

Ecole Nationale Supérieure Louis Lumière

Mémoire de fin d'études

Cédric Payet

Menu Cthulhu, une aventure sonore dont vous êtes le héros !

*Adaptation d'un livre-jeu en binaural
avec suivi des mouvements de la tête.*

Promotion Son 2016

Sous la direction de

Alan Blum
Charles Verron

Rapporteur

Eric Urbain

Remerciements

Je tiens à remercier chaleureusement Alan Blum pour sa présence, ses conseils et ses relectures.

Je remercie également Charles Verron pour sa disponibilité malgré la distance, la mise à disposition de l'ensemble des logiciels qu'il développe ainsi que le temps passé à répondre à l'ensemble de mes demandes.

Je tiens à remercier les éditions Walrus et Neil Jomunsi, l'auteur de Menu Cthulhu, de m'avoir autorisé à adapter ce livre dans le cadre de mon mémoire.

Je voudrais remercier Louise Morel, Sofia Bendia, Jean-Christophe Laurier et Régis Meimoun, les acteurs de ma partie pratique, pour leur disponibilité et leur patience. Sans vous, je suis certain que mon prototype n'aurait pas eu la même saveur.

Je tiens à remercier Lucas Derode, Lidwine Ho et Gaël Nicolas, pour les échanges enrichissants que nous avons eu dans le cadre de ce mémoire mais également Matthieu Aussal, pour l'ensemble de ses conseils et le prêt gracieux de la tête Neumann du CMAP.

Je souhaite aussi remercier Eric Urbain et l'ensemble des personnes ayant écouté mon prototype, pour leur temps et leur critiques constructives.

Je tiens à remercier mes parents pour leur patiente relecture, à traquer mes fautes d'orthographe.

Je remercie également l'ensemble de ma promotion pour leur bonne humeur infaillible pendant cette période d'écriture de mémoire.

Je remercie également Sofia pour sa patience et ses conseils qui m'ont permis de venir à bout de ce mémoire.

Enfin, je vous remercie, chers lecteurs.

Résumé

A l'heure où des dispositifs de réalité virtuelle se développent sur le marché, la technologie binaurale, le plus souvent assimilée à de « *l'Audio 3D* », est plus que jamais d'actualité. La possibilité d'avoir une écoute prenant en compte le suivi des mouvements de la tête de l'auditeur pose en effet de nouvelles questions concernant la production de contenus audio. La société 3D SoundLabs promet par exemple, grâce à son casque Sound One incorporant un système de suivi de la tête, « d'écouter d'une façon totalement nouvelle ». Cependant, les exemples développés dans leur communication ne concernent que l'adaptation de conditions d'écoute sur enceinte qui existe déjà.

En réalité, la technologie n'étant pas encore démocratisée, il n'existe pas aujourd'hui de contenu pensé en binaural avec suivi des mouvements de la tête.

Quel contenu serait-il adapté à cette technologie ?

Ce mémoire s'enracine dans ce questionnement. Le livre audio, en pleine expansion aux Etats-Unis, commence à s'implanter en France grâce à des acteurs numériques tels Audible ou AudioLib. Des initiatives binaurales ont déjà été menées dans le secteur du livre audio, mais ne prenaient pas en compte les mouvements de la tête. La démarche de ce mémoire s'est donc cristallisée autour de l'adaptation sonore d'un livre dont vous êtes le héros en binaural avec suivi des mouvements de la tête.

Comment penser et mettre en œuvre cette adaptation ?

Cette réflexion a donc mené à questionner l'adaptation sonore de ce genre littéraire tout en intégrant des techniques immersives comme le binaural et l'ambisonique dans le processus de création sonore. Ce travail peut permettre de faire émerger de nouvelles pistes de réflexions concernant la création sonore binaurale.

Cette recherche s'adresse à tous les ingénieurs du son intéressés par la technologie binaurale mais également aux concepteurs de produits interactifs, du transmédia au jeu vidéo, qui pourraient rechercher de nouvelles possibilités d'expressions au travers du son.

Mots clefs : Audio 3D, Techniques binaurales, Livre dont vous êtes le héros, contenu interactif, écriture sonore, Suivi des mouvements de la tête

Abstract

At a time when virtual reality devices is a fast developing market, head-tracked binaural technology, known as « 3D Audio », is more than ever in the limelight raising new questions about immersive content production. 3D SoundLabs swears that its headset Sound One allows us to « listen in a whole new way ». But the company merely emulates the listening conditions that already exist on loudspeakers.

In fact, this technology isn't used in consumer electronics yet, and today there is no specific content actually thought of for binaural technology with head-tracking.

What type of content could fit this technology?

This graduate research took its roots in this interrogation. Audio books, widely known in the United States, are beginning to be developed in France thanks to digital companies such as Amazon's Audible or AudioLib. Binaural experiments have already been made in this field, but none of these implement head-tracking. This approach has been focused on the adaptation of a gamebook using head-tracked binaural technology: Menu Cthulhu by Neil Jomunsi, Walrus Edition.

How to think and enforce this adaptation?

This thought has lead me to question the sonification of this kind of book and the integration of immersive technology such as binaural or ambisonics in the process of sound creation. This work might give new perspectives into immersive sound creation.

This research is intended for any sound engineer who shows an interest in binaural technology, as well as interactive content producers who are looking at new means of expression through sound.

Key words: 3D Audio, Binaural Technology, Ambisonic Technology, Audio Book, Gamebook, interactive content, sound writing, head-tracking

Table des matières

Introduction	7
Chapitre I : Des livres-jeux	15
<i>I.1 Petite histoire de la littérature interactive</i>	15
<i>I.2 Le Courage, la Chance et la Destinée, piliers des livres-jeux</i>	18
<i>I.3 Les livres-jeux, des exemples de proto-transmédia ?</i>	20
<i>I.4 Des jeux narrativistes</i>	24
Chapitre II :Vers l'aventure sonore dont vous êtes le héros !	27
<i>II.1 Menu Cthulhu et ses spécificités</i>	27
<i>II.2 Livre audio ou jeu audio ?</i>	30
<i>II.3 Le cahier des charges</i>	33
<i>II.3 Le creative agenda ou les attentes créatives des joueurs</i>	34
<i>II.4 Du texte au son, la nécessaire réécriture</i>	36
Chapitre III : Les techniques de reproduction sonore	39
<i>III.1 La technique binaurale</i>	39
<i>III.1.1 Historique</i>	39
<i>III.1.2 Caractéristiques physiques</i>	41
<i>III.1.3 Les outils</i>	44
<i>III.2 La technique ambisonique</i>	51
<i>III.2.1 Les limites de la synthèse binaurale</i>	51
<i>III.2.2 Historique</i>	52
<i>III.2.3 Principes physiques</i>	53
<i>III.2.4 Les outils</i>	57
Chapitre IV : La prise de son	65
<i>IV.1 Les voix</i>	66
<i>IV.2 Les ambiances</i>	68
<i>IV.3 Les effets</i>	70
<i>IV.4 Bilan de ces prises de son</i>	72

Chapitre V : Le dispositif mis en œuvre	73
<i>V.1 Le suivi des mouvements de la tête</i>	<i>73</i>
<i>V.2 L'interface utilisateur</i>	<i>78</i>
<i>V.3 Le workflow de production</i>	<i>79</i>
<i>V.4 Les outils logiciels</i>	<i>82</i>
Chapitre VI : La post-production	90
<i>VI.1 Un montage « classique »</i>	<i>90</i>
<i>VI.2 Pistes d'analyse d'une création sonore binaurale</i>	<i>91</i>
<i>VI.3 La technologie au service de la narration</i>	<i>95</i>
<i>VI.4 Etude de cas : le début de l'aventure</i>	<i>98</i>
Chapitre VII : Vers un développement du prototype	99
<i>VII.1 Critique du dispositif de diffusion</i>	<i>99</i>
<i>VII.2 Retour d'expérience utilisateur</i>	<i>100</i>
<i>VII.3 Pistes pour le développement du contenu</i>	<i>102</i>
<i>VII.4 Quid des normes ?</i>	<i>103</i>
Conclusion	107
Bibliographie	109
Discographie	115
Table des illustrations	116
Annexes	118
<i>Annexe 1 : Organigramme de Menu Cthulhu, le livre dont vous êtes le héros de Neil Jomunsi</i>	<i>118</i>
<i>Annexe 2 : Interface du jeu audio DeepSea.</i>	<i>119</i>
<i>Annexe 3 : Mon organigramme de Menu Cthulhu, l'aventure sonore dont vous êtes le héros</i>	<i>120</i>
<i>Annexe 4 : Tableau récapitulatif des différentes techniques de spatialisation sonore</i>	<i>121</i>
<i>Annexe 5 : Questionnaire distribuée à l'issue de la présentation publique du prototype</i>	<i>122</i>
<i>Annexe 6 : Auditrice (Amélie Hascoet) de Menu Cthulhu</i>	<i>123</i>

Introduction

L'espace est une invitation au voyage.

Bergame Périaux, Le Son Multicanal

Mon intérêt pour le binaural est né en 2011. Alors en première année de BTS Audiovisuel option son, j'ai été fasciné par la démonstration de QSound Labs « The Barber Shop»¹, qui m'avait été conseillée par un de mes professeurs. Seul dans ma chambre avec mon casque sur les oreilles, j'étais immergé dans un autre univers, parti me faire une nouvelle coupe de cheveux. Les possibilités offertes par cette technologie me semblaient alors infinies.

Mais qu'est-ce donc que le binaural ? « Le terme désigne l'écoute spatialisée au casque bien que la définition exacte du mot binaural soit "ayant trait aux deux oreilles" »². Ainsi, le binaural tend à reproduire une écoute naturelle. Cela m'émerveillait et me frustrait en même temps, car en parallèle de l'enseignement de mes professeurs de son sur la constitution et l'esthétique d'une image stéréophonique, je comprenais qu'il existait une technologie capable de passer outre la notion « d'image », qui ne constitue qu'une fenêtre face à une autre réalité.

Le binaural ouvre la possibilité de nous placer dans une autre réalité. Cela est rendu possible par nos capacités psycho-acoustiques. En effet, dans une écoute naturelle, la sensation de spatialisation d'un objet sonore est induite par des indices binauraux (de niveaux, de temps et de spectre), provenant des deux oreilles.

¹ [The Barber Shop sur Youtube](#)

² **PERIAUX Bergame, OHL Jean-luc, THEVENOT Patrick**, *Le son multicanal*, Paris, Edition Dunod, 2015

En parvenant à reproduire ces indices à la diffusion sur casque, le cerveau est alors dupe et se projette dans une autre réalité.

Elle peut avoir été enregistrée grâce aux technologies binaurales dites natives, qui consistent en un enregistrement stéréophonique grâce à une tête artificielle³ ou via des microphones placés au creux de l'oreille du preneur de son. Mais elle peut également être synthétisée, grâce à un traitement du signal qui émule ces indices binauraux.

Pris d'une véritable boulimie sonore, je cherchais davantage d'expériences binaurales avant de me tourner vers des productions plus conséquentes, avec une véritable narration. Nouvoson⁴, la plate-forme de Radio France sur le son multi-canal et binaural, a particulièrement forgé mon écoute et m'a montré l'étendue des possibilités du binaural sur des contenus variés, de la musique à la fiction en passant par le documentaire.

Pourtant la technologie binaurale ne s'est pas encore réellement démocratisée pour le grand public. Et pour cause, tout n'est pas si simple, car la technique binaurale est loin d'être parfaite !

Le problème majeur concerne l'individualité de l'écoute. En effet, les indices binauraux, nécessaires pour duper le cerveau, sont uniques pour chaque individu. Cela illustre parfaitement la subjectivité de l'écoute. Le binaural natif permet en quelque sorte d'écouter avec les oreilles du preneur de son ou avec celles d'une tête artificielle. Il en résulte des confusions de localisations très marquées, en particulier entre l'avant et l'arrière mais également des imprécisions dans le déplacement des sources ou encore des problèmes d'internalisation⁵ des sources sonores.

³ Une tête artificielle est un dispositif d'enregistrement binaural où les microphones sont placés à l'intérieur du pavillon d'un mannequin humanoïde, ce qui permet de recréer des indices binauraux.

⁴ <http://nouvoston.radiofrance.fr>

⁵ L'internalisation est un phénomène psycho-acoustique où l'auditeur perçoit les sources à l'intérieur même de sa tête. C'est ce qui se produit lorsque l'on écoute un contenu stéréophonique au casque

Au gré de mes études et au cours de mon stage de fin de deuxième année à Louis Lumière, dont l'objectif était de mener une expérience sur le binaural, j'ai eu l'occasion de me documenter puis de tester les avantages apportés par l'intégration d'un dispositif de suivi de mouvement de la tête à la technologie binaurale.

En effet un dispositif de *Head-Tracking*⁶ permettrait de réduire drastiquement les risques de confusions avant/arrière ainsi que le problème d'internalisation⁷. Par ailleurs cela me semblait l'aboutissement même de cette technologie : Reproduire une condition d'écoute quasi-naturelle, où le geste réflexe, inconscient, de la tête lors de la localisation d'une source sonore serait pris en compte.

C'est donc avec grand intérêt que j'ai suivi la campagne de financement participatif du casque 3D Sound One (autrefois appelé NEOH) de la start-up rennaise 3D Sounds Labs. Ce casque incorpore en effet directement un dispositif de mouvement de la tête et promet d'écouter « d'un façon totalement nouvelle »⁸.

Néanmoins, les exemples mis en avant dans la campagne de communication ne m'ont pas convaincu. En effet, la jeune start-up ne promet que des possibilités d'émulation d'enceintes virtuelles, d'un système stéréo pour la musique au 5.1 pour le cinéma et le jeu vidéo.

La « façon totalement nouvelle » d'écouter ne serait donc qu'une émulation au casque qui existe déjà ?

En réalité, en dehors de type de contenu liés à l'image comme la réalité virtuelle, les simulateurs et certains jeux vidéo, il n'y a actuellement aucun contenu permettant d'exploiter pleinement les possibilités offertes par ce dispositif en contenu audio seul.

Afin de faire émerger cette nouvelle écoute, il faudrait donc d'abord penser à de nouveaux types de contenu, qui mettent en valeur la technologie binaurale, particulièrement pour le grand public.

Quel contenu serait-il alors pertinent de proposer en binaural avec suivi des mouvements de la tête?

⁶ - Head-Tracking littéralement « suivre tête » est un technique permettant d'enregistrer les mouvements de la tête

⁷ **BEGAULT, WENZEL, ANDERSON**, « Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source », *Journal of the AES*, 2001

⁸ Présentation de 3D Sound One

Le binaural couplé à un dispositif de head-tracking laisse par définition la possibilité à l'auditeur de tourner la tête au sein d'une scène sonore et ainsi de choisir « vers où » il écoute. L'auditeur est alors dans une situation d'écoute subjective.

Cette notion de subjectivité a donc été mon fil rouge pendant la période de recherche d'un type de contenu. Il était primordial que l'auditeur soit totalement immergé dans une scène sonore pour que mon choix soit pertinent.

Au vu de ma formation et de mes centres d'intérêt très marqués par la production sonore et radiophonique, j'ai décidé de restreindre mon champ des possibles, en ne considérant pas les possibilités offertes par des dispositifs couplés à de l'image, bien que la Réalité Virtuelle et les jeux vidéo immersion semblent être par ailleurs un type de contenu tout indiqué pour cette technologie.

En marge de mon stage de fin de deuxième année, au fil des discussions avec mon maître de stage Matthieu Aussal, j'ai redécouvert le livre audio. Et pour cause, même s'il est en plein essor aux Etats-Unis grâce aux ventes dématérialisées⁹, le livre audio reste assez confidentiel en France. En 2013, seul 8% des Français avait déjà écouté un livre audio¹⁰.

Mon expérience du livre audio se résumait aux *Contes Musicaux* lus par Marlène Jobert, dont je me rappelle avoir été friand lorsque j'étais enfant. Destinés aux plus petits afin de leur faire découvrir la musique classique de façon plus ludique, mon éveil musical a été en grande partie forgé grâce à ces « Livres qui parlent ».

Si la musique a par la suite constitué une part importante de ma vie, c'est sans aucun doute grâce à elle que je me suis orienté vers les métiers du son, j'ai délaissé le livre audio, lui préférant des médiums audiovisuels.

Pourtant aujourd'hui, les grands éditeurs se penchent sérieusement sur le livre audio. Gallimard via sa collection « Ecoutez lire », Hachette et Albin Michel grâce à la maison d'édition Audiolib, Amazon via sa plate-forme Audible, tous ces éditeurs se sont investis dans le livre audio, prévoyant certainement une évolution du marché en France similaire à celle des Etats-Unis.

⁹ - Résultat de l'étude sur les ventes au sein de l'Audio Publishers Association américaine

¹⁰- **MUNZEL Andreas**, *Le livre audio en France et en Allemagne : Motifs, freins et contextes*, Mémoire d'économie, Strasbourg Business School, 2013

RoadBook, un éditeur indépendant, a même déjà exploité la technologie binaurale afin de proposer un livre audio « 3D Son »¹¹ .

Cette émulation autour du livre audio, que je ne soupçonnais pas jusqu'à lors, est devenue le terreau de ma réflexion. Pour autant, elle ne cadrerait pas directement avec le fil rouge que je m'étais imposé, à savoir la nécessité pour l'auditeur d'être dans la scène sonore.

En effet, les livres audio, d'abord pensés comme un moyen pour les mal-voyants d'accéder à la littérature, se résument bien souvent à des livres lus, avec ou sans ajout de musiques et d'ambiances.

Thérèse Raquin, le livre audio binaural de RoadBook se présente quant à lui, comme un « Film Audio », avec un véritable jeu d'acteurs, de la musique, des ambiances, des bruitages. J'ai été immédiatement attiré par cette esthétique sonore, où les frontières entre le livre audio et la fiction radiophonique, dont je suis plus familier, sont floues.

Le « Film Audio » serait-il le contenu que je recherchais ?

Curieux, j'ai contacté cet éditeur qui m'a joyeusement ouvert ses portes. Nous avons pu engager une discussion ouverte et franche et ils n'ont pas hésité à partager avec moi leurs « secrets de fabrication » et leur vision du livre audio. *Thérèse Raquin* a été réalisé presque exclusivement en binaural natif grâce à une tête artificielle.

J'ai pu apercevoir, grâce à des documents de travail, l'importance d'installer des marques très précises pour les comédiens. En effet, de cela résulte la spatialisation des sources sonores, qui se fait directement à la prise de son en binaural natif. De la même façon, afin d'avoir une image binaurale cohérente pour l'ensemble des auditeurs malgré des indices binauraux différents, peu de sources sonores était placées à l'arrière. Cela permettait d'éviter les confusions avant/arrière et les déplacements trop importants.

Ainsi à l'instar du cinéma, ce « Film Audio » est un type de contenu dont l'essentiel des informations sonores sont centrées. L'auditeur est face à un « écran sonore » tel un spectateur de théâtre.

Cette esthétique, bien que très intéressante, ne cadrerait pas avec mon projet, car cette esthétique tend à simuler une situation d'écoute qui existe déjà.

¹¹ - [Présentation de Thérèse Raquin, le livre audio "3D Son"](#)

Le déclic s'est produit un soir, devant mon ordinateur, en quête de nouvelles idées. De fil en aiguille, je finis sur la page de *Dépression Quest*, une fiction interactive partagée par une amie via Facebook.

Avec du texte et une histoire qui se développe grâce à des choix multiples, le joueur se met à la place d'un dépressif et peut donc au fil du jeu se rendre compte de la réelle souffrance des personnes atteintes de sa maladie.

Ce « *Serious Game* »¹² m'a rappelé instinctivement dans sa forme les livres dont vous êtes le héros de mon enfance.

A la fin des années 90, les livres dont vous êtes le héros, ou livre-jeux, étaient plutôt sur le déclin, concurrencés féroce­ment par les jeux vidéo. A l'époque, mes parents contrôlaient drastiquement ma durée d'exposition devant les écrans, particulièrement devant les jeux vidéo. Je considérais ces livres comme des palliatifs aux RPG¹³, les jeux vidéo de rôle qui sont encore aujourd'hui le type de jeux m'attirant le plus.

J'avais enfin trouvé le contenu que je voulais proposer, enfoui au plus profond de mes souvenirs d'enfance. Les livres dont vous êtes le héros me semblaient proposer une littérature subjective qui s'adaptait parfaitement au fil rouge que je m'étais fixé.

Ma démarche s'est donc cristallisée autour de l'adaptation sonore d'un livre dont vous êtes le héros. J'ai découvert qu'elle s'inscrivait par ailleurs dans le sillage d'un projet déjà existant. AudiGame¹⁴ a conçu une solution pour réaliser des adaptations sonores de fictions interactives pour des diffusions via internet ou plateforme mobile. Néanmoins, il ne propose pas de contenus binauraux avec suivi des mouvements de la tête.

Je me suis remis à lire ces livres que je n'avais plus touchés depuis l'enfance, en quête d'une adaptation. Il me fallait trouver un livre qui puisse emporter l'adhésion du

¹² - Le Serious Game est un jeu dont l'intention première est d'informer et d'instruire avant de divertir

¹³ RPG pour Role Playing Game désigne l'ensemble des jeux de rôle. Néanmoins, cet acronyme est le plus souvent utilisé pour désigner les jeux vidéo de rôle.

¹⁴ <http://audigame.fr>

plus grand nombre et qui donne à « entendre » des actions extraordinaires afin d'exploiter au maximum la technologie binaurale.

J'ai ainsi découvert *Menu Cthulhu* de Neil Jomunsi, une histoire déjantée dans laquelle le héros, serveur dans un fast-food, doit se défaire d'une créature hideuse apparue dans le frigo pendant son service. Le ton humoristique, absurde trouvait une résonance en moi et dès la première lecture, je pouvais *écouter* ce livre.

J'avais maintenant un contenu mais de nombreuses questions esthétiques comme techniques se posaient alors à moi.

Comment faudrait-il adapter *Menu Cthulhu* ? Proposer une transcription fidèle au texte original ou à l'inverse, penser à une réécriture du récit ?

Serait-ce encore un livre audio ?

Quelles technologies de prise de son et de post-production faudrait-il utiliser ?

Quel workflow¹⁵ et quel dispositif de diffusion serait-il pertinent de mettre en place ?

Il est possible d'englober ces questions dans la problématique suivante : comment penser et mettre en œuvre une adaptation sonore d'un livre-jeu en binaural avec suivi des mouvements de la tête ?

L'objectif de cette démarche est de faire émerger des nouvelles pistes de réflexion autour de l'écriture sonore binaurale, qui pourraient s'étendre à d'autres types de contenu, de la musique au documentaire radiophonique, en passant par le jeu vidéo ou la réalité virtuelle.

Cette réflexion, dont l'aboutissement est la création d'un prototype de « *Menu Cthulhu*, une aventure sonore dont vous êtes le héros », s'est organisée en sept chapitres.

Tout d'abord, il s'agira d'étudier les livres-jeux, média de base à ma démarche, dans leur singularité et d'en extraire les principes constitutifs de ce genre de littérature.

¹⁵- Workflow, anglicisme pour flux de travaux, est la représentation de la suite des tâches à effectuer dans le but de produire un objet

Ainsi il sera possible d'analyser et de questionner la pertinence de mon choix de sujet de départ, *Menu Cthulhu* ».

Il faudra ensuite s'interroger sur l'adaptation sonore en binaural avec suivi des mouvements en elle-même et analyser ce qu'elle peut apporter en terme de narration et ainsi questionner la réécriture qu'elle engendre. Le format *livre audio*, traditionnellement à destination des personnes malvoyantes, doit également être pris en compte. Ces réflexions permettront de mettre en place un cahier des charges qui sera le fil rouge tout au long de ce mémoire.

Nous pourrons alors nous concentrer sur les techniques de reproduction sonore qu'il serait pertinent d'utiliser et nous pourrons ainsi dresser un état de l'art au travers des outils disponibles sur le marché. Cela me permettra, grâce à des essais, de décider des technologies et des outils à utiliser pour cette adaptation.

Il faudra par ailleurs décider des techniques de prise de son que j'utiliserai, en développant les avantages et les inconvénients inhérents à chaque technologie et ainsi de comprendre et de prévoir ainsi les choix à effectuer lors du tournage.

Nous nous pencherons par la suite sur le dispositif mis en œuvre pour la réalisation de cette adaptation. Je développerai mes choix techniques concernant l'interface, le head-tracker ainsi que le lecteur et le protocole de transmission utilisés. Cela permettra de faire émerger un workflow de production.

La post-production de *Menu Cthulhu* permettra ensuite d'éprouver cette chaîne de travail. Cela nous permettra de nous questionner sur les possibilités et les limites du montage et du mixage dans le cadre de la production d'un contenu *pensé* dès le départ en binaural avec suivi des mouvements de la tête.

Une fois le prototype terminé, les questions relatives à un éventuel développement se poseront. L'esquisse d'un workflow de diffusion, au regard des nouvelles normes qui se développent aujourd'hui sera questionnée. Il s'agira également de vérifier au travers d'écoutes critiques, que les choix effectués aient été perçus par un public.

