Doppler tels que le GRM Tools DOPPLER³⁸ par exemple.

Nous avons déjà amorcé une discussion dans *La spatialisation des musiques électroniques au regard du processus de composition*³⁹ sur la validité ou non de l'emploi de trajectoires dans la musique électronique, une technique de composition de l'espace très répandue mais pas toujours utilisée à bon escient. Nous pouvons à présent nous questionner sur sa validité d'un point de vue perceptif, au regard de l'auditeur. Référons-nous dans un premier temps à l'expérience de Dean [2014] qui consiste à tester la perception de la localisation et du mouvement de sons graves (sinus de 30Hz, 60Hz, et un bruit blanc donc les fréquences aiguës ont été filtrées), chacun joués pendant 0.2, 2 ou 6 secondes, dans l'environnement du studio. Ces sons sont localisés soit à gauche, soit à droite, ou bien bougent de façon constante de droite à gauche ou de gauche à droite. Il en conclut que nous arrivons correctement à localiser les sons graves, mais un peu moins bien à percevoir leur mouvement, surtout lorsque ces sons ne durent que 200ms. Effectivement l'oreille humaine, pour suivre un mouvement avec précision, a besoin que celui-ci dure au moins 172ms pour une alternance gauche–droite et 233ms pour une alternance avant–arrière [Blauert, 1997]. Objectivement, ces expériences réalisées en studios (généralement traités acoustiquement) dans des conditions "idéales" avec des sinus ou des bruits blancs ne prennent pas en compte tous les types de sons et tous les environnements possibles dans lesquels un auditeur peut pourtant se trouver. Moore [1989] déclare d'ailleurs que « nous sommes relativement insensibles au mouvement des sources sonores »⁴⁰. Pour nuancer le propos, nous pouvons citer les résultats obtenus par Payri [2011] suite à son expérience sur les patterns spatiaux, déjà décrite en partie I.2.3 de ce présent mémoire. Il conclut que l'acuité avec laquelle l'oreille humaine est capable de percevoir des patterns spatiaux dépend exclusivement des variations temporelles d'intensité et spectrales du son. On reconnaîtra beaucoup moins une figure spatiale si celle-ci est effectuée par des sons dont la nuance et le timbre varient dans le temps, contrairement aux sons synthétiques comme les sinus. « The recognition rate is very low for the conditions that are closer to real music or audiovisual sound [...] »⁴¹ [Payri, 2011]. Enfin, on peut à nouveau citer Payri [2011] : « Spatial movement patterns should therefore be restricted to sonification tasks

38 Site-web: <http://www.inagrm.com/grmtools>

39 [Nilles, 2016] Pour le lecteur intéressé, se référer à la partie II.2.1.

40 Texte original : « we are relatively insensitive to the motion of sound sources. »

41 « Le taux de reconnaissance est très faible pour des conditions proches de celles de la musique ou de l'audiovisuel. »

where the sound is perfectly stable in time, and the patterns are very simple, possibly imitating feasible sound trajectories found in nature like the regular pendulum movement or the acceleration that can occur with a bouncing object. »⁴².

Bien que le mouvement des sources sonores soit abondamment employé en musiques électroniques, et qu'il s'agisse même d'un moyen usuel de décrire la spatialisation du son à toute personne novice dans ce domaine, les arguments recueillis sur le plan perceptif ne justifient pas forcément un tel recours à cette technique. La perception de l'espace par l'oreille humaine n'est pas toujours précise et est largement influencée par les conditions environnementales et mentales, par le type de son, etc. Il nous paraît essentiel, en tant que compositeur, de privilégier un travail de l'espace qui satisferait à la fois la composition de l'oeuvre musicale et la perception de l'auditeur. C'est pourquoi notre partie II est consacrée à la notion d'immersion que nous considérons comme l'un des défis de la musique spatialisée de nos jours.

I.3.2 Remarques sur différentes situations d'écoute de la musique électronique spatialisée : le concert et l'écoute au casque

Le concert spatialisé

La perception de l'espace dépend de nombreux facteurs, mais est aussi grandement influencée par la vision. Les propriétés visuelles de l'environnement d'écoute sont impérativement à prendre en compte, de même que l'absence ou la présence d'autres membres du public, lors d'un concert de musiques électroniques spatialisées [Ekeberg Henriksen, 2002]. De même, si celui-ci a lieu en extérieur, il ne faut pas négliger l'influence des conditions météorologiques qui peuvent agir directement sur le signal sonore, comme nous l'avons déjà évoqué précédemment en partie I.1.3. Ainsi, si le public est placé à une grande distance des haut-parleurs, les variations de température et d'humidité et les mouvements de l'air créent des perturbations dans le signal sonore atteignant l'auditeur. Ces perturbations peuvent donner lieu à des annulations de phase ou à des pics de fréquences dans le spectre du

42 « Les patterns de mouvements spatiaux devraient être restreints à des tâches de "sonification", où le son est parfaitement stable dans le temps, avec des patterns très simples, imitant des trajectoires sonores que l'on trouve dans la nature comme le mouvement régulier d'un pendule ou des accélérations qui peuvent survenir avec un objet qui jaillit. »

son, conduisant parfois à la perception que la source sonore bouge ou qu'elle semble provenir d'un endroit différent [Ekeberg Henriksen, 2002]. Il est donc difficile de localiser précisément un son situé à une longue distance et ceci peut poser problème lors de concerts en extérieur, mais aussi dans les grandes salles de concert. Les concerts en intérieur, dans de petites salles, seraient donc à privilégier.

Pareillement, lors d'un concert spatialisé, les conditions d'écoute dépendent de la forme et de la taille de la salle ainsi que du type et de l'emplacement du système sonore (cf. partie III.2.2). Il peut y avoir une différence considérable dans les conditions d'écoute entre deux positionnements différents dans l'espace de diffusion. Il est impossible, pour une large audience, que tous les auditeurs aient une écoute égale. Il faut cependant trouver des compromis ; nous tenterons de réfléchir sur ce point en partie III.2. Ces déséquilibres sont principalement dus à l'effet de précedence.

« The precedence effect plays an important role in our perception of everyday sounds. It enables us to locate, interpret, and identify sounds in spite of wide variations in the acoustical conditions in which we hear them. Without it, listening in reverberant rooms would be an extremely confusing experience. Sometimes, however, the effect can be an inconvenience ! An example of this is found in the stereophonic reproduction of music. »⁴³ [Moore, 1989]

Lors d'une écoute stéréophonique, pour que celle-ci soit optimale, l'auditeur doit être placé à égale distance des deux haut-parleurs. S'il dévie de cette position, l'image stéréo est modifiée : l'auditeur perçoit le son comme provenant presque exclusivement⁴⁴ du haut-parleur le plus proche de lui [Moore, 1989]. Ce phénomène est dû aux ITD : on localise le son en direction de l'oreille qui a reçu en premier le signal sonore. Blauert⁴⁵ explique que l'effet de précedence, pour deux signaux sonores identiques, est effectif dès 1ms de retard entre les deux oreilles : la direction perçue n'est déterminée que par le haut-parleur d'où provient le signal sonore arrivé en premier. Puis, à partir de 50 à 80ms de retard entre les deux signaux, on perçoit deux sources sonores : l'une directe (provenant du haut-parleur le plus proche) et son écho (provenant du haut-parleur le plus loin). Ceci implique, lors de concerts spatialisés, un point d'écoute central "idéal", et des déséquilibres en dehors de ce point.

43 « L'effet de précedence joue un rôle important dans notre perception des sons au quotidien. Cela nous permet de localiser, d'interpréter et d'identifier les sons malgré les grandes variations des conditions acoustiques dans lesquelles nous les entendons. Sans cela, l'écoute dans des salles réverbérantes serait une expérience extrêmement déroutante. Cependant, cet effet peut aussi être un inconvénient ! Un exemple de ce phénomène est la reproduction stéréophonique de la musique. »

44 Ceci affecte le matériau sonore diffusé par les deux haut-parleurs à la fois.

45 Jens BLAUERT, *Hearing Of Music In Three Spatial Dimensions*, Mis-en-ligne le 30 Mars 2015 sur https://www.researchgate.net/profile/Jens_Blauert [consulté le 12/10/2017]

Remarques sur l'écoute au casque

Écouter de la musique spatialisée au casque, en binaural, peut sembler une solution pour éviter les problèmes dus au déplacement de l'onde sonore dans un environnement, ou encore les déséquilibres liés au positionnement de l'auditeur dans l'espace de diffusion. Ceci ne perturbe pas non plus l'environnement dans lequel l'auditeur se situe, à moins que le niveau d'écoute soit excessivement élevé. L'utilisation du casque élimine donc les signaux acoustiques résultants de la réflexion, la diffraction et l'absorption des ondes sonores dans l'environnement physique [Ekeberg Henriksen, 2002]. L'écoute peut alors sembler plus claire, moins brouillée, plus précise. Pourtant, l'auditeur ne peut s'appuyer sur les mouvements de la tête pour résoudre certaines imprécisions spatiales puisque les sons bougent avec le casque, à mesure que la tête tourne. Elle est cependant plus "naturelle" qu'une écoute au casque en stéréo puisque dans ce cas précis, l'effet de filtre causé par le pavillon de l'oreille (les indices spectraux) qui varie en fonction de la direction et du volume de la source sonore est absent. Il n'y a alors aucune aide de l'environnement et du pavillon de l'oreille pour détecter la position des sources sonores : le son semble uniquement provenir de l'intérieur de la tête⁴⁶ [Blauert, 1997]. Ceci se produit lorsque la distance de l'auditeur à l'événement sonore est jugée plus petite que le rayon de la tête.

Si elle est pratique pour écouter une musique électronique spatialisée n'importe où, dans n'importe quelles conditions, on peut se demander si l'écoute au casque en binaural peut être immersive ou non, bien qu'elle ne rassemble pas les caractéristiques d'écoute que l'on trouve dans un environnement physique. On peut déjà apporter quelques éléments de réponse en avançant que nous n'écoutons pas seulement avec nos oreilles mais aussi avec notre corps, et ceux-ci sont en fait des récepteurs, puisqu'ils font partie du système nerveux qui transmet les informations au cerveau. L'expérience sonore dite "tactile" n'est alors pas présente lors de l'écoute au casque. Nous reviendrons sur l'ensemble des arguments introduits dans cette seconde moitié de conclusion en partie III.

46 « inside-the-head locatedness » [Blauert, 1997].

II. La notion d'immersion dans la musique électronique spatialisée

« A review of perceptual observations made by composers and survey-based research relating to multi-channel space reveals that a sense of “immersiveness” and “surroundness” are the attributes most commonly perceived. »⁴⁷ [Lynch & Sazdov, 2011]

L'*immersion* est un terme relativement connu, utilisé dans les recherches en acoustique des salles, ou bien en reproduction multicanale [Lynch & Sazdov, 2011]. [Périaux & al., 2015], à propos du système 5.1 : « Il permet de placer le spectateur dans un champ sonore. On parle alors d'effet d'immersion [...] ». Les configurations de haut-parleurs standardisées (5.1, 7.1, Auro3D, ...), dont on se sert le plus souvent dans le milieu du cinéma mais aussi en musique électronique, sont généralement qualifiées par les commerciaux de « Surround Sound System » (pour les systèmes 2D) et « Immersive Sound System » (pour les systèmes 3D). Le terme « immersion » ou « son immersif »⁴⁸ est également extrêmement employé de nos jours en musiques électroniques, à la fois par les compositeurs mais aussi par les auditeurs, pour décrire leur ressenti perceptif. En effet, si l'on prend l'exemple de l'étude menée par Peters [2010], qui a envoyé des questionnaires à tout un panel de compositeurs de musiques électroniques adeptes du travail de l'espace, pour mieux comprendre leurs préférences et leurs envies, on remarque que le facteur appelé « immersiveness », que l'on pourrait traduire en français par « capacité à immerger », est le facteur le plus important en ce qui concerne la spatialisation. Le facteur « elevated sound sources »⁴⁹ est également notable, mais ne semble pas apporter une très grande satisfaction – il donnerait même, toujours d'après

47 « Une revue des observations perceptuelles faites par les compositeurs et par la recherche basée sur des sondages en rapport avec l'espace multicanal, montre que la sensation d'« immersion » et celle d'être « entourée par les sons » sont les attributs les plus souvent perçus. »

48 Hervé Déjardin, Radio France.

49 « élévation des sources sonores »

cette étude, moins de satisfaction que le facteur « virtual sound sources within the audience »⁵⁰, qui est pourtant un résultat extrêmement difficile à obtenir pour l'instant, même avec une technique de pointe telle que la Wave Field Synthesis développée à l'Ircam⁵¹ !

Ainsi, la notion d'immersion, en musiques électroniques spatialisées, prend de plus en plus d'ampleur, et il paraît nécessaire de chercher à la définir, dans le but de donner des pistes au compositeur intéressé, d'y réfléchir sous différents angles de vue. On peut se demander s'il existe une configuration précise de haut-parleurs qui permettrait d'obtenir un espace sonore immersif. Les systèmes 3D ne semblent pas procurer encore beaucoup de satisfaction aux auditeurs du point de vue de la perception ; sont-ils tout de même capables de faire naître un sentiment d'immersion chez l'auditeur ? Y a-t-il une technique de spatialisation, parmi celles déjà étudiées dans le mémoire précédent (cf. [Nilles, 2016] partie II) qui permettrait de créer un espace sonore immersif par le biais de la musique électronique ? « It is proposed that envelopment be investigated, as it is evident that it is the spatial aspect most associated with multi-channel electroacoustic music. »⁵² [Lynch & Sazdov, 2011] La notion d'enveloppement est également très répandue chez les compositeurs et auditeurs de musiques électroniques spatialisées. *Enveloppement* et *immersion* sont-ils la même chose ? Le terme « immersion » fait souvent référence à la mer, au fait d'être dans l'eau. Natasha Barrett dit : « Immersion is a good example of where a sounding object becomes a sounding space. »⁵³ ⁵⁴ Est-ce alors *plonger* l'auditeur dans un véritable *environnement* sonore ? Au vu du nombre d'informations que nous avons pu rassembler sur la notion d'immersion, il paraît évident qu'il y a différentes façons de créer et de ressentir l'immersion.

Pour tenter de définir dans son ensemble la notion d'immersion relative à la musique électronique spatialisée, nous nous appuierons, entre autre, sur diverses expériences et nous considérerons, dans un premier temps, ce qui donne naissance au sentiment d'immersion et, dans un deuxième temps, quelles sont les conséquences perceptibles pour un auditeur placé dans un espace sonore à caractère immersif.

50 « sources sonores virtuelles comprises à l'intérieur de l'audience »

51 Voir [Nilles, 2016].

52 « Nous proposons que des recherches s'intéressent à l'enveloppement, car il paraît évident qu'il est l'aspect spatial le plus associé à la musique électroacoustique multicanale. »

53 « L'immersion est un bon exemple du moment où un objet sonore devient un espace sonore. »

54 Barrett, Natasha, 2010, *Ambisonics and acousmatic space : a composer's framework for investigating spatial ontology*, Proceedings of the Sixth Electroacoustic Music Studies Network Conference, Shanghai.

II.1. Ce qui est à l'origine de l'immersion

II.1.1 La réverbération et le sentiment d'enveloppement

Nous considérons une salle de concert. La réverbération est généralement ce qui caractérise l'acoustique d'une salle. Elle est définie par différents types de réflexions des ondes sonores. Adair, Alcorn et Corrigan [2008] citent Marshall et Barron⁵⁵ pour évoquer la différence entre « les premières réflexions qui créent un élargissement de la source, [ainsi] associé à la largeur apparente de la source » (ASW en anglais), et les réflexions latérales tardives responsables de « l'enveloppement de l'auditeur » (LEV en anglais) et « provoquant un sentiment "d'immersion" »⁵⁶. On apprend ici que la notion d'immersion peut être une conséquence directe de la réverbération d'une source sonore, dans une salle, et notamment des réflexions tardives, à conditions qu'elles surviennent au moins 80ms après la fin de l'évènement sonore [Rumsey, 2001]. On a ainsi une caractéristique d'un *environnement* immersif. En effet, on perçoit les réflexions détachées de la source sonore d'origine. Griesinger⁵⁷, cité dans [Rumsey, 2001], fait état des différentes formes d'impression spatiale d'une salle. Il nomme BSI (pour « Background Spacial Impression ») l'impression spatiale résultante lorsque la majorité de l'énergie réfléchie arrive au moins 50ms après la fin de l'évènement sonore. Le sentiment d'immersion ainsi produit est propre à l'environnement et non à la source d'origine : « Spatially diffuse, late reflected energy of this type results in good envelopment, but BSI is not bound to the source that created it (so it is in effect an environment cue). »⁵⁸ [Rumsey, 2001]. Griesinger explicite également un cas où c'est la source qui est immersive : la CSI (pour « Continuous Spacial Impression »). Lorsque la réverbération d'un son ne peut pas être perçue comme détachée du son, on a alors une source qui semble immersive. Il doit s'agir d'un son continu, que le cerveau ne peut pas séparer en plusieurs évènements (un flux, une nappe sonore, un sinus tenu, ...) pour que ses réflexions

55 Marshall, A.H., Barron, M., *Spatial responsiveness in concert halls and the origins of spatial impressions*, Applied Acoustics, 62, 2001, 91-108

56 Citation originale : « Marshall and Barron differentiate between early lateral reflections that cause the source broadening associated with apparent source width (ASW), and LEV caused by late reverberant sound from the side resulting in the feeling of 'immersion' ».

57 Griesinger, D., 1997, *Spatial impression and envelopment in small rooms*, Presented at 103rd AES Convention, New York, 26-29 September, Pre-print 4638.

58 « Diffusée dans l'espace, il résulte un bon enveloppement de cette énergie réfléchie, mais la BSI n'est pas liée à la source qui en est à l'origine (donc il s'agit d'une propriété de l'environnement). »

soient comprises comme faisant partie intégrante du son. Ce qui nous donne une source directe perçue comme enveloppante, immersive : « [...] when a direct sound source is continuous and cannot be split into separate events, the interaction with reflected energy and the interaural fluctuations in amplitude and time delay that result can give rise to a sense of full envelopment or spaciousness that appears to be connected to the sound (CSI). »⁵⁹ [Rumsey, 2001].

Lorsque l'on s'intéresse de près ou de loin au phénomène de la réverbération, on retrouve souvent les mêmes qualificatifs pour en décrire les effets, et ils se rapportent majoritairement à la notion d'immersion dans le cas des réflexions tardives : comme nous venons de le voir, on parle souvent « d'enveloppement »⁶⁰ ou de « sensation d'espace »⁶¹. Ces expressions « sont d'abord en relation avec l'impression spatiale environnementale, et sont généralement le résultat d'un son réfléchi – quasiment toujours les réflexions tardives du son (en particulier les réflexions latérales après 80ms). »⁶² [Rumsey, 2001]. Le sentiment d'enveloppement est souvent décrit en faisant une analogie avec un concert d'orchestre par exemple dans une salle : si le son semble provenir de partout, on a alors une sensation d'enveloppement, perçue par ailleurs comme positive – voire même comme un gage de qualité du point de vue des auditeurs⁶³. Le terme « spaciousness » est très utilisé pour décrire un sentiment d'espace ouvert ou de *room*, mais aussi pour exprimer le sentiment "d'externalisation" perçue (c'est-à-dire que le son apparaît en dehors de la tête plutôt que dedans ou très proche) [Rumsey, 2001]. Dans [Périaux & al., 2015], il est aussi traduit en français par « l'impression d'être entouré par les sons ». Enfin, ces expressions sont employées aussi bien en relation avec des sources sonores qu'avec des environnements. On a donc souvent un sentiment d'enveloppement mais surtout, il semble que c'est plutôt une image large qui est créée avec les réflexions tardives. Ainsi, elles peuvent produire une sensation d'immersion dans la plupart des cas, mais ne sont pas toujours une condition suffisante.

« [...] subjects often use the term envelopment when they are surrounded by a number of dry sources in surround sound reproduction...this sensation is almost certainly not a property of

59 « [...] quand une source sonore directe est continue et ne peut pas être séparée en plusieurs événements, l'interaction entre l'énergie réfléchie et les fluctuations interaurales d'amplitude et de délai ainsi produites peut être à l'origine d'un sentiment d'immersion totale ou d'un sentiment d'espace qui semble être connecté au son (CSI). »

60 Ou de « LEV » (« Listener envelopment ») chez Griesinger, D., 1997, *op. cit.* et [Rumsey, 2001].

61 En anglais : « spaciousness » (Griesinger, D., 1997, *op. cit.* ; [Rumsey, 2001]).

62 Citation originale : « [...] are primarily related to environmental spatial impression, and are largely the result of reflected sound – almost certainly late reflected sound (particularly lateral reflections after about 80ms). » [Rumsey, 2001]

63 Jens BLAUERT, *Hearing Of Music In Three Spatial Dimensions*, Mis-en-ligne le 30 Mars 2015 sur https://www.researchgate.net/profile/Jens_Blauert [consulté le 12/10/2017]

late reflected sound, as the sources in question can be dry and direct, so it cannot be considered to be LEV in the traditional sense. »⁶⁴ [Rumsey, 2002]

La citation de Rumsey ci-dessus témoigne du fait que la réverbération dans les salles de concert peut créer de l'immersion, mais n'est en aucun cas essentielle. Dans le cas d'une diffusion multicanale, avec des haut-parleurs entourant le public, un amas de sources sonores directes et dites *dry* peut produire un sentiment d'enveloppement chez l'auditeur. En effet, cela revient à créer un *espace environnemental immersif*, tel que décrit en partie II.2.3 de mon mémoire de première année (cf. [Nilles, 2016]). On parle alors d'enveloppement « d'ensemble » ou « environnemental » [Adair & al., 2008].

Il semble important d'ajouter que *immersion* ne signifie en aucun cas *précision*⁶⁵. Bien que, dans la vie de tous les jours, nous sommes "immergés" dans l'environnement qui nous entoure, ressentir l'immersion dans le cas d'un concert ou d'une reproduction sonore ne veut pas dire que l'espace entendu est précis, composé de détails, "comme dans la réalité". Si l'on considère par exemple le phénomène des réflexions tardives d'un son, apparaissant plus de 50ms après la fin de l'évènement sonore, on sait qu'elles participent à la détérioration de la clarté du son, de son intelligibilité [Rumsey, 2001]. Nous allons voir dans ce qui suit les conditions techniques nécessaires pour obtenir un sentiment d'immersion.