Chapitre I : Des livres-jeux

I.1 Petite histoire de la littérature interactive

« - 1. Voulez-vous connaître l'histoire des trois alertes petits pois ?

- Si oui, passez à 4.

- Si non, passez à 2. »¹⁶

Ainsi débute le « conte à votre façon » écrit dès 1967 par Raymond Queneau, alors chef de file de la littérature oulipienne. OuLiPo, acronyme désignant l'Ouvroir de Littérature Potentielle, est une association fondée en 1960 par le mathématicien François Le Lionnais et Raymond Queneau, écrivain et poète.

L'association, d'inspiration surréaliste, se définit avant tout par ce qu'il n'est pas¹⁷ :

Ce n'est pas un mouvement littéraire.

Ce n'est pas un séminaire scientifique.

Ce n'est pas de la littérature aléatoire.

L'OuLiPo, peut être considéré comme un groupe de recherche expérimentale.

En postulant que les contraintes formelles sont des stimulants puissants pour l'imaginaire, le mouvement a permis de voir émerger des nouvelles formes de littérature, dont la littérature interactive.

Le conte à votre façon, avec ses vingt et un « paragraphes » de quelques lignes, semble avoir été la première œuvre à avoir établi un récit dicté par les choix du lecteur¹⁸.

¹⁶ **QUENEAU Raymond**, « Un conte à votre façon », in *La littérature potentielle*, Paris, Gallimard, coll. Folio Essais, 1973, pp.273-276.

¹⁷ **Oulipo**, *La littérature potentielle*, Paris, Gallimard, coll. Folio Essais, 1973, p.8

¹⁸ **BOUYALA-FROSIN Bernadette**, « Les livres dont vous êtes les héros » in *Actes de Lecture* n°21, Mars 1988, p.3

Néanmoins, il faudra attendre 1981 pour que la littérature interactive gagne le cœur du grand public. « *The Warnock of Firetop Mountain* », écrit par les anglais Steve Jackson et Ian Livingstone est le *Gamebook*, littéralement livre-jeu qui démocratisera le genre.

Son succès peut s'expliquer par la reprise des principes des jeux de rôle, pratiqués à l'époque dans des clubs « d'initiés », souvent d'origine universitaire.

En France, le livre sera édité par Gallimard dans la collection Folio-Junior en 1983 sous le titre « *Le Sorcier de la Montagne de feu* ».

Les 15 000 premiers exemplaires ont été épuisés en deux mois et en 1986, 200 000 exemplaires en étaient vendus. Le milieu des années 80 constituait l'apogée de ce genre de littérature et Gallimard en a naturellement profité pour créer sa collection « Les Livres dont vous êtes le héros ».

Cette dénomination est aujourd'hui encore la plus usité en France pour désigner ces *Gamebook*. Ce genre s'est principalement destiné à une population jeune et s'est donc largement implanté chez les adolescents des années 80.

Les années 90 marqueront le déclin du genre. Les éditeurs arrêtent les uns après les autres leur collection de livre-jeu, dû au déplacement d'intérêt des jeunes vers le multimédia émergent grâce à la démocratisation des ordinateurs ainsi que du développement des premiers jeux vidéo de rôles, aussi appelés communément RPG.

Le renouveau s'amorcera doucement dans les années 2000 grâce à au développement des fictions interactives via navigateur internet et se poursuivra dans les années 2010 grâce à l'émergence des e-books, sur ordinateur et supports mobiles mais également celle des réseaux sociaux¹⁹, qui réduisent les coûts de production.

Les livres-jeux, de par leur essence hybride, sont souvent perçus de manière très mitigée par les communautés pédagogiques et littéraires.

¹⁹ <http://www.numerama.com/pop-culture/127741-jouiez-a-livre-dont-etes-heros-twitter.html>

Certains professeurs, dans les années 80, dénonçaient la propriété littéraire de ces œuvres, les accusant d'être « tout au plus un manuel ou un mode d'emploi qui, tare impardonnable, suinte l'ennui », alors que d'autres les percevaient comme un moyen d'accéder à la lecture pour les « mauvais » lecteurs.

Emmanuel Souchier, Professeur des Universités à la Sorbonne, considère cette littérature comme « palpitante et ne manquant pas d'humour »²⁰, mais émet plus de réserves concernant la moralité de ces œuvres, et indique qu'une « adhésion au principe du jeu entraîne une adhésion au discours idéologique du texte avec la caractéristique d'un esprit critique émoussé ». Ces réserves se retrouveront dans les critiques du jeu vidéo narratif qui émergera dans les années 90.

Les analyses sur ce genre se sont d'abord portées sur la notion de livre avant la notion de jeu.

Espen Aarseth, chercheur en littérature et ludologie à l'université de Bergen, qualifie le genre de « littérature ergodique »²¹, du grec ergon, travail, et odos, chemin, puisque le lecteur est invité à « travailler » pour trouver le « chemin » au sein de la narration.

Les livres dont vous êtes le héros peuvent être considérés comme de la littérature métafictionnelle. Cela signifie que l'esthétique du texte écrit n'est plus forcément au centre de la pensée créative et que le livre, cherchant lui-même l'interaction avec le lecteur, permet de poser des questions sur ce médium.

Et pourtant, en héritier des premières parties de jeu de rôle, les premiers livres-jeux étaient sophistiqués. Le lecteur « héros » devait souvent avant même de commencer sa lecture établir une fiche d'aventure. Des dés sont parfois nécessaires et de nombreuses règles doivent être assimilées par le lecteur.

Même si le lecteur se retrouve seul face à l'aventure, il se retrouve dans le même contexte ludique d'une partie de jeu de rôle à plusieurs. Ainsi les livres-jeux ne peuvent pas être simplement résumés à la mécanique d'un « conte à votre façon », où le récit avance au rythme des choix proposés en fin de chapitre.

²⁰ **SOUCHIER Emmanuel**, « Les Livres dont vous êtes le héros », in *Trousse-Livres*, n°58

²¹ **AARSETH Espen**, « The joy of reading role-playing games » in *The Guardian*, 19 juin 2015

I.2 Le Courage, la Chance et la Destinée, piliers des livres-jeux

« – Vous êtes mon épée ? murmure-t-il. »

« – Vous m'avez conquise par votre courage, par votre chance, par votre destinée, dit la voix. » ²²

Cet extrait de « *La horde des démons* » de J. H. Brennan synthétise les principes fondateurs des livres dont vous êtes le héros.²³

Le courage symboliserait l'ensemble des choix effectués par le lecteur-héros.

La chance constituerait le hasard. Hasard du à l'aléatoire d'un lancer de dés dans certains livres ou encore le hasard induit par l'inconnu - le lecteur-joueur ne connaissant pas à priori les chapitres suivants.

La destinée serait quant à elle la pensée de l'auteur. Pensée toute puissante car elle émerge de tous les différents embranchements de l'histoire.

Ainsi, il est plus intéressant d'analyser les effets produits par le livre sur le lecteur que l'étude formelle de la mécanique du livre-jeu, un enchaînement de chapitres grâce à des choix multiples ou au hasard.

Cette mécanique permet aux livres-jeux de générer « plusieurs histoires », qu'il peut être pertinent de représenter sous la forme d'un graphe.

Le genre a une tendance dirigiste sur sa narration, à savoir qu'il n'existe qu'une seule « bonne » fin, le personnage principal ayant un but précis et la seule fin « satisfaisante » est celle où il atteint ce but.

²² **BRENNAN J.H.**, *La Horde des démons*, Paris, Folio Junior, 1984, §31

²³ <http://contentisart.com/fort-mcmoney-le-courage-la-chance-et-la-destinee/>

Certains livres proposent une approche plus convergente des différents chapitres, c'est à dire « qu'un paragraphe auquel on peut accéder par plus d'un paragraphe peut avoir des sens totalement différents selon ce qui s'est passé auparavant »²⁴.

D'autres livres-jeux, comme *Le prisonnier*²⁵ ou *Les Démons des profondeurs*²⁶, essaient d'abolir la dualité réussite/échec et possèdent plusieurs fins « réelles ». Ainsi, le lecteur est invité à se promener à sa guise dans un monde ouvert et la notion même d'avancement est remise en question.

Il est ainsi possible de classer les livres-jeux ²⁷ :

- Par leur linéarité, c'est à dire la variété des chemins proposés. Un récit très linéaire se rapprochera d'un livre normal alors qu'un récit avec de nombreux embranchements proposera une expérience nouvelle à chaque relecture.
- Par leur difficulté, dû au hasard d'un lancer de dés qui détermine la poursuite de l'histoire, ou alors dû à l'exigence du livre d'avoir un « One-true-path », littéralement un seul vrai chemin. Cela consiste en un enchaînement de paragraphes dans un ordre précis, qui s'il n'est pas rempli ne permet pas d'avancer dans le récit.
- Par la qualité intrinsèque du scénario et les qualités littéraires de l'auteur, bien évidemment, qui rendent ainsi la narration du *livre* plus intéressante que le côté ludique du *jeu*. Ainsi, les livres-jeux peuvent également faire partie de la littérature interactive, et donc de plus globalement de la littérature, à savoir qu'ils peuvent faire partie de « l'ensemble des œuvres écrites auxquelles on reconnaît une finalité esthétique ».²⁸

²⁴ **MENELDUR**, « Interview de Paul Mason » in *La Bibliothèque des aventuriers*, 2009

²⁵ **ROSENTHAL Pierre**, *Le Prisonnier*, Paris, Le Livre de Poche, coll. « Histoires à jouer: Missions spéciales », 1988

²⁶ **JACKSON Steve**, *Les Démons des profondeurs*, Paris, Gallimard, coll. « Défis fantastiques », 1986

²⁷ **SAGOT Gildas**, *Jeux de rôle : tout savoir sur les jeux de rôle et les livres dont vous êtes le héros*, Paris, Gallimard, 1986

²⁸ Définition de Littérature dans le Dictionnaire Larousse

I.3 Les livres-jeux, des exemples de proto-transmédia ?

Les livres-jeux sont généralement portés par un récit interactif. Or le récit interactif, dans son essence même, semble relever d'une contradiction²⁹. La narrativité implique de *prendre par la main* le lecteur pour lui délivrer un récit, du début à la fin. L'interactivité, au contraire, *donne la main* au lecteur pour intervenir au cours du récit, de multiples façons (histoire, structure du récit, narration).

Dans le cas de la littérature, le paradoxe du récit interactif paraît exacerbé. Plus le rôle donné aux « lecteurs-acteurs » est important, plus la narration s'affaiblit, plus l'interactivité s'impose, plus la littérature semble *s'absenter*³⁰.

Le récit littéraire interactif présuppose plusieurs points.³¹

D'une part, la présence d'une succession d'événements constituant une *histoire*.

Le mode de représentation principal de cette histoire doit être une *narration*, le lecteur doit alors vivre une histoire avec la médiation d'un narrateur.

Le contenu se doit d'être *interactif*, proposant un dispositif permettant une intervention matérielle du lecteur.

Enfin, la littérature étant « l'usage esthétique du langage écrit »³², le récit littéraire interactif doit comporter une composante textuelle, ce qui exclut les récits fondés uniquement sur l'image.

Le terme regroupe néanmoins des réalités diverses et renvoie à un vaste champ de pratiques, des récits *hypertextuels* aux récits *procéduraux*³³ en passant par les récits *collectifs* et bien sûr les livres dont vous êtes le héros. Les limites sont d'ailleurs floues et renvoient très souvent à d'autres domaines.

²⁹ **BOUCHARDON Serge**, « Le récit littéraire interactif : une valeur heuristique » in *C&L*, Janvier 2008 p.1

³⁰ **CLEMENT Jean**, « La cyberlittérature entre jeu littéraire et jeu vidéo » in *Formules n°10*, Paris, Agnès Vienot, juin 2006.

³¹ **BOUCHARDON Serge**, « Le récit littéraire interactif : une valeur heuristique » in *C&L*, Janvier 2008 p.2

³² Définition du *Trésor de la Langue française*

³³ Un récit procédural est un récit générée par des algorithmes. Ainsi, chaque relecture est entièrement nouvelle.

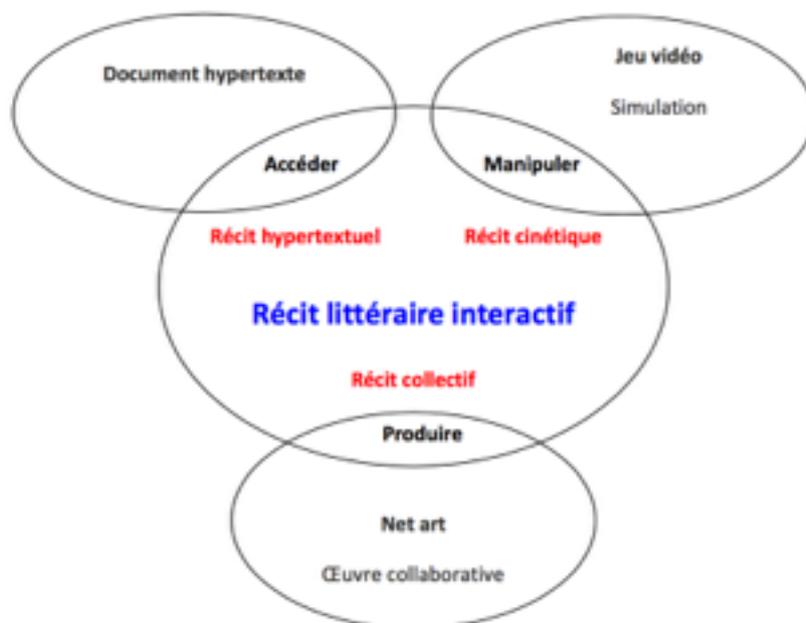


Figure 1. Les frontières du récit littéraire interactif³⁴

Pour Serge Bouchardon, professeur à l'Université de Technologie de Compiègne, le récit littéraire interactif possède une dimension *heuristique*³⁵. « Pour le chercheur, l'intérêt de ces récits littéraires interactifs ne réside pas tant dans la valeur des productions que dans leur faculté d'interrogation : du texte, du récit et de la littérature elle-même. »³⁶.

Ainsi, la linéarité comme la temporalité du récit, évidente dans le cadre d'un récit littéraire classique, se doivent d'être questionnées au travers de chaque forme de littérature interactive.

La « valeur » de ces récits éclatés doit-elle seulement se mesurer à leur qualité *littéraire* ?

La littérature interactive et les livres-jeux en particulier, grâce au caractère paradoxal du récit interactif, permettent de constituer un pacte entre l'auteur et le lecteur.

³⁴ **BOUCHARDON Serge**, « Le récit littéraire interactif : une valeur heuristique » in *C&L*, Janvier 2008 p.12

³⁵ **BOUCHARDON Serge**, *Le récit littéraire interactif. Narrativité et interactivité*, thèse soutenue le 7 décembre 2005 à l'Université de Technologie de Compiègne

³⁶ **BOUCHARDON Serge**, « Le récit littéraire interactif : une valeur heuristique » in *C&L*, Janvier 2008 p.4

Ainsi, un processus narratif et immersif se crée au sein de chaque œuvre, mettant à égalité l'auteur, qui s'est chargé de créer une multitude de possibilités et d'univers possibles, et le lecteur, qui se charge de développer une narration personnelle.

Il me semble que c'est dans l'élaboration de ces processus que réside également la « valeur des productions ».

Les livres-jeux se rapprocheraient ainsi du « transmedia storytelling »³⁷ dans leur fabrication, à la différence près que les livres-jeux sont pensés uniquement pour le médium « livre » alors que le transmédia a vocation à se répandre sur plusieurs types de média.

Par ailleurs, le transmédia, en tant que vecteur artistique et esthétique, a « son point d'origine au tournant du XXI^e siècle »³⁸, alors que les livres-jeux ont eu leur apogée dans le milieu des années 80.

Les livres dont vous êtes le héros ne pourraient-ils pas être considérés comme des *proto-transmédia*³⁹, des exemples de « transmédia d'avant le transmédia »⁴⁰ ?

Dans *Histoire du Transmédia*, Sébastien Denis, chercheur associé au CNRS, pose dès l'introduction une problématique qui englobe mon questionnement :

« Le but de cet ouvrage est d'interroger sur le fait de savoir si le transmédia ou le transmédia *storytelling* (pour insister sur la circulation de contenus narratifs) ont pu préexister ou non aux matériels récents qui ont permis l'émergence de ce qu'on appelle le transmédia. »

En réalité, il y a une foule de réponses différentes à ces questions, pratiquement autant de positives que de négatives.

³⁷ « Le transmédia storytelling est un processus par lequel des éléments autonomes constitutifs d'une fiction se trouvent dispersés à travers plusieurs médias, afin de créer une expérience de divertissement unifiée et coordonnée. » **JENKINS Henry**, *La culture de la convergence*, Paris, Armand-Colin, 2013, p.22

³⁸ **CHATELET Claire**, « Retour vers le futur : le transmédia d'avant le transmédia » in *Histoire du Transmédia*, Paris, L'Harmattan, 2014, p.49

³⁹ **DENIS Sebastien**, « Vers un « proto-transmédia » ? » in *Histoire du Transmédia*, Paris, L'Harmattan, 2014, p.21

⁴⁰ **CHATELET Claire**, « Retour vers le futur : le transmédia d'avant le transmédia » in *Histoire du Transmédia*, Paris, L'Harmattan, 2014, p.29

Frank Rose, un auteur majeur sur la culture digitale aux Etats-Unis, avoue lui-même que le processus narratif et immersif utilisé par les projets transmédia qu'il décrit dans *The Art of Immersion*, peut remonter jusqu'à *Dungeons & Dragons* en 1974 ⁴¹. Or ce jeu de rôle sur plateau a lui-même largement inspiré les livres-jeux, qui ont hérité de sa structure ludique.

D'autres chercheurs, comme Claire Chatelet et Marida Di Crosta, sont plus nuancés. Il est impossible pour elles de parler de transmédia avant le début des années 2000, car l'interactivité n'était pas assez développée dans les autres médias⁴².

La variété des points de vue montre à quel point le transmédia peut faire l'objet de lectures différentes.

Je ne peux que recommander aux lecteurs curieux de lire « *Histoire du Transmédia* » et « *La culture de la Convergence* » de Henry Jenkins, un des pères fondateurs du Transmédia, s'ils souhaitent découvrir davantage les questionnements relatifs au transmédia.

Par rapport à ma démarche d'adaptation sonore d'un livre dont vous êtes le héros, qui peut d'ailleurs être considérée comme une démarche *crossmédia* ⁴³, le caractère *proto-transmédia* des livres-jeux me permet de m'ouvrir à d'autres champs artistiques et de m'inspirer des pratiques transmédiatiques existantes.

En effet, il sera nécessaire d'établir de nouveaux types d'interactions dans le cadre de ce projet et certaines oeuvres transmédia, comme les web-documentaires de Samuel Bollendorff, qui a utilisé l'interaction comme moyen de faire vivre de nouvelles expériences aux spectateurs.

⁴¹ **ROSE Frank**, *The Art of Immersion. How the Digital Generation is Remaking Hollywood, Madison Avenue, and the Way We Tell Stories*, New York, Norton and company, 2011

⁴² **DENIS Sebastien**, « Vers un « proto-transmédia » ? » in *Histoire du Transmédia*, Paris, L'Harmattan, 2014, p.13

⁴³ L'expression *crossmédia* désigne la diffusion d'un même programme sur différents supports et franchises.

I.4 Des jeux narrativistes

Les livres dont vous êtes le héros sont très majoritairement des livres-jeux où le lecteur incarne un rôle déterminé et possède un but précis. A ce titre, les livres-jeux peuvent être analysés de la même façon qu'un jeu de rôle.

Plusieurs théories du jeu de rôle existent et sont souvent menées par des créateurs ou des meneurs de jeu, désireux de connaître les motivations des joueurs afin d'adapter leurs parties, grandeur nature, sur table ou par internet.

Une des théories les plus répandues est la théorie LNS (pour Ludisme, Narrativisme, Situationnisme) de Ron Edwards, un auteur de jeux de rôle indépendant tel que *Sorcerer* ou *Trollbabe*. Il a notamment participé à fonder le forum *The Forge*⁴⁴, qui abrite une communauté virtuelle de soutien à la conception et l'édition de jeux de rôle. Ron Edwards, n'est donc pas un universitaire et ses écrits se retrouvent majoritairement sur *The Forge*.

Edward a remis en cause avec son article «*System Does Matter*»⁴⁵ le modèle tripartite apparu en 1997 au sein des communautés de *rôliste*⁴⁶. En particulier le *Threefold Model*⁴⁷ pensé par Mary Kuhner et représenté graphiquement par Irina Rempt sous la forme d'un triangle :

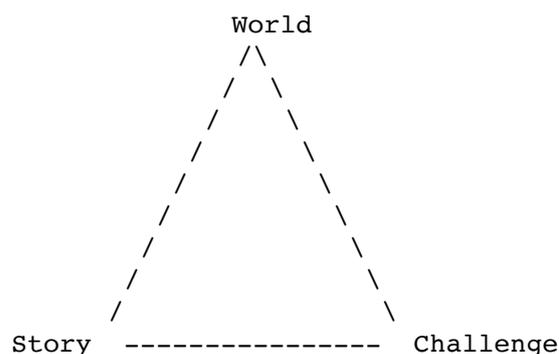


Figure 2. Représentation du Threefold Model par Irina Rempt⁴⁸

⁴⁴ <http://www.indie-rpgs.com/forge/index.php>

⁴⁵ http://www.indie-rpgs.com/articles/system_does_matter.html

⁴⁶ Un rôliste est une personne participant à des jeux de rôles. Le terme est souvent prêté aux personnes participant à des jeux de rôles grandeur nature.

⁴⁷ <http://www.darkshire.net/~jhkim/rpg/theory/threefold/>

⁴⁸ [Conversation du groupe de rôliste](#)

Ron Edward conserve l'organisation tripartite mais au lieu de partir de la construction et de l'orientation ludique du créateur ou du maître du jeu (*les épreuves, l'univers et le récit*), il s'intéresse davantage aux objectifs et aux décisions des joueurs en tant que groupe (*le ludisme, le narrativisme et le simulationnisme*).

En partant du *Contrat social*, à savoir qu'une ou plusieurs personnes se mettent à jouer dans un cadre déterminé, les joueurs partent ainsi en période d'exploration au sein de l'univers du jeu.

Cette exploration comprendra le plus souvent cinq composantes :

- Les personnages
- Le contexte
- Des situations
- Un système au sens large qui permet d'évoluer au sein de ces situations
- Une « couleur », qui serait une perception sensible particulière de la toile de fond

De cette exploration, les joueurs peuvent dégager les *Creative agenda*, les attentes créatives du concepteur du jeu. Elles pourront être satisfaites grâce à des « techniques », qui peuvent être largement de tout type d'actions menant à des conséquences sur le déroulement du jeu.

Enfin, de ces techniques résultent des « *ephemera* », (choses éphémères) qui sont des événements très courts signifiant le bon déroulement du jeu.

Les attentes créatives sont ainsi polarisées autour de trois notions ⁴⁹ :

- Le ludisme, qui est la capacité des joueurs à prendre des risques ou encore à établir une tactique.
- Le narrativisme, qui est la capacité des joueurs à élaborer un récit.
- Le simulationnisme, qui est la capacité des joueurs à s'immerger dans une autre réalité, et où les mécaniques de jeu contribuent à une impression de réalité.

⁴⁹ http://www.indie-rpgs.com/articles/system_does_matter.html

Au regard de cette théorie, les livres dont vous êtes le héros sont essentiellement narrativistes, à savoir que l'ensemble des mécaniques de jeux amène le joueur à user de son imagination afin d'élaborer le récit qui lui correspond le mieux.

Ces livres peuvent également proposer des mécaniques plus ludiques comme l'élaboration d'une fiche de personnage, l'usage de dés et d'énigmes.

Enfin certains livres peuvent avoir une composante situationniste non négligeable en proposant au lecteur-joueur de se déplacer dans un monde « ouvert », où il peut se promener à sa guise.

Entre 1999 et 2005, *The Forge* aura abrité les discussions concernant cette théorie au travers de trois articles spécialisés⁵⁰ ⁵¹ ⁵² et aura permis de faire naître le *Big Model*, qui résume l'ensemble de la théorie.

Pourtant aujourd'hui, Ron Edward considère sa théorie obsolète⁵³ et d'autres modèles sont apparus tels la théorie des niveaux d'immersion multiples⁵⁴ ou encore le modèle intégratif/dissipatif⁵⁵, particulièrement pensés pour les jeux GN (Grandeur Nature).

J'ai néanmoins décidé d'utiliser cette théorie comme cadre de réflexion pour mon adaptation d'un livre-jeu.

Cette grille d'analyse peut s'appliquer sur un ensemble très disparate de jeux de rôle et c'est pour les cas les plus complexes qu'il a proposé de nouvelles théories.