II.1.2 Les conditions techniques pour obtenir de l'immersion (systèmes de diffusion)

« When the orchestra plays at a low volume, listeners experience auditory events which are spatially limited to the visual outline of the orchestra. When the volume is increases it is observed in many concert halls that the auditory events increase in spatial extension - particularly to the sides, in depth, and upwards. It has been known for a long time that listeners

64 « [...] les sujets utilisent souvent le terme enveloppement quand ils sont entourés par un nombre de sources directes lors d'une diffusion multicanale... cette sensation n'est absolument pas une propriété des réflexions tardives d'un son, puisque les sources en question sont *dry* et directes ; ce n'est donc pas un phénomène de LEV au sens traditionnel. »

65 Voir la partie II.2.1 du présent mémoire pour plus de détails à ce sujet.

to music prefer sound fields in which the spaciousness effect can be experienced. »^{66 67}

Lorsqu'un orchestre joue dans une salle de concert, aux nuances *forte* ou *fortissimo*, l'auditeur peut avoir l'impression que le son provient de partout, comme s'il était diffusé par des haut-parleurs placés sur les côtés, à l'arrière, au-dessus. Blauert, dans la citation ci-dessus, utilise le terme « spaciousness » qui, comme nous l'avons vu dans la sous-partie précédente, peut souvent être à l'origine d'un sentiment d'immersion. Ainsi, en faisant une analogie avec la musique électronique, si celle-ci est diffusée à fort volume sur des haut-parleurs entourant le public et même au-dessus, nous pourrions susciter l'impression d'être immergé dans le son. Partant du principe que le sentiment d'enveloppement est tout à fait subjectif, les chercheurs Adair, Alcorn et Corrigan [2008] ont réalisé une expérience afin de comprendre à partir de quel moment un individu perçoit un sentiment d'enveloppement avec la musique électronique et notamment électroacoustique. Pour cela, chaque participant était placé au centre d'un dispositif de huit haut-parleurs et pouvait contrôler à la main, séparément, le volume des haut-parleurs côtés et arrières, les deux haut-parleurs frontaux restant à un volume fixe choisi par les musicologues. Quatre extraits musicaux étaient diffusés à la suite, puis le tout une deuxième fois, pour obtenir une moyenne, et le terme « enveloppement » avait été explicité au préalable. Les résultats obtenus étaient les suivants : la plupart des participants ajustaient très vite après le début de l'extrait musical le niveau des haut-parleurs côtés et arrières, indiquant qu'ils étaient nécessaires dans la perception de l'enveloppement. Ils ont également choisi des niveaux de diffusion plus ou moins semblables, réfutant l'hypothèse que le sentiment d'enveloppement est quelque chose de subjectif. En revanche, les avis divergeaient concernant les fréquences pouvant être responsables de l'impression d'enveloppement, certaines personnes s'étant basées sur les fréquences hautes, d'autres sur les fréquences basses, les chercheurs ne pouvaient donc rien conclure. Plus généralement, le critère semblant indispensable au sentiment d'enveloppement est d'obtenir un son équilibré, provenant de toutes les directions. « We discovered that unlike concert hall acoustics, a high level of sound that does not necessarily consist of late reflections is required from the rear. »⁶⁸ [Adair & al.,

66 « Quand l'orchestre joue à une faible intensité, les auditeurs peuvent expérimenter des événements sonores qui se limitent dans l'espace à la place physique occupée par l'orchestre. Quand le volume augmente on observe, dans beaucoup de salles de concert, que les événements sonores s'étendent dans l'espace – en particulier sur les côtés, en profondeur et en hauteur. C'est connu depuis bien longtemps : les auditeurs, en musique, ont une préférence pour les champs sonores pour lesquels un effet de *spaciousness* [étendue spatiale] peut être expérimenté. »

67 Jens BLAUERT, *Hearing Of Music In Three Spatial Dimensions*, Mis-en-ligne le 30 Mars 2015 sur https://www.researchgate.net/profile/Jens_Blauert [consulté le 12/10/2017].

68 « Nous avons découvert que, au contraire de l'acoustique des salles de concert, un fort volume sonore pas nécessairement attribué aux réflexions tardives est requis à l'arrière. »

2008]. L'immersion sonore, engendrée par le sentiment d'enveloppement, ne dépend pas forcément des réflexions tardives. Comme nous l'avons déjà évoqué en partie II.1.1, une accumulation de sons non réverbérés et de forte intensité à l'arrière et sur les côtés, via le dispositif multicanal, génère une impression d'immersion.

La diffusion de musique électronique sur des systèmes 3D semble également provoquer l'immersion du public. Les écrits de Emmerson [2007], Normandeau [2009] ou encore de Lynch et Sazdov [2011] décrivent le concert de Stockhausen, à Osaka⁶⁹, en 1970, comme « immersif », suscitant un « sentiment d'immersion ». Stockhausen avait lui-même composé la musique et créé le dispositif de diffusion qui consistait en une sphère de haut-parleurs positionnés sur sept niveaux différents, allant du dessous jusqu'au dessus du public [Lynch & Sazdov, 2011]. Emmerson [2007] parle d'une « immersion multicanale ». « The evidence would suggest that a perceptual sense of feeling immersed or enveloped in sound is associated with 3D multi-channel diffusion systems. »⁷⁰ [Lynch & Sazdov, 2011]. Pourtant, une expérience réalisée par Guastavino et Katz en 2004 montre que c'est le dispositif multicanal 2D qui est préféré pour produire une immersion sonore, comparé aux configurations 3D et stéréophoniques. Ces dernières diffusent une image sonore frontale très précise et stable, mais peu présente, et en aucun cas associée à une quelconque sensation d'immersion. « The 2-D configurations were spontaneously described in the open questionnaires as very enveloping, spatially well defined, and providing a good sense of immersion in the scene, equating to a high degree of presence. »⁷¹ [Guastavino & Katz, 2004]. En revanche, suite aux différents tests, les systèmes 3D ont été jugés brouillons, peu enveloppants et ont donné l'impression que les événements sonores diffusés sonnaient loin des auditeurs – ces résultats vont tout à fait à l'encontre des qualificatifs donnés aux dispositifs standardisés par les commerciaux (cf. introduction de la partie II). On peut noter ici le lien entre immersion et présence du son mis en évidence dans les deux cas cités.

« According to Sazdov et al. (2007)⁷², engulfment is perceived as feeling covered over as opposed to being just enveloped or surrounded by sound. Participants rated the perception of engulfment significantly higher when the spatial scenes were presented on elevated

69 Osaka International Exhibition au Japon.

70 « Il semblerait évident de dire que le fait de se sentir immergé ou enveloppé dans le son est associé à la diffusion multicanale 3D. »

71 « Les configurations 2D ont été décrites spontanément, dans les questionnaires, comme très enveloppantes, bien définies spatialement, engendrant une bonne sensation d'immersion, ainsi qu'un haut degré de présence. »

72 Sazdov, R., Paine, G., Stevens, K., 2007, *Perceptual Investigation into envelopment, spatial clarity, and engulfment in reproduced multi-channel audio*, Proceedings of AES 31st International Conference, London.

loudspeakers when compared to presentation on horizontal loudspeakers only. »⁷³ [Lynch & Sazdov, 2011]

Si les systèmes de diffusion 3D ne sont pas capables de procurer une sensation d'immersion – comme les résultats obtenus chez [Guastavino & Katz, 2004] semblent le montrer – c'est qu'ils ne créent pas un sentiment d'enveloppement mais d'« engulfment ». Le terme apparaît pour la première fois dans les écrits de Sazdov⁷⁴ et désignerait le fait d'être « recouvert par les sons ». D'après [Lynch & Sazdov, 2011], l'immersion serait alors un attribut spatial caractéristique des dispositifs multicanaux 2D, tandis que l'« engulfment » serait un attribut spatial correspondant uniquement aux dispositifs multicanaux 3D. Il est néanmoins certain qu'un système multicanal, avec des haut-parleurs côtés et arrières diffusant à un niveau suffisant, est nécessaire pour créer une sensation d'enveloppement avec la musique électronique. Ainsi, la notion d'immersion n'est pas rattachée à un format en particulier (5.1, 7.1, Auro 3D, etc.). Chaque compositeur, selon ses expériences, semble avoir sa préférence. On peut citer ainsi le compositeur Chris Rolfe⁷⁵ qui affirme que le son est enveloppant lorsqu'il est distribué sur 8 haut-parleurs (2D), tandis que Robert Normandeau [2009] pense que le fait de pouvoir se mouvoir dans l'espace d'écoute, indépendamment du dispositif multicanal, fait que l'auditeur est entièrement immergé dans le son (cf. partie III.2.2). Les acousmoniums sont également des moyens de rendre l'écoute d'une œuvre stéréophonique immersive, notamment de par leur configuration : certains haut-parleurs sont orientés de façon à ce que le signal sonore qu'ils projettent soit réfléchi par les murs ou le plafond de la salle de concert avant d'atteindre l'auditeur [Ekeberg Henriksen, 2002]. C'est notamment l'avis de Jonty Harrison⁷⁶, fondateur du système BEAST.

73 « D'après Sazdov et al. (2007), l'« engulfment » est perçu comme le sentiment d'être recouvert, en opposition au fait d'être simplement enveloppé ou entouré par le son. Les participants ont noté que la perception de l'« engulfment » était beaucoup plus significative lorsque les scènes spatiales étaient présentées sur des haut-parleurs élevés, en comparaison à une présentation sur des haut-parleurs uniquement horizontaux. »

74 *Op. cit.*

75 Cité dans [Lynch & Sazdov, 2011].

76 Harrison, J., 1998, *Sound, space, sculpture: some thoughts on the “what”, “how”, and “why” of Sound diffusion*, Organized Sound.

II.1.3 Les « attributs immersifs » et les techniques de spatialisation conduisant à l'immersion

Le sentiment d'enveloppement joue un rôle important dans la création d'une sensation d'immersion. Depuis le début de cette partie II, nous employons cette expression comme directement associée à la notion d'immersion ; ce qui est vrai, mais il est nécessaire de nuancer notre propos. En effet, un espace immersif implique une perception d'enveloppement, mais l'inverse n'est pas toujours juste, si l'on en croit ce que nous avons dit précédemment⁷⁷ : l'attribut « engulfment » d'un système 3D n'aurait rien d'immersif et l'enveloppement, plutôt employé pour les systèmes 2D, semble similaire à cette notion de recouvrement par les sons. On trouve chez [Rumsey, 2002] et [Kendall, 2010] que les « attributs immersifs » sont : l'« enveloppement », décrit comme le fait d'être entouré par les sons, mais aussi la « présence » (se sentir être à l'intérieur d'un espace fermé). Ainsi, ce seraient les deux conditions à réunir pour obtenir une immersion sonore *totale*. Les commentaires des participants de l'expérience décrite chez [Guastavino & Katz, 2004] sont quasi unanimes : « une scène immersive semble proche »⁷⁸. Effectivement, les dispositifs multicanaux 2D, appréciés car, pour une même pièce, l'espace est perçu comme plus immersif qu'avec les autres configurations, procurent également une sensation de présence et une image spatiale semblant plus proche qu'avec des dispositifs 3D par exemple.

L'enveloppement, souvent conséquence des réflexions tardives du son⁷⁹, est caractérisé de différentes façons. Il y a le « LEV » dont nous avons déjà discuté en partie II.1.1, décrit comme « le degré de « fullness » des images sonores autour de l'auditeur »^{80 81}. D'après [Peters, 2010], il dépend du rapport entre l'énergie sonore à l'avant et à l'arrière du public, du volume globale d'écoute, mais aussi du spectre de fréquences de l'évènement sonore, notamment pour celles situées en-dessous de 1000Hz. On retrouve cette importance donnée aux basses fréquences chez [Périaux & al., 2015], dans lequel on apprend d'ailleurs que le critère d'enveloppement repose principalement sur deux paramètres : la décorrélation interaurale et le contenu basses fréquences, en-dessous de 600Hz. « Plus les basses fréquences sont décorrélées, plus la sensation d'enveloppement est grande. » [Périaux & al., 2015]. En effet, « la plupart des recherches en psycho-acoustique suggèrent que les oreilles ont besoin

77 cf. [Guastavino & Katz, 2004], [Lynch & Sazdov, 2011].

78 Texte original : « an immersive scene sounds close ».

79 cf. partie II.1.1 du présent mémoire.

80 Texte original : « degree of fullness of sound images around the listener ».

81 Morimoto, M., Iida, K., Sakagami, K., 2001, *The role of reflections from behind the listener in spatial impression*, Applied Acoustics, 62, 109-124.

de signaux décorrélés pour créer une meilleure sensation d'enveloppement [...] »⁸² [Rumsey, 2001].

« We conclude that, while envelopment in this music genre⁸³ is principally different to concert hall acoustics, a comparable perception of envelopment was achieved by the participants. »⁸⁴ [Adair & al., 2008]

En musiques électroniques, l'enveloppement ne dépend pas toujours de la réverbération (plus précisément des réflexions tardives), mais plutôt du choix du son (enveloppe, fréquences,...) et des techniques de spatialisation utilisées lors du processus de composition de l'oeuvre. Nous pouvons citer la décorrélation micro-temporelle, une technique de spatialisation du son, déjà explicitée dans notre précédent mémoire⁸⁵, qui permet de créer de l'enveloppement voire même de l'immersion [Nilles, 2016]. Alors qu'elle est souvent employée, en stéréo, pour étendre l'image sonore d'une source, elle génère un espace à caractère enveloppant en multicanal et défie l'effet de précédence⁸⁶ [Kendall, 2010] [Lynch, 2014]. À ce moment, nous pouvons affirmer⁸⁷ qu'elle permet de créer un espace immersif, appelé « champ diffus », qui peut être défini comme suit : « En acoustique, un champ diffus est un champ où l'énergie volumique moyenne a la même valeur en tous les points et où l'intensité acoustique est la même dans toutes les directions. La partie tardive de la réverbération d'une salle est un cas concret où nous pouvons noter la présence d'un champ diffus. Du point de vue de l'auditeur on note une sensation d'enveloppement, ce dernier ne peut en effet associer le champ sonore perçu à un nombre déterminé de sources sonores présentes dans l'espace. »⁸⁸ La conséquence d'un champ diffus est généralement la sensation d'un espace immersif [Nilles, 2016].

82 Texte original : « Most of the psychoacoustic research suggests that the ears need to be provided with decorrelated signals to create the best sense of envelopment [...] »

83 Les auteurs font référence au genre « électroacoustique ».

84 « Nous pouvons conclure que, alors que l'enveloppement dans ce genre musical est principalement différent de l'acoustique d'une salle de concert, une perception d'enveloppement comparable a été atteinte par les participants. »

85 cf. partie II de [Nilles, 2016].

86 cf. partie I du présent mémoire.

87 D'après [Nilles, 2016].

88 Colafrancesco, J., 2012, *L'Ambisonie d'ordre supérieur et son approbation par les musiciens : présentation de la bibliothèque Max/MSP hoa.lib*, Actes des Journées d'Informatique Musicales 2012, Mons, Belgique.

La spatialisation timbrale⁸⁹ et les techniques de spatialisation granulaire⁹⁰ - d'ailleurs assimilables à la décorrélation micro-temporelle [Nilles, 2016] – sont d'autres techniques de spatialisation des musiques électroniques qui permettent non seulement d'aboutir à un espace enveloppant, mais aussi de faire ressortir l'« attribut immersif » *présence* de cet espace ainsi créé. Avec ces techniques, c'est souvent le son qui est l'espace. On obtient ainsi un espace sonore opaque, comme un voile, dans lequel l'auditeur peut se déplacer ; il peut quasiment "toucher" le son.

La décorrélation micro-temporelle, la spatialisation timbrale, ou encore les techniques de spatialisation granulaire, permettent donc de créer de l'immersion [Nilles, 2016] : le compositeur qui les utilise peut vraisemblablement faire percevoir à l'auditeur les « attributs immersifs » précédemment cités qui sont l'*enveloppement* et la *présence*. Il nous paraissait important de rappeler ces techniques de composition de l'espace et de les considérer puisqu'elles participent à définir la notion d'immersion. Nous pouvons conclure que, non seulement l'immersion ne dépend pas d'un dispositif-type de haut-parleurs, mais elle n'est pas non plus rattachée à une technique de composition de l'espace particulière.

II.2. Ce qui est conséquence de l'immersion

II.2.1 Immersion et « naturalness » : une expérience sonore *naturelle*

Lors d'une diffusion sonore à caractère immersif, l'auditeur peut percevoir la création d'un environnement sonore dit "naturel". Dans ses écrits⁹¹, Rumsey emploie souvent le terme

89 Pour rappel, on peut donner cette définition, tirée de James, S.G. & Hope, C., 2013, *2D and 3D timbral spatialisation : spatial motion, immersiveness, and notions of space*, International Computer Music Associations (ICMA 2013) : « Whilst timbral spatialisation has emerged recently as a distinct spatial methodology in its own right, it borrows from other existing panning theory, and differs only in that instead of a single point-source approach to a given sound, we have potentially thousands of different point-source locations active at any moment in time for each respective spectral band of a given source sound. ».

90 Voir : Wilson, S., 2008, *Spatial Swarm Granulation*, dans Proceedings of the 2008 International Computer Music Conference, SARC, ICMA: Belfast. et la « Spectral Magnitude based decorrelation » dans Wilson S., Harrison, J., 2010, *Rethinking the BEAST: Recent developments in multichannel composition at Birmingham ElectroAcoustic Sound Theatre*, Organised Sound, 15(03), 239–250.

91 Mason, R., Ford, N., Rumsey, F. and de Bruyn, B., 2000, *Verbal and non-verbal elicitation techniques in the subjective assessment of spatial sound reproduction*, Presented at 109th AES Convention, Los Angeles, 22-25 September, Audio Engineering Society ; Berg, J. and Rumsey, F., 2000, *In search of the spatial dimensions*

« naturalness » pour désigner « le degré de "réalisme" ou de "ressemblance à la nature" avec lequel un sujet perçoit l'expérience spatiale »⁹² [Rumsey, 2001]. La « naturalness » serait plutôt subjective, difficilement mesurable avec des moyens physiques, mais, au vu des commentaires des participants de nombreuses expérimentations, elle semblerait marquer une préférence et devenir un facteur important lorsque l'on cherche à qualifier des espaces sonores. En effet, suite à l'expérience déjà évoquée précédemment, chez [Guastavino & Katz, 2004], les résultats ont montré qu'il y avait une forte préférence pour les configurations multicanales 2D en ce qui concerne la diffusion d'un espace immersif, et que le facteur « naturalness » était largement attribué à ces systèmes et beaucoup plus rarement aux autres également testés. On peut donc supposer que *immersion* et « naturalness » sont directement associées. Les systèmes 2D sont vraisemblablement choisis comme renvoyant au mieux un espace dit "naturel" : « General results [...] show a subjective impression of a more “realistic” or “natural” representation of the soundscape using the 2-D system versus the other systems. A more detailed analysis shows that the subjective ratings depend heavily on the soundscape. »⁹³ [Guastavino & Katz, 2004]. On parle aussi de « réalisme » dans la diffusion des paysages sonores (bruits de trafic, rues piétonnes, ...). Mais, comme le soulignent Guastavino et Katz [2004], ces impressions de « naturalness » dépendent nettement des paysages sonores proposés lors des tests et, plus généralement, du type de sons diffusés. Par exemple, la stéréophonie était préférée et définie comme plus "naturelle" lorsque les extraits audios correspondaient à des enregistrements de concerts de musique instrumentale ; ce qui paraît tout à fait "logique" car les gens ont l'habitude de percevoir la musique instrumentale comme provenant de face, en situation de concert, ou sur leur dispositif domestique stéréophonique ou monophonique. Ainsi, les habitudes d'écoutes de chacun sont très importantes et sont à prendre en compte car elles influencent beaucoup les attentes de l'auditeur, ce qu'il souhaite entendre lors d'une expérience ou d'un concert. Nous sommes d'ailleurs tous habitués à écouter de la musique sur un dispositif stéréophonique qui produit une « image fantôme », ayant pour effet de nous faire oublier en quelque sorte le média haut-parleur entre la source sonore et nos oreilles – à condition que l'on soit positionné à égale distance de chaque enceinte. Par conséquent, nous faisons facilement abstraction du haut-

of reproduced sound : verbal protocol analysis and cluster analysis of scaled verbal descriptors, Presented at 109th AES Convention, Los Angeles, 22-25 September, Audio Engineering Society ; ou encore [Rumsey, 2001].

92 Texte original : « It appears to be relatively independant of other factors and relates to the subject's perception of the degree of "realism" or "trueness to nature" of the spatial experience. »

93 « Les résultats généraux mettent en évidence une impression subjective d'une représentation plus "réaliste" ou "naturelle" du paysage sonore avec un système 2D, par rapport aux autres systèmes. Une analyse plus détaillée montre que les évaluations subjectives dépendent fortement du paysage sonore. »

parleur et l'écoute peut nous sembler "naturelle" même réalisée avec un système stéréophonique [Ekeberg Henriksen, 2002]. D'ailleurs – nous nous éloignons quelque peu du sujet mais nous jugeons cette remarque intéressante – si les participants aux tests réalisés par Guastavino et Katz [2004] ont préféré le système 2D par rapport au système 3D, c'est très certainement aussi parce qu'ils sont plus habitués au premier – qui est en effet plus répandu, et d'autant plus en 2004, époque des tests – contrairement au second qu'il est plus difficile de rencontrer. Donc le facteur « naturalness » dépend bien essentiellement du matériau sonore et des habitudes individuelles d'écoute, plutôt que du dispositif employé.

Hormis les phénomènes "non-naturels" tels la création d'environnements inouïs⁹⁴ – comme le fait d'être "hors-phase" par exemple – , les propriétés spatiales⁹⁵ perçues lors de diffusions sonores sont généralement similaires à celles caractéristiques des environnements dits "naturels". Cependant, il ne faut pas confondre "degré de réalisme" et "précision spatiale". Citons une nouvelle fois [Guastavino & Katz, 2004] qui est riche en exemples : les systèmes 2D et 3D diffusaient les mêmes scènes sonores captées, à l'origine, en Ambisonie, puis décodées pour chaque configuration. Ainsi, on peut dire que les dispositifs 2D projetaient un espace sonore "incomplet" par rapport aux systèmes 3D puisqu'il "manquait" la hauteur ; c'est pourtant la 2D qui a été préférée par les participants. Ce n'est donc pas la précision spatiale – au détail près – qui caractérise le facteur « naturalness » mais le degré de réalisme qui est plutôt un ressenti, c'est-à-dire en quoi l'espace entendu nous semble naturel.

« The illusion of “being there” is not necessarily related to true spatial fidelity »⁹⁶ [Rumsey, 2002]. L'impression « d'être là » décrit encore une façon de ressentir l'immersion. Comme nous l'avons déjà souligné, *immersion* ne signifie pas haute précision spatiale. Nous pouvons dire également que le sentiment d'immersion renvoie à un environnement naturel. Nous sommes naturellement baignés dans les sons au quotidien. Un environnement immersif engendre donc une « impression de naturel » – « naturalness ». L'auditeur a la sensation que les haut-parleurs n'existent pas, il se sent *immergé* dans ce nouvel espace créé. Il est littéralement plongé, par le biais de l'immersion, dans une autre réalité : il vit une *expérience* sonore.

94 cf. [Nilles, 2016] partie II.

95 La distance, la direction, le mouvement, et la réverbération (Zelli, B., 2009, *Spatialization as a Musical Concept*, dans « IEEE International Symposium on Mixed and Augmented Reality 2009 Arts, Media and Humanities Proceedings », Orlando, FL, USA, pages 35–38).

96 « L'illusion d' "être là" n'est pas nécessairement liée à une réelle fidélité spatiale. »

II.2.2 Immersion et ressenti par le corps de l'auditeur

Lors d'une écoute ou dans un concert de musique électronique sur dispositif multicanal, l'auditeur, pour percevoir les sons et l'espace, doit être placé au cœur du système. Nous avons introduit, dans la partie précédente, l'« impression de naturel » que peut ressentir quelqu'un lors d'une diffusion sonore immersive, mais aussi le sentiment « d'être là », directement lié à la notion d'immersion. Nous savons également que les attributs d'un espace immersif sont l'enveloppement et la présence. L'auditeur est alors "plongé" dans l'espace sonore dont il ressent les vibrations, les mouvements, il peut s'y mouvoir, le toucher... Ainsi, être à l'intérieur d'un dispositif de haut-parleurs projetant un espace sonore immersif, c'est littéralement *être dedans*, comme être dans l'eau ou dans un avion en plein vol. On utilise d'ailleurs souvent les expressions "être plongé", "être enveloppé", "être immergé", qui rappellent le milieu aquatique. Dans le dictionnaire⁹⁷, on trouve la définition suivante du terme "immersion" : « action d'immerger, de plonger dans un liquide ; résultat de cette action ». Dans un espace immersif, le son n'est donc pas seulement autour de l'auditeur mais partout. Nous venons de faire une analogie avec l'eau, mais celle d'un avion en plein vol, que nous empruntons à [James & Hope, 2013]⁹⁸, est tout aussi pertinente : lorsque nous sommes assis dans la cabine, on perçoit et on ressent toutes les vibrations sonores des moteurs, des hélices, des matériaux, ... on ne peut pas s'en échapper, si l'on se déplace dans l'avion, on entend toujours ces sons puisque nous sommes *dedans* ; nous sommes contenus dans l'espace physique de l'avion délimité par ses parois, nous sommes également contenu dans un espace sonore. On peut comprendre mieux pourquoi les systèmes 3D ne nous font pas forcément sentir immergés dans le son car, comme l'ont dit Lynch et Sazdov [2011], une caractéristique de cette configuration de haut-parleurs est l'« engulfment », c'est-à-dire être *recouvert* par les sons. Or l'immersion, c'est le fait que l'espace sonore dans lequel on se trouve soit *rempli* par le son, comme le fait d'être dans de l'eau qui se trouve elle-même partout, qui nous transperce presque le corps. C'est pourquoi l'enveloppement de l'auditeur n'est pas toujours une raison suffisante pour créer de l'immersion. On peut être enveloppé par les sons – ou recouvert, avec le système 3D – et sentir pourtant que le son est éloigné, distant [Guastavino & Katz, 2004], alors que la présence est un « attribut immersif » [Rumsey, 2001]. Il s'agit d'un aparté, mais

97 Rey-Debove, J. et Rey, A., 2003, *Le Petit Robert*, Éditions Le Robert.

98 « One could consider the auditory scene on an aircraft: it is immersive in a sense that the entire scene we experience is in fact the aircraft itself resonating and vibrating, and this sound does not have one direction, we experience it as circumspace. » [James & Hope, 2013]. Ici, le terme « circumspace » est emprunté à Denis Smalley, il désigne un espace qui est partout, autour de l'auditeur. Pour lire d'avantage à ce propos : Denis Smalley, Interview with Davies, 15 October 2006, « Rêverie », southern Quebec.

nous pouvons dire ainsi que l'emploi de l'expression « surround » pour certains dispositifs de haut-parleurs ne signifie pas que le dispositif en question est immersif, puisque l'on sous-entend uniquement le fait d'*entourer* l'auditeur... Le terme « fullness » est à ce propos une autre expression qui se rapporte à la notion d'immersion. On pourrait le traduire, en français, par l'*ampleur*, l'*entièrete*, ou le *remplissage*. Il est, lui aussi, très employé par les participants lors d'expériences⁹⁹ sur la perception de l'espace sonore, réalisées avec des dispositifs multicanaux comportant des haut-parleurs sur les côtés et à l'arrière, pour décrire ce qu'ils ressentent [Rumsey, 2001].