Le livre-jeu, à l'inverse, reste une situation de jeu de rôle simple, le lecteur étant seul face à l'auteur/maître du jeu. Avec la théorie LNS, il me semble déjà possible d'avoir une vision globale et critique sur les mécaniques de jeux que je vais mettre en oeuvre dans le cadre de mon adaptation. Il me sera également possible d'interroger les attentes créatives des joueurs sur le contenu que je propose.

⁵⁰ **EDWARD Ron**, « The right to Dream » in *The Forge*, 2003

⁵¹ **EDWARD Ron**, « Narrativism: Story Now » in *The Forge*, 2004

⁵² **EDWARD Ron**, « The whole model - this is it » in *The Forge*, 2003

⁵³ <http://indie-rpgs.com/archive/index.php?topic=12794.0>

⁵⁴ <http://ptgptb.fr/theorie-des-niveaux-d-immersion-multiples/>

⁵⁵ <http://ptgptb.fr/index.php/jdr-et-theorie-du-chaos/>

Chapitre II :Vers l'aventure sonore dont vous êtes le héros !

II.1 *Menu Cthulhu* et ses spécificités

Plus conscient maintenant de l'ensemble des spécificités et de l'héritage des livres dont vous êtes le héros, analysons le livre-jeu que je souhaite adapter : *Menu Cthulhu*⁵⁶ de Neil Jomunsi.

Menu Cthulhu est un livre-jeu numérique sorti en 2012, pendant l'esquisse du renouveau du genre grâce aux technologies mobiles. Il a été écrit par Neil Jomunsi⁵⁷, un écrivain berlinois résolument contemporain, puisqu'il ne tient à publier que sous format numérique.

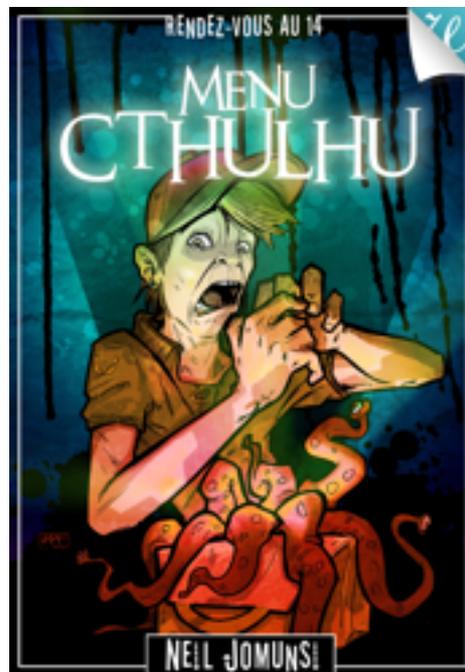


Figure 3. Artwork du livre *Menu Cthulhu*, propriété des éditions Walrus

⁵⁶ <http://www.walrus-books.com/menu-cthulhu-une-nouvelle-dont-vous-etes-le-heros/>

⁵⁷ <http://page42.org/a-propos/>

Il collabore le plus souvent avec les Editions numériques Walrus, éditeur de littérature alternative tel que les romans noirs, les *pulps*⁵⁸ et bien sûr les livres dont vous êtes le héros.

Le format *ebook*, proposé par Walrus, permet de passer d'un chapitre à l'autre d'un simple clic. Les livres dont vous êtes le héros numérique se rapprochent ainsi de la littérature hypertextuelle⁵⁹.

Le résumé du livre est simple : Le *lecteur-héros* est un employé de fast-food qui devra gérer, alors qu'il est seul responsable de la fermeture, l'arrivée d'un monstre à tentacule qui a surgi du réfrigérateur du restaurant. Ainsi, le lecteur devra faire de son mieux pour *régler* ce problème.

L'originalité de ce livre réside dans sa gestion du récit. En effet, quels que soient les choix effectués par le narrateur, la mort est inévitable. Ainsi, il n'y a pas de *bonne* fin à trouver, la fin la « moins pire » étant celle où le héros a survécu mais l'apocalypse a eu lieu.

Le lecteur, libéré de la contrainte de trouver La fin , pourra alors relire ce livre afin de découvrir les nombreuses façons de passer l'arme à gauche.

Le lecteur est alors en perpétuelle exploration au sein de ce livre. L'intérêt réside donc plus dans les détails et « l'atmosphère » du livre, qui correspondrait au contexte et à la couleur de l'univers du livre-jeu dans la théorie LNS.

L'intérêt premier est donc niché dans la capacité du récit à surprendre, émouvoir ou faire rire le lecteur. La mécanique de jeu, quand à elle, reste celle *classique* des livres-jeux - Un enchaînement de paragraphes à choix multiples -, et n'offre, au regard de la théorie LNS, qu'une part limitée de ludisme.

⁵⁸ Les pulps sont historiquement des publications, populaire aux Etats-Unis au début du XXe siècle, peu coûteuses et centré sur des fictions. Le style de ces fictions, souvent présentée comme la narration de faits réels, a fait du terme pulp un genre littéraire.

⁵⁹ **CAUVIN Aurélie**, *La Littérature Hypertextuelle, analyse et typologie*, Maitrise de Lettres Modernes 2001, Université de Cergy Pontoise

Néanmoins cet enchaînement propose une approche convergente de certains chapitres (le chapitre 6, 5 ou 11 par exemple⁶⁰), c'est à dire qu'il y a plusieurs *chemins* pour arriver au même chapitre et que les chapitres seront perçus différemment selon le chapitre choisi précédemment. Cela permet de décupler les possibilités d'explorations au sein du récit.

Ce livre-jeu a donc essentiellement une dimension narrative bien que l'aventure racontée puisse être considérée comme une « projection » ou une simulation. En effet, le lecteur a un rôle et un but précis : échapper au monstre.

Le lecteur peut se projeter au sein de ce rôle mais sera limité par la narrativité induite par les différents choix et chapitres du livre.

Il y a donc une tendance simulationniste au sein de *Menu Cthulhu*, qui est incorporée à la dimension narrative globale et dominante.

La composante ludisme est néanmoins plus faible, et je souhaite dans mon adaptation équilibrer les différentes composantes de la théorie LNS, afin d'offrir une expérience globale à l'auditeur. Il faudrait donc créer davantage d'interactions dans mon adaptation. Le binaural avec suivi des mouvements de la tête, point de départ de ma réflexion, pourrait être utilisé comme un élément interactif.

Menu Cthulhu possède tous les atouts pour être adapté de façon sonore.

Il questionne la notion de fin et de progression au travers de son récit interactif, il possède donc une dimension *heuristique*, typique de la littérature interactive.

Il possède une empreinte narrative très marquée, signe distinctif des livres-jeux, mais connaît une tendance simulationniste qui pourrait me servir dans le cadre de mon adaptation.

⁶⁰ Voir Annexe 1 : Organigramme de « Menu Cthulhu » de Neil Jomunsi

II.2 Livre audio ou jeu audio ?

Maintenant, il me faut commencer à imaginer *Menu Cthulhu* de façon sonore.

Une des premières questions préliminaires à l'adaptation de *Menu Cthulhu* qui s'est posée à moi a concerné la *forme* qu'elle prendrait. En effet, ma démarche n'était pas si simple à classer.

D'une part, on peut considérer qu'elle est une adaptation d'un livre et qu'à ce titre, mon adaptation serait un livre audio. C'est en effet, une « "œuvre de l'esprit" comportant un support sonore »⁶¹.

Néanmoins traditionnellement, un livre audio est considéré comme un flux, stocké et transmis de façon dématérialisée (en format numérique) ou matérialisée (en CD par exemple). Or l'adaptation d'un livre-jeu nécessite la mise en place d'un dispositif d'interaction avec le lecteur, pratique impossible avec les supports physiques historiques des livres audio.

A l'inverse, les jeux audio, qui s'adressent également majoritairement à un public déficient visuel, ont permis l'émergence de dispositif d'interaction entre le joueur et le jeu.⁶²

Certains jeux, à l'instar de *A Blind Legend*⁶³ du studio Dowino, reprennent des interfaces déjà existantes (surface tactile de smartphone, clavier et souris) et les adaptent de façon à ce que les contrôles du jeu soit accessibles aux personnes malvoyantes.

D'autres jeux tels que *Deep Sea*⁶⁴ ont développé une interface de contrôle dédiée, ce qui permet de développer leur *Gameplay*⁶⁵ jusqu'à dans leur interface.

⁶¹ Définition du livre audio par le Syndicat National de l'Édition française

⁶² **DIENER Elliot, TOULON Arnaud**, *Réalisation d'un jeu audio*, Mémoire de fin d'étude de Louis-Lumière, 2015, p.48

⁶³ <http://www.ablindlegend.com>

⁶⁴ **ARNOTT Robin**, *Deep Sea*, 2006

⁶⁵ Le Gameplay correspond à l'ergonomie du jeu, la façon dont on y joue.

Dans le cas de *Deep Sea*, le jeu étant centré sur la privation sensorielle sous la mer, l'interface a même pris la forme d'un masque à gaz afin de rendre le jeu plus immersif encore.⁶⁶

Je m'aperçois qu'afin de continuer ma réflexion sur la classification de mon adaptation, je dois bien distinguer la notion d'immersion et la notion d'interactivité.⁶⁷

Jacques Tisseau, professeur des universités à l'ENIB dont les recherches portent principalement sur la réalité virtuelle, a proposé un schéma situant différentes applications numériques suivant trois axes principaux :

- Autonomie du monde simulé
- Immersion de l'utilisateur dans et avec le monde simulé
- Interaction de l'utilisateur dans et avec le monde simulé

Une sensation de *présence* peut être caractérisée par des combinaisons des critères d'immersion et d'interaction.

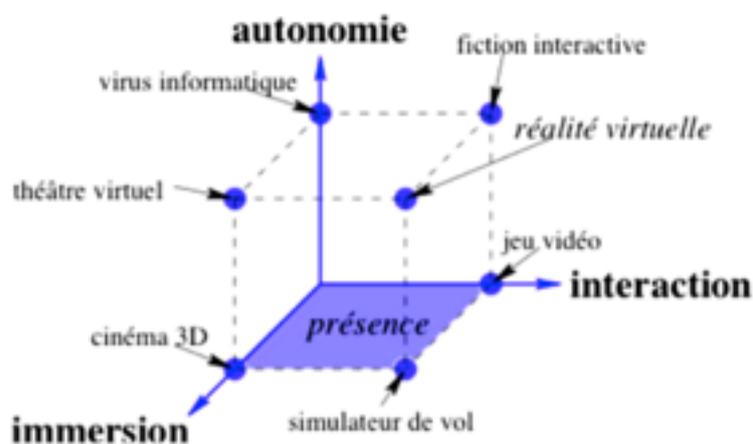


Figure 4. Immersion, interaction et présence selon Jacques Tisseau ⁶⁸

Quelle place aurait mon adaptation dans le le schéma de Jacques Tisseau ?

⁶⁶ Voir Annexe 2

⁶⁷ ZENOUDA Hervé, "Son, interaction, immersion et effets de présence" in *Journées Neptune : Environnements immersifs*, Colloque à Toulon, 17 avril 2012

⁶⁸ TISSEAU Jacques, *Réalité virtuelle et complexité*, Manifeste scientifique du CERV, 2004

Mon approche possède deux facettes, intimement liées au format hybride du livre-jeu.

Tout d'abord celle d'une démarche immersive qui se caractérise par l'exploitation de la technologie binaurale avec suivi des mouvements de la tête. En effet, ma volonté de départ est de proposer un contenu où l'auditeur *est* dans un autre espace.

Pour cela, je pense pouvoir approfondir la tendance simulationniste déjà présente dans *Menu Cthulhu* afin de donner à l'auditeur une sensation de réalité.

J'aimerais ainsi que l'auditeur-joueur soit littéralement *à la place* du héros.

L'adaptation d'un livre-jeu m'a également amené à m'interroger sur l'interaction et la démarche *ludique* de ce type de littérature, relativement limités par rapport à d'autres formes de jeux comme le jeu vidéo.

Je souhaite donc, au sein de mon adaptation, ajouter d'avantage d'interactions et créer des phases qui n'existent pas dans le livre-jeu. Le jeu audio où la conception du *Game design*⁶⁹ et du *Level design*⁷⁰ sont primordiaux⁷¹, me servira d'exemple afin de baliser les interactions homme/machine que je peux proposer.

Je considère que les interactions se doivent d'être facilement comprises et assimilées. Elles devraient également être accessibles si possible aux personnes malvoyantes, audience historique des livres et des jeux audio.

Néanmoins, Les livres dont vous êtes le héros s'inscrivent traditionnellement dans une logique narrativiste. *Menu Cthulhu*, n'échappant pas à cette tendance, ne possède en effet aucune règle ludique spécifique, mais une histoire très aboutie.

Ma démarche pour cette adaptation n'est décidément pas facile à classer. Elle tend à incorporer des codes des jeux audio tout en étant naturellement plus proche de la fiction interactive et possède dans le même temps une volonté première d'immersion de l'auditeur.

C'est pourquoi, j'ai décidé d'appeler mon adaptation « Une aventure sonore dont vous êtes le héros », évitant ainsi l'écueil de la dualité entre livre et jeu, présent dans l'essence même du projet.

⁶⁹ Le Game Design est le processus de création et de mise au point des éléments d'un jeu.

⁷⁰ Le Level Design est le processus de création des niveaux au sein d'un jeu.

⁷¹ **DIENER Elliot, TOULON Arnaud**, *Réalisation d'un jeu audio*, Mémoire de fin d'étude de Louis-Lumière, 2015, p.53

II.3 Le cahier des charges

Il convient maintenant d'écrire un cahier des charges du contenu que je vais proposer dans ma partie pratique.

Cela constituera ainsi un fil rouge pour le reste du mémoire et toutes les décisions prises au cours de cette démarche devront répondre à ce cahier des charges.

Menu Cthulhu, une aventure sonore dont vous êtes le héros sera un contenu sonore binaural librement inspiré du livre-jeu éponyme.

Tel un « Film pour les oreilles », ce contenu ne sera pas une version lue du livre mais une réécriture avec des acteurs placés dans une scène sonore, comprenant des ambiances, des effets et des bruitages ainsi que de la musique.

L'esthétique de ce projet se rapprochera donc de celle que l'on peut retrouver dans les fiction radiophoniques.

L'écoute de ce contenu se fera grâce à un casque associé à un dispositif de suivi des mouvements de la tête. L'auditeur sera donc libre de déplacer sa tête et de découvrir les différentes scènes comme il le souhaite.

L'auditeur aura accès à une interface permettant d'effectuer l'ensemble des interactions dans l'aventure.

L'aventure sonore pourra se décomposer en trois types de phases différentes :

- Une phase narrative inspirée du livre d'origine où l'auditeur sera immergé dans le récit.
- Une phase de choix où l'auditeur, grâce au narrateur et à l'interface fournie, devra choisir entre deux suites possible de l'aventure. Il y aura ainsi quatre fins différentes à cette histoire.
- Une phase plus ludique *d'exploration* où l'auditeur pour avancer dans le récit, devra tourner la tête dans la direction d'une source sonore. S'il n'arrive pas à se repérer dans l'espace, un bouton s'affichera sur l'interface et lui permettra d'avancer. Cela dans le but de ne pas frustrer les auditeurs en cas de difficultés.

II.3 Le *creative agenda* ou les attentes créatives des joueurs

Les *attentes créatives* (ou *creative agenda*) de la théorie du jeu LNS me permettront de baliser mon contenu.

Elles prennent naissance dans le « Contrat social » établi entre le joueur et le jeu ainsi que dans l'exploration de l'univers diégétique du contenu.

L'auditeur s'attend dès le départ à participer à une « aventure dont vous êtes le héros ». Cette information aura une incidence fondamentale sur la façon dont les auditeurs peuvent appréhender ce contenu.

Le joueur s'attend donc à incarner un héros. La spatialisation du héros doit donc être unique et fixe afin d'éviter les confusions avec les autres personnages et de plonger l'auditeur *dans* ce personnage.

De plus, le livre *Menu Cthulhu* inclut un narrateur très présent et qui permet la médiation entre le joueur et le récit. Ce narrateur devra lui aussi être spatialisé de façon unique et sa place ne devra soulever aucune ambiguïté.

Les termes *héros* et *aventure* portent en eux des références à des situations et des contextes extraordinaires. Les situations et le contexte développé au sein de ce contenu devront se montrer *dignes* de ces termes, au risque de décevoir l'auditeur.

La référence aux livres-jeux par le groupe verbal *dont vous êtes le héros*, pose des attentes concernant le système mis en place au cours de ce contenu. En effet, le dispositif interactif ainsi que la progression au sein de ce contenu doivent pouvoir être comparables aux livres dont vous êtes le héros.

Au regard de la théorie LNS, le livre d'origine possède une forte composante narrative et c'est véritablement la découverte des différentes façons de mourir au travers des chapitres qui constitue l'intérêt majeur.

Néanmoins, je pense que mon contenu devra équilibrer les différentes composantes. La liberté de tourner la tête dans cette scène sonore à 360° permettra d'apporter un élément simulationniste, en plus des éléments présents à l'intérieur du récit.

La phase d'exploration que j'ai rajouté dans mon aventure a elle pour objectif d'apporter une composante ludique supplémentaire. A la manière des QTE⁷² dans les jeux vidéo, cela peut permettre d'accrocher l'attention du joueur pendant la lecture des *chapitres*, qui ne sont alors plus des séquences totalement passives.

Des « *techniques* », à savoir les différentes phases de mon aventure, font également partie du creative agenda.

Il est alors possible de prévoir des profils types d'auditeurs.

En ce qui concerne les choix à faire dictés par le narrateur, certains privilégieront les choix *raisonnables* et suivront une histoire puis une autre. D'autres encore voudront tester l'ensemble des possibilités d'une traite et s'amuseront à faire des choix dans l'optique de découvrir l'ensemble des fins.

Concernant la phase d'exploration, on pourra déceler trois grands types d'auditeurs : ceux qui adhèrent au principe du jeu et réussissent, ceux qui adhèrent mais qui ne réussiront pas et ceux qui refusent le principe et qui attendront que la phase d'exploration avec les mouvements de la tête se terminent.

Ainsi, on peut identifier deux types d'*ephemera*, à savoir des moments d'action de l'auditeur, au sein de cette aventure :

- Tout d'abord à la fin d'une séquence lorsque l'auditeur devra faire un choix .
- D'autre part, les *ephemera* liées aux phases d'explorations.

L'introduction de ces phases plus ludiques est conditionnée par une réécriture de *Menu Cthulhu*, car le livre n'incorpore pas cette mécanique de jeu à l'origine.

⁷² QTE, acronyme de Quick Time Event, est une phase dans un jeu vidéo où l'exécution d'indications affichés dans un temps imparti conditionne l'exécution d'une issue pré-calculée.

II.4 Du texte au son, la nécessaire réécriture

« *Le monde sonore ne détient aucune place dans les arts [...]. C'est un monde encore sauvage.* »⁷³

Daniel Deshays

La démarche de sonorisation⁷⁴ d'un livre nécessite forcément une réécriture.⁷⁵

Même dans le cadre de la plus simple et de la plus fidèle adaptation sonore d'un livre, à savoir une personne seule lisant le livre, l'interprétation et les talents oratoires du narrateur constituent déjà une forme de réécriture.

Les modalités perceptives ne sont pas les mêmes (*la vue et l'ouïe*) et la manière de s'immerger et de parcourir le contenu est tout à fait différente. L'écoute suggère une unité temporelle - l'auditeur écoute toujours un contenu d'une *durée donnée* -, qui n'existe pas à l'écrit. Le support écrit permet alors de prendre plus facilement de la distance, de marquer des arrêts, de faire des retours en arrière...

Ainsi, une adaptation sonore mot à mot d'un livre peut devenir un challenge à écouter s'il contient des termes techniques ou savants.

Par ailleurs, à l'inverse du texte et de la musique qui peuvent être préconçus sur papier, il n'existe pas de système écrit formel de notation pour le son. En effet, l'écriture sonore trouve ses racines à la prise de son⁷⁶ et s'élabore ensuite plus lentement, au fil du montage et des ajouts possibles. Le son peut permettre également, par des procédés de superpositions, l'apparition d'une « *autre vérité de langage, une autre voie, un autre chemin ...* »⁷⁷. Il ne faudrait donc pas négliger les possibilités narratives propres au son lors d'une adaptation d'un récit écrit.

⁷³ **DESHAYS Daniel**, *Pour une écriture du son*, Paris ,coll. 50 questions, éditions Klincksieck, 2008, p.14

⁷⁴ Sonorisation doit ici être bien évidemment compris comme *rendre sonore* ce qui n'était que visuel.

⁷⁵ **MALLO Eléonore**, *Etude pour la création d'une bande dessinée sonorisée*, Mémoire de fin d'étude de Louis-Lumière, 2015

⁷⁶ **DESHAYS Daniel**, *Pour une écriture du son*, Paris ,coll. 50 questions, éditions Klincksieck, 2008, p.72

⁷⁷ **PARENTHOEN Yann**, *Propos d'un tailleur de son*. Editions phonurgia nova, 2002, P.22

Je pense pouvoir dégager trois formes principales de réécriture qui se sont posées à moi durant ma démarche :

- Une réécriture *économique*. En effet, *Menu Cthulhu* est un livre de 225 pages avec une arborescence étoffée⁷⁸ et de nombreux personnages. Dans le cadre de ce mémoire, je savais pertinemment que j'avais une durée et un budget limités et qu'il faudrait ainsi réduire le nombre de chapitres et de personnages au sein du récit. J'ai donc enlevé deux personnages au récit - *Un adolescent et un routier* - tout en réadaptant les dialogues en veillant à la personnalité des personnages encore présents. Ainsi, cette adaptation a pu être réalisée grâce à l'aide bénévole de quatre comédiens. J'ai également revu le nombre de *chapitres* à la baisse et je ne propose dans ce prototype que trois choix et quatre fins, que je trouve personnellement les plus comiques.⁷⁹ Les choix proposés offrent des récits de durées différentes et permettent ainsi des rythmes de jeu différents à chaque relecture. Je suis conscient qu'ainsi, l'aventure proposée est plus linéaire que l'originale mais cela permet tout de même de rendre compte des spécificités de ce type de littérature.
- Une réécriture *médiatique*. En effet, j'ai décidé dès le début de mon projet d'avoir une esthétique où le livre serait interprété par des acteurs et non lu par un seul narrateur. Ainsi, une délimitation claire devait être faite entre les dialogues joués par les acteurs, le texte lu par le narrateur et d'autres informations qui seraient intégrés plutôt par l'écriture sonore. Le passage du texte au son s'est accompagné d'une réflexion à chaque réplique. En effet, au moment du tournage, certaines phrases très drôles et pertinentes à l'écrit, apparaissaient plus lourdes une fois jouées par des acteurs. Durant le tournage, je n'ai ainsi pas hésité à raccourcir certaines phrases afin de gagner en fluidité, nécessaires à l'oral, tout en tâchant de ne pas nuire à la compréhension et à l'humour émanant du texte. De la même manière, il était nécessaire de donner une conclusion plus nette à la fin principale du récit. Le narrateur, présent au cours de l'aventure devient alors le policier à qui le héros raconte tout ce qui s'est déroulé. Cela permet de clore et d'expliquer l'omniprésence de ce narrateur.

⁷⁸ Voir Annexe 1

⁷⁹ Voir Annexe 3

- Enfin, une réécriture *méta-médiatique*. Cette aventure étant un contenu inédit, il était nécessaire de poser le concept dès le début et de créer un didacticiel afin que l'auditeur puisse pleinement prendre en main le jeu. De plus, il était nécessaire de structurer de façon sonore la phase d'exploration. Afin d'aider à la repérer, j'ai sonifié à l'aide d'un son de cloche l'entrée dans cette phase. J'ai également intégré des indices verbaux permettant à l'auditeur de savoir vers quelles sources se diriger. Par ailleurs, j'ai abandonné le mot chapitre. En effet, l'auditeur va *vivre* cette aventure au travers d'une interface numérique, et ne va plus *aller* à un chapitre comme dans un livre-jeu.

La réécriture a été pour moi un moment délicat, car il fallait assumer pleinement sa part de création, et ne pas la cacher derrière l'œuvre originale. Bien qu'ayant tenté de garder au maximum l'esprit et l'atmosphère du livre, je sais également qu'un autre réalisateur aurait sûrement adapté le livre d'une façon totalement différente.

Une fois cette réécriture réalisée, je me suis mis à la recherche des solutions techniques permettant de remplir le cahier des charges que j'ai mis en place.

Chapitre III : Les techniques de reproduction sonore

sonore

Au cours de ce chapitre, nous allons nous intéresser aux deux technologies d'enregistrement et de reproduction sonore que j'ai envisagées pour la réalisation de *Menu Cthulhu* : La technique binaurale et la technique ambisonique.

III.1 La technique binaurale

III.1.1 Historique

Le binaural connaît ses balbutiements à la fin du 19e siècle en France grâce à Clément Ader qui invente le Théâtrophone. Ce dispositif permet à des auditeurs situés jusqu'à deux kilomètres de l'Opéra de Paris d'écouter la pièce jouée au sein de celui-ci.