Ce ressenti très *physique* induit par l'immersion sonore, ces analogies avec l'eau qui touche et même traverse la peau, ou la cabine d'un avion en vol, peuvent nous faire supposer que les vibrations sont caractéristiques du *milieu* sonore résultant. Le son est un phénomène acoustique, c'est un déplacement d'onde dans l'air. Il faut être *là*, être présent, pour percevoir le son avec les oreilles et le cerveau, et que le corps ressente aussi le son via des vibrations produites par les fréquences les plus basses. En effet, le corps agit comme un récepteur puisque c'est le système nerveux qui envoie les informations au cerveau. Il peut donc être considéré comme un "deuxième cerveau". « The second component rests on the low frequency parts of early lateral reflections and leads to a greater depth of the auditory events up to the impression of being enveloped. »^{100 101} La partie basse-fréquence de la réverbération d'un événement sonore est donc responsable de l'enveloppement de l'auditeur dans une salle ou un environnement, mais aussi les fréquences graves composant le spectre sonore d'une source directe. Comme nous l'avons vu en partie II.1.2, si ces fréquences sont décorréliées, l'auditeur percevra une sensation d'immersion. De plus, à fort volume de diffusion, la sensation d'enveloppement est renforcée mais aussi les fréquences graves sont ressenties par l'auditeur sous forme de vibrations qui traversent le corps. C'est d'ailleurs de cette façon que les personnes sourdes ont la possibilité de percevoir la musique : par les vibrations – c'est le cas, par exemple, de la percussionniste Evelyn Glennie, sourde depuis l'âge de 12 ans. Placé au sein d'une diffusion sonore via un dispositif multicanal, l'auditeur est plongé dans un véritable *milieu sonore immersif* qu'il entend, ressent, qu'il peut presque toucher, et dans lequel il peut se déplacer. Pour ces raisons, avec une écoute au casque, même en binaural, on ne peut pas parler d'expérience sonore immersive *totale* : le corps de l'auditeur est absent du

99 Nakayama & al., 1971, *Subjective assessment of multichannel reproduction*, J. Audio. Eng. Soc.

100 « Le second composant à considérer repose sur les fréquences graves des réflexions latérales précoces qui apportent de la profondeur aux événements sonores, jusqu'à l'impression d'être enveloppé. »

101 Blauert, J., *Hearing Of Music In Three Spatial Dimensions*, Mis-en-ligne le 30 Mars 2015 sur : https://www.researchgate.net/profile/Jens_Blauert [consulté le 12/10/2017]

milieu sonore¹⁰². Pour pallier ce problème, on pourrait associer l'écoute au casque en binaural à une diffusion de la partie basse du spectre via des subwoofers extérieurs, ou bien à un fauteuil vibrant¹⁰³, sur lequel l'auditeur serait allongé. Les vibrations ressenties ne seraient pas "naturelles" puisqu'elles ne proviendraient pas de la source sonore directement, mais pourraient être un moyen de le simuler.

II.2.3 L'espace sonore immersif comme moyen d'expression et de communication

« [...] it is clear that hearing contributes to a wide range of experiences and functions. Hearing, together with its active complement, listening, is a means by which we sense the events of life, aurally visualize spatial geometry, propagate cultural symbols, stimulate emotions, communicate aural information, experience the movement of time, build social relationships, and retain a memory of experiences. To a significant but under-appreciated degree, aural architecture influences all of these functions. »¹⁰⁴ [Blessner & Salter, 2009]

Les espaces sonores peuvent nous faire ressentir des émotions, influencer notre humeur. Ils sont comme des stimulus auxquels on réagit inconsciemment [Blessner & Salter, 2009]. Blessner et Salter expliquent en effet dans leur livre [2009] que les espaces que l'on perçoit avec le son renvoient différentes impressions (un espace "froid", "chaud", ...), nous indiquent si le lieu est public, privé, etc. Les auteurs donnent l'exemple d'un restaurant chaleureux et bien décoré au premier abord, qui s'avère au final très bruyant (avec la résonance des couverts et des conversations) et nous procure du stress et de l'anxiété. Ils font émerger la notion d'« architectes du sonore » – « Aural Architects » – que nous avons déjà évoquée en partie I.2.2, qui désigne des personnes « qui se concentrent sur la façon avec

102 cf. échanges avec Eric Lyon en ANNEXE F.

103 La société Aurasens développe des fauteuils vibrants pour que l'auditeur, allongé dessus, écoute de la musique au casque ressente en même temps les vibrations spécialement composées. Voici un extrait de la présentation du projet : « Le projet AURASENS consiste à vivre une expérience musicale par l'ouïe et par le toucher. L'équipe a développé un fauteuil équipé d'un réseau de vibreurs intégrés, afin d'écouter de la musique tout en la ressentant simultanément avec son corps tout entier. » Site internet : <http://www.aurasens.com/>

104 « Il est clair que le fait d'entendre contribue à un large éventail d'expériences et de fonctions. Entendre, avec son complément actif, écouter, est un moyen par lequel nous percevons les événements de la vie, nous visualisons de façon auditive la géométrie spatiale, nous propageons des symboles culturels, nous stimulons nos émotions, nous communiquons des informations auditives, expérimentons le mouvement du temps, construisons des relations sociales, et gardons en mémoire nos expériences. De façon importante mais pourtant sous-estimée, l'architecture auditive influence toutes ces fonctions. »

laquelle les auditeurs expérimentent l'espace (acoustique culturelle). »¹⁰⁵ Ces « architectes du sonore » sont capables, par le biais de l'espace sonore, d'influencer l'humeur de l'auditeur, de stimuler ses émotions, de générer des images dans sa tête, de propager des symboles culturels, ... L'auditeur est alors en quelque sorte immergé "mentalement" dans cet espace sonore, puisque ses émotions et son esprit sont sollicités. Les espaces sonores influent également beaucoup sur notre rapport à la société : « Rehabilitation workers often report that blindness is less socially and emotionally burdensome than deafness. »¹⁰⁶ [Blessner & Salter, 2009]. Il s'agit d'un moyen de communiquer, de s'exprimer [Ekeberg Henriksen, 2002]. "L'immersion mentale", par le biais des espaces sonores, existe bien, mais peut-on dire qu'une musique électronique spatialisée, ressentie comme immersive, peut aussi influencer notre humeur, nos émotions ?

« In this context, music has traditionally constituted an item *in* a place. Over the last two decades, development of "spatial music" has been within the prevailing engineering paradigm, informed by psychophysical data ; here, space is an abstract, Euclidean 3-dimensional 'container' for events. The emotional consequence of spatial arrangements is not the main focus in this approach. »¹⁰⁷ [Lennox, 2009]

Le contexte dont parle Lennox est le fait de baser la pensée de l'espace en musique sur celle des scientifiques : la musique est donc *dans* un espace, un espace contient des événements sonores, mais l'on ne tient pas vraiment compte des conséquences émotionnelles d'un tel espace créé. D'après lui, il faut une approche écologique pour considérer la musique spatialisée comme un environnement et prendre en compte la façon dont elle agit sur nos émotions¹⁰⁸. « [...] from 'music-in-a-place' to 'music-as-a-place' »¹⁰⁹ [Lennox, 2009]. La musique spatialisée, à partir du moment où elle est comprise comme un environnement, peut donc influencer nos émotions. C'est alors le cas d'une musique électronique spatialisée qui, diffusée, crée un espace sonore immersif pour toutes les raisons que nous avons pu voir jusque maintenant. On peut donc affirmer qu'avec la musique électronique spatialisée, on

105 Texte original : « [they] focus on the way that listeners experience the space (cultural acoustics). »

106 « Les travailleurs en réinsertion rapportent souvent que le fait d'être non-voyant est moins pénible socialement et émotionnellement que le fait d'être mal-entendant. »

107 « Dans ce contexte, la musique a traditionnellement constitué un objet *dans* un lieu. Au cours des deux dernières décennies, le développement de la "musique spatiale" a été inséré dans le paradigme dominant de l'ingénierie, informé par des données psycho-physiques ; ici, l'espace est un "contenant" abstrait et euclidien à trois dimensions pour des "événements". La conséquence émotionnelle des arrangements spatiaux n'est pas l'objectif principal de cette approche. »

108 « An ecological approach to music-as-environment intrinsically treats the emotional consequences of spatio-musical arrangement holistically. » [Lennox, 2009].

109 « [...] de la musique contenue dans un lieu à la musique qui *est* le lieu. »

pourrait arriver à créer un espace immersif *total* car "naturel", qui stimule le corps de l'auditeur mais aussi son esprit... Franck Ekeberg Henriksen, dans sa thèse [2002], donne divers exemples qui nous semblent très pertinents sur la façon dont la spatialisation des sons peut provoquer, sur l'auditeur, différentes réactions, des sauts d'humeurs, ... et qui se rapportent tout à fait à la spatialisation des sons en musique. Il parle notamment du mouvement, de la vitesse à laquelle se déplace un son, de si celui-ci nous semble proche ou éloigné : « Sounds approaching from the distance, for example, lead to different psychological reactions compared with sounds that are moving into the distance or passing by. Especially when sounds are experienced as violating the body territory, a strong reaction may be provoked. Speed of movement is an influential factor, particularly when sounds are moving towards the listener. Sounds approaching at a high speed can provoke human emergency reactions, such as fear and the mobilisation for fight or flight. A similar reaction can occur when facing sounds of great magnitude or sounds that refer to dangerous, threatening or uncomfortable situations. »¹¹⁰. L'auteur montre également comment chaque auditeur peut interpréter différemment un même espace sonore selon son passé, ses expériences, son "bagage socio-culturel", ... : « Particularly when working in surround sound, spatial cues known from our cultural and natural environments can be utilised in electroacoustic works with a high degree of accuracy and realism. The combination of referential features of the sound material and spatial behaviour in composed space may lead to spatial experiences that are extra vivid due to familiarity of the sound environment gained from real-world experience. Visualisation is key in this regard. Spaces and sounds that are known from personal real-life experience are much easier to visualise than unfamiliar material, and relationships among known spaces are easier to understand. However, relying on such cultural knowledge as spatio-structural functions in a work will almost certainly ensure a great variety of listener interpretations. »¹¹¹ [Ekeberg Henriksen, 2002].

110 « Par exemple, les sons venant de loin conduisent à des réactions psychologiques différentes par rapport aux sons qui se déplacent dans le lointain ou qui passent. Surtout quand les sons sont perçus comme transgressant les limites imposées par le territoire du corps, une réaction forte peut être provoquée. La vitesse du mouvement est un facteur influent, en particulier lorsque les sons se déplacent vers l'auditeur. Les sons qui approchent à une vitesse élevée peuvent provoquer des réactions humaines d'urgence, telles que la peur et la préparation pour le combat ou la fuite. Une réaction similaire peut se produire lorsque l'on fait face à des sons de grande magnitude ou à des sons qui se réfèrent à des situations dangereuses, menaçantes ou inconfortables. »

111 « Particulièrement lorsque l'on travaille en son surround, les indices spatiaux connus provenant de nos environnements culturels et naturels peuvent être utilisés dans des œuvres électroacoustiques avec un degré élevé de précision et de réalisme. La combinaison des caractéristiques référentielles du matériau sonore et du comportement spatial dans l'espace composé peut mener à des expériences spatiales qui sont plus vives en raison de la familiarité de l'environnement sonore acquise à partir de l'expérience du monde réel. À cet égard, la visualisation est la clé. Les espaces et les sons qui sont connus à partir de l'expérience personnelle sont beaucoup plus faciles à visualiser que les matériaux non familiers, et les relations entre les espaces connus sont plus faciles à comprendre. Cependant, s'appuyer sur de telles connaissances culturelles pour en faire des

II.3. Conclusion

Tout au long de ce chapitre, nous avons vu que la sensation d'immersion, lors d'une diffusion sonore spatialisée, peut être obtenue grâce aux réflexions tardives de la réverbération dans une salle de concert, à la décorrélation interaurale du contenu basses-fréquences, via diverses techniques de spatialisation du son faisant partie du processus de composition de l'oeuvre ou employées lors d'une interprétation acousmatique ; les critères essentiels pour qu'un espace sonore soit ressenti comme immersif étant que le dispositif de diffusion soit multicanal, comportant obligatoirement des haut-parleurs sur les côtés et à l'arrière [Rumsey, 2001] [Guastavino & Katz, 2004] afin qu'une sensation à la fois d'*enveloppement* et de *présence* de cet espace sonore soit éprouvée.

La notion d'immersion, appliquée aux musiques électroniques, n'est alors pas rattachée à un format de diffusion (5.1, octophonie, ...) , ni à une technique de spatialisation, elle ne semble pas non plus dépendante d'une technique de reproduction du champ sonore¹¹² (VBAP, Ambisonie, WFS, ...) ¹¹³. Il est important de noter que le son peut être *enveloppant*, il peut y avoir *immersion*, sans que l'oeuvre diffusée ne soit nécessairement multicanale dès l'étape de composition. L'oeuvre peut très bien avoir été composée en stéréo – voire mono – et l'interprétation acousmatique, ou bien simplement le fait de diffuser la stéréo (ou mono) sur plusieurs points sur un système multicanal, peut amener au même résultat. Bien entendu, le contenu sonore, ainsi que les habitudes d'écoute de l'auditeur jouent beaucoup sur la perception auditive de l'espace. *Pour l'instant*¹¹⁴, les configurations 2D sont, dans l'ensemble, préférées par les auditeurs afin de percevoir le caractère immersif de l'espace sonore [Guastavino & Katz, 2004], mais cela peut être amené à évoluer au fil du temps et des expériences de chacun.

Enfin, les termes ou expressions : « enveloppement », « impression de naturel » ou

fonctions spatio-structurelles dans une oeuvre engendrera presque certainement une grande variété d'interprétations par les auditeurs. »

112 Nous reviendrons sur ce point en partie III, en considérant des cas plus pratiques, comme celui du concert de musiques électroniques spatialisées.

113 Se référer à [Nilles, 2016] pour plus détails sur ces techniques de reproduction du champ sonore et leurs applications.

114 Nous insistons sur l'expression « pour l'instant » car, comme nous l'avons déjà dit, la perception auditive et le ressenti des gens sont très influencés par leurs habitudes d'écoute, qui elles-mêmes varient selon les époques.

« naturalness », « spaciousness », « fullness », « être là », ou encore « circumspace »¹¹⁵, se rapportent tous de près ou de loin à l'immersion sonore. Le sentiment « d'être là » n'est pas en rapport avec la précision – l'acuité – spatiale, le côté "naturel" de l'espace sonore immersif ainsi créé renvoie à un environnement sonore dit "naturel". Nous pouvons alors rapprocher les notions d'immersion et d'expérience. En effet, le corps de l'auditeur est traversé par les ondes sonores, il ressent des vibrations, il peut se déplacer dans le son. Percevoir un espace sonore immersif en écoutant de la musique électronique peut être une expérience pour l'auditeur puisqu'il stimule ses émotions, lui renvoie des images, presque tous les sens de l'auditeur sont en éveil ; le côté exceptionnel de cet événement peut, en quelque sorte, "marquer" l'auditeur.

C'est l'ensemble de tous les points que nous venons de rappeler qui définit la notion d'immersion dans le cas de la musique électronique spatialisée. L'immersion mène inéluctablement à l'expérience, puisqu'à la fois notre mécanisme de la perception auditive, notre corps, nos émotions, notre mémoire, ... sont stimulés par le son. Une expérience sonore immersive, c'est alors être plongé dans un bain de sons. Cependant, il est important de noter qu'il est rare que l'immersion soit *totale* ou *globale*, c'est-à-dire que l'ensemble de ces points soit effectif au même moment. C'est aussi le côté psychologique de l'auditeur qui joue sur le fait de ressentir ou non de l'immersion : par exemple le fait de visualiser le système de haut-parleurs, de se trouver dans le noir, de se tenir assis ou debout, de partager un moment avec d'autres, etc. ; tout cela agit et influe sur notre ressenti immersif. On peut très bien sentir de l'immersion avec une diffusion stéréophonique si on aime cette musique, qu'on la joue assez fort, que l'on se sent enveloppé, ... Un concert immersif, c'est aussi un concert qui parvient à capter l'attention de l'auditeur, comme un lecteur peut être "absorbé" par un roman. *Immersion* et *concentration* sont donc tout à fait associables.

115 Denis Smalley, Interview with Davies, 15 October 2006, « Rêverie », southern Quebec.

III. Le concert spatialisé vécu comme une expérience par l'auditeur. Le cas pratique du concert *Immersion Sonore*.

« There is a considerable contextual difference between solitary listening in a relaxed environment at home and making the effort to go to a concert hall to listen to music in the presence of other people. In addition to differences in sound system and acoustics between home and concert hall, the social situation of listening in a public space affects the listener's receptivity to spatial information and perception of the communicational aspects of space in the music. The architectural layout and seating arrangement of the concert hall and the social rules for the concert situation are important factors here. »¹¹⁶ [Ekeberg Henriksen, 2002]

On observe trois facteurs importants liés à la situation de concert spatialisé : le dispositif de haut-parleurs, la place de l'auditeur (et de l'artiste) et les règles sociales imposées par cette situation de concert [Ekeberg Henriksen, 2002]. Comment conjuguer ces trois paramètres essentiels ? Quelle configuration haut-parleurs – public choisir afin de diminuer l'effet « sweet-spot » et ainsi avoir une meilleure perception auditive de l'espace ? L'objectif ici est de comprendre s'il est possible de créer une expérience à la fois musicale, spatiale et sociale pour l'auditeur par le biais du concert de musique électronique spatialisée, une relation de partage entre les gens, et de stimuler leur écoute et donc leur attention.

Le concert *Immersion Sonore* va nous servir de cas pratique, d'essai, de tout ce que nous avons vu jusqu'à présent, aussi bien d'un point de vue compositionnel : les compositeurs y participant ont pu travailler certaines techniques de spatialisation avec différentes

¹¹⁶ « Il y a une différence contextuelle considérable entre le fait d'écouter seul dans un environnement calme, à la maison, et de faire l'effort d'aller dans une salle de concert pour écouter de la musique avec la présence des autres personnes. En plus des différences liées à la sonorisation et à l'acoustique entre la maison et le concert, la situation sociale de l'écoute dans un espace public affecte la réceptivité de l'auditeur à l'information et à la perception spatiale des aspects communicationnels de l'espace dans la musique. La disposition architecturale et le placement des sièges dans la salle de concert et les règles sociales imposées par la situation de concert sont des facteurs importants ici. »

technologies de reproduction de l'espace, que d'un point de vue perceptif : le public, via ses réponses au questionnaire mis à disposition, ou de part ses retours de vive-voix, nous a communiqué ses impressions quant à la perception auditive de l'espace et la notion d'immersion. Aussi, le choix de présenter divers types de musiques électroniques à un public hétérogène, au sein d'un même concert, ajoutait une difficulté supplémentaire. En effet, chaque type de musique demande une écoute différente, peut-être même une concentration différente de la part de l'auditeur. Est-ce possible alors de respecter les divers types d'écoute que requiert chaque genre ? Est-ce que la spatialisation du son est un moyen de rassembler des publics divers et de capter leur attention et ainsi, d'agir sur l'écoute du temps ? Cet événement nous a permis d'obtenir certaines réponses mais surtout, a soulevé d'autres questions.

Dans une première partie, nous présenterons le concert *Immersion Sonore*, les difficultés rencontrées, les problèmes résolus et non résolus, d'un point de vue technique, compositionnel et perceptif. Dans une deuxième partie, nous verrons les questions soulevées suite au concert, et les différents points importants à considérer dans un concert de musique électronique spatialisée. Nous allons parler d'Ambisonie et de WFS, déjà explicitées dans [Nilles, 2016] que nous invitons le lecteur à consulter.

III.1. Le cas pratique du concert spatialisé *Immersion Sonore*

III.1.1 Présentation du concert : le déroulement, les aspects techniques

La première édition de *Immersion Sonore* s'est déroulée le 7 Décembre 2017 dans l'Auditorium de la Maison des Sciences de l'Homme Paris Nord (MSH), à Saint-Denis. Ce concert spatialisé, d'une durée d'1h30 environ, a rassemblé divers types de musiques électroniques autour de la thématique de l'espace, et de l'immersion du public. Il a été réalisé dans un cadre étudiant, avec l'aide de l'Université Paris VIII, du CICM (Centre de recherche Informatique et Création Musicale), de Frank Gillardeaux, enseignant en son à l'ENS Louis Lumière, de Quentin Nivromont, réalisateur en informatique multimédia, et de la société

Euphonia qui commercialise les processeurs Wave Field Synthesis de Sonic Emotion¹¹⁷. Un questionnaire papier (également disponible en ligne) était distribué pour connaître l'avis des spectateurs sur le concert et sur leur rapport à l'immersion, adressé aussi bien à des personnes novices qu'à des personnes expérimentées en son spatialisé (cf. ANNEXE A).

En effet, la programmation était voulue diversifiée, pour mélanger les publics ; sur les 6 sets présentés, nous avons : une pièce acousmatique quadriphonique de Jérémie Nicolas¹¹⁸, une pièce mixte de Jean-François Ducher¹¹⁹ pour saxophone ténor et électronique en temps réel, spatialisée en Ambisonie avec la bibliothèque HOA¹²⁰ sur Max/MSP, un extrait d'un poème sonore¹²¹ de Léonore Mercier et Arthur H, mixé en WFS Sonic Emotion par Hervé Déjardin à Radio France et projeté tel quel lors du concert, un live-électronique expérimental de capob spatialisé en Ambisonie avec le Spat¹²² de l'Ircam, un live-électronique-jazz expérimental d'Amélie Nilles, spatialisé à la fois en Ambisonie avec la bibliothèque HOA mais aussi en WFS Sonic Emotion¹²³, et enfin le live-électronique ambient expérimental de Matthieu Ruben, spatialisé en Ambisonie avec la bibliothèque HOA sur Max/MSP (cf. ANNEXE B). Les artistes avaient la possibilité, les semaines précédant le concert, de faire des essais dans le studio du CICM sur une demi-sphère de 16 haut-parleurs. Pour le concert dans l'Auditorium, nous avons installé 21 haut-parleurs au total : 6 à l'arrière, 4 sur chaque côté, et 5 à l'avant plus 2 subwoofers de part et d'autre de l'espace scène (cf. la modélisation de cette configuration dans l'Auditorium en ANNEXE C, *Figure I*).

Le choix de cette configuration n'était pas anodin : l'Auditorium possédait déjà les 4 haut-parleurs installés de chaque côté du public, les 3 haut-parleurs frontaux et les 2 subwoofers. Nous avons d'ailleurs l'habitude, avec le CICM, d'organiser des concerts de musiques mixtes et acousmatiques sur ce système, en utilisant la bibliothèque HOA et en corrigeant les angles des haut-parleurs pour simuler un cercle parfait, ce qui fonctionne très bien. Puisque nous connaissions la WFS Sonic Emotion et que nous l'apprécions pour son rendu "naturel", nous avons eu envie de la tester en configuration 360° sur un large auditoire.

117 Site-web de Sonic Emotion : <http://www2.sonicemotion.com/>

118 Lien pour écouter la réduction binaurale de *Sons Seuls* de Jérémie Nicolas : <http://jeremienicolas.com/compositions/sons-seuls-acousmatique/>

119 Lien pour écouter la réduction stéréophonique de *Cinq souffles de Tchernobyl* de Jean-François Ducher : <https://soundcloud.com/jfducher/cinq-souffles-de-tchernobyl-carl-emmanuel-fisbach-live-dec-2017>

120 Conçue en partenariat avec le CICM et l'Université Paris 8, par des enseignants-chercheurs du CICM et des étudiants de Paris 8. La bibliothèque HOA est libre d'accès et disponible sur le site : <http://www.mshparisnord.fr/hoalibrary/>. Elle est pensée flexible pour les musiciens [Guillot, 2013]).

121 Lien pour écouter les réductions binaurale et 5.1 de *Cauchemar Merveilleux* de Léonore Mercier et Arthur H : <http://hyperradio.radiofrance.fr/son-3d/le-merveilleux-cauchemar-sonore-darthur-h/>

122 Site-web: <http://forumnet.ircam.fr/fr/produit/spat/> ou encore le Spat Flux qui vient de sortir (janvier 2018) : <https://fluxhome.com/project/spat-revolution/>

123 Voir la partie III.1.2 pour plus de détails sur ces choix de techniques de reproduction du champ sonore.

Grâce à la société Euphonia, nous avons donc pu bénéficier du processeur WFS de Sonic Emotion. La Wave Field Synthesis (WFS)¹²⁴ est basée sur la simulation de la propagation d'une onde sonore dans l'air. Il faut généralement un très grand nombre de haut-parleurs pour employer cette technique de reproduction du champ sonore¹²⁵. Bien que Sonic Emotion développe des processeurs WFS capables de travailler avec un nombre très raisonnable de haut-parleurs (minimum 4 enceintes en frontal) – pour ce faire, la société se base sur la perception humaine de l'espace, et non sur la théorie acoustique pure –, nous avons tout de même ajouté une ligne de 6 haut-parleurs à l'arrière, et 2 haut-parleurs à l'avant, en renforcement – car « plus on a de haut-parleurs, mieux c'est ! », dixit Arnault Damien, acousticien chez Euphonia. Ainsi, un large éventail de configurations et de techniques de reproduction du champ sonore s'offrait aux artistes : ils pouvaient choisir de créer une pièce, ou de réaliser un live spatialisé, en utilisant 4 haut-parleurs, 8 haut-parleurs ou plus, et employer la WFS Sonic Emotion, l'Ambisonie, etc. Outre le poème sonore de Léonore Mercier, déjà mixé en WFS Sonic Emotion à Radio France, j'étais la seule à travailler *avec* la WFS – et non à l'utiliser uniquement pour la diffusion¹²⁶ – et, bien entendu, également *avec* HOA. Jean-François Ducher était le seul à diffuser sur l'octophonie d'origine, avec HOA, sans passer par le processeur Wave1. Deux autres compositeurs employaient l'Ambisonie, mais les haut-parleurs étaient placés dans l'espace virtuel du processeur WFS (cf. ANNEXE D, *Figure1*, pour visualiser cette configuration dans l'interface WavePerformer). Les 4 haut-parleurs nécessaires à la pièce de Jérémie Nicolas – spatialisée avec le plugin de Reaper¹²⁷ Reasurround¹²⁸ – étaient également virtuels et passaient par le processeur Wave1.

La programmation de ce concert était donc riche et variée en terme de genre musical, mais aussi en terme de choix de spatialisation et, par conséquent, de configurations techniques. Le synoptique de l'installation du concert, réalisé par Frank Gillardeaux, est disponible en ANNEXE E, pour mieux visualiser le dispositif. Chaque artiste live ayant besoin de son ordinateur sur scène, la pièce mixte de Jean-François Ducher n'étant pas diffusée via le processeur Wave1 mais directement via l'octophonie "circulaire", il nous fallait

124 Ouvrages de référence conseillés pour en apprendre plus sur la WFS : Berkhout, A.J., 1988, *A Holographic Approach to Acoustic Control*. Journal of the Audio Engineering Society, 36(12), pages 977–95 et Berkhout, A. J., 1998, *A Wavefield Approach to Multichannel Sound*, Proceedings of the 104th Convention of the Audio Engineering Society, Preprint 4749.

125 L'Espace de Projection de l'Ircam comptait 264 haut-parleurs régulièrement répartis sur une couronne horizontale pour la WFS.

126 Voir la partie III.1.2 de ce présent mémoire dans lequel j'explique justement mes choix de techniques de reproduction du champ sonore pour ce concert.

127 Site-web : <http://www.reaper.fm/>

128 Plugin directement intégré au séquenceur audio Reaper (depuis 2011). Il permet la spatialisation en 2D et 3D et peut gérer jusqu'à 64 canaux en entrées et sorties.

passer d'une configuration à une autre sans problème, le plus rapidement et facilement possible, sans avoir besoin de brancher/débrancher des câbles entre chaque set. La solution qu'il nous a paru la plus simple était d'utiliser le réseau DANTE¹²⁹, déjà installé dans la salle. Il nous a permis de relier tous les canaux sans trop de problèmes, ni de branchements supplémentaires. Chaque ordinateur disposait de la carte son DANTE Virtual Sound, nous permettant ainsi de relier les ordinateurs de scène à l'ordinateur principal de la régie-gradins (placée dans le public), via un câble Éthernet. Celui-ci disposait du logiciel WavePerformer. Après une conversion en MADI¹³⁰ via la carte son MADIface, la régie-gradins renvoyait les informations de spatialisation au processeur Wave1, située dans la régie-cabine en haut des gradins.

Si nous souhaitons réitérer ce type d'événements, dans différents lieux, avec des contraintes de temps, il est important de considérer sérieusement les aspects techniques et de trouver des solutions pour obtenir la plus grande fluidité possible. Nous avons eu, par exemple, qu'un très court moment pour régler les enceintes et leur réponse fréquentielle dans la salle, ce qui a impacté le rendu final¹³¹. Aussi, lorsque l'on présente des artistes différents, avec des choix de spatialisation différents, il serait judicieux d'avoir un seul ordinateur, pour gérer la spatialisation de chacun : il ferait office de "spatialiseur" ou "spatialisateur"... Nous y reviendrons en partie III.2.1.

III.1.2 Ma participation au concert en tant qu'artiste : mes choix de spatialisation

En plus d'être en charge de l'organisation du concert, j'ai souhaité y participer en tant qu'artiste pour proposer un live intitulé « Laïka project », qui est l'adaptation en live spatialisé à 360 degrés d'un projet personnel studio mélangeant musiques électroacoustiques et jazz. Quentin Nivromont a travaillé avec moi pour mettre en place techniquement ce live. J'ai présenté en effet deux standards de jazz entièrement ré-arrangés : *Strange Fruit*¹³² et *Don't*

129 DANTE (Digital Audio Network Through Ethernet) a été développé par Audinate Pdy. Ltd. Il s'agit d'un système combinant matériel informatique et logiciels, utilisé pour transporter de l'audio numérique multicanal sur un réseau IP standard ou Ethernet, jusqu'à 512 canaux avec un câble Éthernet.

130 Le processeur Wave1 fonctionne en MADI.

131 Quentin Nivromont, avec Les Ateliers des Lutheries Numériques, vient justement (Mai 2018) de développer un outil permettant d'effectuer la mesure de la réponse fréquentielle d'une enceinte de façon très rapide. Site-internet : <https://www.ateliersdeslutheriesnumeriques.com/>

132 Écrit et composé par Abel Meeropol. Lien vers ma version (stéréo) : <https://www.youtube.com/watch?v=cix202h6IG0>

*Explain*¹³³, que j'interprète également au chant. Le matériau sonore est essentiellement composé de sons captés de l'environnement, d'instruments et de voix parlée ou chantée, le tout ayant subi des processus de transformation sonore. Le défi principal était de « mettre en espace » des morceaux travaillés en studio, en stéréophonie, en essayant de ne pas tomber dans le piège de l'espace comme moyen d'ornementation¹³⁴.

Bien que ces morceaux aient été composés en stéréo, leur possible adaptation pour un live spatialisé a toujours été pour moi une finalité. J'ai donc depuis longtemps en tête de travailler, le moment venu, l'espace sonore pour ces morceaux ; la composition en studio avait donc été effectuée dans le but qu'une future spatialisation multicanale soit "aisément" réalisable. Lorsque la possibilité du live s'est présentée, chaque morceau possédant un environnement sonore différent, j'ai eu la volonté de créer un espace sonore à caractère immersif, propre à chacun d'eux.

Ainsi, j'ai imaginé, pour le morceau *Strange Fruit*, que la voix soit frontale et localisée, présente mais lointaine à la fois, avec une réverbération assez longue. Les instruments (la basse, l'orgue et la grosse-caisse) sont également positionnés frontalement, la basse étant projetée de part et d'autre de l'espace scène, les deux canaux légèrement décorrélés pour tenter d'intensifier le sentiment d'immersion. Les éléments rythmiques qui arrivent en seconde partie sont soit précisément localisés, soit en mouvement, mais en veillant à ne pas "brouiller" le rythme. En version stéréo, c'est déjà les nappes sonores et le fourmillement de petits éléments qui créent une atmosphère particulière. Dans la version live, l'accumulation de toutes ces textures spatialisées, au fil du morceau, participe au côté immersif de l'espace créé.

Le morceau *Don't Explain* est plus intimiste. La voix est plus proche ; les sons, le rythme, semblent "partir" de la voix. J'ai donc imaginé, contrairement au morceau précédent, une voix *immersive*, qui serait partout, de plus en plus proche de l'auditeur ("comme s'il ne pouvait s'en débarrasser"). Ainsi, j'ai tenté de créer un environnement sonore avec la voix. Le reste des sons et instruments seraient simplement localisés, ou effectueraient des trajectoires simples décidées par rapport à leurs caractéristiques sonores. L'important est qu'ils semblent "jaillir" de cet espace-voix, qu'ils soient *contenus* dans cet espace-voix.

Pour cela, je disposais de plusieurs possibilités que j'avais déjà expérimentées auparavant avec des pièces acousmatiques ou des lives¹³⁵ :

133 Écrit et composé par Billie Holiday et Arthur Herzog Jr. Lien vers ma version (stéréo) : <https://soundcloud.com/amelie-nilles/dont-explain-demo>

134 Je ne sais pas si j'ai réussi à remplir cette mission. Le résultat était plutôt satisfaisant, mais pas encore assez abouti. Cela demande encore du temps de travail.

135 Voir la partie II de [Nilles, 2016] pour plus d'applications sur mon emploi de ces techniques. Il y a bien entendu une multitude de possibilités pour spatialiser une œuvre, autres que celles présentées. Je parle uniquement de celles que j'utilise le plus fréquemment.

- travailler en VBAP¹³⁶, spatialiser mes sources avec le plugin Reasurround dans Reaper;
- travailler avec l'Ambisonie (High Order Ambisonics) dans Max/Msp avec la Bibliothèque HOA ;
- travailler avec la WFS de Sonic Emotion dans le logiciel WavePerformer.

Nous l'avons vu dans [Nilles, 2016], on peut limiter notre pensée de l'espace si l'on se contraint à n'utiliser qu'une seule technologie. Comme pour tout instrument, il est essentiel de garder en tête que l'on peut tout à fait mélanger les techniques de composition et les technologies de reproduction du champ sonore, selon nos besoins, car elles ont chacune des propriétés intéressantes qui peuvent être utiles selon nos intentions de composition [Baalman, 2010]. Ainsi, pour mettre en application ce que j'avais en tête, j'ai choisi d'utiliser à la fois l'Ambisonie, avec la Bibliothèque HOA pour Max/Msp, et la WFS Sonic Emotion. L'usage de l'Ambisonie et de la Wave Field Synthesis au sein d'une même œuvre permet de tirer partie de ces technologies de reproduction de l'espace en fonction de leurs caractéristiques propres : par exemple, je vais privilégier la WFS pour localiser précisément des sons tout en gardant une écoute dite "naturelle" (écoute du son et non du haut-parleur), tandis que je vais préférer l'Ambisonie pour jouer avec les composantes-mêmes de l'espace sonore et créer des champs diffus (grâce aux fonctionnalités de la Bibliothèque HOA). Aussi, j'ai employé diverses techniques de spatialisation telles que la spatialisation timbrale et la décorrélation micro-temporelle qui permettent d'obtenir des espaces immersifs [Nilles, 2016], et la localisation.

Spatialisation de *Strange Fruit*

Avant le début du morceau, j'ai inséré une sorte d'introduction électroacoustique stéréo, pour habituer petit à petit l'auditeur à l'atmosphère de *Strange Fruit*, et pour qu'il sente se construire un espace immersif autour de lui. Je choisis donc de partir de la stéréo et de faire grossir l'espace sonore petit à petit (c'est également ce qu'il se passe au niveau des couches de sons qui s'accumulent) pour qu'il commence à s'approcher des auditeurs et à les envelopper. À la fin de l'introduction, le matériau sonore étant relativement riche, je le découpe alors en bandes de fréquences et les disperse dans l'espace de la salle¹³⁷. À ce moment, toutes les

¹³⁶ Voir [Nilles, 2016]. Le lecteur souhaitant plus d'informations sur la technologie VBAP est invité à se référer à : Pulkki, V., 2001, *Spatial Sound Generation And Perception By Amplitude Panning Techniques*, Thèse de Doctorat, Helsinki University of Technology.

¹³⁷ Technique de spatialisation timbrale.

sources sonores sont spatialisées avec le logiciel WavePerformer, y compris la stéréo de départ sous forme de haut-parleurs virtuels.

Le morceau *Strange Fruit* apparaît peu à peu en *fade* pendant l'introduction, plutôt frontalement. Je procède ensuite de la même façon qu'avec l'introduction : je réalise un élargissement progressif, par les côtés, les sons arrivent également par l'arrière et entourent le spectateur. Globalement, je choisis la WFS Sonic Emotion pour localiser précisément les sons et effectuer des trajectoires. J'utilise la Bibliothèque HOA pour travailler avec le module de décorrélation sur certains sons importants, présents dès le début du morceau (cf. ANNEXE D, *Figure3*). Les harmoniques circulaires vont être décorréliées de plus en plus, petit à petit, pour tendre vers un champ diffus vers le milieu de la pièce. Je modifie, pour cela, la valeur du "samples delay" au fur et à mesure de l'avancée du morceau, le facteur de diffusion restant fixe, à 1. Par exemple, le son principal de craquements de branches, que l'on entend clairement au début, devient par conséquent de plus en plus diffus. L'effet est intéressant car, dans la salle de concert, on avait presque l'impression que le son se trouvait au-dessus de nos têtes, oppressant. J'ai également dirigé un son de type "nappe" dans l'harmonique 0 pour que celui-ci provienne de tous les haut-parleurs.

Spatialisation de *Don't Explain*

Pour obtenir une voix *immersive*, qui crée un espace, j'ai tout d'abord pensé à la spatialisation timbrale sur les harmoniques circulaires avec la Bibliothèque HOA. Mais cela ne fonctionnait pas, le spectre de la voix n'étant pas assez riche sur toute son étendue, la voix sonnait trop "téléphone". J'ai donc opté de nouveau pour la décorrélation micro-temporelle, ce qui a très bien fonctionné : le facteur de diffusion restait à 1, seul la valeur du "samples delay" variait, afin de se rapprocher ou de s'éloigner du champ diffus, et créer ainsi un sentiment de contraction/dilatation entre les A et les B¹³⁸. Nous obtenions donc une voix peu claire, dans laquelle le public était immergé. Les autres sons gravitaient autour en étant localisés précisément ou en se déplaçant à l'aide du WavePerformer. J'ai de nouveau eu recours à la technique de spatialisation timbrale sur un son de scie transformé, dont l'énergie spectrale était très étendue, afin de renforcer cette idée d'espace immersif, englobant. De même que précédemment, les bandes de fréquences étaient disposées dans l'espace, par l'intermédiaire du logiciel WavePerformer. Les sons de types nappes, pour ne pas perturber l'espace créé par la voix, étaient envoyés dans l'harmonique 0 grâce à la Bibliothèque HOA.

¹³⁸ Sachant que la forme du morceau est : AABABA.

Dans les deux morceaux, je ne souhaitais pas attribuer de trajectoires aux instruments, mais plutôt les localiser de façon précise ; cela me semblait plus juste, les gens étant habitués à voir et entendre les instruments à l'avant. Globalement, on obtenait alors des espaces tantôt contractés, tantôt étendus, englobants, tantôt mouvants... Un autre objectif était de créer un espace immersif et *organique*¹³⁹, où les sons seraient comme "vivants", palpables...

D'un point de vue technique, il faut noter que Quentin Nivromont a réalisé un objet Max for Live permettant de lier les modules HOA à Live Ableton 9 (cf. ANNEXE D, *Figure4*). Nivromont a d'ailleurs développé des modules *hoa.Live*¹⁴⁰, compatibles avec Live 10, pour ne pas avoir à faire ce pont entre les deux logiciels.

Le passage du studio à la salle de concert

Outre un travail au casque, en binaural, à la maison, j'avais la chance de pouvoir répéter également au studio du CICM sur une couronne de 8 haut-parleurs. J'avais, dans le studio, un très bon rendu, dû à une bonne disposition des haut-parleurs, un bon traitement acoustique de la salle, mais aussi de part l' "intimité" que l'on entretient dans ce cas avec les sons. Le passage du studio intimiste au grand Auditorium, avec 21 haut-parleurs, n'a pas été fait que de bonnes surprises, bien que je connaisse la salle (la scène, et donc les haut-parleurs avant, sont contenus dans une sorte de renforcement), mais aussi mes outils qui sont HOA et la WFS de Sonic Emotion. Le défi, à ce stade, est de réussir à s'adapter à ce changement notable d'espace.

Le problème le plus important a été que l'ensemble des sons sonnait *dry* dans la grande salle, l'acoustique n'étant pas suffisante pour lier le tout. Chaque son, indépendamment, ne possédait pas de réverbération – sauf si celle-ci faisait partie du timbre du son – pour ne pas obtenir une accumulation de réverbérations différentes et brouiller le tout. Or, dans le processeur WFS final, il n'y a pas de moyen pour ajouter une réverbération d'ensemble. Ne pouvant compter sur l'acoustique du lieu, j'ai dû ajouter un peu de réverbération via la Bibliothèque HOA, mais cela ne concernait pas tous les sons puisqu'ils n'étaient pas tous traités en Ambisonie. Et surtout, les sons utilisés dans la Bibliothèque HOA servaient à créer

139 Horacio Vaggione, à propos de la décorrélation micro-temporelle : « Cette absence d'absolu quant à la simultanéité des événements musicaux n'est pas cependant quelque chose de gênant ; au contraire, nous pouvons tirer bénéfice de ce fait pour créer des sons « vivants », comportant des mouvements divers et complexes. » Vaggione, H., 2002, *Décorrélation microtemporelle, morphologies et figurations spatiale*, Actes des Journées d'Informatique musicale (JIM 2002), Marseille, France.

140 Disponibles sur le site-web des Ateliers des Lutheries Numériques :

<https://www.ateliersdeslutheriesnumeriques.com/>

un champ diffus, et l'ajout d'une réverbération n'était pas forcément nécessaire à cela. Il a donc fallu ajouter un peu de réverbération sur certaines sources évoluant dans l'espace du processeur WFS, toujours en prenant le temps d'écouter le résultat pour ne pas détériorer le rendu sonore. De plus, nous étions également limités par des problèmes de CPU.

Dans le studio, le sentiment de proche, d'intimité, ne mettait pas en évidence ce problème, au contraire : le tout sonnait "juste bien". Mais dans une grande salle comme celle de l'Auditorium, on manquait de ce "son de salle" auquel on est habitué et qui lie le tout lorsqu'il est maîtrisé, ou brouille l'ensemble dans le cas contraire. J'étais donc confrontée au problème inverse de celui auquel je m'attendais : là où souvent il y a trop de réverbération en concert, ici, je n'en avais justement pas assez. La question du portage d'une œuvre spatialisée d'une configuration de haut-parleurs à une autre, d'un lieu à un autre, est très intéressante, et nous pensons que ceci est un vrai sujet lorsque l'on s'intéresse à la spatialisation. Nous reviendrons dessus en partie III.2.1.

III.1.3 Les retours du public : réponses au questionnaire et commentaires

Afin de mieux cerner le public présent à ce concert, outre les étudiants et enseignants-chercheurs du Département Musique de l'Université Paris VIII qui ont plutôt l'habitude de ce type d'évènement, nous avons proposé un petit questionnaire papier au public, à remplir pendant ou après le concert, également disponible en ligne pour plus de facilité. Il y a eu 26 participants sur environ 120 personnes présentes au concert, soit 17% au total, ce qui est relativement peu. Leur position dans la salle étant demandée, nous pouvons dire que l'ensemble des répondants représente tout de même un bon échantillon, car ils étaient placés de façon dispersée dans la salle (à l'avant, à l'arrière, au milieu, sur les côtés, ...) (cf. ANNEXE C, *Figure2*). Beaucoup d'étudiants et enseignants-chercheurs présents au concert nous ont communiqué leurs impressions de vive-voix, ce qui explique certainement le peu de réponses au questionnaire. De même, sur les 26 répondants, 67% se présentait comme totalement novice en son spatialisé (même si certains d'entre eux étaient musiciens), le reste étant expérimenté voire même professionnel dans le domaine. Les questions posées étaient générales afin d'être comprises par tout un chacun¹⁴¹. Elles portaient sur la sensation ou non

¹⁴¹ NB : Les questions posées étaient peut-être même trop "simples", ce qui a freiné certains professionnels du son spatialisé, ne sachant finalement que répondre. Ceci explique également pourquoi la majorité des répondants étaient novices en la matière.

d'immersion des auditeurs, sur la cohérence entre la spatialisation et la musique présentée, sur la comparaison entre un concert "classique" (diffusion stéréophonique) et ce type de concert multicanal, mais aussi sur la différence entre la réalisation finale de ce concert *Immersion Sonore* et leurs attentes de départ (cf. questionnaire en ANNEXE A pour voir l'ensemble des questions). Nous n'allons pas regarder ici en détail les réponses aux questions posées, mais plutôt choisir les remarques récurrentes ou bien qui nous paraissent pertinentes et qui peuvent apporter des éléments de réponses concernant le point de vue des auditeurs sur la perception d'un concert de musiques électroniques spatialisées ou, au contraire, soulever d'autres questions. Il est important de préciser que le public était assis, orienté face à l'espace-scène. Lors des lives, les artistes se tenaient donc sur scène, face au public.

Globalement, on peut dire que les spectateurs ont apprécié le concert et l'ont même préféré à un concert plus "classique" avec diffusion stéréo. Ils ont trouvé l'ensemble spatialisation et musique tout à fait cohérent (65% des répondants), les novices n'ont pas ressenti de "gêne" avec la perception d'un espace sonore musical.