Cela est rendu possible grâce à une rampe de microphones placée devant la scène et un téléphone sur chaque oreille des auditeurs. L'écoute est donc pour la première fois *binaurale*, ce qui signifie en latin deux oreilles.

Les auditeurs ont ainsi eu pour la première fois une impression d'espace dans un dispositif de reproduction sonore.



Figure 5. Affiche d'époque faisant la promotion du Théâtrophone ⁸⁰

⁸⁰ - Source de l'image : lesnumeriques.com

C'est Léon Blumlein, ingénieur en électronique, qui utilise pour la première fois en 1931 le terme *binaural*, pour décrire un enregistrement sonore, alors stéréophonique.

Lors de l'exposition universelle de 1933 à Chicago, la compagnie de téléphone américaine AT&T a dévoilé un prototype de tête mécanique nommé Oscar, avec deux microphones en lieu et place des oreilles. Le public pouvait alors mettre un casque et écouter *ce qu'écoute* Oscar, la tête artificielle. C'est ainsi la première expérience en *binaural natif*, à savoir que la spatialisation est possible grâce au dispositif de prise de son qui est alors fixe.

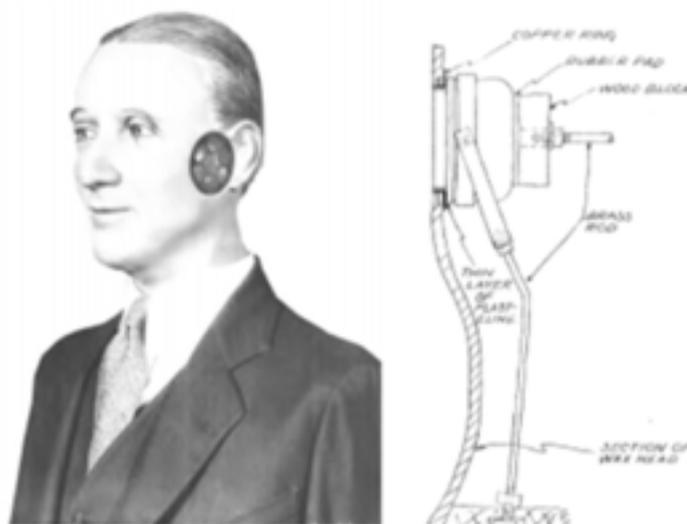


Figure 6. Schéma de présentation de la tête mécanique Oscar d'AT&T⁸¹

Les recherches sur le binaural ne se développent réellement qu'à partir des années 70 dans des laboratoires tel que l'IRCAM en France, l'Institut de Recherche et Coordination Acoustique Musique fondé par Pierre Boulez en 1969. En effet, la technologie binaurale représente un cadre de restitution « d'une importance primordiale pour la situation de laboratoire »⁸².

Ces années de recherche ont permis de développer des outils numériques tels que le Spat, le spatialisateur de l'IRCAM dont le projet a commencé en 1991. Ces outils permettent de synthétiser, à savoir recréer artificiellement, le placement d'une source sonore dans un espace virtuel pour une écoute au casque. Ils ont pu voir le jour grâce à des études précises des caractéristiques physiques de l'écoute binaurale.

⁸¹ Source de l'image : lesnumeriques.com

⁸² [Présentation de la technologie binaurale par l'IRCAM](#)

III.1.2 Caractéristiques physiques

La restitution d'un espace sonore au casque, grâce à la technologie vulgairement appelée « *Audio 3D* », est possible grâce à certains mécanismes psycho-acoustiques, mis en œuvre lors la localisation d'un objet sonore.

La localisation est possible grâce aux indices monauraux et binauraux, qui se basent sur des phénomènes temporeux, d'intensité et spectraux. Ces indices binauraux sont uniques et personnels et c'est pour cette raison qu'il peut y avoir des aberrations de localisation différentes selon deux auditeurs lors d'une écoute d'un contenu binaural.

- Les différences entre les deux oreilles dites « *interaurales* » de niveau et de temps

Ces différences sont liées à la distance interaurale, d'environ 17cm en moyenne, et à la morphologie de l'auditeur. Cette distance module le retard et l'intensité d'un son entre les deux oreilles. Ces différences ont d'abord été utilisées dans le cadre de la restitution stéréophonique.

Un son est également localisé du côté de l'oreille l'ayant perçu au niveau le plus élevé⁸³. De plus, la conjugaison des différences de niveaux d'intensité et de temps⁸⁴ entre les deux oreilles permet à l'auditeur de localiser de manière efficace -jusqu'à 3-4°, selon l'angle d'incidence- des sources situées sur le plan azimuthal.

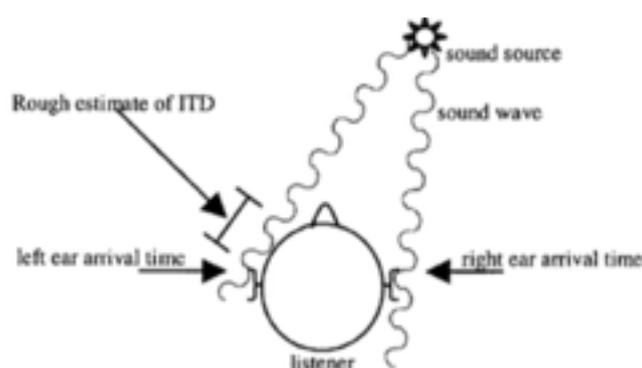


Figure 7. Schéma représentant la différence interaurale de temps (ITD) ⁸⁵

⁸³ - HUGONNET & WALDER, *Théorie et pratique de la prise de son stéréophonique*, Paris, Eyrolles, 1995

⁸⁴ - En anglais ITD pour *Interaural Time Difference* et ILD pour *Interaural Level Difference*

⁸⁵ - Cheng, Corey ,Wakefield, Gregory , « Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space » in *JAES*, 2001

En élévation, la localisation est moins précise : il n'existe plus de différences interaurales de temps ou d'intensité. De la même façon, des cônes de confusion apparaissent autour des oreilles dues au fait que ces positions dans l'espace possèdent les mêmes différences de temps et d'intensité. C'est ce qui peut expliquer les confusions avant/arrière lors d'une écoute d'un contenu binaural.

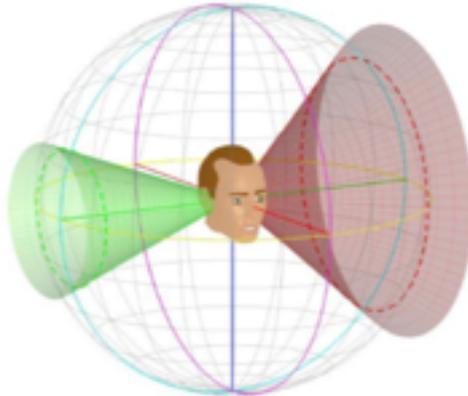


Figure 8. Représentation des cônes de confusion relatifs aux différences interaurales de temps et d'intensité.⁸⁶

Ce sont les indices spectraux présents dans les HRTFs qui définissent particulièrement la localisation en élévation et apportent une solution aux confusions avant/arrière.

- Les HRTFs

Les fonctions de transfert de la tête ou Head-Related Transfer Function, regroupent l'ensemble des déformations subies par le son depuis la source jusqu'au canal auditif, en modélisant également le filtrage du son par le corps. Elles sont propres à chaque individu dans la mesure où nous disposons tous de morphologies différentes, et ainsi autant de sources de déformations, déphasages, diffractions diverses du son.

⁸⁶ SALMON François, *Monitoring au casque d'un outil de spatialisation sonore*, Rapport de stage de Louis-Lumière, 2015, p.10

Les HRTFs comprennent à fortiori les différences interaurales de niveau et de temps tout en permettant la discrimination avant-arrière et la perception de l'élévation⁸⁷.

Les HRTFs peuvent être mesurés dans une chambre anéchoïque à l'aide d'une ou de plusieurs enceintes et de microphones placés dans les oreilles d'un auditeur ou recréés numériquement⁸⁸.

En synthèse binaurale les filtres HRTFs, issus de réponses impulsionnelles sont donc employés afin de filtrer numériquement le signal sonore pour le situer dans un point de l'espace.

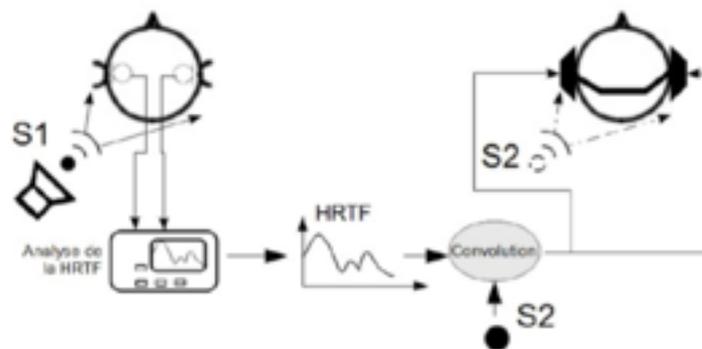


Figure 9. Schéma représentant l'acquisition et l'utilisation des HRTFs⁸⁹

La synthèse binaurale est donc possible grâce à l'opération de convolution⁹⁰ qui permet ainsi de replacer un objet sonore (S2) dans l'espace grâce à la mesure dans un laboratoire d'une source-test placée dans un point précis de l'espace (S1) lors de l'acquisition des HRTFs.

Le lecteur curieux pourra approfondir dans Spatial Hearing de Jens Blauert, toutes les considérations psycho-acoustiques évoquées dans ce chapitre et Traitement du signal audiovisuel de Laurent Millot donnera toutes les précisions mathématiques concernant l'opération de convolution.

⁸⁷ **BLAUERT Jens**, *Spatial Hearing*, MIT Press, 1997

⁸⁸ **AUSSAL Matthieu**, *Méthodes numériques pour la spatialisation sonore, de la simulation à la synthèse binaurale*, Thèse de l'École Polytechnique, 2014

⁸⁹ Source de l'image : www.fks.oeaw.ac

⁹⁰ **MILLOT Laurent**, *Traitement du signal audiovisuel : Applications avec Pure Data*, Paris, Dunod, 2008

III.1.3 Les outils

La technologie binaurale a permis l'émergence de deux techniques :

- Le binaural natif, où les indices binauraux sont pris en compte à la prise de son et ils ne peuvent pas être modifiés par la suite. L'image est donc fixe et il n'est pas possible de réaliser un suivi de mouvement de la tête.
- La synthèse binaurale, où les indices binauraux sont *ajoutés* par traitement du signal en post-production. L'objet sonore devient alors une source virtuelle, qui peut être placée dans l'espace virtuel. Un capteur peut être adjoint à la synthèse binaurale afin d'effectuer une compensation des mouvements de la tête et de garder une source sonore au même endroit de l'espace virtuel, quelque soit la position de la tête.

Le binaural natif

Le binaural natif est donc constitué d'un ensemble de solutions de prise de son permettant à l'enregistrement la prise en compte des indices binauraux.

Elles peuvent être décomposées en deux catégories :

- *Les prises de son avec tête artificielle*

L'enregistrement des objets sonores se fait grâce à deux micros placés au creux des oreilles d'une tête artificielle. Cette tête possède donc des indices binauraux plus ou moins proches de ceux d'un humain et le dispositif peut se mettre sur perche ou sur pied. L'ensemble des grandes marques de microphones a déjà sorti des têtes artificielles (Neumann, AKG, Sennheisser, Schoeps..).

Certains modèles, plus sophistiqués, incorporent même un buste supplémentaire alors que d'autres se limitent à une demi-sphère. De ces spécificités dépendront la qualité des indices binauraux produits.

Dans le cadre de ce mémoire, j'ai pu avoir accès, grâce à la bienveillance de Matthieu Aussal et du CMAP, à la tête Neumann KU-100 qui est une référence en matière de tête artificielle.



Tête Neumann KU-100 gracieusement prêté par le CMAP

- *Les prises de son à tête réelle*

Les microphones peuvent également être placés à l'entrée du conduit auditif du preneur de son. Le dispositif est alors plus léger et discret, se confondant avec des écouteurs. Néanmoins, pour une prise de son d'ambiance, il est nécessaire pour le preneur de son de ne pas bouger sa tête sous peine d'avoir une restitution binaurale *flottante*. Les microphones DPA 4060 sont des références pour ce type de prise de son.

Dans le cadre de ma démarche, je n'ai pas vu l'intérêt d'utiliser ce type de prise de son. Néanmoins, de par ses aspects pratiques et économiques, il présente un intérêt certain pour la production de contenu autant documentaire⁹¹ que fictionnel.

⁹¹ **CHEVRIER Léa**, *Expérimentation des techniques binaurales appliquées au documentaire radiophonique*, Mémoire de fin d'étude de Louis-Lumière, 2015

Microphone DPA 4060 équipé sur une preneuse de son (Léa Chevrier)⁹²



Avantages et inconvénients

Le binaural natif possède bien des avantages. Une mise en place légère voire discrète pour une prise de son à tête réelle, une possibilité d'avoir une écoute de contrôle facile à la prise de son. La technique présente plus d'inconvénients lors de la post-production. En effet, le montage de son en binaural natif consiste en une superposition de pistes stéréophoniques ce qui peut contribuer à augmenter drastiquement le bruit de fond si celui-ci est incontrôlé. En effet, le binaural natif ne permet pas d'isoler une source mais enregistre l'ensemble d'une scène sonore. Par essence, il fige la scène sonore et ne permet pas de retouche de spatialisation à postériori⁹³. Il est donc impossible d'adjoindre un dispositif de head-tracking à une prise de son en binaural natif. C'est pour cette raison que dans le cadre de ma démarche, je me suis d'avantage intéressé aux solutions de synthèse binaurale.

La synthèse binaurale

Les dispositifs de synthèse binaurale sont le plus souvent des plug-ins⁹⁴ que l'on insère dans un séquenceur et qui permettent de spatialiser une source sonore lors de la post-production. Les sources sonores sont le plus souvent un signal pré-enregistré qui peut être monophonique, stéréophonique ou multicanal (5.1, Decca Tree,...). Ainsi le dispositif de synthèse binaurale est capable, si les enceintes

⁹² Ibid

⁹³ Voir Annexe 4

⁹⁴ En informatique, un plug-in ou module d'extension est un paquet logiciel qui complète un logiciel hôte.

virtuelles sont correctement placées, d'effectuer un *downmix*⁹⁵ de fichiers multicanaux pour une écoute au casque.

C'était cette possibilité qui était particulièrement mise en avant dans la campagne de communication du Casque Sound One de 3D SoundsLabs.

Par ailleurs, il est donc possible de faire des prises de son multicanales *classiques* pour réaliser un contenu binaural. L'ensemble des possibilités des techniques de prise de son sont donc conciliables avec un contenu binaural.

La solution la plus connue est le Spat, le spatialisateur de l'IRCAM.

Le projet est né en 1991 et possède des domaines d'applications très divers, de la musique aux films en passant par de la recherche et développement. A la base, le Spat s'est présenté sous la forme d'une suite d'objets pour le logiciel de programmation orientée objet Max/MSP et une version plugin a pu voir le jour grâce à la collaboration avec Flux, qui a ainsi édité les Ircam Tools. Son originalité réside dans sa gestion d'une acoustique virtuelle, qui permet de définir l'espace sans avoir pour autant de vocabulaire architectural.



Figure 10. Vision d'ensemble du SPAT, transcodé en plug-in par Flux⁹⁶

⁹⁵ Un downmix consiste en une réduction de canaux audio afin d'adapter un contenu à un autre système de diffusion

⁹⁶ http://www.fluxhome.com/products/plug_ins/ircam_spat-v3

Les paramètres principaux, que l'on peut retrouver dans l'ensemble des moteurs de synthèse binaurale concernent la position du signal dans l'espace virtuel en trois dimensions - *Azimut, Elévation, Distance* -. Le SPAT propose également des fonctions permettant de simuler la distance et de gérer le rayonnement de la source.

Il propose également des fonctionnalités poussées concernant la gestion de l'espace virtuel (réverbération et égalisation). Ce logiciel permet également de prendre en compte la position des enceintes réelles qu'il est possible de déplacer sur un cercle autour de l'auditeur.

Ainsi, il permet de s'adapter à des conditions d'écoutes normées (5.1,7.1) comme non normées.

Il existe de nombreuses solutions pour faire de la synthèse binaurale⁹⁷ qui peuvent se différencier selon plusieurs critères :

- La qualité intrinsèque du moteur binaural⁹⁸ ainsi que la possibilité de moduler les HRTFs, qui peut engendrer des différences audibles, particulièrement au niveau de la restitution du timbre.
- Les possibilités d'édition de l'espace et des sources virtuelles, dans le placement comme dans le traitement.
- La possibilité d'adjoindre au programme un dispositif de suivi des mouvements de la tête.
- La prise en main pour l'opérateur qui peut être intuitive ou laborieuse selon la dénomination des fonctions.

⁹⁷ <http://multiphonie.free.fr/plugins4.htm>

⁹⁸ Le moteur binaural constitue en un ensemble d'algorithmes permettant l'ajout d'indices binauraux à un signal

Dans le cadre de ce mémoire, j'ai eu l'occasion de découvrir le plug-in *Binauralizer*⁹⁹ de Noise Makers, la société de Charles Verron, mon directeur externe.

Ce plug-in, intuitif, permet de placer une source monophonique ou multicanal (2.0,5.1,7.1) dans un espace horizontal tout autour de l'auditeur en permettant de contrôler sa position dans l'espace (*Angle*) et de contrôler sa largeur (*Width*).

Il permet l'importation de HRTFs au format SOFA, qui est un format de transmission d'HRTFs élaboré au sein du programme de recherche BiLi¹⁰⁰, un projet de recherche collaboratif sur l'écoute binaurale faisant intervenir des entreprises privées (Orange, Arkamys) et des partenaires publics (Radio France, IRCAM, CNSMD).



Figure 11. Plug-in Binauralizer¹⁰¹

⁹⁹ <http://www.noisemakers.fr/binauralizer/>

¹⁰⁰ <http://www.bili-project.org/le-projet/>

¹⁰¹ - <http://www.noisemakers.fr/binauralizer/>

J'ai eu l'occasion de prendre en main ce plug-in et de faire des tests comparatifs et il s'est révélé très satisfaisant.

Malheureusement il ne permet pas de placer les sources en élévation.

J'ai tout de même décidé d'utiliser ce plug-in afin de réaliser mon projet, et j'ai décidé de ne pas prendre en compte le plan zénital lors de la post-production.

L'auditeur sera donc uniquement libre de ses mouvements sur un plan horizontal.

Néanmoins, il ne possède pas dans sa version commerciale de dispositif de head-tracking. Je voudrais donc particulièrement remercier Charles Verron d'avoir pris le temps de coder une possibilité d'incorporer un head-tracker à son programme.

III.2 La technique ambisonique

III.2.1 Les limites de la synthèse binaurale

La synthèse binaurale montre ses limites par rapport à mon adaptation d'un livre-jeu. En effet, afin de réaliser l'ensemble de mon projet en synthèse binaurale, il est nécessaire que chaque source sonore possède une itération de plug-in de spatialisation. Or, effectuer en temps-réel et en parallèle une spatialisation par synthèse binaurale est couteux en ressources processeurs.

Une solution aurait pu être de constituer un *bed* d'un format multicanal normé (5.1, 7.1). Néanmoins, ces normes de diffusions, pensées pour le son à l'image, possèdent un certain nombre d'aberration de localisation et de *trous* dans l'image sonore à l'arrière et sur les cotés. Ce n'est pas gênant pour du son à l'image, car l'esthétique frontale - où la majorité des sources sont à l'avant, là où est l'écran- est prépondérante.

Or je souhaitais une technologie me permettant d'avoir le moins d'aberration de spatialisation possible pour l'environnement sonore immersif que je souhaite développer dans mon adaptation. La technologie ambisonique, à l'inverse du binaural et des formats multicanaux normés, n'est pas un système *channel-based*¹⁰², où la technologie est fonction du nombre de canaux de diffusion. Au contraire, l'ambisonique est *scene-oriented*, son format de transmission permet de considérer un enregistrement ou un mixage ambisonique comme une scène sonore.

De fait, l'ambisonique n'a pas un système de diffusion attiré, et est donc adaptable sur l'ensemble des systèmes.

Cette notion de scène sonore est très intéressante dans le cadre de mon adaptation.

En effet, elle permet de contourner les limites de la synthèse binaurale, car l'ensemble des sources de l'aventure n'ont plus à être synthétisées directement en binaural. Je peux donc créer une scène sonore complexe, avec le nombre de sources sonores que je souhaite, pour un coût processeur indépendant du nombre de sources, et ne dépendant plus que de l'ordre de l'ambisonique choisie. (cf. la partie qui suit sur les principes de l'ambisonie).

¹⁰² Un format de prise de son channel-based est une configuration de prise de son pensée spécifiquement par rapport à un système de diffusion possédant un certain nombre de canaux.

III.2.2 Historique

L'ambisonie est un ensemble de techniques de synthèse, d'enregistrement et de reproduction de champs sonores à 360° basées sur une décomposition du champ acoustique en harmoniques sphériques¹⁰³. La technique est née dans les années 70 sous l'impulsion d'un groupe universitaire anglais dont faisait notamment partie Michael Gerzon de l'Institut de Mathématiques d'Oxford.



Photo de Michael Gerzon, inventeur de la technologie ambisonique

L'ambisonie de premier ordre, aussi appelé B-Format a ainsi été développée¹⁰⁴. Il permet à partir de quatre canaux d'encoder les informations directionnelles d'un champ sonore en trois dimensions.

A l'inverse de la quadriphonie qui a connu une exposition éphémère dans les années 80, l'ambisonique n'a jamais été déployée pour le grand public et est restée une technique confidentielle.

Le renouveau s'est amorcé dans les années 2000 grâce au développement du High Order Ambisonics (plus communément HOA)¹⁰⁵, qui permet en augmentant le nombre de canaux d'avoir un encodage plus précis d'une scène sonore.

A l'instar du binaural, l'ambisonique peut être natif grâce à une prise de son ambisonique, soit synthétisé, ce qui permet de recréer une scène sonore complexe grâce aux traitements du signal.

¹⁰³ <http://www.mshparisnord.fr/hoalibrary/ambisonie/les-champs-sonores/>

¹⁰⁴ <http://www.mshparisnord.fr/hoalibrary/ambisonie/formats/>

¹⁰⁵ <http://www.mshparisnord.fr/hoalibrary/ambisonie/ambisonie-ordre-superieur/>

III.2.3 Principes physiques

Contrairement à d'autres formats surround multicanaux que sont *channel-based*¹⁰⁶, son format ne contient pas les signaux de chaque canal. Au lieu de cela, il est constitué d'un nombre de canaux en fonction de l'ordre de l'encodage, qui est ensuite décodé en fonction de la configuration des enceintes de l'auditeur. Ceci permet alors de penser en termes de directions de source plutôt que les positions de haut-parleurs, et offre à l'opérateur un degré considérable de souplesse quant à la mise en place et le nombre de haut-parleurs utilisés pour la lecture.

Ainsi l'ambisonique peut être converti en d'autres formats de diffusion normés (2.0, 5.1, 7.1)¹⁰⁷ et en binaural.

Afin de mieux comprendre le principe de l'ambisonie, prenons l'exemple de l'encodage et du décodage de l'ambisonique d'ordre 1.

Le B-Format est composé des canaux W, pour la pression acoustique, X, Y et Z pour les gradients de pression acoustique dans les trois directions de l'espace. Le signal W correspond à un microphone omnidirectionnel, alors que X, Y et Z sont des composantes qui seraient captées par des capsules bidirectionnelles orientées selon les trois axes de l'espace.

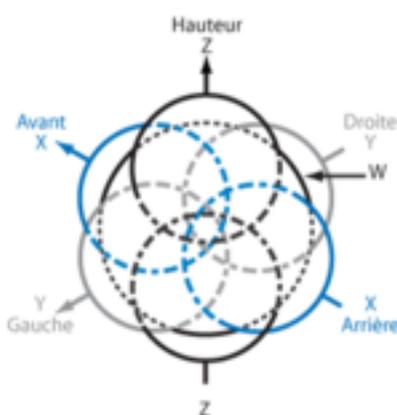


Figure 12. Représentation des composantes W,X,Y et Z¹⁰⁸

¹⁰⁶ Un format de prise de son channel-based est une configuration de prise de son pensée spécifiquement par rapport à un système de diffusion possédant un certain nombre de canaux.

¹⁰⁷ GERZON Michael, « Ambisonics in multichannel broadcast and video » in *JAES*, 1984

¹⁰⁸ - <http://www.surround-library.com/white-papers>

On peut d'ailleurs assimiler ce format à un équivalent d'empilement de couples MS pour les trois directions de l'espace.