Nous nous sommes souvent demandé si l'emploi de la spatialisation était pertinent pour tout style de musique électronique ([Nilles, 2016]). Si la composition de l'espace sonore a un "sens" en musiques électroacoustiques (acousmatiques et mixtes), en a-t-elle un dans les autres genres de musiques électroniques ? Ne serait-elle pas "réduite" à un effet d'ornementation ? Chaque type de musique demande une écoute différente, peut-être une concentration différente. Comment respecter les divers types d'écoute que requiert chaque genre ? Nous avons donc décidé, au sein d'un même concert, de tester différents genres de musiques électroniques, spatialisées de différentes façons, et nous avons invité les auditeurs à réagir à ce propos. J'ai moi-même, avec le « Laïka project », présenté un projet qui pourrait paraître "ambigu", teinté de musiques acousmatiques, jazz et pop, à la fois chanté, spatialisé... C'était un énorme défi pour ma part, comme je l'ai déjà évoqué, et un point d'interrogation quant à la réception devant un public diversifié. En général, le projet a été apprécié – la présence de la voix chantée y étant sûrement pour quelque chose du point de vue des auditeurs novices à la fois en spatialisation sonore, mais aussi vis-à-vis de l'écoute de musiques dites "expérimentales". Généralement, les utilisations de la voix dans les œuvres proposées, avec la spatialisation, ont plu aux auditeurs : « L'emploi de la voix est pertinent d'une manière générale », ainsi que le rythme. D'après l'ensemble des commentaires positifs, les techniques de spatialisation du son pourraient donc être employées sans problème pour tout type de musiques électronique. Un professionnel du son 3D note à ce propos : « Le

schéma de la musique tonale mérite d'être aussi exploré via la spatialisation. ». Certains auditeurs parlent de « vraie plus-value par rapport à un concert agencé de manière plus classique », d'autres soulignent l'importance de la création d'un espace sonore « à des fins artistiques » et semblent y avoir trouver leur compte. Certaines remarques montrent cependant que le défi n'est pas encore tout à fait relevé et qu'il y a beaucoup de travail en perspective pour mettre à nouveau en place un événement qui plairait à la fois au plus grand nombre, tout en étant pertinent en ce qui concerne l'usage des techniques de spatialisation : « Peut-être un manque global de sens donné à la spatialisation. Souvent, on ressent que la musique cherche à démontrer le procédé et non que le procédé sert la musique. » Il y a, en effet, toujours un équilibre à trouver, pour les compositeurs, entre création de mouvements et emploi de trajectoires précises pouvant sembler peu "naturelles", création d'un champ diffus, ou simple élargissement de l'espace, ... Au vu de l'ensemble des commentaires, on peut dire que la question de la cohérence entre genre musical et spatialisation est subjective, c'est à l'auditeur de se faire sa propre idée. On sait que 77% des participants pensent que la spatialisation du son les a "aidés" à apprécier un contenu musical qu'ils n'ont peut-être pas l'habitude d'écouter autrement.

Une autre question essentielle dans ce type de concerts est celle du rapport à l'image avec la spatialisation sonore. Nous avons volontairement souhaité une scénographie épurée pour que le public se concentre sur le son et non sur ce qu'il voit, car nous pensons que le visuel influence l'écoute (cf. partie I.2.1). Mais peut-être l'influence-t-il de façon positive et que le manque de repères visuels agit négativement sur l'écoute de l'auditeur ? Seulement deux commentaires réclament « une relation à l'image visuelle » ou bien « une pièce avec du mapping vidéo afin de voir/entendre un spectacle total, même si il est vrai que fermer les yeux pendant les pièces me permettait d'avoir une expérience enrichie (entre autres avec "visions" de fractales). » Même s'il a peut-être manqué un rapport marqué entre son, espace et images, cette scénographie épurée a vraisemblablement réussi à générer des images personnelles, ce qui est très positif car il s'agissait de l'un des buts recherchés. En revanche, il est important de remarquer que 19% des répondants se sont plaints d'un manque de centre : « surround très bien. Pas convaincu par la "face". », « certaines compositions étaient trop latéralisées et manquaient de centre », « Très latéral, pas assez de centre. ». Serait-ce la faute des compositeurs qui auraient délaissé l'avant pour privilégier le reste ? Un mauvais paramétrage entre le dispositif et l'acoustique de la salle ? Tout est possible. Mais il faut noter que l'ensemble de ces commentaires proviennent tous de personnes se présentant comme

"novices" dans le son spatialisé. Ils ont donc l'habitude d'assister à des concerts, ou d'écouter de la musique avec un dispositif stéréo. Ils sont également fortement influençables : le fait de *voir* un artiste sur la scène porte inévitablement leur attention vers l'avant. De plus, cinq haut-parleurs avant sur les sept au total (si l'on compte les deux subwoofers) étaient encastrés dans le mur en fond de scène donc totalement cachés à la vue des spectateurs ! Seuls les deux haut-parleurs que nous avons ajouté de chaque côté de la scène étaient visibles. Ainsi, ne voyant quasiment pas de haut-parleurs avant – à part les deux sur scène – comparé aux treize autres, sur les côtés et à l'arrière, qui étaient complètement visibles, l'écoute des auditeurs était influencée, ce qui explique certainement pourquoi ils percevaient beaucoup plus le son comme provenant des côtés et de l'arrière. Ce commentaire : « Aurait mérité d'ajouter des enceintes sur scène. » vient justement appuyer le propos. On constate alors le "pouvoir" du visuel sur le son, chez les novices. C'est tout à fait ce qu'avait jadis observé Rumsey [2001], et nous en avons déjà parlé en partie I.2.1 de ce présent mémoire : lors d'une écoute binaurale, sans aucun stimuli visuel, les gens ont tendance à placer les sons à l'arrière de la tête. Et ceci démontre l'importance de la vision dans la localisation. Au quotidien, on localise l'arrière uniquement au son puisque nous n'avons pas d'yeux derrière la tête ! Nous faisons donc la même analogie lorsque nous ne visualisons pas les sources sonores car nous faisons appel à notre mémoire et donc à notre expérience (cf. partie I.2.). Mais Rumsey ajoute que cela est plutôt le cas avec les personnes n'ayant pas l'habitude d'écouter du contenu audio spatialisé. « [...] it is interesting to note that most people, when played binaural recordings of sound scenes without accompanying visual information or any form of head tracking, localise the scene primarily behind them rather than in front. [...] This may be because one is used to using the hearing sense to localise things where they cannot be seen, and that if something cannot be seen, it is likely to be behind. » [Rumsey, 2001]¹⁴² Il faut tout de même nuancer notre propos puisque le public, lors du concert Immersion Sonore, avait la possibilité de bouger la tête pour résoudre certains conflits de perception. Il s'agit donc certainement à la fois de problèmes dans la composition de l'espace pour les compositeurs, mais aussi de l'influence de la vision chez les novices sur leur perception spatiale, sans oublier leurs habitudes d'écoute de la musique en général qui sont majoritairement frontales. Ces retours nous permettent de poursuivre nos réflexions avant d'organiser un nouveau concert de ce type. L'emprise du visuel, chez les novices, qui trompe leur perception auditive est belle et bien

142 « [...] il est intéressant de noter que la plupart des gens, lors d'une écoute en binaural de scènes sonores non-accompagnées d'informations visuelles ou d'aucune forme de trackeurs, localisent la scène surtout derrière eux plutôt devant. [...] Cela peut être dû au fait que nous sommes habitués à utiliser le sens auditif pour localiser les choses que nous ne pouvons pas voir, et dans ce cas, il est plus probable qu'il s'agisse de quelque chose qui se trouve à l'arrière. »

réelle et il est essentiel de ne pas négliger ce point. Nous pouvons également envisager une autre raison à ces déséquilibres dans la perception de l'espace lors du concert : le problème du « sweet-spot », sujet qui sera abordé en partie III.2.2.

III.2 Réflexions sur le cas général du concert de musiques électroniques spatialisées, suite à l'évènement *Immersion Sonore*

III.2.1 De la nécessité de pouvoir adapter son œuvre spatialisée à tout type de systèmes de diffusion et de salles

Parmi les nombreuses questions soulevées en préparant l'évènement *Immersion Sonore* mais aussi suite au concert, il y a notamment celle du portage des œuvres diffusées ou jouées en live sur un système différent, dans une salle différente. Plusieurs solutions sont possibles, avec leurs points positifs et leurs contraintes. Nous allons également évoquer la question du "up/down-mixing".

Les orchestres de haut-parleurs et la « ré-orchestration spatiale »

Nous en avons déjà discuté dans [Nilles, 2016], si l'on considère une oeuvre stéréo ou multipiste diffusée sur un Acousmonium (ou orchestre de haut-parleurs), c'est l'interprète acousmatique, à la console, qui va permettre l'adaptation de l'oeuvre sur le dispositif de haut-parleurs, dans la salle donnée. L'interprète acousmatique est généralement expérimenté, il connaît l'oeuvre (s'il ne s'agit pas du compositeur lui-même) et l'Acousmonium sur lequel il "joue", il peut être habitué à la salle également ; il est comme un musicien qui interprète une oeuvre dans différentes salles de concert, il doit être attentif à comment sonne son instrument dans le nouveau lieu et adapter sa façon de jouer. Cependant, comme pour une pièce instrumentale, l'interprète acousmatique se place entre le compositeur et son oeuvre – à moins que ce ne soit le compositeur lui-même qui interprète sa propre pièce mais là encore, le point négatif peut être son manque d'expérimentation dans ce domaine –, il y a donc un facteur

"humain" qui entre en jeu : le compositeur n'est pas entièrement maître du résultat. Stefani et Mooney parlent, dans leur article [2009], du module Resound qui se présente comme une sorte de console "intelligente" (qui associe logiciel et interface hardware) qui permet de faire exactement le même travail que l'interprète acousmatique, en utilisant beaucoup moins de faders. En effet, il est possible de programmer – *mapper* – les entrées/sorties des faders, de réaliser une spatialisation « semi-automatisée » via l'enregistrement, au préalable, de mouvements par exemple. Il s'agit donc d'une console plus simple à utiliser, et surtout qui ne se limite pas à la dextérité de l'interprète ou du compositeur. Pour les mêmes raisons que celles citées plus haut, Resound permet donc au compositeur d'adapter son œuvre au dispositif et à la salle de concert, « tout en respectant le travail multicanal »¹⁴³ [Stefani & Mooney, 2009] déjà composé.

Le compositeur-chercheur Eric Lyon parle de « ré-orchestration spatiale » [2014] lorsqu'il s'agit de porter sa pièce *Spirits* sur trois dispositifs de haut-parleurs différents – trois dômes – , sans perdre l'essence même de la pièce, bien que celle-ci repose sur sa spatialisation. En effet, il a conçu sa pièce avec le Klangdom (ZKM) constitué de 43 haut-parleurs (les plus bas étant situés au niveau de la tête des auditeurs) et 4 subwoofers au sol, puis il l'a rejouée au SARC Sonic Lab¹⁴⁴ possédant 42 haut-parleurs disposés tout à fait différemment par rapport au Klangdom (les haut-parleurs les plus bas étant situés sous la tête des auditeurs), et enfin sur le système LSU iCAST¹⁴⁵, présentant 24 haut-parleurs additionnés de 24 sources virtuelles diffusées via 4 haut-parleurs hexagonales positionnés au sol de façon centrale. Il emploie l'expression « ré-orchestration spatiale » tout à fait dans le sens d'une orchestration instrumentale. Il considère en effet le dispositif de haut-parleurs (pour cette pièce : un dôme) comme un véritable orchestre de haut-parleurs puisque son processus de composition de l'espace repose sur le regroupement des enceintes selon leur position dans l'espace et leur réponse fréquentielle, exactement comme les différents regroupements d'instruments dans l'orchestre. On observe donc une pensée qui se rapproche de celle de l'Acousmonium ici, sauf qu'il s'agit de dômes. Pour le compositeur, les trois différentes versions de *Spirits* sont des « ré-orchestrations spatiales » qui conservent ses volontés initiales de composition de l'espace ; il s'agit donc de la même pièce (cf. ANNEXE F pour lire l'échange de mails réalisé avec Eric Lyon). D'après lui, ce ne serait pas le cas s'il avait dû adapter sa pièce à un dispositif de 8 haut-parleurs circulaires (2D), car l'élévation est essentielle et serait alors complètement annihilée. Enfin, la question du temps que doit prévoir

143 Texte original : « [...] with respect to multi-channel work. » [Stefani & Mooney, 2009]

144 Queen's University, Belfast.

145 Université de l'État de Louisiane.

un compositeur pour réaliser ce nouvel agencement spatial (qui est une forme d' "up/down-mixing") est bien entendu très importante. Lyon dit à ce propos : « Following the experience of spatial orchestration for SARC, preparing a new 48-channel version for LSU took one full day. The striking decrease in the time required to prepare a new spatial orchestration indicates that, although still labor-intensive, the process of spatially re-orchestration a multichannel work for large numbers of speakers will become more streamlined and efficient with practice. »¹⁴⁶ [Lyon, 2014]. Ainsi, le compositeur pense que le travail de portage de l'oeuvre peut devenir de plus en plus rapide avec le temps, pour une même pièce, car l'on acquiert des sortes d'automatismes. C'est finalement ce à quoi j'ai dû me confronter pour adapter mes morceaux stéréo au système de haut-parleurs du concert *Immersion Sonore*. Pour ma part, il s'agissait plutôt d'un *élargissement*, d'un *démasquage*, plutôt que d'une réelle « ré-orchestration spatiale » ; mais aussi un moyen d'appuyer la dramaturgie de la pièce et d'aller au bout de mes idées de composition, la notion d'espace sonore étant déjà au cœur de ma pensée à l'étape du studio. Cependant, il faut noter qu'il s'agit ici d'une oeuvre électroacoustique, et donc d'une musique dite "de création" dans le cas d'Eric Lyon, composée et diffusée dans un cadre de recherche. On peut également considérer qu'il s'agit d'une expérience similaire pour ma part puisque je n'avais pas d'énormes contraintes de temps. Lorsque l'on souhaite organiser des événements tels *Immersion Sonore*, qui prennent place dans des lieux toujours différents, avec des musiques électroniques de tout genre et donc avec des artistes dont l'approche vis-à-vis de la spatialisation est variée, il serait peut-être plus judicieux d'avoir la possibilité d'adapter ses morceaux en très peu de temps, voire de façon immédiate (sans compter les modifications dues à l'acoustique du lieu).

Encodage – décodage : l'Ambisonie

Les techniques ambisoniques sont caractérisées par les étapes d'encodage dans le domaine des harmoniques (circulaires ou sphériques) et de décodage ; celui-ci va dépendre du système employé. Le choix du dispositif de diffusion est alors très flexible, bien qu'il soit préférable que les haut-parleurs soient placés au minimum en forme de cercle (ou arc-de-cercle) ou de sphère (ou demie-sphère / dôme). La Bibliothèque HOA intègre des modules permettant un décodage optimal selon le dispositif de haut-parleurs utilisé, ainsi que des

146 « Suite à l'expérience de l'orchestration spatiale pour le SARC, la préparation d'une nouvelle version de 48 canaux pour la LSU a pris une journée entière. La diminution notable du temps requis pour préparer une nouvelle orchestration spatiale indique que, bien que nécessitant encore énormément de travail, le processus de ré-organisation spatiale d'un travail multicanal pour un grand nombre de haut-parleurs deviendra plus simple et efficace avec la pratique. »

manipulations du champ sonore applicables au moment de l'encodage (décorrélation des harmoniques, granulation, convolution, ...). Jean-François Ducher, pour sa pièce *Cinq souffles de Tchernobyl*, a utilisé la Bibliothèque HOA pour réaliser la spatialisation. Il a ensuite décodé le contenu audio pour 8 haut-parleurs en corrigeant les angles¹⁴⁷ si nécessaire. S'il joue de nouveau sa pièce sur un système circulaire de 16 haut-parleurs par exemple, il pourra très facilement décoder sa pièce pour ce nombre de haut-parleurs, de même pour 4 haut-parleurs, etc. Travailler dès le studio en Ambisonie s'avère donc être un choix judicieux en ce qui concerne le portage de l'oeuvre sur divers systèmes par la suite. Le compositeur et musicologue Raphaël Néron-Baribeau témoigne : « Finalement, c'est aussi beaucoup l'aspect compatibilité qui a motivé le choix d'utiliser cette méthode. La possibilité d'exporter une pièce vers tous les types de formats est un avantage particulier à l'Ambisonie qui peut s'avérer extrêmement pratique. En effet, durant tout le processus de composition de la bande, j'ai alterné entre mon studio personnel, qui était configuré en quadraphonie, et celui octophonique de l'Université de Montréal. Un système d'encodeurs et décodeurs en Ambisonie de troisième ordre, créé avec les plugiciels B2X, m'a permis d'adapter les configurations de sortie en fonction du lieu où je travaillais, pour ultimement procéder au rendu final en octophonie. »¹⁴⁸ Le fait d'avoir des encodeurs / décodeurs portables, sous forme de plugins, logiciels ou patches Max, est un atout majeur en ce qui concerne la mobilité.

Le « mixage objet »

Les interfaces développées pour la technologie WFS sont pensées pour la pratique de ce que l'on appelle le « mixage objet » ; c'est-à-dire que l'on ne va pas envoyer un même nombre de pistes (ou canaux) vers un même nombre de haut-parleurs, mais l'on va manipuler directement chaque source (ou objet) sonore dans l'espace de façon individuelle (ou en groupe), indépendamment du nombre de haut-parleurs. Ce type d'interfaces repose sur la visualisation des sources sonores (représentées graphiquement dans l'espace). C'est ce que propose Sonic Emotion avec son logiciel WavePerformer. Ainsi nous avons pu, pour le concert *Immersion Sonore*, diffuser sans problèmes des oeuvres quadraphoniques ou encore octophoniques sur un dispositif de 21 haut-parleurs. De même, si j'utilise de nouveau le processeur WFS de Sonic Emotion, je pourrais jouer mes morceaux spatialisés sur n'importe

147 Angle correspondant à une déviation de l'orientation d'un haut-parleur par rapport à son orientation idéale pour constituer un cercle, par exemple.

148 Néron-Baribeau, R., 2014, *Méthodes de spatialisation sonore et intégration dans le processus de composition*, Mémoire de Master 2, Université de Montréal.

quel système de diffusion – même uniquement sur un banc de haut-parleurs frontaux ! L'inconvénient est qu'il faudrait toujours transporter le processeur avec moi, ou bien jouer uniquement dans des salles disposant de ce processeur¹⁴⁹. Actuellement, des étudiants, chercheurs et RIM sont en train de mettre en place des processeurs WFS virtuels, opensource, qui permettraient de pallier ce problème. On peut noter que la Bibliothèque HOA, gratuite, dispose de l'objet *hoa.map* qui rend également possible le mixage objet.

Enfin, l'outil idéal permettant à la fois le portage facile d'une œuvre sur divers systèmes, mais aussi le travail de l'espace, dès le processus de composition, avec différentes techniques de reproduction du champ sonore, le tout ne nécessitant qu'un ordinateur, serait le très connu Spat de l'Ircam. Il consommerait, en revanche, beaucoup de CPU¹⁵⁰. Ainsi vient l'idée de consacrer un ordinateur à la gestion de la spatialisation lors des concerts – un "spatialiseur" ou "spatialisateur". C'est ce que tentent de lancer plusieurs entreprises dont la marque d'enceintes Amadeus avec le système HOLOPHONIX¹⁵¹ développé, entre autres, par Thierry Coduys. L'objectif est d'installer, dans les salles de concert, un dispositif "prêt-à-emploi" pour faire des concerts spatialisés, comprenant : haut-parleurs (Amadeus), réseau¹⁵², et module de spatialisation (en collaboration avec l'Ircam) laissant la possibilité à l'artiste d'utiliser, en même temps, diverses technologies de spatialisation comme la WFS, HOA, VBAP, etc. Il possède également un moteur de réverbération algorithmique multicanal.

Écoute domestique d'oeuvres spatialisées : toucher un plus large auditoire

Pour Schmele [2011], un moyen d'éviter tout problème de portage est les systèmes mobiles : le compositeur joue toujours sur le même dispositif qu'il connaît bien, et l'auditeur n'a pas besoin de se déplacer dans un endroit particulier pour pouvoir écouter de la musique spatialisée : « Concerning the accessibility and availability of spatial music today, the concept of mobile sound dispersion systems and technology forms an alternative to the establishment of standardized spaces around the world. Instead of installing multiple systems in fixed places, these systems, like BEAST, can reach the listener instead of requiring the audience to travel to them. »¹⁵³ [Schmele, 2011]. On peut citer également « The Game Of Life

149 C'est le cas, par exemple, de l'Institut du Monde Arabe (Paris), du Théâtre de Chaillot (Paris), de La Marbrerie (Montreuil), du ZKM (Allemagne), ...

150 Mais je ne vais pas m'étendre sur ce point ne l'ayant pour ma part encore jamais utilisé en situation de concert.

151 Site-web: <http://holophonix.xyz/>

152 DANTE mais compatible si besoin avec d'autres formats : MADI, RAVENNA ou AES67. Il permet la gestion des 128 entrées et 128 sorties, extensibles à 256 ou 384.

153 « Concernant l'accessibilité et la disponibilité de la musique spatiale de nos jours, le concept des systèmes de dispersions des sons mobiles et la technologie forment une alternative à l'établissement d'espaces standardisés

Foundation »¹⁵⁴ qui a établi un système mobile utilisant la WFS.

Nous pensons, de notre point de vue, que c'est justement la diversité des lieux, des systèmes et des techniques de reproduction du champ sonore qui fait que le travail de l'espace en musiques électroniques est si intéressant, à la fois pour le compositeur qui possède, à chaque nouvelle pièce ou à chaque concert, un nouveau défi à relever, et pour l'auditeur qui ne peut se lasser puisqu'il vit, à chacun de ces événements, une nouvelle découverte auditive riche et inouïe. C'est pourquoi il est important de discuter de la question du portage des œuvres spatialisées, car il est essentiel de préserver cette non-standardisation que certaines entreprises souhaiteraient pourtant imposer. Une solution pour l'écoute domestique est le format binaural, qui permet de conserver une écoute spatialisée au casque, bien qu'il manque la possibilité de se mouvoir dans l'espace, ou même simplement de bouger la tête pour résoudre quelques imprécisions. Mais si une œuvre est composée pour un système particulier comme le Klangdom du ZKM, ou encore pour un lieu précis, la réduction binaurale ne fonctionnera pas : on perdra la majorité des caractéristiques spatiales de l'œuvre. Pour ce type de pièces, il est certainement préférable de ne pas avoir de mixages binauraux ou stéréo. Ainsi la diffusion de ces œuvres conçues pour un lieu ou un système précis devient un moment inédit, une *expérience*, qui peut donner l'envie au public de venir assister au concert pour le côté *unique* de l'évènement. Pour le reste, réaliser un "down-mix" binaural ou stéréo d'une œuvre normalement conçue pour un autre format peut permettre de toucher un plus large public et lui donner envie de venir aux concerts spatialisés pour découvrir la version originale. Comme le disait Stockhausen : « [...] Il vaut mieux voir une photographie d'une sculpture que ne rien voir ; l'envie de la connaître dans son original naîtra peut être ainsi. »¹⁵⁵

III.2.2 Repenser l'organisation globale du concert spatialisé : la place de l'artiste, la place de l'auditeur

Avec le concert *Immersion Sonore*, nous avons pu observer, lors d'un cas *pratique*, l'influence de la vue sur la perception auditive de l'espace (cf. partie III.1.3), que nous avons déjà évoquée de façon plus *théorique* en partie I.2.1. Elle peut venir en effet perturber la

dans le monde. Au lieu d'installer de multiples systèmes à des endroits fixes, ces systèmes, comme le BEAST, peuvent atteindre directement l'auditeur au lieu d'avoir besoin que l'audience voyage jusqu'à eux. »

154 En savoir plus sur The Game Of Life Foundation et son système : <http://gameoflife.nl/en/>

155 Stockhausen cité chez Menezes, F., 1998, *La spatialité dans la musique électroacoustique*, dans Chouvel, J.M. et Solomos, M., 1998, *L'espace : Musique/ Philosophie*, l'Harmattan, Paris.

perception de l'auditeur, et ce dernier recherche parfois une certaine cohérence entre image et son, qu'il ne retrouve pas forcément lors d'un concert de musique spatialisée. Dans l'Auditorium de la MSH Paris-Nord, l'espace-scène étant placé à l'avant, face aux gradins, les sièges orientés vers lui, nous avons donc jugé "logique" de placer les artistes live sur la scène lors de leur set. Nous avons envisagé les placer dans les gradins, mais le fait que les sièges soient fixes nous en a dissuadés. Nous pouvons nous demander également si les différences de perception entre les auditeurs viennent de leur placement dans la salle. De nombreuses questions se posent alors : le public doit-il *voir* l'artiste, ou en tout cas se trouver face à lui ? Cela pourrait perturber son écoute car dirigerait son attention entièrement sur lui. Les incohérences entre son spatialisé et image fixe, à l'avant, pourraient le gêner. Il est néanmoins essentiel que la personne qui contrôle l'espace ou les sons déjà spatialisés, en direct, soit placée dans le cercle de haut-parleurs. L'artiste doit-il alors se trouver obligatoirement au centre du système ? Le public doit-il forcément être positionné face à la scène, ou bien de façon à *voir* l'artiste ? Nous ne prétendons pas apporter des réponses exactes à ces questions que se posent déjà de nombreux artistes et musicologues depuis plusieurs décennies, mais plutôt à poursuivre la réflexion, le sujet étant particulièrement important lorsque l'on considère un concert de musique spatialisée.

Lors des concerts de musiques électroacoustiques sur Acousmonium, l'interprète acousmatique est placé au centre du dispositif, parmi les spectateurs. Ceux-ci étant généralement orientés vers l'avant, les personnes positionnées derrière lui peuvent le voir, les autres non. Le système lui-même est pensé et installé en fonction de cette disposition des spectateurs et de l'interprète acousmatique à la console. La musique diffusée n'ayant pas de causes visuelles, les auditeurs n'ont finalement "rien à voir", ils peuvent choisir de fermer les yeux ou non, regarder les gestes de l'interprète. La configuration de ce type de concerts fonctionne plutôt bien. Notre réflexion est principalement portée sur les cas de concerts spatialisés avec des artistes live, des instrumentistes, ... Une solution souvent adoptée¹⁵⁶ est de placer l'audience en cercle, tout autour de l'artiste qui se trouve au milieu : les sièges sont en quelque sorte groupés, puis placés de part et d'autre du centre, un peu face à face, et donc orientés vers le milieu. Le dispositif de haut-parleurs est alors installé tout autour du public, en cercle, ou bien les enceintes sont réparties autour mais aussi à l'intérieur (plutôt dans le cas d'un Acousmonium). « This has serious implications for the perception of spatio-structural elements in the music. With the audience facing in different directions, what is in front for

156 Par exemple par le Festival Présences électronique organisé par le GRM.

some listeners becomes behind for others and, with a circular seating arrangement, what is a side-to-side movement for some listeners is a front-to-back movement for others. The social and metaphorical significance of certain directions of placement and movement of sound material can no longer be relied upon as structural means. »¹⁵⁷ [Ekeberg Henriksen, 2002]

L'auteur soulève ici les problèmes dus à l'orientation des spectateurs dans ce cas de figure, puisque ceux-ci sont chacun orienté de façon différente (par groupe de personnes) et donc ne perçoivent pas la même chose. Ainsi, les choix « spatio-structurels » du compositeur ne sont pas compris par l'ensemble des auditeurs, dû au positionnement de certains qui sont dans la "mauvaise direction", mais aussi dû à nos capacités de perception de l'espace. Par exemple, nous n'avons pas la même facilité à percevoir l'espace à l'arrière de la tête qu'à l'avant (cf. partie I) et cela peut s'avérer être "injuste" pour certains auditeurs, car ils ne vont pas tous vivre l'expérience de façon optimale, du point de vue de la perception de l'espace sonore. Doit-on privilégier la vision de l'artiste, peut-être au détriment de la perception auditive ? Ou bien l'inverse ? On pourrait dire que la précision spatiale ne devrait pas être l'enjeu lorsque l'on a ce type de placement. Il faudrait plutôt privilégier un contenu *immersif*, que l'auditeur peut percevoir, *ressentir*, peu importe son placement, et non des détails précis dont la compréhension dépendrait de l'endroit où l'on est placé.

Le « sweet-spot »

Lorsque le public est en configuration assise avec les haut-parleurs disposés tout autour, la perception de l'espace dépend bien de l'endroit où l'auditeur est placé. Les remarques que nous avons obtenues, via le questionnaire, suite au concert *Immersion Sonore*, à propos du manque de centre, de trop de latéralisation, ou bien d'un manque d'arrière (cf. partie III.1.3), provenaient d'auditeurs assis à des sièges plutôt excentrés du point central de l'espace de diffusion¹⁵⁸. On peut remarquer également ce commentaire : « Malheureusement pour être vraiment immergé dans le son, il faut se placer au milieu de la salle. Ce n'était pas mon cas. ». La place idéale se trouverait donc au milieu du dispositif ; en dehors de celle-ci, l'écoute serait perturbée et le sentiment d'immersion serait même fortement diminué. On

157 « Cela a de sérieuses implications quant à la perception des éléments spatio-structurels dans la musique. Le public se faisant face dans des directions différentes, ce qui est à l'avant pour certains auditeurs est à l'arrière pour les autres et, avec une disposition des sièges circulaire, ce qui apparaît comme un mouvement latéral pour certains auditeurs, est un mouvement avant-arrière pour les autres. La signification sociale et métaphorique de certaines directions de placement et de mouvement du matériel sonore ne peut plus être invoquée comme un moyen de structure de l'oeuvre. »

158 Se référer à l'ANNEXE C *Figure2* pour visualiser le positionnement de ces personnes dans l'Auditorium : sièges D21, F1, F2, F4, G5.

appelle « sweet-spot » ou « Position d'Écoute Centrale (CLP) » ce point d'écoute de référence « où tous les haut-parleurs sont équidistants et également calibrés en Niveau de Pression Sonore (SPL). »¹⁵⁹ [Peters, 2010]. Les autres sièges sont situés dans la « Position d'Écoute Hors-centre » – OCP en anglais chez [Peters, 2010] pour « Off-center Listening Position » – et les auditeurs assis dans cette zone peuvent percevoir une dégradation du son.

Qu'entend-t-on par dégradation du son ? Peters, dans sa thèse [2010], décrit une expérience qu'il a réalisée cherchant à comprendre ce que l'auditeur perçoit lorsqu'il se trouve dans la zone d'écoute hors-centre, comparée à un placement au point d'écoute de référence. Les résultats montrent que les artefacts entendus sont liés à la perception du positionnement des sources sonores, à leur distance et profondeur, à la réverbération et au sentiment d'enveloppement, à la largeur de la source et enfin, à son timbre. Ces observations témoignent de la complexité de cette dégradation du son dans la zone hors-centre, le compositeur lui-même pourrait donc difficilement y remédier en amont, lors de l'étape de composition. Et surtout, ces artefacts dépendent, bien entendu, du matériau sonore employé puisque les réponses des participants à l'expérience de Peters variaient en fonction. Par exemple, on perçoit plus de différences hors du « sweet-spot » sur des sons percussifs que sur des sons longs, tenus. Ce qu'il faut retenir principalement : « [...] the perception of sound degradation is first due to spectral and spatial alterations in the higher frequency range at and above 2.5 kHz and that the listening environment and properties of the audio material both affect the strength of the off-center sound degradation. »¹⁶⁰ [Peters, 2010] Aussi, la sensation d'immersion peut être affectée car le LEV – conséquence des réflexions latérales tardives (cf. partie II.1) – peut devenir instable¹⁶¹.

Il est également intéressant de noter que, même dans le cas d'une écoute positionnée au « sweet-spot », le simple fait de faire bouger un son ne produit pas le même résultat dans toutes les directions, et l'ajout d'haut-parleurs pour diminuer les « trous d'espace » ne serait pas une solution [Kendall, 2010]. Kendall fait ici référence aux dispositifs multicanaux utilisés avec les techniques stéréophoniques. On peut supposer qu'en employant l'Ambisonie (HOA) ou la WFS par exemple, le problème serait atténué. Enfin, nous l'avons souligné plus haut, la perception d'une dégradation dépend du matériau sonore mais aussi de nos habitudes d'écoute. Si un son nous est familier, on arrivera mieux à le percevoir dans l'espace, même si

159 Texte original : « where all loudspeakers are equidistant and equally calibrated in Sound Pressure Level (SPL). »

160 « [...] la perception de la dégradation du son est tout d'abord due à des altérations spectrales et spatiales dans la plage de fréquences supérieures à 2,5 kHz et l'environnement d'écoute et les propriétés du matériau audio affectent l'intensité de la dégradation du son hors du centre. »

161 Toole, F. E., 2008, *Sound reproduction: loudspeakers and rooms*, Oxford, UK: Focal Press.

nous sommes placés hors du « sweet-spot ». De même, on peut avoir l'impression de percevoir moins de dégradations sonores si l'on écoute avec un dispositif qui nous est familier. Ainsi, dans leur expérience, Guastavino et Katz [2004] ont demandé aux participants, placés au centre de chaque système, de s'en éloigner progressivement tout en poursuivant l'écoute et les résultats sont sans appel : c'est la stéréo qui a été jugée la « plus stable » comparée aux configurations 2D et 3D. On reconnaît ici les habitudes d'écoute que les gens ont avec le système stéréo puisque, lors d'une écoute à la maison, il est fréquent que les auditeurs se déplacent tout en continuant d'écouter la musique, ce qui peut expliquer ce type de réponses.

Immersion et élargissement du « sweet-spot »

« Wilson and Harrison (2010)¹⁶² propose a number of spatial decorrelation techniques, which they define as including any approach that produces decorrelated versions of a signal across two or more channels. The authors use the decorrelation techniques to 'enhanced diffusion', 'increased volume perception' of a source and create the perception of 'envelopment'. Also, the authors state that the decorrelation approaches are relatively effective across a number of different listening positions with signals maintaining their relative spatial locations across a large sweet spot. This indicates that decorrelation approaches could be used to negate the influence of the precedence effect. »¹⁶³ [Lynch, 2014]

Quelles solutions peuvent être employées pour élargir le « sweet-spot » ? Nous l'avons vu en partie I.3, c'est notamment l'effet de précedence qui est responsable de l'existence de la zone d'écoute dite « hors-centre ». D'après [Kendall, 2010] et [Lynch, 2014], l'emploi des techniques de décorrélation peut servir à diminuer voire à contrer l'effet de précedence. Elles permettent en effet de créer un état spatial où le matériau sonore n'est ni fusionné ni localisé : « [...] they¹⁶⁴ point out that when one noise signal is presented with another noise signal

162 L'auteur fait référence à : Wilson S., Harrison, J., 2010, *Rethinking the BEAST: Recent developments in multichannel composition at Birmingham ElectroAcoustic Sound Theatre*, Organised Sound, 15(03), 239–250.

163 « Wilson et Harrison (2010) proposent un certain nombre de techniques de décorrélation spatiale qu'ils définissent comme incluant toute approche produisant des versions décorrélées d'un signal sur deux canaux ou plus. Les auteurs utilisent les techniques de décorrélation pour « améliorer la diffusion », « augmenter la perception du volume » d'une source et créer la perception de « l'enveloppement ». En outre, les auteurs déclarent que les approches de décorrélation sont relativement efficaces à travers un certain nombre de positions d'écoute différentes avec des signaux conservant leurs emplacements spatiaux relatifs à travers un grand point d'écoute. Ceci indique que les approches de décorrélation pourraient être utilisées pour annuler l'influence de l'effet de précedence. »

164 L'auteur fait référence à : Perrott, D.R., Strybel, T.Z. et Manligas, C.L., 1987, *Conditions Under Which the Haas Precedence Effect May Or May Not Occur*. Journal of Auditory Research 27: 59–72.

uncorrelated to the first, the listener perceives not one, but two identical sounding sources which are hardly affected by ATDs. »¹⁶⁵ [Kendall, 2010]. Nous avons vu dans [Nilles, 2016] que les techniques de décorrélation micro-temporelle appliquées à la stéréo servent à étendre la source sonore. Sur un dispositif multicanal, elles sont aussi capables de susciter un sentiment d'*enveloppement* qui lui-même peut induire de l'immersion. Pour le compositeur, créer un contenu musical "immersif", en usant des techniques de décorrélation du son, peut aider à élargir le « sweet-spot » lors de la diffusion. De même, jouer dans un lieu modérément réverbérant plutôt que dans un petit espace moins réverbérant, permet d'agrandir le point d'écoute optimal [Peters, 2010]. D'ailleurs, dans l'expérience de Peters [2010] présentée plus haut, les caractéristiques spatiales pour lesquelles les participants ont perçu le moins de dégradation, placés à une position hors du « sweet-spot », sont le « degré d'immersion » et la « spaciousness ». Encore une fois, pour le compositeur, il faut privilégier le ressenti et non la précision spatiale pour que l'auditeur "mal placé" puisse tout de même percevoir l'oeuvre telle que le compositeur l'a conçue.

La déambulation

« J'aurais préféré que la scène soit au centre avec la possibilité de me balader autour, histoire de pouvoir aller à différents endroits du mix pour pouvoir vraiment apprécier la 3D, mais bon, je suppose que c'était compliqué. » (témoignage d'un participant au concert *Immersion Sonore*, "novice" et "musicien", placé en A2, récolté via le questionnaire distribué)

Les concerts spatialisés, où le spectateur est invité à se déplacer dans l'espace de diffusion, sont tout à fait communs. Certaines œuvres spatialisées, pour être entièrement perçues, demandent de la part de l'auditeur qu'il change souvent de place pour apprécier l'oeuvre dans son intégralité. Nous pensons notamment à la pièce *StrinGDberg* (2001-03) de Robert Normandeau, basée sur la technique de spatialisation timbrale. Le compositeur souhaite que le spectateur se déplace librement à l'intérieur du son qui crée l'espace, afin de ressentir une immersion totale. Cela semble judicieux puisque, nous l'avons vu en partie I, les mouvements de tête et même les déplacements du corps tout entier favorisent la perception auditive de l'espace. Certains auditeurs, lors du concert *Immersion Sonore*, auraient d'ailleurs aimé être debout pour avoir la possibilité de se mouvoir. Les sièges étant fixes dans

165 « [...] ils soulignent que, lorsqu'un signal *noise* est présenté avec un autre signal *noise* non corrélé au premier, l'auditeur ne perçoit pas une, mais deux sources sonores identiques qui sont à peine affectées par la différence de temps d'arrivée (ATD). »

l'Auditorium de la MSH Paris-Nord, cela était donc impossible à mettre en place. Mais nous aimerions proposer cette configuration pour une prochaine édition. Eric Lyon, lors de notre échange de mails¹⁶⁶, m'a d'ailleurs confié qu'il n'appréciait guère l'alternative de l'écoute au casque en binaural, puisqu'elle enlève cette dimension inhérente à un espace sonore : le fait de pouvoir se déplacer *dans* le son.

La déambulation pourrait-elle être proposée comme une alternative à l'effet « sweet-spot » ? Certainement, puisqu'elle permettrait à l'auditeur de résoudre certains conflits de perception par le déplacement. En revanche, il ne faut pas négliger que la position debout peut, à la longue, fatiguer l'auditeur et diminuer sa concentration. Une autre option serait de permettre à l'auditeur, à la fois de se mouvoir dans l'espace de diffusion, mais aussi de s'asseoir en lui mettant à disposition des sièges ou des poufs. La position assise, en concert, est d'ailleurs associée à « l'écoute attentive » [Ekeberg Henriksen, 2002].

III.2.3 Les concerts spatialisés participent à la formation de « l'oreille spatiale » et au développement de l'attention

Localiser avec les sons est le résultat d'un apprentissage amorcé avec notre rapport à l'environnement depuis la naissance (cf. partie I.2), mais aussi avec les nouvelles technologies : les expériences auditives spatiales sont partout et facilement accessibles avec le cinéma, la réalité virtuelle, les radios proposant du contenu binaural, etc. L'ensemble de ces expériences favorisent notre travail de « l'oreille spatiale ». Les concerts de musiques spatialisées peuvent donc également participer à cet apprentissage. En effet, ceux-ci sont vécus comme des expériences auditives par l'auditeur puisque leur perception, leurs sens sont stimulés (création d'images propres à nous, sons "vivants" presque palpables, ...). Ces concerts permettent d'apprendre à entendre l'espace de façon consciente pour certains, ou bien sans même s'en rendre compte pour d'autres, plus novices. L'essentiel est que les auditeurs vivent une expérience qu'ils vont ensuite garder en mémoire.

Derrière la scénographie très sobre du concert *Immersion Sonore* se cachait une volonté de tenter de stimuler l'écoute des auditeurs en les poussant à se créer leurs propres images. Sans pour autant les questionner à ce sujet, nous avons eu un répondant qui s'est exprimé de lui-même à ce propos : « Il est vrai que fermer les yeux pendant les pièces me

¹⁶⁶ Voir ANNEXE F pour la restitution complète de notre échange.

permettait d'avoir une expérience enrichie (entre autres avec "visions" de fractales). » Nous pensons que ces images, que chacun forme dans sa tête grâce à la spatialisation des sons, ont le pouvoir de stimuler la concentration de l'auditeur. « L'occultation visuelle permet de vivre l'expérience de l'intérieur. Le fait de fermer les yeux pour l'exploration facilite le contact avec nos sensations internes, en évitant le danger du contrôle externe, du « je » du mental, de la volonté... »¹⁶⁷. D'ailleurs, dans un concert dit "immersif", celui-ci est comme "absorbé" par ce qu'il entend (cf. partie II.). C'est de cette façon que nous pouvons tenter de proposer des concerts de musiques électroniques diverses à un public hétérogène composé d'initiés et de non-initiés à la notion d'espace dans la musique, de personnes amatrices de musiques de création tandis que d'autres préfèrent les musiques d'exploitation. En effet, c'est en partie l'écoute du temps qui va poser problème à l'auditeur habitué aux musiques d'exploitation lorsqu'il entend des musiques dites de création. L'espace sonore peut-il alors modifier l'écoute du temps ? Nous pensons que la réponse est « oui » puisque le concert spatialisé peut aider à capter l'attention de l'auditeur qui est "immergé", "absorbé", ... Dans ce cas, nous pouvons dire que le temps peut sembler évoluer différemment. Denis Smalley [1991] parle d'« espace temporel » : « The relationship between space and the psychological experience of time – *temporal space* – is also intrinsic to music. Temporal space is an impression of space, invariably large and beyond personal space, created through relative stability and continuity in time. »¹⁶⁸ Lorsque l'auditeur n'est pas concentré sur des détails mais sur la globalité d'une oeuvre spatialisée, il peut ressentir une sorte d'« existence perpétuelle », assimilable à la contemplation d'un paysage visuellement permanent. D'autre part, l'espace structure, comme nous l'avons vu dans la citation d'Ekeberg Henriksen, partie III.2.2. Il peut donc aider à mieux "comprendre" la musique, il guide l'écoute. La spatialisation crée également des effets de surprise [Nilles, 2016], tout comme le provoquent les sauts de registres, les jeux de nuances, ou encore les modulations, qui permettent de relancer l'écoute.

Enfin, un concert de musique électronique spatialisée, c'est surtout un moment d'échange, de partage, de sociabilisation. « A major advantage that music for large numbers of speakers will continue to hold over stereo music is that it is inherently social in nature, requiring an audience to come to a space, rather than be able to consume the music privately

167 Lisa Nelson, danseuse-chercheuse, citée dans l'article : *Sur les explorations menées les yeux fermés* de Pascale Gille, p.68, dans : Corin, F. (sous la direction de), 2001, *Vu du corps, Lisa Nelson. Mouvement et perception*, Nouvelles de Danse n°48/49 (automne-hiver 2001), Éd. Contredanse, Bruxelles ; cité chez [Mahec, 2010].

168 « La relation entre l'espace et l'expérience psychologique du temps – *l'espace temporel* – est également intrinsèque à la musique. L'espace temporel est une impression d'espace, invariablement grand et au-delà de l'espace personnel, créé par la stabilité relative et la continuité dans le temps. »

at home. »¹⁶⁹ [Lyon, 2014]. L'attention du spectateur se développe en quelque sorte "naturellement" lorsqu'il entre dans la salle de concert, que les lumières sont tamisées, que les gens autour chuchotent. Dans un concert de musiques spatialisées, l'auditeur entre généralement dans un espace délimité par les haut-parleurs, ce qui peut "forcer" sa concentration puisqu'il y a ici une influence visuelle importante à considérer. Savoir qu'il y a du monde et que toutes ces personnes partagent la même expérience, dans le silence, renforce également la concentration de l'auditeur. Le fait de choisir de se déplacer à un concert de musique spatialisée, que l'on ne peut aisément écouter autrement, rend le moment unique et ne peut qu'influer positivement sur l'attention de l'auditeur.

III.3. Conclusion

« Human experience of space – how we feel *in* it and how we feel *about* it – has therefore become a new dimension of musical experience. »¹⁷⁰ [Smalley, 1991]

Organiser l'évènement *Immersion Sonore* a été une expérience enrichissante puisqu'elle nous a permis de travailler la spatialisation, d'expérimenter avec différentes techniques de reproduction du champ sonore, et surtout, de mélanger différents styles de musique électronique au sein d'un même concert, et donc de diversifier les auditeurs. Nous avons pu également récolter leurs avis, ce qui est essentiel pour comprendre comment chacun a perçu le concert selon ses connaissances ou non en matière de son spatialisé, quelles étaient leurs attentes et quelles sont leurs remarques. 65% des répondants au questionnaire se sont sentis "immergés" dans la musique, dans le son – le reste ayant répondu "moyennement". La question portait sur l'immersion au sens de l'« enveloppement » *physique* de l'auditeur par les sons, de ressentir des vibrations, ... Le public semble avoir été en majorité convaincu. Mais les quelques problèmes de placement observés précédemment ont sans doute influé sur les réponses de certains auditeurs. Pourtant, nous l'avons vu en partie II, l'immersion n'est pas

169 « Un avantage majeur que la musique pour un grand nombre de haut-parleurs continuera à avoir sur la stéréo est qu'il est intrinsèquement de nature sociale, nécessitant de la part du public que celui-ci vienne dans un espace, plutôt que de pouvoir consommer la musique en privé, à la maison. »

170 « L'expérience humaine de l'espace – comment nous nous sentons *dedans* et qu'est-ce que nous ressentons à *propos* de cet espace – est donc devenue une nouvelle dimension de l'expérience musicale. »

uniquement définie par des sensations *physiques*, mais aussi par les émotions qu'elle procure, par le pouvoir de capter l'attention de l'auditeur. Le concert spatialisé "immersif" semble alors stimuler l'écoute et l'attention du public [Ekeberg Henriksen, 2002]. L'auditeur est au centre du dispositif, il a la possibilité de marcher, d'être assis, ou simplement de rester debout, il est en quelque sorte *acteur*. Pour toutes ces raisons, nous pouvons affirmer qu'il *vit* une expérience auditive spatiale.