Un encodage ambisonique simple consiste à prendre un signal S , d'azimut θ et d'élévation ϕ . La source est positionnée à l'angle désiré en distribuant le signal sur les composantes Ambisonic avec des gains différents :

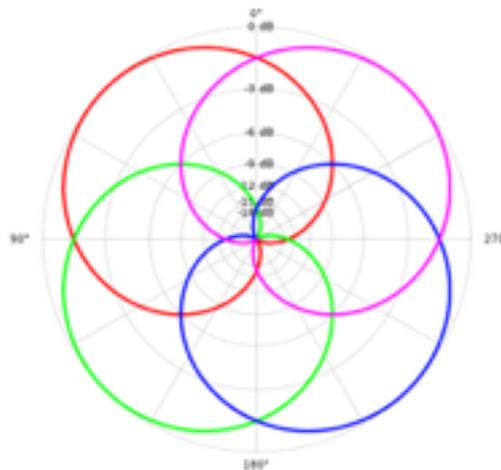
$$W = \frac{S}{\sqrt{2}}$$

$$X = S \cdot \cos \theta \cos \phi$$

$$Y = S \cdot \sin \theta \cos \phi$$

$$Z = S \cdot \sin \phi$$

Pour un dispositif d'ordre 1 parfaitement régulier, on a les quatre canaux (LF, LB, RB, RF) suivants pour une restitution en quadraphonie :



$$LF = (2W + X + Y)\sqrt{8}$$

$$LB = (2W - X + Y)\sqrt{8}$$

$$RB = (2W - X - Y)\sqrt{8}$$

$$RF = (2W + X - Y)\sqrt{8}$$

Pour encoder un signal avec un ordre plus élevé, il est nécessaire d'incorporer de nouvelles harmoniques sphériques, qui permettent de décrire plus précisément un espace sonore.

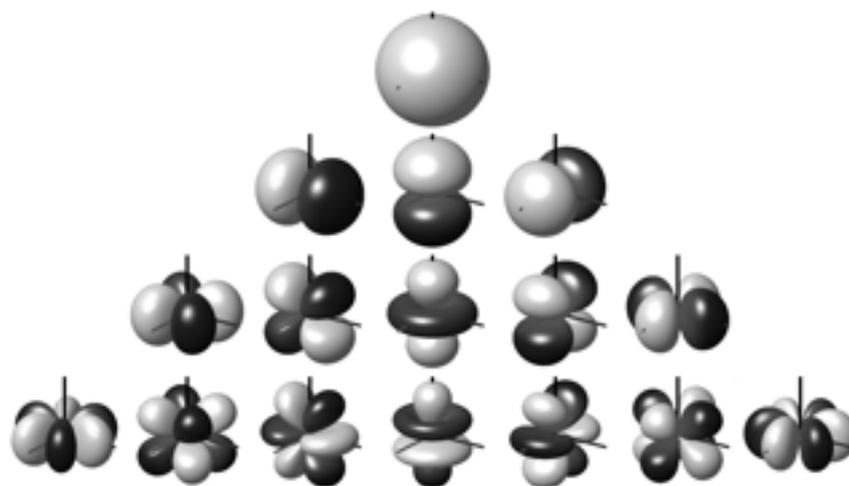


Figure 13. Représentation des harmoniques sphériques des ordres 0 à 3

La technologie ambisonique permet une grande souplesse en post-production. En effet, l'ambisonique, à l'instar d'un couple MS, se fait grâce à des matrices de signaux et peut donc être modifié en aval de la prise de son. Il est ainsi possible de faire des opérations de zoom, de mouvement et de rotation dans la scène sonore.¹⁰⁹

C'est l'opération de rotation qui m'intéresse particulièrement dans le cadre de ma démarche. Ainsi le suivi des mouvements de la tête traditionnellement appliqué aux enceintes virtuelles lors d'une synthèse binaurale, s'appliquerait ici directement sur le signal ambisonique.¹¹⁰

¹⁰⁹ MALHAM Dave, *Spatial Heading Mechanisms and Sound Reproduction*, University of York, England, 1998

¹¹⁰ NOISTERNIG, MUSIL, SONTACCHI, HOLDRICH, « 3D binaural sound reproduction using a virtual ambisonic approach. » In *Virtual Environments, Human- Computer Interfaces and Measurement Systems*, Symposium international, 2003, p. 174–178

Cela présente plusieurs avantages non négligeables :

- Les enceintes virtuelles resteraient alors fixes autour de l'auditeur. Il n'y a donc pas besoin de faire une interpolation d'HRTFs entre deux positions d'enceintes virtuelles en temps réel, qui est une opération lourde en ressources¹¹¹.
- Le nombre de sources sonores n'est plus corrélé aux nombres d'enceintes virtuelles passant par un encodage en ambisonique. Il est donc possible de constituer des scènes sonores avec de nombreuses sources.
- Il est possible d'encoder les informations spatiales en élévation avec un ordre moindre, la perception humaine étant moins précise. C'est ce que l'on appelle de l'ambisonique d'ordre mixte.

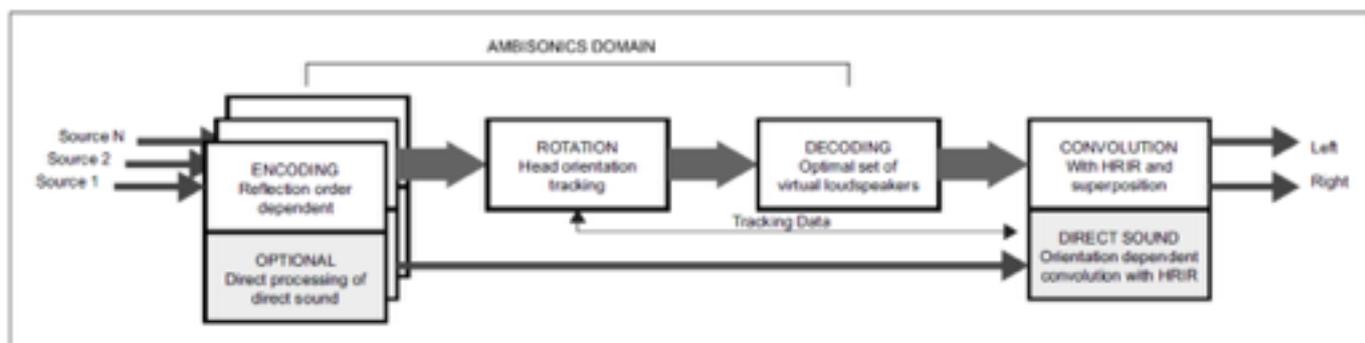


Figure 14. Principe d'encodage Ambisonique destiné à une restitution binaurale avec head-tracking¹¹²

C'est cet ensemble de raisons qui m'a poussé à exploiter la technologie ambisonique dans le cadre de mon projet.

¹¹¹ Ibid

¹¹² Ibid

III.2.4 Les outils

L'ambisonique, à l'instar du binaural, a vu l'émergence de plusieurs outils de production qui peuvent être classés ainsi :

- L'ambisonique natif, ou la prise de son ambisonique, qui permet l'encodage d'une scène sonore *réelle*.
- L'ambisonique synthétisé, qui à partir de traitements du signal, encode, à l'aide d'un format ambisonique, un ensemble de sources afin de créer une scène sonore nouvelle.

La technologie ambisonique native

Elle est donc constituée de l'ensemble des solutions de prises de son ambisoniques. Elles peuvent être distinguées en deux catégories : Le format A (ou ambisonique de premier ordre) et l'ambisonique d'ordre supérieur.

Le format A

Un microphone ambisonique au format A est forcément d'ordre 1. C'est la solution la plus utilisée aujourd'hui pour faire des prises de sons d'ambiance en ambisonique, les seules proposées à la location dans les grandes entreprises de location de matériel.

Michael Gerzon propose dès les années 80 un capteur ambisonique de 1^o ordre composé de 4 capsules microphoniques à directivité cardioïde disposées suivant les 4 faces d'un tétraèdre régulier.

C'est ce que l'on appelle un microphone ambisonique au format A.



Photo d'un SoundField STS200, un microphone ambisonique au format A

Une fois correctement matricés et traités, les signaux bruts issus de ces 4 capsules constituent le format B, identifiés chacun par une lettre : W, X, Y et Z. Il est donc nécessaire d'avoir un convertisseur logiciel ou matériel du format A au format B.

La marque SoundField est aujourd'hui, celle qui s'est le plus investie dans la prise de son ambisonique en proposant plusieurs modèles de microphone¹¹³ (STS200, ST450,...).



Valise du SoundField ST350, avec convertisseur Hardware format A vers format B, mis à la location par Tapages et Nocturnes

Néanmoins, d'autres marques, telles Core Sound avec son Tetra Mic¹¹⁴ ou le constructeur indien Brahma avec son Brahma Mic¹¹⁵ ont émergé sur le marché et offre des solutions moins chères afin de démocratiser la prise de son ambisonique. Néanmoins, ces marques restent plus confidentielles et ne possèdent pas un circuit de distribution comparable à SoundField.

¹¹³ <http://www.tslproducts.com/soundfield-type/soundfield-microphones/>

¹¹⁴ <http://www.core-sound.com/TetraMic/1.php>

¹¹⁵ <https://www.kickstarter.com/projects/1569945514/brahma-affordable-ambisonics-microphone>

La prise de son ambisonique d'ordre supérieur

Des systèmes ont été développés dans les laboratoires afin de pouvoir effectuer des prises de son ambisoniques d'ordre supérieur. A l'instar des systèmes holophoniques¹¹⁶, ces solutions multiplient les capteurs autour d'une sphère et grâce à un traitement du signal adéquat permettent d'avoir l'ensemble ou une partie des harmoniques sphériques à un ordre donné.

Le microphone ambisonique d'ordre supérieur le plus connu est l'Eigenmike¹¹⁷ de MHAcoustics. Il permet grâce à ces 32 capsules et à un traitement du signal adéquat, de capter une scène sonore en ambisonique d'ordre 3.



Figure 15. Plusieurs microphones Eigenmike¹¹⁸

Ces microphones sont donc coûteux et encore en développement. Ils sont surtout utilisés dans les laboratoires de recherche tels le CNAM afin de valider certaines hypothèses de recherche. Il n'a donc pas été possible d'avoir accès à ce type de microphones dans le cadre de mon mémoire mais à l'avenir, il est possible que cette technologie devienne plus accessible aux ingénieurs du son.

¹¹⁶ <http://www.holophone.com>

¹¹⁷ <http://www.mhacoustics.com/products>

¹¹⁸ Ibid

La synthèse ambisonique

La scène sonore peut être également, grâce à un traitement du signal, encodée en format ambisonique. Ainsi n'importe quelle source, monophonique ou multicanale peut être placée dans une scène ambisonique.

La synthèse ambisonique permet d'effectuer par la suite les mêmes opérations sur le champ sonore que le natif, à savoir les rotations comme les zooms.

La synthèse ambisonique, comme la synthèse binaurale, se présente le plus souvent sous la forme d'un plug-in, que l'on rajoute à un séquenceur.

Plusieurs critères distinguent les différents logiciels :

- L'ordre de la synthèse ambisonique. En effet, plus l'ordre sera élevé, meilleure sera la restitution spatiale mais les ressources processeurs nécessaires sont plus importantes (augmentation du nombre de canaux lus en parallèle,...)
- La possibilité de faire une rotation de la scène ambisonique. Dans le cadre de ma démarche, il est absolument nécessaire d'avoir cette possibilité afin d'intégrer plus facilement le dispositif de suivi des mouvements de la tête.
- La gestion des sources multicanales sur la sphère ambisonique. En effet, à l'instar des logiciels de synthèse binaurale, il s'agit de placer des enceintes virtuelles sur une sphère englobant l'auditeur. La gestion du *panning*¹¹⁹ doit donc être souple afin d'avoir une utilisation rapide.
- La prise en main et l'ergonomie du plug-in qui permette à l'opérateur d'avoir une utilisation plus intuitive.

¹¹⁹ Le panning est un argot technique désignant le fait de placer une source dans un espace grâce à la fonction panoramique.

Des solutions, gratuites ou payantes, ont aujourd'hui émergé sur le marché¹²⁰. La suite de plug-in gratuites et open source ambiX¹²¹, développée par Matthias Kronlachner, un ingénieur autrichien diplômé de l'université de Graz, est aujourd'hui une des suites logicielles les plus complètes.

Cette suite¹²² permet de travailler sur de l'ambisonique de troisième, cinquième ou septième ordre et dispose d'outils permettant facilement d'effectuer des opérations sur le flux ambisonique, de la gestion de la directivité (Ambix directionnel loudness) à la rotation de l'espace (Ambix Rotator). Le logiciel utilise le « *ACN Channel ordering* » qui permet de facilement tronquer le flux ambisonique afin de sélectionner des ordres inférieurs à l'ordre synthétisé. Ainsi, la sélection de la qualité de restitution spatiale peut être facilement implémentable.

Néanmoins, dans le cadre de mon mémoire, je lui ai préféré le plug-in Ambi-Pan de Noise Makers, la société de mon directeur externe Charles Verron.

Il permet de réaliser un placement d'une source monophonique ou stéréophonique dans une scène ambisonique d'ordre 1. En effet, je souhaitais rester dans un ordre ambisonique faible, car les différents lecteurs de flux ambisonique binauralisé qui ont commencés à émerger -Youtube 360° et Facebook - ne lisent aujourd'hui que de l'ordre 1. De plus, les tests préliminaires effectués m'ont permis d'estimer que l'ordre 1 était suffisant pour ce projet. Cela a permis de rendre la conversion binaurale moins coûteuse en calcul.



Figure 16. Plug-in AMBI PAN de Noise Makers

¹²⁰ https://en.wikipedia.org/wiki/List_of_Ambisonic_Software

¹²¹ <http://www.matthiaskronlachner.com/?p=2015>

¹²² <http://lac.linuxaudio.org/2013/papers/51.pdf>

Ce plug-in possède l'ensemble des fonctionnalités dont j'ai besoin pour réaliser mon aventure sonore. Il gère facilement la distance, l'azimuth et l'élévation d'une source dans le flux ambisonique.

Même s'il est moins précis en termes de localisation lorsque la source est proche, il représente une excellente solution pour des sources diffuses et lointaines.

Il me faut donc trouver une solution afin de binauraliser ce flux ambisonique.

De l'ambisonique vers le binaural

La solution la plus utilisée pour passer de l'ambisonique vers le binaural consiste à reproduire un ensemble d'enceintes virtuelles autour de l'auditeur et d'avoir, par traitement du signal, un décodage approprié pour chaque position d'enceinte. Le plus souvent, les logiciels permettent de passer d'un ambisonique de premier ordre (B-Format) vers du binaural.

Comme pour la synthèse ambisonique, il existe de nombreuses solutions logicielles, payantes ou gratuites exploitant cette solution technologique.

Elles se différencient par la qualité de leur moteur binaural, la gestion d'un éventuel système de head-tracking et leur facilité d'utilisation.

La suite ambisonique *HoaLibrary*¹²³ développé par le CICM, bien qu'open source et ayant été déjà utilisée dans un mémoire de l'école¹²⁴, ne permettait pas d'intégrer de façon immédiate un dispositif de suivi de mouvement de la tête.

Harpex-B, développé par la société danoise Harpex¹²⁵, permet une binauralisation avec de nombreux paramètres afin d'affiner le décodage ambisonique en amont. Cela permet ainsi d'avoir une plus grande précision lors de la localisation des sources.

¹²³ <http://www.mshparisnord.fr/hoalibrary/>

¹²⁴ **MATAHRI Hanna**, *Adaptation d'une pièce de jazz en ambisonique*, Mémoire de Louis-Lumière, 2014

¹²⁵ <http://harpex.net/about.html>

Pour ce projet, j'ai finalement opté pour B-Format-2-Binaural, version beta de Ambi-Head¹²⁶, développé également par Noise Makers, qui après des tests s'est avérés très satisfaisant.

Ce logiciel propose une gestion du suivi du mouvement de la tête sur les trois axes et permet de changer les HRTFs utilisés grâce à l'importation d'HRTFs au format SOFA. Au niveau paramétrique, il ne possède qu'un paramètre *Width* qui permet de gérer la largeur de l'image sonore ambisonique. Cela le rend très simple d'utilisation et tout en restant paramétrable.

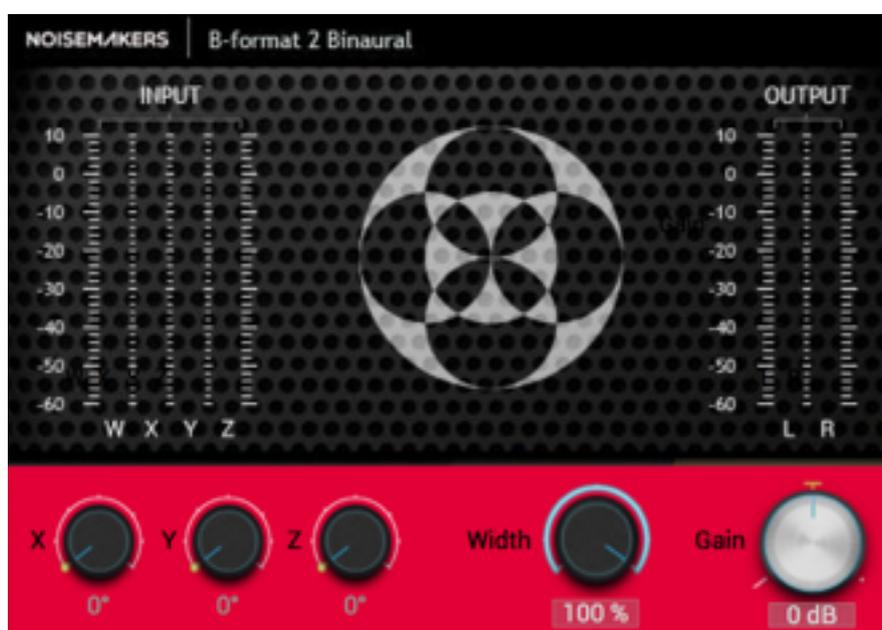


Figure 17. Le Plug-in B-format 2 Binaural de Noise Makers

J'ai ainsi pu dégager les différentes technologies et les différents outils logiciels que je compte utiliser pour la réalisation de mon adaptation. Néanmoins, il y a également des choix à effectuer en amont de la prise de son, car elle conditionne dès le départ une certaine esthétique sonore.

Il est possible de récapituler l'ensemble de mes choix concernant les solutions techniques explorées dans ce chapitre dans un tableau. Il est possible de noter que l'ensemble des techniques de synthèse a été assuré par la suite de plug-in de

¹²⁶ <http://www.noisemakers.fr/ambi-head/>

NoiseMaker. Comparé à d'autres suites équivalentes, elle ne représente pas un investissement trop coûteux.

	Prise de son	Synthèse
Technologie binaurale	Tête Neumann KU-100 	Plug-in Binauralizer 
Technologie ambisonique	SoundField ST-350 	Plug-in Ambi-Pan et B-Format-2-Binaural 

Solutions de prises de son et de post-production retenues pour ce projet

Chapitre IV : La prise de son

En son multicanal, vous n'écoutez plus les échos d'un reportage, mais vous faites partie de l'expédition.

Guy Senaux, ingénieur du son et mixeur à Radio France

Dans ce chapitre, j'expose mes choix concernant la prise de son. La prise de son conditionne une certaine esthétique. En effet, l'acte de prise de son n'est pas gratuit¹²⁷ et produit aussi un certain nombre d'artefact (timbre, localisation) qu'il convient de connaître. Ma volonté première était de produire des scènes sonores *équi-énergies* et non polarisées vers un axe de l'espace, une esthétique très commune au cinéma. En effet, il me semble que la sensation de subjectivité et d'immersion ne peut être atteinte que s'il y a un nombre de sources ponctuelles et diffuses réparties tout autour de l'auditeur. C'est donc avec cette volonté de départ que j'ai effectué mes prises de son.

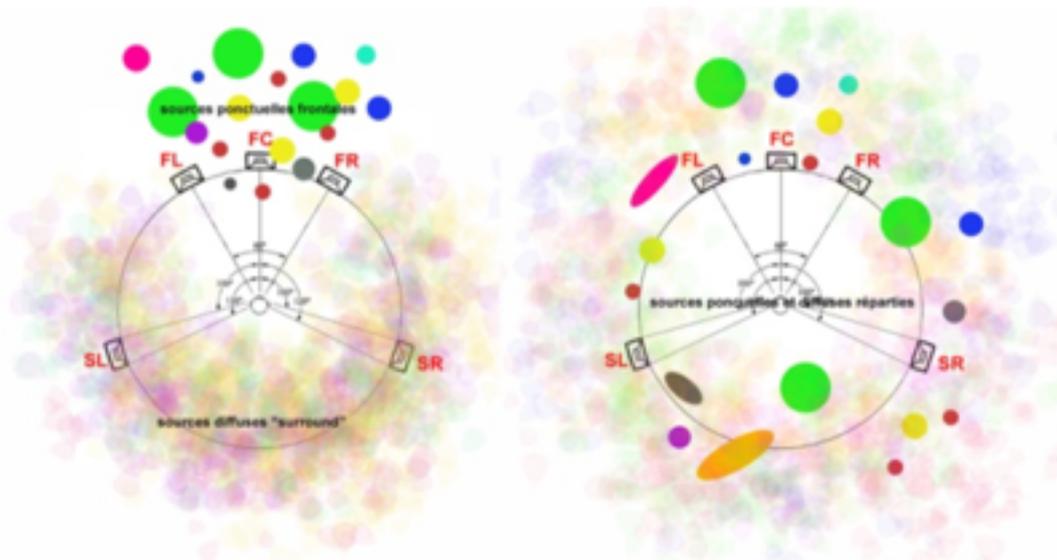


Figure 18. Scène sonore polarisée et scène sonore équi-énergie ¹²⁸

¹²⁷ DESHAYS Daniel, *Pour une écriture du son*, coll. 50 questions, Paris, 2006, p.65

¹²⁸ PERIAUX Bergame, OHL Jean-luc, THEVENOT Patrick, *Le son multicanal*, Paris, Edition Dunod, 2015

IV.1 Les voix

Les dialogues étant prépondérant dans cette adaptation, je souhaitais que les voix aient un timbre satisfaisant et qu'elles puissent se déplacer de façon fluide dans l'espace sonore. De plus, je voulais avoir un traitement spécifique sur le narrateur et le héros afin de pouvoir très facilement les distinguer des autres personnages.

Ainsi, la technologie ambisonique comme la technologie binaurale aurait pu s'appliquer pour la prise de son des voix.

Néanmoins, je souhaitais que les personnages puissent jouer tous ensemble afin de donner un caractère spontané à leur jeu. Ainsi le héros et les autres personnages devaient se trouver physiquement dans la même pièce. De plus la prise de son ambisonique aurait impliqué une prise de son in-situ, ce qui était très compliqué à mettre en place.

C'est pour cette raison que j'ai décidé d'utiliser la synthèse binaurale et d'effectuer la prise de son des voix en cabine speak à l'école, où chaque acteur avait son microphone. Je souhaitais à la base faire mon enregistrement en studio musique à la manière des fictions radiophoniques réalisées en troisième année. En effet, le studio était plus spacieux et moins bruyant, mais il était en maintenance les deux jours de l'enregistrement des voix, que je ne pouvais pas déplacer à cause d'un planning serré des comédiens.

Je ne voulais pas effectuer de prise de son au couple car je voulais me garder la possibilité de déplacer les voix dans l'espace sonore lors de la post-production. J'ai donc utilisé deux Neumann TLM-103 pour les voix féminines et un Neumann U-87 pour le héros et le narrateur.

J'ai donc aménagé la cabine speak, afin que les trois comédiens puissent me voir lors de l'enregistrement tout en gardant un minimum de distance entre eux afin qu'il y ait le moins possible de diaphonie entre les différents microphones.



Comédienne enregistrant une voix en cabine speak

Par souci économique, je n'ai eu que quatre comédiens bénévoles. Même après réécriture, je ne pouvais pas adapter ce livre avec seulement quatre personnages, trois des comédiens ont donc joué plusieurs rôles. J'ai néanmoins décidé, afin d'éviter les confusions, que l'acteur ayant le rôle du héros ne joue pas d'autres rôles.

Au cours de ces deux séances d'enregistrement, j'ai décidé d'être seul à la technique et d'assumer ainsi le rôle d'opérateur et de directeur artistique.

Il m'était nécessaire d'avoir donc deux écoutes :

- Une écoute analytique d'ingénieur du son, afin de juger la qualité du timbre et ainsi de mettre en œuvre des solutions pour garder une *qualité sonore* constante
- Une écoute globale et artistique, que je devais conserver afin de guider les comédiens et d'ajuster leur jeu d'acteurs.

Ces prises de sons constituent la matière principale de ce projet. Je pense qu'à ce stade, il est important de ne jamais fermer de porte concernant le jeu des acteurs.

En effet, entre la prise de son et le mixage, j'ai changé de façon d'écouter et de penser ce contenu. Si c'était à refaire, je pense que j'aurais pris plus de temps à l'enregistrement afin d'avoir plusieurs versions de certains dialogues.

IV.2 Les ambiances

En ce qui concerne les ambiances, je souhaitais immerger le plus possible l'auditeur dans un lieu donné et avoir une captation à 360°. La prise de son ambisonique, captant globalement une scène sonore m'apparaissait alors comme la solution privilégiée afin de réaliser des prises de son d'ambiance. Je savais néanmoins que je n'étais pas limité à cette technologie, et que je pouvais utiliser tous les systèmes de prises de son multicanal *channel-oriented*, plus facile à obtenir à l'école, en le couplant avec un dispositif de synthèse binaural.

J'ai donc cherché à louer un système de prise de son ambisonique. Les principaux loueurs de matériel audio, partenaires de l'école proposent assez peu de solutions. DCAudiovisuel propose un microphone au format-A SoundField STS200 et Tapages le SoundField ST-350, un ensemble microphone format-A et un matricieur hardware Format-A vers Format-B.

Frank, un commercial de DCA, m'a déconseillé de louer le microphone STS200 car il est très fragile et n'a jamais eu une reconstitution du timbre satisfaisante. Philippe Chenevez, fondateur de Cinéla et enthousiaste de la technologie ambisonique, m'a permis d'avoir une vision de l'utilisation de l'ambisonique dans le monde professionnel parisien. Il y a quelques professionnels travaillant en ambisonique à Paris et ils ont tous investi dans leur matériel ambisonique, car il est extrêmement fragile.

Au final, les ingénieurs du son intéressés par l'ambisonique mais ne possédant pas de matériel, se tournaient le plus souvent vers un système double MS, qui correspond à un « *ordre 0* », appellation professionnelle abusive qui correspond en fait à de l'ambisonique d'ordre 1 uniquement horizontal.

Je me suis donc naturellement tourné vers Tapages et j'ai pu ainsi louer une journée le SoundField ST-350.

Le convertisseur format-A vers format-B de ce SoundField n'est en rien un enregistreur et nécessite une batterie externe ou un raccord au secteur afin de fonctionner. Ayant réservé un kit stéréo avec un enregistreur Roland R44, j'ai dû donc réarranger la valise Portabrace afin de rendre mon dispositif de prise de son transportable.

J'ai ainsi compris l'engouement autour du microphone ambisonique indien Brahma¹²⁹ qui se rajoute au Zoom H2 tout en conservant une taille identique.



Ma disposition du sac Portabrace afin de rendre « portable » le dispositif de prise de son ambisonique

Avec cette organisation, j'ai été à même d'enregistrer l'ensemble de mes ambiances, en extérieur comme en intérieur, de manière autonome. J'avais besoin notamment d'enregistrer des ambiances de fast-food, de chambre froide et d'extérieur, qui sont les principaux lieux où se déroule *Menu Cthulhu*.

J'ai également fait un test comparatif sur une scène sonore quotidienne, afin d'avoir des éléments de comparaison sur une scène sonore complexe.

Bien que moins précis sur l'ensemble de la sphère, la solution d'enregistrer des ambiances multicanales en quadraphonie au zoom reste relativement satisfaisante et plus facilement accessible en cas d'enregistrement d'ambiance complémentaire.

Je n'ai donc pas abandonné cette solution. Néanmoins, j'ai toujours superposé une ambiance ambisonique *raccord*, afin d'éviter d'avoir un changement de répartition d'énergie sonore entre deux scènes.

¹²⁹ <http://www.embracecinema.com/gear/product-view.php?slug=brahma-in-zoom>

IV.3 Les effets

Menu Cthulhu étant avant tout une histoire fantastique et surréaliste, de nombreux effets sonores sont nécessaires afin de sonifier¹³⁰ correctement le monstre qui sort du frigo. Voulant utiliser des effets sonores en proximité comme en profondeur, je devais pouvoir maîtriser suffisamment ma matière sonore de base. C'est pour cette raison que je ne suis pas seulement contenté d'utiliser des sons de sonothèque. Il me fallait donc enregistrer certains effets.

Il est possible de classer ces prises de sons en deux catégories : la texture sonore et les bruitages en ultra-proximité.

La texture sonore

Dans le cadre de ce projet, la description du monstre et du lieu se fait uniquement avec du son. Ainsi, il convient de donner le maximum d'informations sonores possible afin de donner une impression *vraisemblable* à l'auditeur. C'est dans cette optique que je me suis mis à enregistrer différentes étapes d'une préparation culinaire.

Ainsi, je pouvais dans cette cuisine récupérer diverses sonorités de fritures ainsi que celle d'un réfrigérateur et également profiter d'une préparation de viande hachée pour avoir la texture sonore *visqueuse*.

Ne pouvant pas assurer la gestion de l'enregistrement ainsi que celle des bruitages, un camarade de classe m'a donc assisté sur la partie technique.

J'ai donc ainsi pu me concentrer uniquement sur le bruitage et les textures que je voulais obtenir. N'ayant pas de formation de bruiteur, cette session m'a permis de parfaire mes compétences dans ce domaine. N'étant pas en studio, il fallait donc optimiser le rapport signal sur bruit et chercher des textures *significatives*. L'enregistrement s'est fait avec un couple Oktava mk-12 en XY et un Roland R44 et m'a permis d'avoir une matière satisfaisante. Afin de gagner du temps, j'ai pu écouter les prises de son dans la foulée ce qui m'a permis également de me rendre compte de l'influence des microphones sur la matière sonore que je voulais enregistrer. C'est un paramètre primordial à prendre en compte.

¹³⁰ La sonification est la représentation et l'émission de données sous forme de signaux acoustiques non verbaux

Les bruitages en ultra-proximité

Dans une situation d'écoute naturelle, l'auditeur se retrouve parfois en situation d'ultra-proximité avec certaines sources sonores. De plus, la perception d'un son est différente si elle se fait en ultra-proximité, car le comportement physique du son est différent à une distance inférieure à 20 cm¹³¹.

Menu Cthulhu possède des moments où le héros se retrouve très proche de sources sonores, par exemple lorsque l'on le recouvre de sel ou lorsque la jeune femme lui fait un baiser. Il était donc primordial pour moi de recréer cette sensation de proximité.

Avec le recul, c'est également cette sensation qui m'avait attiré lors de ma première écoute du *Barber Shop* de QSound. En effet, j'avais vraiment le sentiment que quelqu'un me chuchotait dans l'oreille pendant cette séance virtuelle de coiffure.

J'avais l'intuition que le binaural natif, grâce aux indices binauraux intégrés dès la prise de son, pourrait donner de meilleurs résultats que la synthèse binaurale. La synthèse binaurale comme la technologie ambisonique révèle en effet des aberrations audible lorsque l'on place des sources en ultra-proximité¹³².

J'ai donc effectué un parti-pris : En ultra-proximité, les mouvements de la tête sont négligeables car les sources très proches de la tête et du corps se déplacent avec la tête de l'auditeur. Seule compte alors la sensation de proximité.

Ainsi, je pouvais donc intégrer à mon projet des prises de son en binaural natif.

Grâce à Matthieu Aussal, mon ancien maitre de stage, et le CMAP, j'ai pu avoir accès une matinée à une tête artificielle Neumann KU-100, une référence en matière de binaural natif.

¹³¹ **LLAVE Adrien**, *Simulation et modélisation de l'éloignement d'une source*, Mémoire de Louis-Lumière, 2015 p.75

¹³² **MENZIES, AL-AKAIDI**, « Nearfield binaural synthesis and ambisonics » in *JAES*, 2007



Enregistreur Roland R44 et tête Neumann KU-100 sur perche, gracieusement mis à disposition par le CMAP

En posant la tête sur un pied, j'ai pu donc faire ces bruitages d'ultra proximité seul tout en ayant une écoute de monitoring. Cela a été un challenge car il fallait à la fois se concentrer sur la qualité de l'enregistrement et sur la qualité de la texture sonore que je voulais.

IV.4 Bilan de ces prises de son

J'ai donc choisi d'enregistrer la majorité des ambiances en ambisonique, qui seront binauralisées.

En ce qui concerne les effets et le sound design du monstre, j'ai préféré opté pour de la prise de son monophonique ou stéréophonique qui va être synthétisée directement en binaurale.

Enfin pour les bruitages en ultra-proximité, j'ai opté pour un enregistrement en binaural natif, qui offre de meilleur résultats en terme de présence.

Une fois les prises de son effectuées, il faut alors mettre en place un dispositif de diffusion ainsi qu'un workflow de production approprié, qui exploiteront la variété de ces prises de son.

Chapitre V : Le dispositif mis en œuvre

Ce projet nécessite une mise en place d'un dispositif de production et de diffusion qui est à inventer. En effet, il n'y a aujourd'hui pas encore de *format* ni de *workflow* étendu pour le binaural avec suivi du mouvement de la tête. Comme a pu me témoigner Gaël Nicolas, monteur son pour le cinéma et enthousiaste de l'ambisonique et de la réalité virtuelle¹³³, « c'est aujourd'hui le far-ouest de l'audio ». Il fallait donc structurer chaque étape de la production de ce contenu. Ce chapitre s'intéressera donc à l'ensemble des éléments constitutifs du dispositif mis en œuvre.

V.1 *Le suivi des mouvements de la tête*

Mes premières réflexions ont d'abord porté sur le dispositif de suivi des mouvements de la tête. En effet aujourd'hui, aucun dispositif ne fait l'objet d'une large adhésion. Pourtant aujourd'hui, il est possible d'implémenter ce genre de dispositif à partir d'objets du quotidien possédant un gyroscope, du smartphone à la manette de Wii en passant par une souris informatique. Louis Anglionin, ancien étudiant en section son à l'école, avait même déjà consacré un mémoire entier sur ce sujet ¹³⁴.

Au vu des résultats de ce mémoire et en sachant que 3D SoundLabs a déjà sorti Sound One, un casque possédant son système de *tracking*¹³⁵ intégré, j'ai donc décidé de me concentrer sur une solution dédiée. En effet, avoir un head-tracker dédié, à poser sur le casque, permet non seulement de contrôler totalement la transmission des données et présente des meilleurs résultats, selon les capteurs utilisés.

¹³³ <http://www.spheric.fr>

¹³⁴ **ANGLIONIN Louis**, *De nouveaux outils pour le suivi des mouvements de la tête*, Mémoire de Louis-Lumière, 2013

¹³⁵ Le tracking est un anglicisme définissant tout système permettant de localiser une position ou une direction dans l'espace.

De plus, c'est une pratique qui se retrouve sur le marché aujourd'hui, 3D SoundLabs prévoyant de sortir un capteur à part, qui pourrait s'ajouter sur tous les casques. Aujourd'hui de nombreux détaillants électroniques proposent des ensembles de capteurs permettant d'effectuer le suivi des mouvements de la tête.

Pour ce projet, mon choix s'est porté sur la solution proposée par le détaillant électronique Sparkfun pour deux raisons. Tout d'abord, il existe sur internet un logiciel open source¹³⁶ très documenté qui peut s'embarquer sur la carte électronique et qui permet de transformer la centrale de mesure inertielle (IMU pour Inertial Measurement Unit) en AHRS (Attitude and Heading Reference System).

A la différence d'une IMU seule, un AHRS délivre les informations, tirées des données brutes des capteurs de l'IMU, sur les déplacements de la tête à travers trois axes, appelés Yaw, Pitch et Roll. Ainsi, il est possible de mesurer les déplacements de la tête à 360° et d'obtenir cette information à la sortie de la carte électronique.

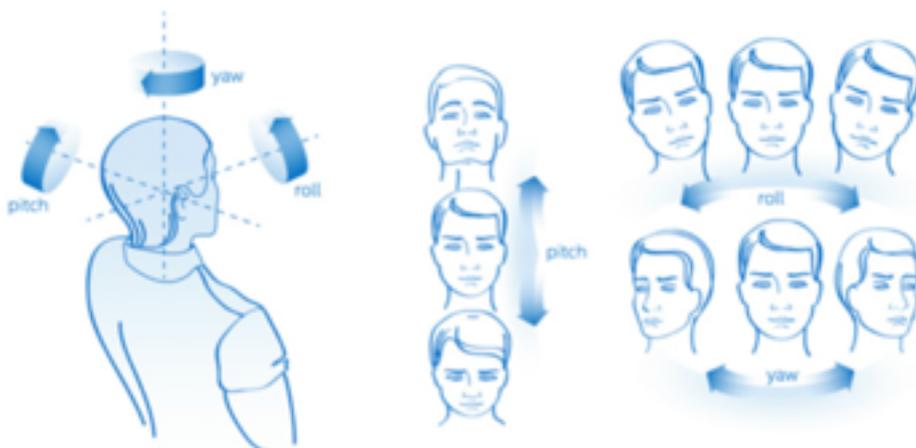


Figure 19. Illustration du Yaw, Pitch, Roll ¹³⁷

Par ailleurs, j'avais été familiarisé à cette solution de head-tracking au cours de mon stage de fin de deuxième année en recherche audio au CMAP.

¹³⁶ <https://github.com/ptrbrtz/razor-9dof-ahrs/wiki/Tutorial>

¹³⁷ <https://software.intel.com/en-us/articles/realsense-overview>

La réalisation de ce head-tracker¹³⁸ en liaison bluetooth a nécessité l'achat de trois cartes électroniques SparkFun.

- *Razor IMU - 9 Degrees of Freedom (9DOF)*

Une centrale à inertie, ou Inertial Measurement Unit (IMU), est un dispositif électronique qui mesure la vitesse, l'orientation et les forces gravitationnelles d'un objet mobile. Cet outil utilise une combinaison d'accéléromètres, de gyroscopes et de magnétomètres. Ceux-ci fournissant des données dans les trois axes de l'espace, le système est constitué de neuf degrés de mesure inertielle, d'où le nom de la carte *9 Degrees of Freedom*. Les sorties des capteurs sont traitées par le micro-contrôleur ATmega328 de Atmel. Une IMU est généralement employée en aéronautique ou en astronautique.

C'est sur cette carte électronique que s'installera le logiciel embarqué permettant de constituer un AHRS. Ce programme permet la fusion des données entre les différents capteurs en s'appuyant sur un algorithme DCM pour *Direction Cosine Matrix*. Le lecteur curieux trouvera dans le tutorial de réalisation du tracker ¹³⁹, les bases mathématiques utilisées au sein du programme.



Figure 20. Carte électronique Razor IMU ¹⁴⁰

¹³⁸ Head-tracker est un anglicisme désignant une solution de suivi des mouvements de la tête

¹³⁹ <https://github.com/ptrbrtz/razor-9dof-ahrs/wiki/Tutorial#mathematical-background-and-firmware-internals>

¹⁴⁰ sparkfun.com

- *FTDI Basic Breakout*

Ce module intègre la puce FT232RL de Future Technology Devices International permettant la sérialisation des données afin d'établir une liaison filaire en micro-USB entre l'IMU et le port série de l'ordinateur.



Figure 21. Carte FTDI ¹⁴¹

- *Bluetooth Mate Silver*

Ce module est constitué d'une puce *RN-42* de *RovingNetworks* qui permet de transmettre à courte portée un flux série de 2400 à 115200 bps entre l'émetteur et le récepteur bluetooth. L'utilisation de cette carte pour le head-tracker nécessite donc une alimentation sur batterie de l'ensemble. Ces cartes peuvent être alimentées par des batteries LiPo, utilisées majoritairement en modélisme, délivrant une tension de 3,7V.

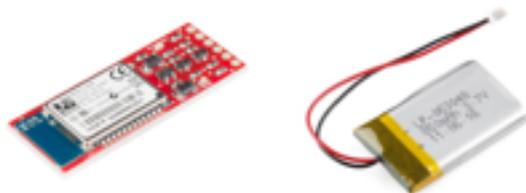


Figure 22. Bluetooth Mate Silver et Batterie LiPo 850mAh/3,7V

¹⁴¹ Ibid

Mise en fonctionnement du Head-tracker

La mise en fonctionnement du head-tracker s'est déroulée selon plusieurs étapes.

Tout d'abord, à la réception des composants électroniques, il fallait souder les connecteurs sur les cartes afin de pouvoir aisément effectuer les branchements. La programmation de ces cartes se fait dans l'univers Arduino. Cela m'a permis d'approfondir mes connaissances à propos de cet univers de programmation.

Il fallait ensuite installer le logiciel AHRS ainsi que le protocole permettant la communication entre l'IMU et la carte Bluetooth.

Enfin, il était nécessaire d'effectuer une calibration des trois capteurs de l'IMU - Gyroscope, Accéléromètre, Magnétomètre -.

En suivant méticuleusement le tutorial proposé par Peter Bartz¹⁴², il est possible d'avoir un head-tracker fonctionnel en une après-midi.



Le casque Sennheiser HD-25 équipé du head-tracker en liaison filaire utilisé pour ce projet

¹⁴² <https://github.com/ptrbrtz/razor-9dof-ahrs>

V.2 L'interface utilisateur

Mon attention s'est ensuite portée sur l'interface utilisateur que je pourrais utiliser. En effet, penser un contenu audio interactif implique de penser à une interface qui serait ergonomique et intuitive, afin de ne pas détourner l'attention de l'auditeur du flux audio.

Le jeu audio ayant déjà eu cette réflexion, je me suis donc tourné vers ce genre. Historiquement, le jeu audio s'est développé sur PC et l'interface la plus utilisée était le duo Clavier/Souris. Néanmoins, quelques jeux audio ont produit des contrôleurs propres¹⁴³, ce qui leur a permis de développer un gameplay unique mais difficilement reproductible car l'interface ne s'est pas démocratisée.

Aujourd'hui avec l'avènement du tout-tactile grâce aux smartphones et aux tablettes, des jeux audio comme *A Blind Legend* utilisent ces outils afin de faire interagir les joueurs.

Les avantages sont multiples. Tout d'abord, l'interface s'est généralisée et peut donc paraître particulièrement intuitive. De plus, elle possède un éco-marché propre -App Store et Google Play- qui permet de distribuer et de promouvoir les jeux audio.

C'est donc en toute logique que j'ai cherché à reproduire ces conditions dans le cadre de ce mémoire. Ainsi, l'auditeur, portant un casque avec un head-tracker, interagira avec une tablette en *touchant* certains boutons afin d'effectuer certaines actions.

¹⁴³ Voir Annexe 2

V.3 *Le workflow de production*

Le dispositif que je voulais mettre en œuvre dans ce mémoire n'a de sens que lorsqu'il est compris dans une chaîne de travail, appelé communément *Workflow*. Or c'est précisément ce cadre qui n'est pas normalisé aujourd'hui et qu'il convient de définir.

Néanmoins en son multicanal, des workflows spécifiques et normés ont émergé, particulièrement dans les systèmes de diffusion orientés objet. A partir des techniques évoquées dans le chapitre IV, il est possible de distinguer plusieurs stratégies :

- La technologie ambisonique, native ou de synthèse, permet de créer une scène sonore prenant en compte les mouvements des sources. Pour ce projet, il est donc possible d'utiliser cette technique en rajoutant un traitement de rotation de la scène sonore lié en temps réel au head-tracker. Par analogie avec la technologie Dolby Atmos, le format ambisonique pourrait donc être considéré comme un *bed*¹⁴⁴ dans le cadre de cette adaptation.
- La synthèse binaurale, à l'inverse, ne permet pas *d'inscrire* dans un format de lecture les mouvements des sources. Il est donc nécessaire d'avoir un double travail de placement des sources dans l'espace lors de la lecture. Tout d'abord, un placement dans l'espace, qui serait effectué par le mixeur lors de la post-production mais également un placement de la source par rapport à la tête de l'auditeur, qui serait lié en temps réel au head-tracker lors de la lecture. Les sons placés en synthèse binaurale pourraient alors être considérés comme des objets, au regard des workflows orientés objet.
- Enfin le binaural natif ne permet pas la prise en compte des mouvements de la tête. Cette technologie reste cohérente avec mon projet esthétique pour des sons en ultra-proximité de l'auditeur. Cela est possible pour des sources qui sont littéralement sur le héros, et où le déplacement de la tête ne permet de déplacer la source sonore.

¹⁴⁴ Un Bed est un mixage traditionnel de type 5.1,7.1,9.1 orienté *channel*, constituant la base principale d'un contenu audio orienté objet - Le son Multicanal, Glossaire

- Afin de créer des effets d'internalisation¹⁴⁵, il est possible d'utiliser dans ce projet des sources monophoniques ou stéréophoniques. Ces considérations peuvent se rapprocher dans l'esprit du canal zénital « *Voice of God* » placé au centre de la salle cinéma en Dolby Atmos.

Une fois ce cadre posé, il devient primordial de réfléchir en fonction de la nature des sources sonores.

- La technologie ambisonique bien que pratique à mettre en oeuvre pour ce projet possède des limitations¹⁴⁶. Avec un ordre peu élevé, la qualité de la localisation spatiale sera altérée et un ordre élevé nécessite un nombre important de haut-parleurs virtuels. En revanche, il permet une bonne description du champ diffus même avec un ordre peu élevé. Les ambiances, les réverbérations artificielles et les effets se situant dans le champ diffus sont les sources les plus adaptées à cette technologie.
- La synthèse binaurale permet, à l'inverse, de mieux localiser les sources ponctuelles, si les HRTFs utilisés dans ce cadre correspondent approximativement à ceux de l'auditeur. En revanche, on ne peut pas utiliser cette technologie sur l'ensemble des sources, car la gestion du head-tracking pour des sources synthétisées en binaural est plus consommateur de ressources processeurs¹⁴⁷. Il convient donc de choisir des sources qui doivent rester précises dans l'espace sonore et qui peuvent être dans le champ direct. Les voix, les bruitages et les effets dans le champ direct me semblent être de parfaits candidats pour cette technologie.

¹⁴⁵ L'internalisation est la perception de sources sonores à l'intérieur de la tête

¹⁴⁶ **SALAUN Stéphanie**, *Formats Audio 3D Hiérarchiques : Caractérisation objective et perceptive des systèmes Ambisonics d'ordres supérieurs*, Thèse de l'INSA Lyon, 2009 p.123

¹⁴⁷ **NOISTERNIG Markus** [sous la dir.] , « A 3D Ambisonic Based Binaural Sound Reproduction System » , *AES 24th International Conference on Multichannel Audio*, p.1

- Les sources enregistrées en binaural natif doivent être en ultra-proximité du Héros-Auditeur, afin de pouvoir négliger les mouvements de la tête. Les bruitages permettant la sensation de présence ainsi que les chutes d'objet sur le héros me semblent être les sources à privilégier.
- Enfin, les sources monophoniques ou stéréophoniques, non placées dans l'espace binaural permettent de créer une distance avec ce même espace et offre ainsi une dimension supplémentaire. La voix du narrateur ainsi que la musique extradiégétique peuvent ainsi se placer dans cet autre espace, plus intimiste et déconnecté du *réalisme* que peut offrir le binaural.

Il est possible de schématiser ainsi :

			
Ambiances	Voix	Narrateur	Sources en Ultra-proximité
			
Champs Diffus	Bruitages et Effets	Musique Extradiégétique	
Technologie Ambisonique	Synthèse Binaurale	Source monophonique	Binaural Natif
<i>Ambi-Pan BFormat2Binaural SoundField ST350</i>	<i>Binauralizer</i>	<i>Microphones « traditionnels »</i>	<i>Tête Neumann KU-100</i>
<i>Remis à jour par le head-tracker</i>	<i>Remis à jour par le head-tracker</i>	<i>Non remis à jour par le head-tracker</i>	<i>Non remis à jour par le head-tracker</i>

Figure 23 Classification des sources sonores en fonction de la technologie de reproduction utilisée

Une fois cette réflexion mise à plat, il est alors nécessaire de choisir les outils permettant la bonne exploitation du head-tracker et de l'ensemble des technologies étudiées.

V.4 Les outils logiciels

Ma volonté première a été de regrouper les différents besoins techniques au sein d'une seule application. Néanmoins, je pressentais que le prototype de mon aventure sonore nécessiterait une puissance de calcul importante. Je n'avais par ailleurs ni les compétences ni le temps de coder une application fluide pour tablette. L'application recherchée serait donc une application sur ordinateur et le prototype pourrait être présenté sous la forme d'une installation, où l'ordinateur, bien que non accessible à l'auditeur, effectuerait l'ensemble des calculs nécessaires au projet. Néanmoins, l'auditeur aura la même *sensation* qu'une application sur tablette, car il interagira uniquement avec celle-ci. Cela a donc représenté pour moi un bon compromis dans le cadre de ce mémoire.

Ce projet nécessite plusieurs fonctions travaillant de concert :

- Le logiciel doit pouvoir être un lecteur de sources sonores. Il doit être capable de lire l'ensemble des sons produits et de gérer les plug-in de spatialisation sonore en temps-réel.
- Le logiciel doit gérer les différents signaux. Il doit être bien sûr capable de recevoir les données du head-tracker et de les traiter, mais également de pouvoir gérer l'interface iPad ainsi que la barre de transport (Lecture/Pause, passage d'un chapitre à un autre...)
- Je souhaitais enfin que le logiciel transmette le flux audio à la tablette afin que l'auditeur branche le casque directement sur la tablette, recréant ainsi la sensation d'une application développée sur iOS.