Nous nous sommes d'ailleurs demandé s'il y avait une configuration optimale haut-parleurs – public – artiste pour le concert de musiques électroniques spatialisées, des systèmes de diffusion qui tendraient vers une écoute où les corps sont au centre du système, mais où il n'y aurait pas ou peu de déséquilibre audible entre ceux placés au milieu et ceux qui seraient sur les côtés. L' "idéal" serait peut-être que l'auditeur ait le choix, pendant le concert, de sa position. Quant à l'artiste, il n'est pas obligatoire qu'il soit placé "devant", "face" à l'auditeur, pour ne pas donner une orientation à l'écoute s'il n'y en a pas besoin, ou encore pour ne pas accaparer l'attention du spectateur uniquement avec la vue. Ce qui est certain, c'est qu'il n'y a pas un seul système standard qui se détacherait et qui mériterait d'être exploité plus que les autres. C'est un défi pour le compositeur, et toujours une nouvelle expérience pour l'auditeur, de changer de dispositif puisque cela permet une diversification de l'écoute. Il est vrai que les systèmes 2D sont plus simples à transporter et à installer, ils sont polyvalents et paraissent plus satisfaisants aussi bien pour les compositeurs que pour les auditeurs [Guastavino & Katz, 2004] [Lynch & Sazdov, 2011]. Les systèmes 3D répondent moins facilement aux besoins créatifs des compositeurs [Peters, 2010], les auditeurs apprécient moins l'écoute sur de tels dispositifs, et il est plus difficile d'avoir la possibilité d'installer des haut-parleurs en hauteur ou d'avoir accès à un tel dispositif [Ontondo, 2008]. Mais il ne faudrait pas pour autant délaisser les systèmes 3D. Il faudrait créer des pièces *pour* ces configurations, développer de meilleurs outils et des studios pour avoir la possibilité d'expérimenter, comme l'a fait Eric Lyon avec sa pièce *Spirits*, transposable presque exclusivement sur des dômes, l'élévation étant une dimension essentielle de sa pièce (cf. ANNEXE F).

Enfin, les concerts de musique spatialisée présentent une configuration plus "intimiste" car le public est généralement entouré de haut-parleurs, dans un espace plus réduit que pour un concert "classique", il est souvent libre de ses mouvements et de sa position – ce qui peut donner l'impression d'être "comme à la maison". Les haut-parleurs sont placés plus proches de l'auditeur et leur niveau global est moins fort. Il semble important de considérer le niveau d'écoute car de plus en plus d'auditeurs sont victimes d'un niveau sonore élevé, au quotidien. Ainsi, le concert de musique spatialisée permettrait d'allier sons et confort de l'écoute. De plus

amples recherches et expériences sont bien entendu à réaliser pour avoir la possibilité d'affirmer ce dernier point.

Conclusion

« [...] a musical environment in which perceivers can immerse themselves, exploring and interacting in plausible ways. »¹⁷¹ [Lennox, 2009]

« The exciting challenge for composers in this “musical garden” is to discover plausible experiences, entities and events that do not have exact analogues in the ‘real’ world. »¹⁷² [Lennox, 2009]

Composer l'espace n'est pas uniquement lié à la formulation de concepts, c'est aussi et surtout expérimenter et écouter le rendu sonore [Smalley, 1991]. L'objectif de ce second mémoire de recherche est donc d'apporter au compositeur de musique électronique spatialisée un regard tout à fait différent du précédent écrit : on tente de s'intéresser à la réception de l'oeuvre spatialisée par l'auditeur en dirigeant notre attention sur la perception humaine de l'espace sonore, le sentiment d'immersion et la situation de concert. Nous ne prétendons pas donner des solutions au compositeur qu'il doit appliquer à la lettre, mais plutôt suggérer des pistes de réflexions qui pourront l'aider lors du processus de création d'une œuvre spatialisée, ou lorsqu'il prépare un concert. Comme le dit Smalley [1991] : « Musical spatial experience can therefore not be systematically controlled by the composer. »¹⁷³. Assurément, l'expérience de l'espace sonore, en musique, est dépendante de nombreux facteurs que nous nous efforçons d'explorer dans ce présent mémoire.

Nous avons ainsi tâché, dans la première partie, de comprendre certains principes du fonctionnement de la perception auditive humaine pour en avoir conscience au moment du travail de l'espace dans la musique. Si l'on considère l'espace comme un ensemble de trois plans dont le point de croisement est le centre de la tête de l'auditeur, et si, par conséquent, la

171 « [...] un environnement musical dans lequel les auditeurs peuvent s'immerger, explorer et interagir de façons convaincantes. »

172 « Le défi passionnant pour les compositeurs dans ce "jardin musical" est de découvrir des expériences convaincantes, des entités et des événements qui n'ont pas d'analogues exacts dans le monde "réel". »

173 « L'expérience spatiale musicale ne peut donc pas être systématiquement contrôlée par le compositeur. »

position d'une source sonore peut être définie par des coordonnées, en pratique, la précision avec laquelle l'auditeur peut localiser cette même source sonore dépendra de propriétés temporelles, d'amplitudes et spectrales, elles-mêmes influencées par l'environnement et la morphologie de l'auditeur. Nous avons également mis en évidence la notion d'apprentissage et donc la possibilité pour l'auditeur de travailler son « oreille spatiale » de par ses expériences quotidiennes et potentiellement par l'écoute « active » et décomplexée de musique électronique spatialisée. Nous avons rassemblé ces informations pour mettre en perspective certaines techniques de spatialisation reposant sur la manipulation du spectre sonore, et remettre en cause celles qui consistent à définir des trajectoires précises pour le mouvement des sons. Cette partie I nous a semblé essentielle pour démarrer notre étude et considérer comme acquises certaines connaissances sur lesquelles nous nous appuyons par la suite.

Avant de poursuivre, nous souhaitons revenir brièvement sur nos débuts de recherches de Master pour mieux faire comprendre au lecteur les différents cheminements qui ont donné lieu à ce présent document, et justifier notre intérêt relatif à la notion d'immersion, traitée en partie II. Au commencement, notre volonté était de s'interroger sur l'existence d'un dispositif de haut-parleurs "idéal", aisément transportable, dans lequel la zone d'écoute hors du « sweet-spot » serait réduite voire annihilée, où les personnes seraient au centre du système pour partager une expérience de musique électronique spatialisée de façon égale. Au fur et à mesure de nos lectures, écoutes, et expérimentations, nous avons compris qu'il n'existait pas à proprement parler de système idéal et que surtout, il y avait une dimension bien plus importante à prendre en compte : la façon avec laquelle seraient projetés les sons via ce dispositif de haut-parleurs. Nous nous sommes alors intéressés aux différentes techniques de reproduction du champ sonore : la stéréophonie, le VBAP, l'Ambisonie, la WFS, etc. Toutes sont grandement intéressantes, selon nos envies de composition, aucune n'est idéale. À moins que l'idéal ne soit de les combiner ! (cf. [Baalman, 2010] et [Nilles, 2016]). Dans tous les cas, le plus important semble finalement le désir d'expérimentation du compositeur et la réception de l'oeuvre par l'auditeur. C'est ainsi que nous sommes arrivés au point de nous intéresser à la perception auditive de l'espace par l'être humain. D'autre part, dès [Nilles, 2016], nous avons commencé à faire ressortir une aspiration commune aux compositeurs de musique spatialisée et aux auditeurs : se sentir immergés dans le son. Vraisemblablement, l'immersion est la propriété spatiale la plus appréciée par les compositeurs [Peters, 2010] [Lynch & Sazdov, 2011] et la plus significative pour les auditeurs [Guastavino & Katz, 2004]. C'est pourquoi nous avons cherché, dans la deuxième partie de ce présent mémoire, à définir le plus largement possible la notion d'immersion dans la musique électronique spatialisée en nous

interrogeant sur ses origines, dans un premier temps, puis sur ses conséquences perceptibles par l'auditeur. Nous avons alors répertorié divers critères techniques et découvert que non seulement l'enveloppement, mais aussi la sensation de présence, sont essentiels pour créer un sentiment d'immersion et sont d'ailleurs appelés « attributs immersifs » [Rumsey, 2001] ; les conséquences de l'immersion étant une écoute « naturelle », telle que décrite dans l'introduction de ce présent mémoire, et une perception tactile. Une musique électronique spatialisée immersive agit aussi sur nos émotions, elle est capable d'accaparer notre attention. L'immersion *totale* englobe l'ensemble de ces caractéristiques. D'autre part, elle dépend également du type de son, s'il nous est familier ou non, et n'est pas rattachée à un format de diffusion, ni à une technologie de reproduction du champ sonore : une diffusion stéréo peut nous sembler immersive car c'est tout de même la psychologie de l'auditeur qui domine. Enfin, une expérience sonore immersive peut être assimilée au fait d'être plongé entièrement dans de l'eau – dans un bain de sons – puisque le son n'est pas simplement autour : il est partout.

La troisième partie rassemble les informations accumulées tout au long de cette étude, elle pourrait être considérée comme une première synthèse. Elle présente en effet une mise en pratique de nos recherches et expose nos observations qui font suite à l'organisation du concert *Immersion Sonore*. Ce dernier est à appréhender comme une sorte de "concrétisation" de l'ensemble de nos réflexions, au cours de nos deux années de recherches, sur le concert de musique électronique spatialisée. Par conséquent, nous avons pu confirmer notre hypothèse sur l'influence de la vision et des habitudes d'écoute de l'auditeur sur la perception de la musique électronique spatialisée (le fait de voir ou non le dispositif de haut-parleurs, les artistes sur scène, etc.). Le choix d'une scénographie minimaliste avait pour but de laisser au spectateur la possibilité de se former ses propres images. Nous pensons en effet que l'occultation visuelle a la capacité d'augmenter l'attention à l'écoute des phénomènes sonores et donc de faire travailler l'imagination de l'auditeur pour ainsi contrebalancer la surstimulation visuelle que connaît notre époque. Nous avons également réfléchi sur des questions laissées en suspens dans [Nilles, 2016] concernant le portage d'une œuvre spatialisée dans divers lieux, sur divers systèmes (le "up/down-mixing"), et le « sweet-spot ». Nous avons alors conclu qu'il n'existe pas de système idéal ni de configuration haut-parleurs – artiste – spectateurs modèle. Une solution serait finalement de laisser le choix, à l'auditeur, de se positionner librement dans l'espace de diffusion, pendant toute la durée du concert : il peut s'asseoir, rester immobile debout, ou décider de déambuler. De même, plus la diversité des systèmes et des lieux sera grande, plus les créations seront riches, et les expériences spatiales

sonores variées. En outre, si ces expériences spatiales sonores sont à caractère immersif, elles seront moins dépendantes d'une configuration de haut-parleurs, d'une disposition rigoureuse du public, puisque le sentiment d'immersion ne résulte aucunement de la précision de la reproduction spatiale. Enfin, si l'espace sonore est ressenti comme immersif par l'auditeur, cela signifie que l'attention de ce dernier est démultipliée. Ceci laisse entrevoir qu'au sein d'un même concert spatialisé à caractère immersif, divers genres de musiques électroniques peuvent être proposés dans l'intention de mélanger les publics et de diversifier leurs habitudes d'écoute, en ayant comme point commun l'écoute de l'espace.

Nous faisons ainsi le lien entre les nombreuses notions qui ressortent du présent écrit et nos intentions de départ. L'immersion semble donc un facteur "universel" : elle est peut-être la clé pour monter des concerts mélangeant tout type de musiques électroniques spatialisées, peu importe le système choisi ou la technique de reproduction du champ sonore utilisée, et elle répond à la problématique de l'écoute « naturelle »¹⁷⁴ formulée dans l'introduction. Bien évidemment, nous ne prétendons pas que toutes les pièces doivent être à caractère immersif – la diversité musicale et spatiale est absolument nécessaire ! Il serait dommage, par exemple, de délaisser les dispositifs 3D sous prétexte qu'ils ne sont pas assez "immersifs" [Guastavino & Katz, 2004]. Comme le remarquent Lynch & Sazdov¹⁷⁵ [2001], ils permettent tout de même une écoute proche des conditions d'écoute dans notre environnement naturel. Il est important de continuer à réaliser des créations spécialement *pour* et *avec* ces systèmes, et de chercher des moyens de parvenir à une diffusion immersive, si les intentions compositionnelles de l'oeuvre s'y prête. En revanche si l'on souhaite, comme énoncé dans l'introduction, que les concerts de musiques électroniques spatialisées se développent davantage et s'étendent à tous les genres de musiques électroniques, ainsi qu'à un grand nombre de lieux et d'auditeurs (novices et initiés), favoriser l'immersion au niveau du rendu perceptif de l'oeuvre spatialisée se révèle être un bon compromis. Par ailleurs, le travail de l'« oreille spatiale » étant effectif, les concerts de musiques électroniques spatialisées sont autant de moyens pour aiguïser sa sensibilité à l'écoute de l'espace sonore.

En somme, le concert spatialisé immersif stimule à la fois la perception auditive de l'espace, mais aussi l'attention du public et donc ses émotions, puisqu'il est immergé – *absorbé* – dans le son. Nous pensons que l'écoute du temps musical est alors altérée de différentes façons. Nous retenons ici la notion d'« espace temporel » introduite par Smalley [1991] qui peut être comprise comme une sorte de *contemplation* dans laquelle est plongé l'auditeur, à travers l'écoute de l'espace. Ceci fait contrepoids aux habitudes de vie et d'écoute

174 « naturalness »

175 cf. citation présentée en introduction du mémoire.

dans une société hyperactive, et peut répondre à un besoin de lenteur, de *prendre le temps*. D'après nous, l'expérience musicale spatiale immersive est capable d'extraire l'auditeur de son quotidien le temps d'un concert, et de sensibiliser son écoute à d'autres genres de musiques électroniques. L'écoute de l'espace dans la musique, des « spatiomorphologies »¹⁷⁶, nous semble la dimension importante sur laquelle concentrer notre attention et nos explorations futures.

« Je pressens l'avènement d'une musique d'après le langage, de pure sensation, qui s'affranchit des codes, comme le rythme ou la mélodie qui nous font croire qu'il faut comprendre la musique en même temps qu'on la ressent. Je milite pour une expérience libérée. »¹⁷⁷

176 Voir : Smalley, D., 1997, *Spectromorphology : explaining sound shapes*, dans *Organised Sound* 2(2), p.107–26, Cambridge University Press, Royaume-Uni ; et déjà évoquées dans [Nilles, 2016].

177 Interview de François Bonnet, directeur artistique du GRM, dans : Lamm, Olivier, *Musiques électroniques : les forces en prescience*, publié le 19 mars 2018 dans le journal *Libération*.

URL : http://next.liberation.fr/musique/2018/03/19/musiques-electroniques-les-forces-en-prescience_1637318 [consulté le 21/03/18].

ANNEXE A

Questionnaire proposé par Amélie Nilles pour le concert *Immersion Sonore* du 07/12/17 à la MSH Paris Nord, dans le cadre de son Master2 Théorie et Pratique à l'Université Paris 8.

À remplir pendant ou après le concert et à déposer dans la boîte prévue à cette effet en sortie de l'Auditorium. Vous pouvez également remplir ce formulaire en ligne, disponible à cette adresse : <https://lc.cx/YzGw>

N'hésitez pas à soumettre le questionnaire même si celui-ci est incomplet ! Chaque réponse peut servir par la suite.

Le questionnaire porte sur l'ensemble du concert. Néanmoins, vous pouvez tout à fait citer des artistes ou œuvres présenté(e)s lors du concert pour appuyer vos réponses, cela sera d'autant plus utile !

***NB :** Les expressions "espace sonore" et "spatialisation des sons" se réfèrent à la façon avec laquelle la musique jouée ou diffusée lors du concert, est déployée sur l'ensemble des haut-parleurs : les mouvements des sons, ou bien le fait d'entendre la musique à la fois face à soi, mais aussi sur les côtés et à l'arrière.*

Ces termes se réfèrent en aucun cas à l'espace physique de la salle de concert (sous-entendu le lieu).

Rangée (lettre) : Numéro de place :

=> Votre mail (pas obligatoire, si vous souhaitez être tenu(e) informé(e) des futures programmations des concerts *Immersion Sonore*) :

Vous êtes :
(Entourer. Plusieurs choix possibles)

- musicien(ne) - novice en matière de son en 3D
- expérimenté(e) en matière de son en 3D - professionnel(le) du son 3D

I. Musique et espace sonore

(entourez vos réponses)

- D'après vous, est-ce qu'il y avait une sorte de lien entre la musique et l'espace sonore de chaque oeuvre ou live présenté(e) ? Est-ce que l'on peut dire que le rendu sonore semblait cohérent ?

- Pas du tout - Moyennement - Tout à fait

- Est-ce que la spatialisation des sons vous a « aidé » à apprécier le contenu musical?

- Pas du tout - Moyennement - Tout à fait

- Au contraire, est-ce que la spatialisation des sons gênait l'écoute de la musique ?

- Pas du tout

- Moyennement

- Tout à fait

Éventuelles impressions/remarques :
.....
.....

II. Votre perception de l'espace sonore

(entourez vos réponses)

Globalement, sur la totalité du concert, vous avez perçu/ressenti :

- une impression de déséquilibre

- Pas du tout

- Moyennement/Ça dépend

- Tout à fait

- Rien de spécial, comme un concert habituel (c'est-à-dire qui diffuse les sons uniquement face au public)

- Pas du tout

- Moyennement/Ça dépend

- Tout à fait

- une écoute moins agréable que lors des concerts habituels (qui diffusent les sons uniquement face au public)

- Pas du tout

- Moyennement/Ça dépend

- Tout à fait

- Finalement, vous êtes-vous senti(e) "immergé(e)" dans le son ?

- Pas du tout

- Moyennement/Ça dépend

- Tout à fait

Si possible : pourquoi ?

(donnez vos impressions, n'hésitez pas à vous appuyer des oeuvres et performances présentées lors du concert)

.....
.....

III. Questions générales

- Y a-t-il eu un ou des moment(s) plus marquant(s) que d'autres, notamment en ce qui concerne la spatialisation des sons, pendant le concert ?

.....
.....
.....

- Qu'auriez-vous aimé entendre ou ressentir du point de vue de l'espace sonore, et qui vous aurait manqué lors de ce concert Immersion Sonore ?

.....
.....
.....

- Souhaiteriez-vous retenter l'expérience d'un concert de musiques électroniques spatialisées à 360 degrés ou en 3 dimensions ? *(Soyez honnête !)*

.....
.....

ANNEXE B

Programme Immersion Sonore

Jeudi 7 Décembre 2017 à la MSH Paris Nord

***Sons seuls* [2017, création] – Jérémie Nicolas**

Pièce acousmatique

***Cinq souffles de Tchernobyl* [2017, création] – Jean-François Ducher**

Pièce mixte pour saxophone ténor et électronique en temps réel

« Let it blaze against the mists » – capob

Live

Extrait de *Cauchemar Merveilleux* [2015] (introduction et premier poème) – Léonore Mercier et Arthur H

Poème sonore

« Laïka project » – Amélie Nilles

Live

« CHERENKOV » – Matthieu Ruben

Live

Organisé par Amélie Nilles

En partenariat avec la Maison des Sciences de l'Homme Paris Nord, le Centre de recherche Informatique et Création Musicale (CICM), Euphonia, Sonic Emotion, l'Université Paris 8 et l'Université Paris 13.

Et avec l'aide précieuse de Frank Gillardeaux, Quentin Nivromont et Filippos Sakagian.

ANNEXE C

Immersion Sonore : modélisation du dispositif de haut-parleurs dans l'Auditorium et visualisation du placement des répondants au questionnaire

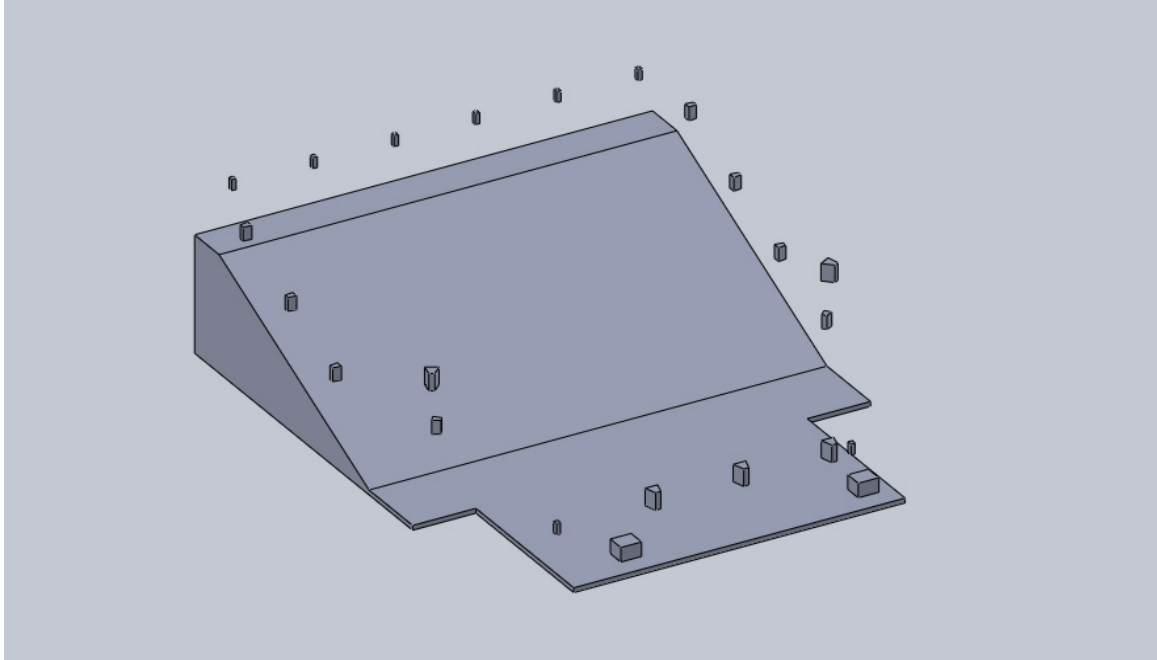


Figure1 : (ci-dessus) Disposition des haut-parleurs dans l'Auditorium de la MSH Paris-Nord pour le concert Immersion Sonore. Modélisation 3D réalisée par Frank Gillardeaux.

Figure2 : (ci-dessous) Photographie de l'Auditorium de la MSH Paris-Nord (vue depuis la scène). Mise en évidence des places occupées par les répondants au questionnaire (en rose).



ANNEXE D

Outils de spatialisation utilisés pour le concert *Immersion Sonore*

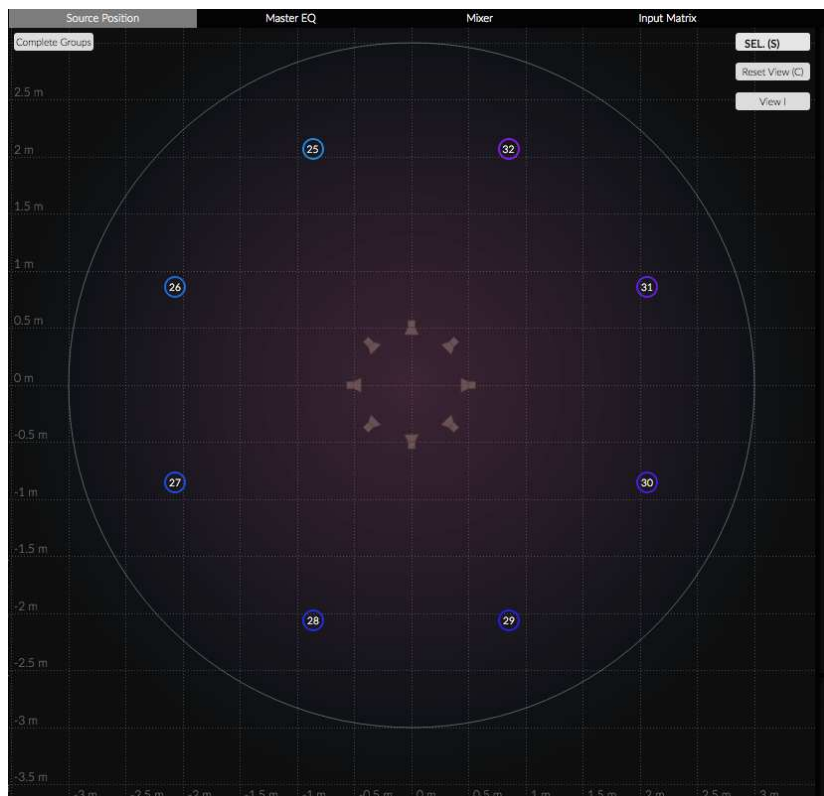
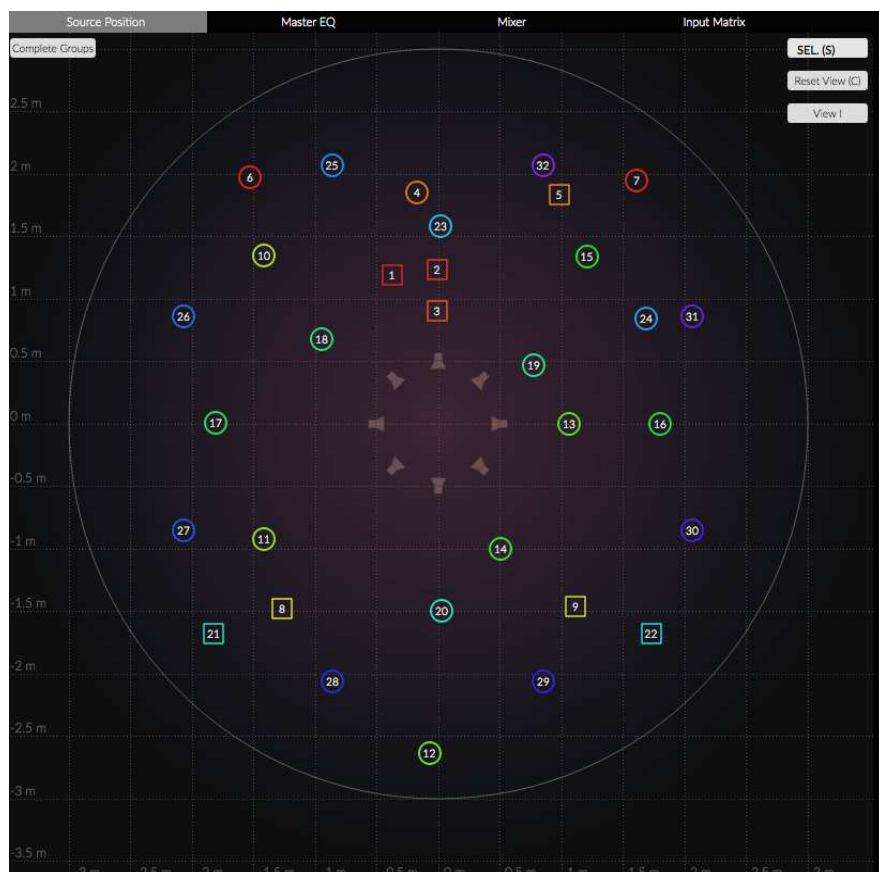


Figure 1 : Haut-parleurs virtuels
 Capture d'écran du logiciel Wave Performer. On y voit le placement des 8 haut-parleurs virtuels correspondants aux 8 canaux HOA. Cette session a été employée pour la pièce de Matthieu Ruben, par exemple.