Certaines de ces fonctions, en particulier la gestion des signaux, nécessitent une part de programmation. Je me suis donc naturellement tourné vers les logiciels de programmation haut-niveau dédiés à la création sonore.

Dans le cadre de mes études à l'école, j'avais déjà pris en main le logiciel de programmation open source Pure Data. Néanmoins, il ne permet une gestion facile des plug-ins de spatialisation sonore.

Je me suis donc tourné vers une des solutions les plus connues du marché : Max.

Max (anciennement Max / MSP)

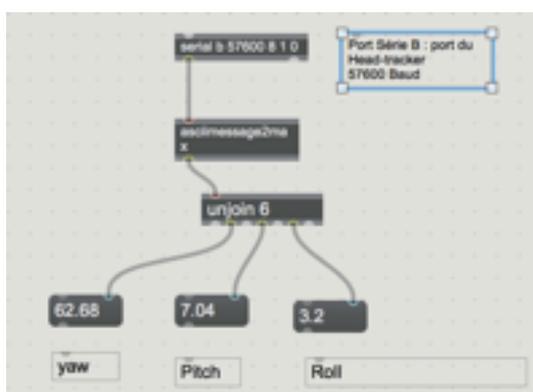
Max est un logiciel musical permettant de faire de la programmation haut-niveau développé par l'IRCAM dans les années 80. Inventé par Miller Puckette, également créateur du logiciel open source Pure Data, il est distribué depuis 1999 par Cycling'74.

Aujourd'hui Max est un acteur majeur du domaine de la création sonore interactive. Une importante communauté d'utilisateurs et de développeurs existe autour de Max, permettant une collaboration et le développement de projets et de logiciels.

Max est à l'origine un logiciel permettant de faire des calculs mathématiques, et donc par extension contrôler des signaux MIDI. Depuis de nombreuses fonctionnalités ont été ajoutées et permettent une utilisation intuitive de plug-in VST par exemple.

Ce logiciel est donc parfaitement adapté pour effectuer le traitement des signaux venant de l'interface utilisateur et du head-tracker.

Les cartes électroniques Sparkfun se développent dans l'univers Arduino. En liaison filaire comme un bluetooth, ces cartes délivrent un flux de données séries. Max possède un outil [*serial*] qui permet de lire les données séries d'un port donné. Néanmoins le flux de données séries donné par les cartes est un flux ASCII¹⁴⁸. Grâce à l'objet externe [*asciimessage2max*]¹⁴⁹, il est possible de convertir ces messages en donnée. Les messages sortant du head tracker se présentent sous la forme : YPR = 0.4 , 0.5 , 0.6 , qui correspondent aux données du Yaw, du Pitch et du Roll récupérables séparément à l'aide de l'objet [*unjoin*].

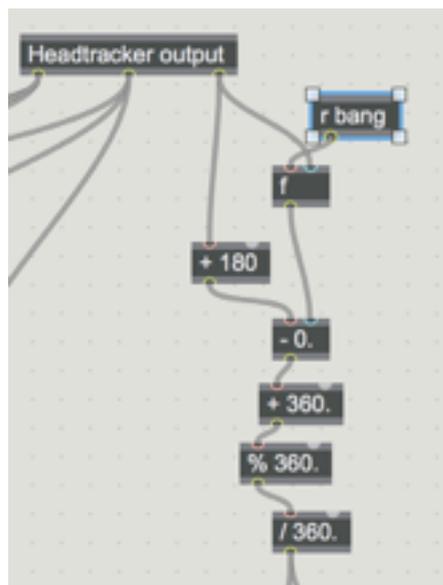


Mise en forme des données du Head-tracker sur Max

¹⁴⁸ American Standard Code for Information Interchange est un format d'encodage des données. Il permet notamment de coder les caractères alphanumériques des claviers informatiques.

¹⁴⁹ https://github.com/mccartyn/Arduino_Dev/blob/master/libraries/SimpleMessageSystem/MaxMsp%20Example/asciimessage2max.mxb

Par ailleurs, il était également nécessaire de coder une mise à zéro du head-tracker. En effet, l'auditeur doit pouvoir à tout moment réinitialiser la position de sa tête dans l'espace. L'idée est donc de soustraire une position de la tête à l'instant de la réinitialisation (un bang) et d'appliquer ce changement par la suite sur l'ensemble des données du tracker. Etant limité entre -180 et 180, il était nécessaire d'ajouter un outil « modulo » afin de pouvoir soustraire sans « clipper » la valeur du head-tracker. Enfin cette valeur divisée par 360 permet d'agir sur les plug-in de localisation sonore qui sont limités entre 0 et 1.



Implémentation de la remise à zéro du Head-tracker pour le Roll

De plus, Max peut transmettre et envoyer des données via Open Sound Control (OSC). Il peut donc communiquer avec une tablette possédant une application de contrôle OSC. Par ailleurs, Max possède une interface dédiée à l'iPad, Mira, qui permet une recopie d'écran du patch utilisé. Il est donc très facile et intuitif de réaliser l'interface homme/machine.

Enfin, Max possède une gestion intuitive des plug-ins VST ce qui constitue un atout pour le lecteur des sons de mon projet. Malheureusement, la gestion de la lecture et des signaux entraîne une surcharge des ressources du processeur. Il était donc nécessaire de trouver une alternative.

Les données du head-tracker doivent être envoyées vers un autre logiciel, qui sera uniquement le lecteur.

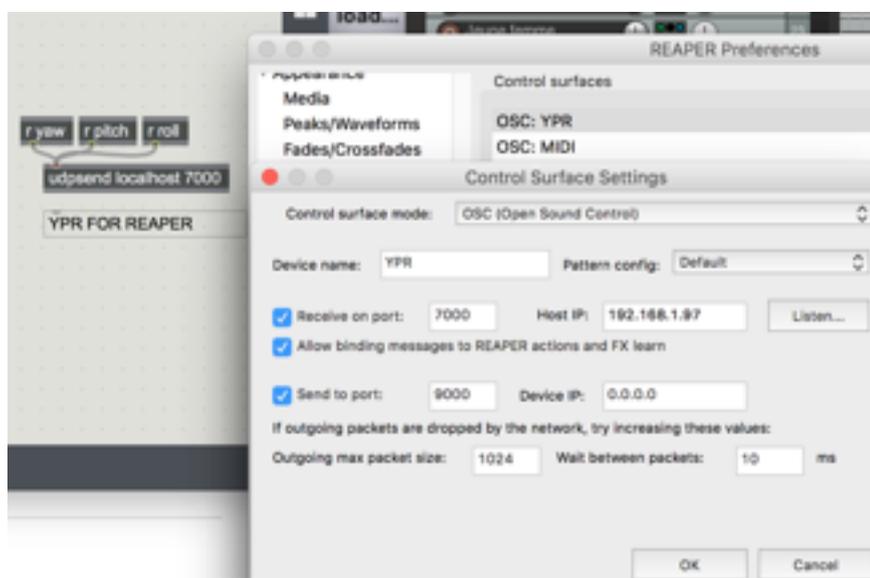
Le séquenceur Reaper, que j'utilise pour la post-production de ce projet, permet de récupérer les données du tracker et d'envoyer des données via le protocole OSC.

Le protocole OSC (Open Sound Control)

L'Open Sound Control est un format de transmission de données entre logiciels, ordinateurs, synthétiseurs ou tout autre matériel compatible. Il a été développé en 2004 par Université of California, Berkeley Center for New Music and Audio Technology. Il a été conçu pour le contrôle et la gestion de signaux en temps-réel. Il utilise la technologie réseau au travers des protocoles UDP¹⁵⁰ et permet une amélioration nette en fluidité et rapidité, par rapport à la norme MIDI, qui remplit la même fonction.

Des plug-ins open source tel que Tosca¹⁵¹ de l'IRCAM, permettent à n'importe quel séquenceur d'envoyer et de recevoir des données en OSC. Max par ailleurs possède des outils tels *udpsent* et *udpreceive* qui permettent de traiter l'OSC. Le séquenceur Reaper permet lui aussi de façon native d'envoyer et de recevoir des données en utilisant le protocole OSC.

Une fois les ports d'entrée et de sortie bien paramétrés, la communication entre Max et Reaper peut se faire sans problème. Il n'y a plus qu'à assigner ces données OSC au plug-in de spatialisation sonore (*Binauralizer* et *BFormat2Binaural*).



Configuration du protocole OSC entre Reaper et Max

¹⁵⁰ L'User Datagram Protocol (UDP) est un des principaux protocoles de télécommunication sur internet.

¹⁵¹ Thibaut's OpenSoundControl Automation est un plug-in permettant à un séquenceur d'envoyer et de recevoir des données en OSC. Il a été développé principalement pour la gestion du Spat.

La programmation de l'aventure

Grace à Max, Reaper et Tosca, il m'a été possible de coder l'ensemble des interactions nécessaires à cette aventure.

Afin de coder les choix que l'auditeur doit faire avec la tablette, j'envoie une information de Reaper vers Max grâce à Tosca, afin de faire apparaître les boutons sur l'iPad.

J'ai ensuite assigné le raccourci Reaper « Lire au marqueur X » à une information OSC, qui est transmises uniquement si l'auditeur appuie sur le bouton. Les marqueurs sont ainsi placés au début de chaque chapitre.

Cette technique est également utilisée pour le retour au début de l'aventure.

Ces informations OSC permettent également de faire disparaître les boutons sur l'interface graphique.



Interface iPad lors des phases de choix

Concernant la phase d'exploration, une information ToscA permet de l'activer dans Max. Au début de chaque chapitre où se déroule une phase d'exploration, l'angle de la source sonore à trouver (α) est chargé.

Une tolérance de 40° est donné autour de cette angle. Si les données provenant du head-tracker sont supérieure à la limite basse ($\alpha - 20^\circ$) et inférieure à la limite haute ($\alpha + 20^\circ$), un compteur se déclenche et se remet à zéro dès qu'une des deux conditions n'est plus remplie.

Au bout de 3 secondes, le compteur délivre une information OSC transmise à Reaper lui permettant de lire la suite de l'aventure.

Une information ToscA est envoyée à Reaper si le joueur ne parvient pas à réussir cette phase au bout de 10 secondes. Cela déclenche la fin de cette phase de jeu et l'apparition d'un bouton action sur l'iPad, permettant à l'auditeur de passer à la suite sans être frustré.



Interface iPad lorsque l'auditeur n'a pas réussi la phase d'exploration

La transmission du flux audio de l'ordinateur à la tablette

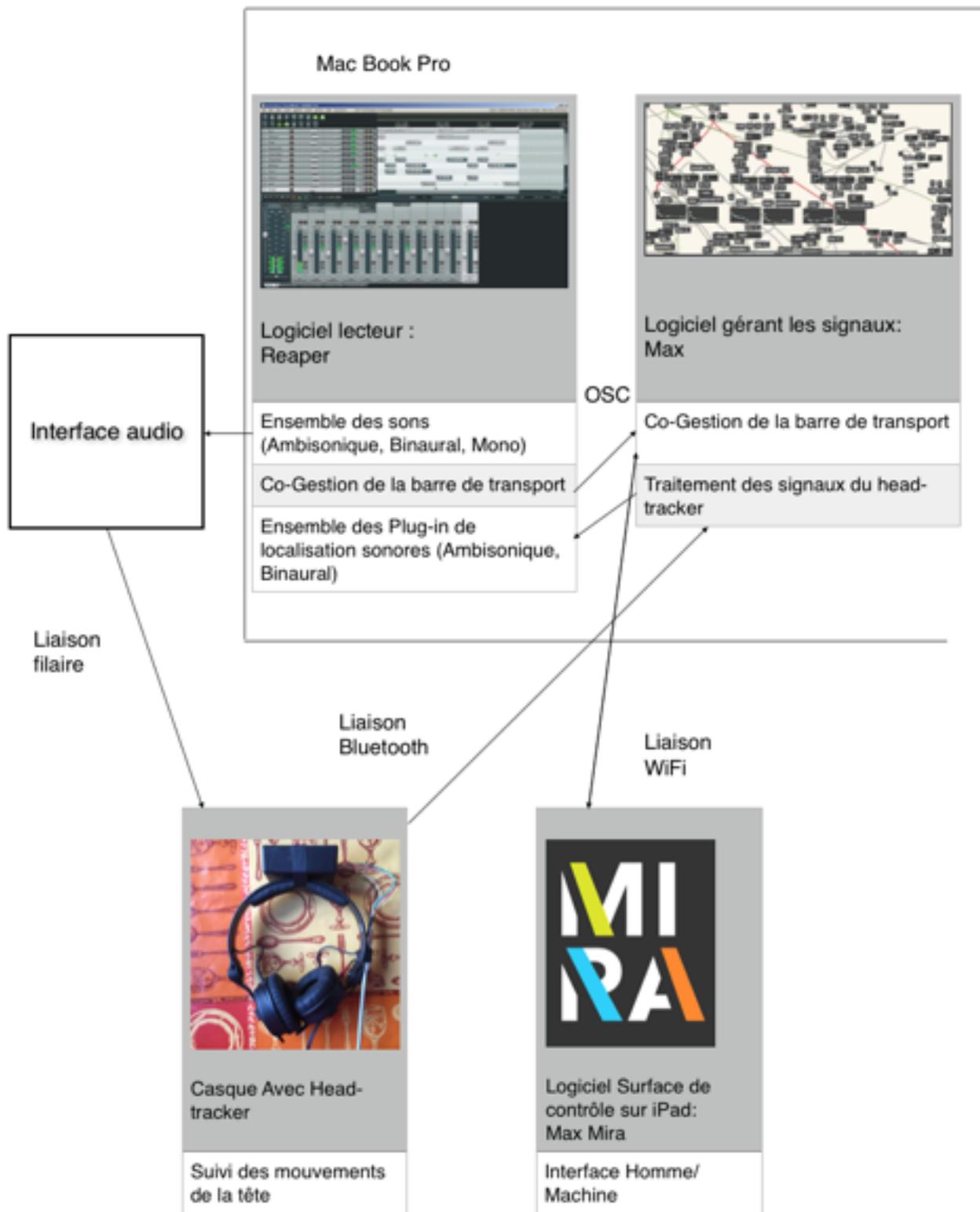
La post-production de cette aventure sonore a été réalisée sur un Mac. C'est un iPad qui sera l'interface Homme/Machine. Il fallait donc trouver un logiciel tournant dans l'univers Mac, capable de transmettre de l'audio d'un Mac vers un iPad. De nombreux programmes permettant de réaliser cette fonction ont émergé.

Malheureusement aucun d'entre eux ne m'a permis de faire transiter les deux canaux audio en WiFi en utilisation parallèle avec Mira. Il aurait été nécessaire d'avoir deux cartes réseaux distinctes et de créer deux réseaux WiFi afin Mira ne prennent pas le débit nécessaire à la transmission de l'audio.

N'ayant pas les moyens d'effectuer cela, j'ai donc mis en place une liaison filaire entre la carte son de l'ordinateur et l'auditeur.

Une fois l'ensemble des outils et des protocoles choisis et définis, il est possible de schématiser l'ensemble de ma réflexion dans un workflow de diffusion pour le prototype, qui représente la Partie Pratique de mon Mémoire et de pouvoir commencer pleinement la post-production de cette aventure dont vous êtes le héros.

Figure 24. Workflow de diffusion du prototype



Chapitre VI : La post-production

C'est la matière qui doit décider, c'est pas toi. Parce que si tu fais déjà le cadre au niveau du minutage, tu vas passer ton temps à couper ou à rallonger. ¹⁵²

Dans ce chapitre, il sera question de la post-production de cette adaptation, du montage au mixage. C'est à cette étape que j'ai pu concrètement mettre en œuvre l'esthétique sonore de mon adaptation. Cette post-production s'est effectuée au casque, le même que celui utilisé lors de la diffusion du prototype.

VI.1 Un montage « *classique* »

Le montage, à savoir l'agencement temporel des différentes sources sonores, s'est déroulé de la même façon que celle d'un film. J'ai d'abord monté le dialogue que l'on peut assimiler au montage des *directs* au cinéma. Etant l'adaptation d'un livre, il était particulièrement nécessaire d'avoir des dialogues fluides, afin de faire oublier le côté *écrit* de ceux-ci. A partir des différentes sessions d'enregistrement, j'ai donc pu constituer une première continuité dialoguée, véritable *ossature* de cette aventure. Avec un support de diffusion n'ayant pas de contrainte de durée, j'ai pu avoir un *ours*¹⁵³ d'une trentaine de minutes.

A partir de cette base, j'ai pu rajouter les différentes ambiances et les effets. Cela s'est fait avec un aller-retour avec les dialogues. En effet, à certains moments, il me semblait plus naturel de laisser vivre l'ambiance sonore. De la même façon, certaines phrases se sont avérées redondantes par rapport à l'ambiance sonore créée.

Contrairement à une chaîne de post-production cinéma, le montage son a largement continué lors du mixage, car je savais que le binaural peut démasquer certaines sources et faire apparaître des *trous* dans l'espace sonore créé.

En effet, cette technologie permet une discrimination spatiale plus importante.

¹⁵² PARENTHOEN Yann, *Propos d'un tailleur de son*, Paris, Editions phonurgia nova, 2002

¹⁵³ Un ours au cinéma est un pré-montage juxtaposant l'ensemble des plans qui vont être utilisés, afin d'avoir une première idée du résultat

VI.2 Pistes d'analyse d'une création sonore binaurale

L'ouvrage collectif *Le son multicanal* propose dans son chapitre 4 un ensemble de critères objectifs qui peuvent servir à analyser n'importe quelle création sonore multicanale.

La création binaurale, étant spatialisée, peut également être analysée selon ses critères.

Image réelle ou virtuelle ?

Un des premiers critères concerne la caractérisation de l'image sonore. Selon Guy Laporte¹⁵⁴, « l'image sonore est définie comme étant la représentation mentale d'un sujet et de son acoustique »¹⁵⁵.

Elle est peut être réelle si elle est corrélée à un lieu d'enregistrement. Cela peut être le cas dans un concert de musique classique ou un événement sportif. En résulte le plus souvent une esthétique frontale, l'action se déroulant généralement face au spectateur.

L'image sonore est virtuelle si elle crée « un espace imaginaire, qui en multicanal, met en scène les sons tout autour de l'auditeur »¹⁵⁶. Ainsi l'image virtuelle possède alors généralement une esthétique équirépartie où le champ direct et le champ réverbéré sont situés de façon égale autour de l'auditeur.

Dans le cadre de cette aventure dont vous êtes le héros, c'est une image sonore virtuelle ainsi qu'une esthétique équirépartie qui ont été mises en place. En effet, le texte original faisant la part belle au fantastique, il me semblait permettre une répartition équilibrée des sources dans l'espace.

¹⁵⁴ Guy Laporte est un preneur de son au cinéma

¹⁵⁵ **PERIAUX Bergame, OHL Jean-luc, THEVENOT Patrick**, *Le son multicanal*, Paris, Edition Dunod, 2015, p.147

¹⁵⁶ Ibid

Constitution d'une image sonore

Une image sonore peut être pensée et analysée selon plusieurs critères :

- La localisation

Une image est caractérisée par la localisation des différentes sources qui la composent.

Cette localisation peut être perçue dans la distance.

La profondeur est un critère traduisant la distance maximale ressentie à l'écoute d'un enregistrement. Une image sonore profonde permet de mieux rendre compte de l'acoustique dans laquelle l'auditeur est immergé.

Le relief, quant à lui, « repose sur l'existence de différents plans sonores »¹⁵⁷. Une image avec du relief permet de donner un certain réalisme à l'image sonore.

Dans le cadre de ce mémoire, j'ai essayé de privilégier une image profonde avec du relief avec d'essayer de recréer des conditions quasi-naturelles.

La localisation peut être également analysée d'un point de vue latéral.

En esthétique frontale, la localisation des sources sonores reste le plus souvent sur l'avant, par exemple sur les trois enceintes L C R en diffusion 5.1.

En esthétique équirépartie, les sources sont placées en cercle autour de l'auditeur.

Ayant choisi cette esthétique, c'est donc une localisation circulaire qui a été privilégiée.

- L'impression spatiale

L'impression spatiale renvoie à la sensation de l'auditeur *d'être* dans un espace.

Elle peut être donnée par la largeur apparente des sources. Il est possible de définir la largeur apparente des sources comme « la distance maximale sur laquelle s'étend un secteur du multicanal »¹⁵⁸. Dans le cas d'une esthétique équirépartie, il est possible de définir trois secteurs : frontal, latéral et arrière. La largeur et la perception de la taille d'un objet sonore sont corrélées de fait.

¹⁵⁷ Ibid p.149

¹⁵⁸ Ibid p.150

Au cours de la post-production de cette adaptation, j'ai essayé de moduler cette largeur des sources afin de servir le récit.

L'impression d'espace peut être également donnée par l'enveloppement de l'auditeur, à savoir la sensation d'être entouré par les sons. A priori, une esthétique équirépartie, avec un rapport frontal/arrière équilibré, procure un enveloppement optimal. L'enveloppement de l'auditeur peut permettre l'immersion de l'auditeur dans la scène virtuelle « comme s'il se trouvait réellement dans ce lieu »¹⁵⁹.

Enfin l'impression spatiale peut être définie par la réverbérance des sources sonores, à savoir « le prolongement temporel des événements »¹⁶⁰, que ce soit grâce à une réverbération naturelle enregistrée dès la prise de son ou artificielle ajoutée au mixage.

Pour ce projet, j'ai utilisé majoritairement des réverbérations artificielles que j'ai rajoutées aux voix et aux effets enregistrés en proximité.

Néanmoins lorsque la narration s'y prête, j'ai également utilisé des ambiances ambisoniques dans des lieux très réverbérés (la Nef de la Cité du Cinéma, les escaliers ultra-réverbérants de l'école) afin de créer de nouveaux espaces.

- *Les équilibres*

Une scène sonore peut être également étudiée en termes d'équilibre.

L'esthétique équirépartie m'a guidé vers un équilibre frontal/arrière relativement strict. J'ai essayé de constituer une image sonore possédant une continuité de localisation. J'ai également essayé d'avoir une scène sonore équilibrée spectralement et d'avoir un rapport champ direct/ champ réverbéré constant quelle que soit l'orientation dans l'espace des sources sonores.

Néanmoins, certaines sources comme le monstre sont placées dans l'espace et possèdent un équilibre spectral et une réverbérance propre. Je n'ai donc pas cherché un équilibre strict mais plutôt une sensation globale cohérente entre l'avant et l'arrière.

¹⁵⁹ Ibid p.151

¹⁶⁰ Ibid

- *La couleur de l'image et la qualité globale du signal*

Une image sonore peut enfin être analysée au regard de sa couleur et de la qualité du signal audio.

La qualité de restitution des timbres des différentes sources est forcément approximative, puisque le timbre de la majorité des sources se retrouve altéré par le processus de synthèse binaurale, qui effectue un filtrage en fonction de son placement dans l'espace. De plus, le casque est également un transducteur déformant. La reproduction à l'identique d'une source sonore semble donc impossible.

Néanmoins, j'ai essayé de conserver une certaine clarté permettant de facilement identifier les timbres des différentes sources constituant la scène sonore.

La qualité du signal audio rassemble un ensemble de caractéristiques objectives, qui peut se décomposer ainsi :

- La définition « désigne la pureté du signal sonore, sa finesse et la précision de ses transitoires avec laquelle on la perçoit »¹⁶¹.
- L'intelligibilité, désigne plus spécifiquement le degré de compréhension d'une voix.
- La transparence est la « capacité de mettre en relation immédiate les sources sonores et l'auditeur, en l'absence de tout autre élément »¹⁶².
- Enfin la dynamique représente la différence entre les moments les plus forts d'un contenu et les moments les plus faibles. Etant dans le cadre d'une diffusion au casque, il était particulièrement nécessaire de faire attention à ce critère. Le transducteur étant au plus proche de l'oreille de l'auditeur, il est nécessaire pour son confort de ne pas avoir de sons avec un dynamique instantanée trop importante. Néanmoins, l'écoute au casque permet de distinguer des sons à volume très faible, qui restent encore compréhensible par l'auditeur.

J'ai donc pris en compte l'ensemble de ces éléments afin d'avoir une qualité la plus optimale possible.

¹⁶¹ Ibid p.154

¹⁶² Ibid p.155

VI.3 La technologie au service de la narration

J'ai tenté lors de la spatialisation des sources sonores de mettre les différents outils au service du récit.

VI.3.1 Le placement des sources dans l'espace

Un des enjeux principaux lors de la post-production de mon aventure dont vous êtes le héros a été le placement des sources sonores dans l'espace. En effet, je recherchais une esthétique sonore équi-énergie. Il s'agit donc de placer autant de sources à l'avant et à l'arrière de l'auditeur. Pour autant, je ne souhaitais pas déstabiliser l'auditeur avec des voix placées de façon diamétralement opposée.

Sauf dans des cas exceptionnels dictés par le récit, les voix ont donc été placées majoritairement à l'avant et les réactions du monstre, caché dans le frigo, à l'arrière.