Figure 2 : Mixage objet
 Capture d'écran de la session de travail du morceau Strange Fruit dans le logiciel Wave Performer. Les sources 1 à 24 correspondent à des sources sonores isolées et directement spatialisées avec la WFS de Sonic Emotion (mixage objet). Les sources 25 à 32 représentent les haut-parleurs virtuels correspondant aux 8 canaux HOA.



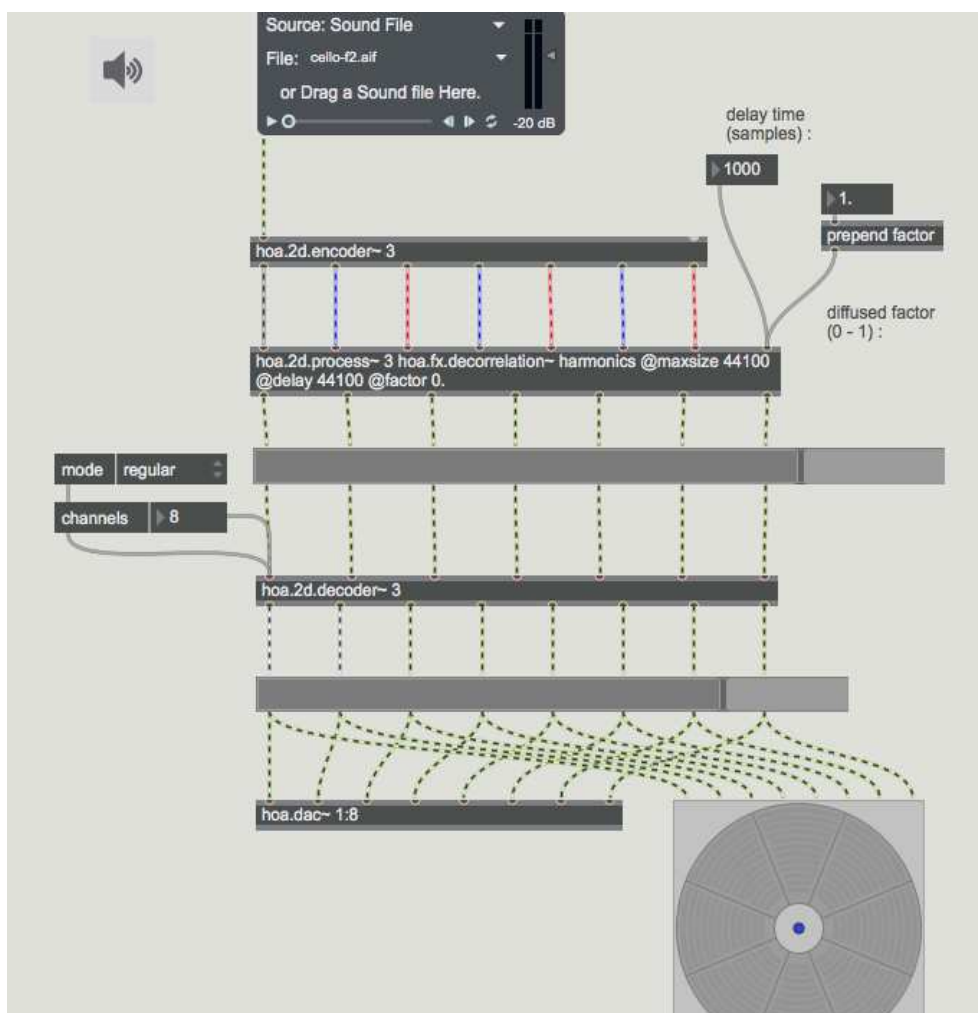
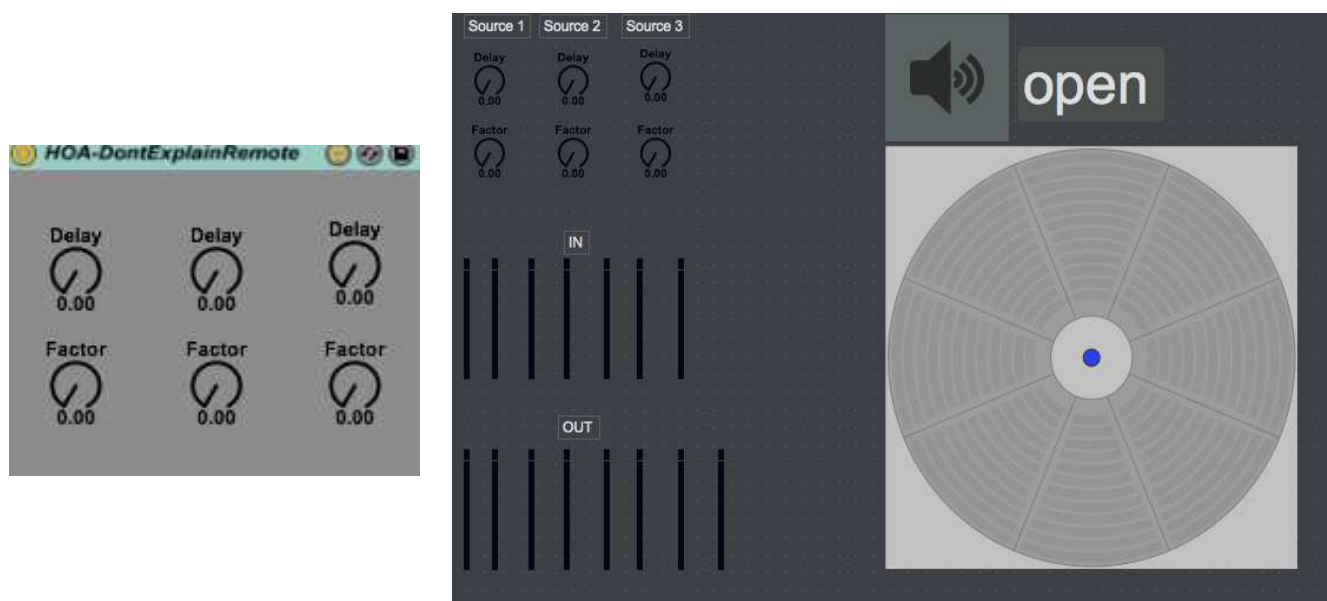


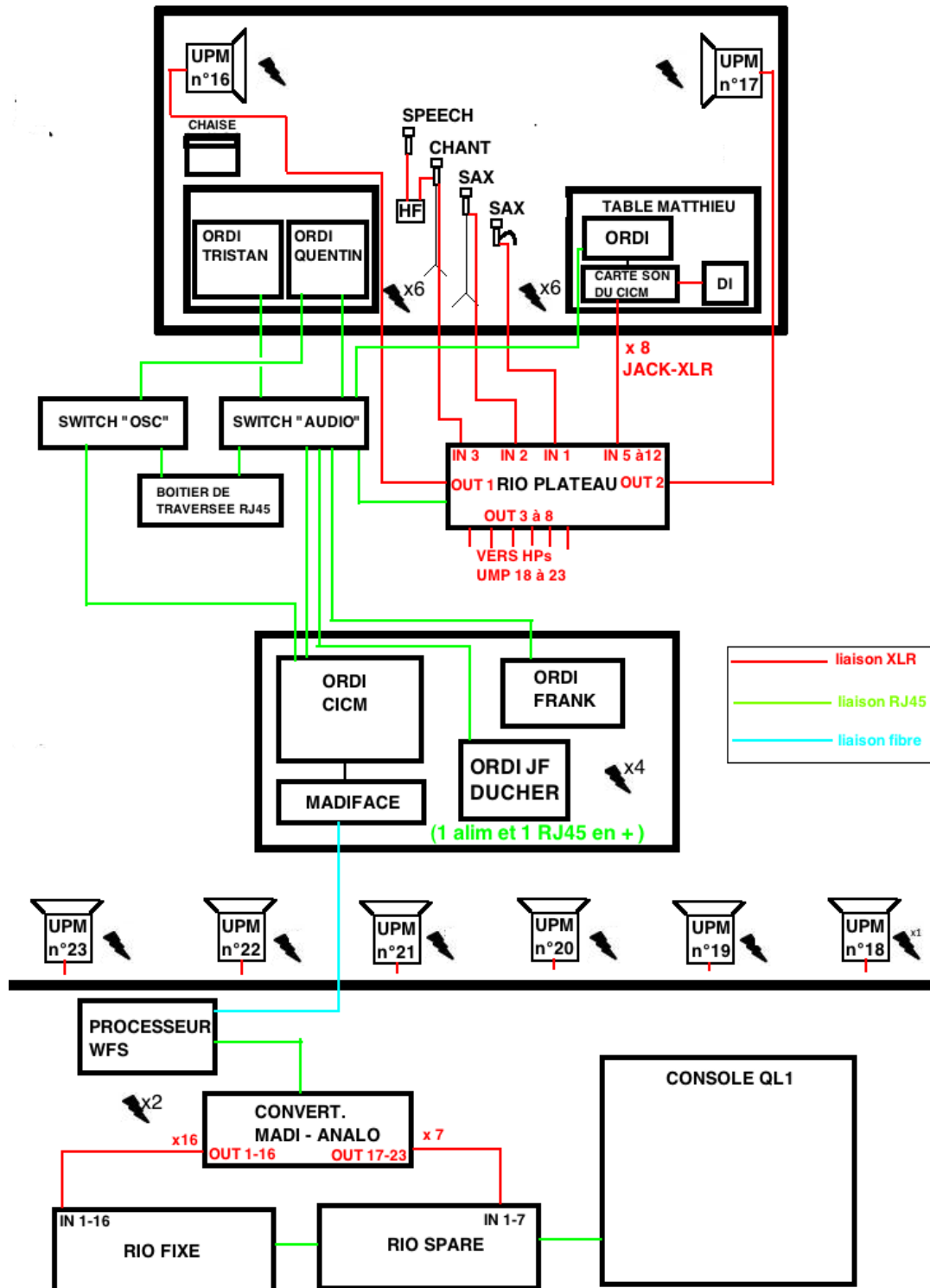
Figure3 : (ci-dessus) Capture d'écran depuis le logiciel Max/MSP d'un exemple de patch simple de décorrélation appliquée aux harmoniques circulaires, avec un facteur de diffusion à 1 et un samples delay à 1000, pour un décodage sur une octophonie circulaire, via la Bibliothèque HOA.

Figure4 : (ci-dessous) Patch Max/MSP (à droite) contrôlant les différentes valeurs du facteur de diffusion et de delay du patch de décorrélation, pour les sources sonores du morceau Don't Explain spatialisées dans HOA. Interface de l'objet MaxForLive (à gauche) utilisé dans le logiciel Live Ableton, servant à communiquer les informations entre les deux logiciels.



ANNEXE E

Synoptique¹⁷⁸ du dispositif technique du concert *Immersion Sonore*



¹⁷⁸ Réalisé par Frank Gillardeaux.

ANNEXE F

Interview de Eric Lyon à propos du portage de sa pièce *Spirits* sur 3 différents systèmes

Réalisée le 5 Juin 2018 par échanges de mails.

Amélie Nilles : *If composing the sound space is the most important "composition process" in "Spirits", could you say it is the same piece if it is played on a different configuration of speakers (and maybe with a different sound field reproduction technique) ?*

Eric Lyon : « This is an important question, and my answer might not generalize to other multichannel music. In the case of "Spirits" it was the same piece on all three systems. I did not use a generalized sound field technique like Ambisonics ; rather the work was spatially orchestrated through selection of specific loudspeakers and speaker groups. This is something like playing an organ composition on different organs, and using different registrations to express the piece. »

A.N. : *What about the binaural recordings ? What could be missing from the original ?*

E.L. : « A lot is missing. For a start, the possibility to move around the space is lost. Even the ability to move one's head for different perspectives is lost, though some newer 360 degree recording and reproduction techniques for binaural sound can bring some of this back. Finally, headphones close you off from your environment, which I prefer to avoid. »

A.N. : *You use the expression "spatial orchestration". For you, is it the same to do a spatial orchestration from the ZKM Klangdom to a 8 channel system, than in instrumental music, a piece written for a great ensemble reduced to piano ?*

E.L. : « Downmixing from the Klangdom to 8-channel could be a big challenge. The usual 8-channel configuration is planar, but the Klangdom has significant elevation built into its system. »

A.N. : *According to you, in which way it could not work and completely "denaturate" the piece ?*

E.L. : « You would lose a lot of density, and would have fewer loudspeakers to express the

complexity of the piece. It could be done, but you might get a result like this:
<https://youtu.be/pKPD-BZEEU0>¹⁷⁹ »

A.N. : Finally, what are the limits of the spatial re-orchestration ?

E.L. : « As an approach, rather than a specific technique (like Ambisonics or WFS), I think there's a lot of range for spatial orchestration, but you need an orchestra! By that, I mean, you need to be working with a high-density loudspeaker array to have the full range of possibilities for creating spatial musical experiences. »

¹⁷⁹ Il s'agit du *Sacre du Printemps* d'Igor Stravinsky arrangé pour deux trompettes par Edward Hoffman et joué à University of Minnesota School of Music par David Baldwin et Loren Endorf.

Bibliographie

ADAIR, Sarah, ALCORN, Michael & CORRIGAN, Chris, 2008, *A study into the perception of envelopment in electroacoustic music*, Proceedings of the 2008 International Computer Music Conference, Belfast, UK, ICMA.

AUSTIN, Larry, 2000, *Sound Diffusion in Composition and Performance: An Interview with Denis Smalley*, Computer Music Journal 24(2):10–21.

BAALMAN, Marije A.J., 2010, *Spatial Composition Techniques and Sound Spatialisation Technologies*, Organised Sound 15(3), pages 209–218, Cambridge University Press.

BIGAND, Emmanuel & POULIN-CHARRONNAT, Bénédicte, 2005, *Are we « experienced » listeners ? A review of the musical capacities that do not depend on formal musical training*, In « Cognition » numéro 100, 2006, pp.100-130.

BLAUERT, Jens, 1997, *Spatial Hearing : The Psychophysics of Human Sound Localization*, Cambridge, MIT Press.

BLESSER, Barry & SALTER, Linda-Ruth, 2009, *Spaces Speak, Are You Listening? Experiencing Aural Architecture*, MIT Press.

CAMUS, Jean-François, 1996, *La psychologie cognitive de l'attention*, Éditions Armand Colin/Masson, Paris.

CITTON, Yves (sous la direction de), 2014, *L'économie de l'attention. Nouvel horizon du capitalisme ?*, Éditions La Découverte, Paris.

CLARKE, Eric F., 2005, *Ways of Listening. An Ecological Approach to the Perception of Musical Meaning*, Oxford University Press.

DEAN, Roger T., August 2014, *Low frequency spatialization in electro-acoustic music and*

performance : composition meets perception, Vol. 42, No.2, Acoustics Australia.

DOBREVA, Marina S., O'NEILL, William E. & PAIGE, Gary D., 2011, *Influence of aging on human sound localization*, Journal of Neurophysiology 105, pp. 2471–2486.

EKEBERG HENRIKSEN, Frank, 2002, *Space in Electroacoustic Music : Composition, Performance and Perception of Musical Space*, Thèse de Doctorat sous la direction de Denis Smalley, City University, Londres.

EMMERSON, Simon, 2007, *Living Electronic Music*, Ashgate Publishing Limited.

FREY, Aline, 2009, *Traitement perceptif et cognitif de l'information auditive et musicale : segmentation des flux et figures sonores*, Thèse de Doctorat sous la direction de Charles Tijus, Université Paris VIII, Saint-Denis.

GINOT, Isabelle (sous la direction), 2014, *Penser les somatiques avec Feldenkrais. Politiques et esthétiques d'une pratique corporelle*, Lavérune, L'Entretemps, coll. Lignes de corps.

GREGORY, Richard, 1998, *Brainy Mind*, Brit. Med. Journal. Article en ligne : http://www.richardgregory.org/papers/brainy_mind/brainy-mind.pdf [consulté le 4/04/2018].

GROTHE, Benedikt, PECKA, Michael & McALPINE, David, 2010, *Mechanisms of Sound Localization in Mammals*, Physiol Rev VOL 90. American Physiological Society.

GUASTAVINO, Catherine & KATZ, Brian F. G., 2004, *Perceptual evaluation of multi-dimensional spatial audio reproduction*, Acoustical Society of America.

GUILLOT, Pierre, 2013, *Les traitements musicaux en Ambisonie*, Mémoire de Master2 sous la direction d'Anne Sèdes et Alain Bonardi, Université Paris VIII, Saint-Denis.

HOFMAN, Paul M., VAN RISWICK, Jos G. A. & VAN OPSTAL, A. John, 1998, *Relearning sound localization with new ears*, Nature Neuroscience 1: 417–421.

KENDALL, Gary S., 1995, *A 3-D Sound Primer: Directional Hearing and Stereo Reproduction*, In Computer Music Journal, Cambridge, Massachusetts: MIT Press.

KENDALL, Gary S., 2010, *Spatial Perception and Cognition in Multichannel Audio for Electroacoustic Music*, Organised Sound 15(3): 228–238 & Cambridge University Press.

KO, Yi-Chun, 2013, *Espace sensible : une proposition d'expérience somatique à partir de la musique interactive*, RFIM (revue en ligne), n° 3 - automne 2013. URL: <http://revues.mshparisnord.org/rfim/index.php?id=262> [consulté le 12/12/2016].

LOSSIUS, Trond, 2007, *Sound - Space - Body. Reflections on Artistic Practice*, Submitted in partial fulfillment for Research Fellowship in the Arts, Bergen National Academy of the Arts.

LYNCH, Hugh, & SAZDOV, Robert, 2011, *An Investigation Into Compositional Techniques Utilized For The Three-Dimensional Spatialization Of Electroacoustic Music*, Proceedings of the Electroacoustic Music Studies Conference, New York.

LYNCH, Hugh Adam, 2014, *Space in Multi-channel Electroacoustic Music : Developing Sound Spatialisation Techniques for Composing Multi-channel Electroacoustic Music with Emphasis on Spatial Attribute Perception*, Thèse de Doctorat, University of Limerick.

LYON, Eric, 2014, *The Future of Spatial Computer Music*, Article disponible en ligne : <http://disis.music.vt.edu/eric/main/wp-content/uploads/2015/10/The-Future-of-Spatial-Computer-Music-Full.pdf> [consulté le 12/03/2017].

MAHEC, Marie, 2010, *Les sensations corporelles et l'interprétation musicale*, mémoire de fin d'études, sous la direction de Georges Lambert, CEFEDM Bretagne-Pays de la Loire.

McADAMS, Stephen & BIGAND, Emmanuel, 1994, *Penser les sons : psychologie cognitive de l'audition*, Presses universitaires de France.

MIDDLEBROOKS, John C., 1992, *Narrow-band sound localization related to external ear acoustics*, Journal of the Acoustical Society of America.

MOORE, Brian, 1989, *An introduction to the psychology of hearing*, Emerald, 2013, 6ème édition.

NILLES, Amélie, 2016, *La spatialisation des musiques électroniques au regard du processus de composition*, Mémoire de Master1 sous la direction d'Anne Sèdes, Université Paris VIII.

NORMANDEAU, Robert, 2009, *Timbre Spatialisation : The medium is the space*, dans *Organised Sound*, 14(03), pages 277-285.

PAYRI, Blas, 2010, *Limitations in the recognition of sound trajectories as musical patterns*, Proceedings of SMC Conference 2010, Barcelona.

- PÉRIAUX, Bergame, OHL, Jean-Luc & THÉVENOT, Patrick, 2015, *Le Son Multicanal*, Éditions Dunod.
- PETERS, Nils, 2010, *Sweet [re]production: Developing sound spatialization tools for musical applications with emphasis on sweet spot and off-center perception*, Thèse, McGill University, Montréal, Canada.
- RIBOT, Théodule, 1889, *Psychologie de l'attention*, publié aux Éditions L'Harmattan, 2007.
- RUMSEY, Francis, 2001, *Spatial Audio*, Focal Press, coll. Music Technology Series.
- RUMSEY, Francis, 2002, *Spatial quality evaluation for reproduced sound : terminology, meaning and a scene-based paradigm*, J. Audio Eng. Soc. 50, 651–666.
- SCHMELE, Timothy, 2011, *Exploring 3D Audio as a New Musical Language*, Mémoire de Master2, Universitat Pompeu Fabra, Barcelone.
- SMALLEY, Denis, 1991, *Spatial experience in electro-acoustic music*, dans *L'Espace du Son II*, sous la direction d'Annette Vande Gorne, Éditions Musique & Recherches, éditions numériques, 1991.
- STEFANI, Ewan & MOONEY, James, 2009, *Spatial composition in the multi-channel domain: aesthetics and techniques*, Proceedings of the International Computer Music Conference 2009, Montréal, Canada.
- VANDE GORNE, Annette, 2002 (mis à jour le : 09/10/2014), *L'interprétation spatiale. Essai de formalisation méthodologique*, Revue DEMéter, Université de Lille-3. URL : <http://demeter.revue.univ-lille3.fr/lodel9/index.php?id=466> [consulté le 20/12/2015].
- WAKUDA, Akiko, FURUYA, Hiroshi, FUJIMOTO, Kazutoshi, ISOGAI, Kenji & ANAI, Ken, 2003, *Effects of arrival direction of late sound on listener envelopment*, Acoust. Sci. & Tech. 24.
Article en ligne : <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.536.4467&rep=rep1&type=pdf> [consulté le 23/06/2017].
- ZAHORIK, Pavel, 2002, *Assessing auditory distance perception using virtual acoustics*, Journal of the Acoustical Society of America.
- ZWIERS, M. P., VAN OPSTAL, A. J. & CRUYSBURG, J. R. M., 2001, *A Spatial Hearing Deficit in Early-Blind Humans*, The Journal of Neuroscience, Vol. 21 RC142.