Cela m'a permis de construire une scène sonore équilibrée tout en plaçant les dialogues de façon privilégiée.

Néanmoins, s'il le souhaite, il a la possibilité de tourner la tête afin de percevoir plus frontalement les bruits produits par le monstre.

En ce qui concerne les effets, je souhaitais que le monstre puisse se déplacer et que l'auditeur perçoive le plus distinctement possible ses mouvements. J'ai donc privilégié le déplacement du sound design du monstre d'un seul côté, afin de ne pas provoquer de confusion de spatialisation.

Les sources sonores où je me suis posé le plus de questions concernant leur placement sont la voix du narrateur et celle du héros.

En effet, le narrateur étant extra-diégétique, il me semblait absolument nécessaire que l'auditeur le comprenne relativement intuitivement. La diction, plus proche du texte lu ainsi que le timbre de la voix permettent de déjà de distinguer le narrateur.

Néanmoins j'ai donc décidé de ne pas binauraliser la voix du narrateur afin qu'il marque une différence supplémentaire avec les autres voix.

La voix du héros m'a également questionné. Je souhaitais en effet dans cette aventure que l'auditeur ait la *sensation* d'être le héros. J'ai choisi donc de la dupliquer sur deux pistes. Une piste n'est pas binauralisée et j'ai choisi de renforcer légèrement les bas médiums. L'autre piste a été réverbérée et binauralisée mais ne répond pas au dispositif de mouvement de la tête. Ainsi la réverbération de la voix du héros se situe toujours à l'avant de l'auditeur.

VI.3.2 L'internalisation

De façon plus générale, l'internalisation de certaines sources m'a permis de jouer la rupture avec les autres sources spatialisées. Ainsi, elles acquièrent un statut différent.

L'internalisation m'a permis de facilement sonifier les pensées du héros. Lorsque le héros évoque son expérience dans un restaurant japonais, l'ambiance est monophonique et non binauralisée et contraste avec la scène sonore.

De la même façon lorsque le héros évoque son plan d'action face au monstre, l'ensemble des étapes est sonifié et n'est pas placé dans la scène sonore binaurale.

La rupture entre un son internalisé et la scène sonore binaurale me semble d'autant plus forte.

VI.3.3 La gestion du Width du Binauralizer

Le paramètre Width¹⁶³, présent dans le plug-in de synthèse binaurale, Binauralizer permet de gérer la largeur d'une source sonore dans l'espace binaural. Il est possible d'automatiser ce paramètre et ainsi de créer un effet d'englobement d'une source qui peut progressivement occuper l'ensemble de l'espace.

J'ai particulièrement utilisé cet effet pour transcrire le rapprochement du monstre lorsque celui-ci s'apprête à dévorer le héros. Ainsi, les sons des tentacules, ponctuels dans l'espace, deviennent très présents et se retrouvent sur l'ensemble de la scène sonore.

VI.3.4 La réverbération libérée

La gestion de la réverbération, qui permet de créer de la profondeur et du relief a été un point de réflexion crucial lors de la post-production. En effet, je souhaitais créer de la profondeur et du relief, afin de donner à entendre un espace *virtuel*.

J'ai choisi de rajouter une réverbération artificielle sur l'ensemble des sources et de placer la réverbération dans la sphère ambisonique au même endroit que la source, grâce au plug-in AmbiPan. Ainsi, j'ai conservé un nombre de sources synthétisées en binaural relativement faible tout en rajoutant l'ensemble des réverbérations principales ou secondaires que je souhaitais. La réverbération étant de nature diffuse, le passage par l'ambisonique de 1er Ordre ne crée pas d'aberration de localisation.

¹⁶³ Width signifie littéralement largeur en anglais

VI.4 Etude de cas : le début de l'aventure

L'ouverture de mon aventure dont vous êtes le héros est symptomatique des problématiques de spatialisation et de mixage que j'ai eues pour l'ensemble de ce projet. C'est un moment d'autant plus important que le début de l'aventure donne le ton et pose les bases de l'esthétique sonore développée par la suite.

Le contenu s'ouvre donc sur une ambiance ambisonique d'un fast-food où le héros, au comptoir, doit prendre des commandes toute la journée. Le héros répète, machinalement, les mêmes actions et les mêmes phrases. Il s'agit pour moi dès le début de faire prendre conscience à l'auditeur de la lenteur et de l'ennui que peut ressentir le personnage.

La spatialisation de la caisse enregistreuse et de la voix, ralentie par moments afin de sonifier le fait que « la vie passe au ralenti », permet de faire vivre à l'auditeur cette répétition.

La caisse enregistreuse, spatialisée précisément dans l'espace au début, prend davantage de place, grâce au paramètre Width et se transforme petit à petit en une machine à sous.

La post-production de cette adaptation a permis de me poser des questions sur l'élaboration d'une image sonore à 360°. Ces réflexions, loin d'être absolues et exhaustives, peuvent constituer également des pistes pour d'autres types de contenu.

Par ailleurs, elles ne peuvent être véritablement validées qu'après une écoute critique, qu'il a été convenu d'organiser.

Chapitre VII : Vers un développement du prototype

Le peu qu'on peut faire, le très peu qu'on peut faire, il faut le faire.

Théodore Monod, scientifique naturaliste français

Une fois la mise en place du workflow de diffusion et la post-production sonore terminée, il me semblait nécessaire de présenter mon prototype à un public. J'ai présenté pendant quatre jours mon prototype à l'écoute sur réservation. Ainsi chaque personne disposait d'une demi-heure de prise en main du prototype et avait l'occasion de faire deux des quatre fins possibles.

En plus de récolter des retours d'expérience sur le contenu en lui-même, cette présentation a également été un moyen de tester la viabilité de mon prototype.

VII.1 Critique du dispositif de diffusion

Au bout de ces quatre jours de présentation, je me suis aperçu de certaines limites de mon prototype. La liaison bluetooth entre l'ordinateur et le head-tracker ainsi que la liaison Wi-Fi entre l'iPad et l'ordinateur peuvent parfois être capricieuses. Il convient de s'y reprendre à plusieurs fois avant de pouvoir connecter ses appareils.

Le head-tracker se révèle fragile à l'usage. En effet, il a au fil des journées introduit des aberrations de placement de la tête. Après des tests, il s'avère que l'accéléromètre de la carte SparkFun doit être recalibré régulièrement afin que les mouvements de la tête, même rapides, n'introduisent plus d'aberration.

Le calibrage de l'accéléromètre étant fastidieux, je n'ai pas, par manque de temps, pu recalibrer le capteur à chaque début de journée.

La conséquence est que certains auditeurs ont ressenti des passages brutaux de gauche à droite de certaines sources, lorsqu'ils déplaçaient rapidement leur tête.

VII.2 Retour d'expérience utilisateur

Au cours de ces quatre jours, 30 personnes ont pu prendre en main le prototype ¹⁶⁴. Ils en ont disposés pendant 30 minutes, ce qui leur a permis de vivre au moins une fin sur les quatre possibles. Ils ont eu en amont un ensemble d'indications concernant le dispositif. A la suite de leur passage, j'ai pu recueillir leurs impressions au travers d'un petit questionnaire¹⁶⁵.

Le but de celui-ci était de s'assurer que ma réflexion autour de ce prototype avait été perçue.

En ce qui concerne le profil des auditeurs, un peu plus d'un tiers des personnes étaient habituées à écouter du contenu binaural. Un tiers des participants jouent régulièrement aux jeux vidéo et la quasi-majorité ne jouent jamais à des jeux audio.

L'ensemble des auditeurs était enthousiaste à propos de cette aventure dont vous êtes le héros et la quasi-majorité a pris en main facilement l'ensemble des fonctionnalités.

Néanmoins, environ un quart des participants, soit ne s'est pas senti totalement à l'aise avec le casque avec le capteur. Il est donc évident que pour un hypothétique développement, il conviendrait d'améliorer l'implémentation du capteur de mouvement au sein du casque et d'arriver à une forme aboutie.

Concernant les points d'intérêts de cette aventure, seulement 3 personnes se sont déclarées peu ou pas intéressées par le récit. Il y a donc 90% des auditeurs qui ont adhéré à cette histoire. J'avais choisi d'adapter *Menu Cthulhu*, parce que j'avais le sentiment qu'il pouvait rassembler un maximum de gens, grâce à son humour absurde et son ton décalé.

Par ailleurs, 5 personnes se sont déclarées être peu ou pas intéressées par les phases d'explorations, où l'auditeur devait bouger la tête afin de faire évoluer le récit. Cette phase de jeu a tout de même intéressé une majorité de participants.

¹⁶⁴ Voir Annexe 6

¹⁶⁵ Voir Annexe 5

La critique principale que j'ai eue était que ces phases n'étaient pas forcément intégrées à la narration et pouvait frustrer l'auditeur qui se retrouve dans une situation de jeu qu'il n'a pas forcément choisi.

Concernant l'immersion, environ 2/3 des auditeurs a eu la sensation *d'être* le héros. C'était un objectif de cette adaptation et les auditeurs l'ont majoritairement perçu.

Concernant l'interaction, seulement deux personnes ont préféré l'interaction avec l'iPad à l'interaction grâce à leur tête, possible durant les phases d'*exploration*.

Par ailleurs, l'interaction grâce au dispositif de suivi des mouvements de la tête est ce qui semblait le plus novateur dans ce contenu pour la majorité des auditeurs.

J'ai également demandé aux auditeurs de classer certains points d'intérêts - histoire, interaction et liberté de découvrir l'espace - de 1 à 3 (1 étant ce qui leur a semblé le plus intéressant).

L'objectif de cette question était de savoir si les différentes composantes de la théorie LNS - Ludisme, Narrativisme, Simulationnisme - étaient perçues de la même manière par les auditeurs et si une hiérarchie émergerait de ce contenu.

Ayant cherché à équilibrer les différentes composantes, je me suis aperçu à la lecture des résultats qu'aucune tendance n'a véritablement réussi à émerger.

La composante simulationniste recueille une légère faveur, 12 personnes l'ayant trouvé plus intéressante que les deux autres. Les deux autres composantes ont 9 personnes les ayant préférées.

Concernant les attentes, 24 auditeurs sur 30 auraient souhaité que l'aventure soit plus longue et cette même proportion d'auditeurs aurait souhaité d'avantage d'interactions.

Il serait donc possible dans un hypothétique développement de réfléchir à des adaptations plus longues et plus interactives.

C'est la monétisation de ce genre de contenu s'il était totalement abouti qui a le plus divisé les auditeurs. Sur les 30 auditeurs, 16 personnes seraient prêtes à payer pour ce genre de contenu. En vue d'un hypothétique développement, il serait sûrement à prévoir une monétisation publicitaire ou encore un modèle *freemium*, où il serait nécessaire de payer pour débloquer l'ensemble des contenus.

VII.3 Pistes pour le développement du contenu

Ces écoutes critiques ont été également l'occasion d'esquisser plusieurs pistes concernant le développement du concept de *l'aventure sonore dont vous êtes le héros*.

Concernant les fonctionnalités, certains auditeurs ont été frustrés de devoir repartir du début s'ils souhaitaient enchaîner deux fins différentes. Je n'avais pas conçu le contenu dans ce sens, mais il s'avère nécessaire d'ajouter une fonction « Retour au choix précédent » pour ceux qui souhaitent enchaîner l'ensemble des fins.

Une majorité d'auditeurs interrogés souhaitait encore d'avantage d'interactions.

Il serait donc souhaitable de mener une réflexion sur les possibilités d'interactions grâce à la technologie binaurale. Une des critiques des *phases d'explorations* est qu'elles ralentissent la narration et que la phase de jeu n'apporte rien à l'histoire.

Comment peut-on mieux lier la narration et l'interaction ?

Une des pistes possibles serait de rajouter, au lieu de substituer, une phase d'enquête à la phase d'exploration. L'auditeur pourrait aller via ses mouvements de la tête, discuter avec différents personnages et ainsi en apprendre d'avantage sur l'histoire, selon ses choix. L'auditeur serait alors acteur de la narration.

Une autre remarque que j'ai eue concerne la prise en main des fonctionnalités. Certains auditeurs ont préféré, afin de maximiser leur expérience, se lever et fermer les yeux. Ils ont alors été frustrés de devoir ouvrir les yeux afin de choisir sur la tablette.

De plus, le livre comme le jeu audio étant traditionnellement destiné à un public malvoyant, il est impératif dans un hypothétique développement d'inclure une possibilité de vivre cette aventure les yeux fermés. Ainsi, le suivi des mouvements de la tête pourrait faire office de tablette. Certains gestes, comme le hocher « oui » ou « non » de la tête, pourraient être pris en compte comme des commandes de l'interface.

Ces pistes sont évidemment loin d'être exhaustives et montrent qu'il est possible d'envisager de nombreux développements à ce type de contenu.

VII.4 Quid des normes ?

Le workflow de production que j'ai proposé n'est pas adapté aux normes standardisées de diffusion. En effet, aucune de ces normes ne propose de gérer les sources *objets* et ne prévoient pas une synthèse binaurale à l'écoute.

Néanmoins, deux normes prenant en compte ces paramètres commencent aujourd'hui à émerger. Il est alors intéressant de se pencher sur ces normes et de voir si elles seraient adaptées à la diffusion de cette aventure.

VI.4.1 Le MPEG-H

Le MPEG-H désigne un ensemble de standards en cours de développement par le ISO (International Organization for Standardization), l'IEC (International Electrotechnical Commission) et le MPEG (Moving Picture Experts Group). Il comprend principalement un nouveau container de fichiers multimédia ainsi que de nouveaux standards de compression audio et vidéo.

Officiellement, le MPEG-H est appelé « ISO/IEC 23008 - High efficiency coding and media delivery in heterogeneous environments »¹⁶⁶. Il a donc été pensé comme une norme unique, où l'ensemble des supports de diffusion (TV, tablettes, ordinateur,...) aurait été pris en compte. L'objectif est de produire du contenu, d'origine télévisuelle, sous un format unique et la norme permettra d'adapter le contenu au système de diffusion¹⁶⁷.



Figure 25. Graphique de présentation du MPEG-H 3D Audio par Fraunhofer IIS¹⁶⁸

¹⁶⁶ http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=62835

¹⁶⁷ <http://www.mpeg-haa.com/presentations/MPEG-H%20Fraunhofer%20HPA%20February%202015.pdf>

¹⁶⁸ http://www.mpeg-haa.com/presentations/NeuendorfM_NAB2014_MPEGH_Status_v013.pdf

Le MPEG-H Part 3, plus connu sous le nom de MPEG-H 3D Audio, correspond à la partie audio de cette norme. Il a été impulsé par des entreprises européennes telles Fraunhofer IIS, créateur du MP3 ou encore Technicolor et Qualcomm.

La partie audio, à l'instar de l'ensemble de la norme, a été pensée afin d'inclure l'ensemble des technologies de reproductions sonores. Il est ainsi annoncé que cette norme serait capable de gérer de la synthèse binaurale, des objets audio, comme de l'ambisonique d'ordre supérieur (HOA)¹⁶⁹.

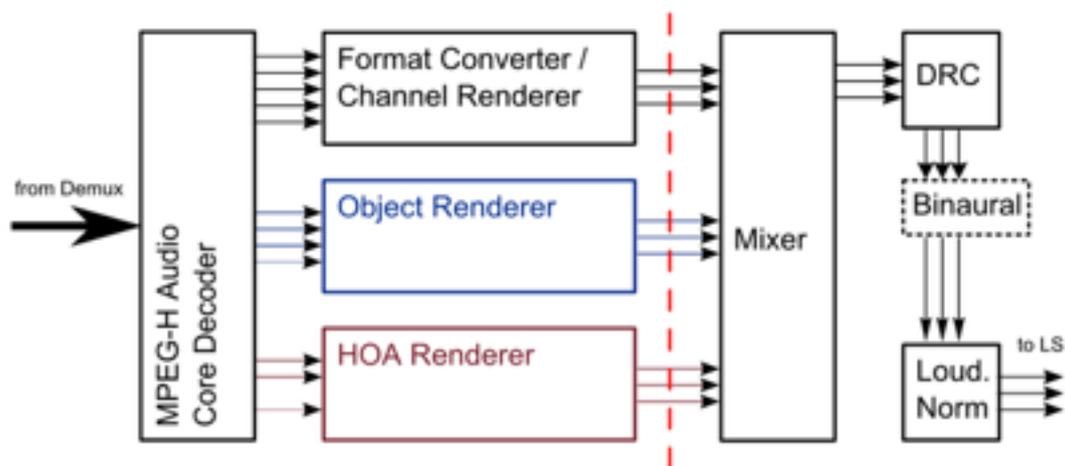


Figure 26. Présentation du décodeur audio du MPEG-H, un synthétiseur binaural peut donc être rajouté en fin de chaîne de décodage.¹⁷⁰

Cette norme semble très adaptée à mon contenu, car le décodeur effectue l'ensemble des tâches effectuées par mon prototype.

Néanmoins, cette norme, toujours en développement, n'a pas encore l'adhésion des grands groupes de télévision et de radiodiffusion. Par ailleurs, le décodeur MPEG-H, audio et vidéo, semble gourmand en ressources et la diffusion avec un tel standard impacterait forcément ceux qui ne possèdent pas un terminal multimédia assez puissant.

Par ailleurs, il est concurrencé par une autre norme d'origine américaine, qui peut déjà s'appuyer sur la renommée de la marque qui la commercialise.

¹⁶⁹ Ibid

¹⁷⁰ Ibid

VI.4.2 Le Dolby at home et Dolby on the go

Dolby Laboratories, Inc est une société fondée par Ray Dolby en 1965. D'abord spécialisée dans la réduction de bruit de fond, elle produit des technologies de diffusion multicanale, d'abord pour le cinéma puis pour l'ensemble des médias. Ainsi le surround aujourd'hui est souvent associé à la marque Dolby.

En 2012, Dolby a développé pour le cinéma une nouvelle technologie hybride de reproduction du son surround, le Dolby Atmos. Les salles Dolby Atmos se voyaient rajouter un nombre considérable de haut-parleurs (entre 35 et 54 enceintes).

Cette technologie a permis pour la première fois d'introduire la notion *d'objet* sonore au cinéma. Ainsi, une source n'est plus encodée sur un canal, mais est diffusée avec ses métadonnées de localisation. Aujourd'hui, Dolby s'est imposé comme un leader face au DTS-X et Auro-3D, même si l'ensemble des salles de cinéma est loin d'être équipé par de tels systèmes de diffusion.

Fort de cette expérience au cinéma, Dolby a développé une version Atmos At Home, pour les systèmes de diffusion home cinema. Les systèmes Atmos At Home disposent de plus d'enceintes qu'un système 5.1 classique en rajoutant des haut-parleurs au plafond. Afin de signifier le nombre d'enceintes au plafond, Dolby a rajouté un deuxième point dans la description du système de diffusion. Dolby propose ainsi des systèmes 5.1.2, 5.1.4, 7.1.2, 7.1.4 et même 9.1.2 pour les plus fortunés.¹⁷¹

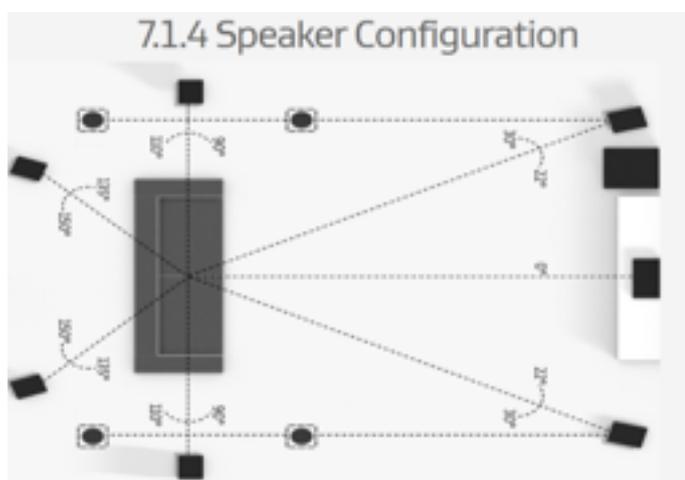


Figure 27. Exemple de configuration 7.1.4 de Dolby At Home

¹⁷¹ - <http://www.dolby.com/us/en/guide/speaker-setup/index.html>

De nombreux types de contenus sont compatibles avec cette technologie, Dolby ayant soigneusement choisi ses partenaires (Google Play Movies, Amazon fire TV, Netflix et Sony)¹⁷².

Afin d'étendre son marché, Dolby a développé une version mobile de sa technologie : Dolby on the go. Possédant les mêmes partenaires que pour sa version domestique, Dolby propose ainsi d'expérimenter l'Atmos sur smartphone et tablette possédant la technologie, grâce à la technologie binaurale.¹⁷³

Dolby semble donc prendre en charge une synthèse binaurale sur le terminal de diffusion ainsi qu'une gestion des objets. La gestion de l'ambisonique n'a pas été évoquée dans la communication autour de ces technologies. Elles ne semblent donc pas être capables de gérer un flux ambisonique.

Par rapport à mon projet, il sera donc légèrement plus compliqué mais possible d'adapter mon contenu au Dolby Atmos. Je devrais enregistrer le flux ambisonique sur un ensemble de canaux traditionnels (5.1, 7.1 ou 9.1).

Il est possible de voir clairement la différence de philosophie entre le MPEG-H et le Dolby on the go. L'un est ouvert et prend son cahier des charges de chaînes de télévision traditionnelles alors que l'autre propose des produits avec une technologie brevetée pensée pour des jeux vidéo comme pour les nouveaux dispositifs de diffusion télévisuelle, avec l'appui de poids lourds du secteur.

Néanmoins, ces propositions sont toutes les deux émergentes et il me paraît impossible de savoir à l'avance laquelle va s'imposer à l'avenir.

Le concept d'*aventure sonore dont vous êtes le héros*, est donc encore perfectible que ce soit au niveau du contenu que de l'interaction. Néanmoins, on constate, à travers ces normes en développement, qu'il y a de nouvelles tendances qui sont favorables à la diffusion de ce genre de contenu.

¹⁷² <http://www.dolby.com/us/en/home/index.html?onInk=pl-sc>

¹⁷³ <http://www.dolby.com/in/en/technologies/dolby-atmos/mobile.html>

Conclusion

Ce mémoire avait pour but de penser et de mettre en oeuvre une adaptation sonore de Menu Cthulhu, un livre dont vous êtes le héros en binaural avec suivi des mouvements de la tête, qui s'est concrétisé avec un prototype présenté à une trentaine de personnes.

Les réflexions développées sur ce genre littéraire peuvent permettre d'avoir une première base de travail pour l'adaptation d'un autre livre-jeu. Cela m'a permis d'apercevoir plus généralement, le travail préliminaire nécessaire à une démarche *cross-média*.

L'état de l'art technologique réalisé au chapitre III donne une première idée de l'ensemble des outils disponibles aujourd'hui pour réaliser un contenu en binaural prenant en compte les mouvements de la tête.

L'ensemble des exemples sonores livrés avec ce mémoire permet, je l'espère, de se faire une idée plus concrète des avantages et des inconvénients inhérents aux différentes technologies.

Au cours de ma recherche, j'ai pris conscience qu'il ne fallait pas se limiter à une seule technologie et qu'il était même préférable d'avoir une utilisation combinée de l'ensemble des technologies, en fonction de la nature des sources sonores.

J'ai ainsi pu noter que l'ambisonique binauralisé était particulièrement adapté pour des ambiances et des sources très diffuses alors des sources plus ponctuelles ont besoin d'être synthétisées directement en binaural. De plus le binaural natif, bien que ne permettant pas de prendre en compte les mouvements de la tête, a permis d'intégrer des éléments en ultra-proximité dans ce contenu.

Malgré les limites de la technologie binaurale (problème d'individualisation des HRTFs, confusion avant/arrière ou en élévation), la très grande majorité des personnes ayant pris en main le prototype réalisé dans le cadre de la partie pratique de ce mémoire, se sont senties immergées dans l'aventure proposée.

Le binaural, comme les autres systèmes de reproduction sonore possède donc ses avantages et ses inconvénients. De plus, il ne se suffit pas à lui même. Il reste un outil à mettre au service d'un propos et d'une narration.

Ce mémoire a d'ailleurs été l'opportunité d'ébaucher une esthétique sonore à 360°, inspirée par l'esprit de l'oeuvre originale, et il présente des pistes de réflexion qui peuvent être reprises pour n'importe quel contenu dit « *immersif* ».

Aujourd'hui, des dispositifs immersifs, et particulièrement les lunettes de réalité virtuelle, se développent sur le marché. Des poids lourds d'internet, tels Youtube ou Facebook, commencent à proposer des lecteurs de vidéo à 360°, où le spectateur a la possibilité de tourner la tête.

L'émergence de ces contenus promet de nouvelles perspectives pour la technologie binaurale, et particulièrement avec un dispositif de suivi des mouvements de la tête.

Bibliographie

Ouvrages

Relatif à l'ingénierie ou l'esthétique sonore

BLAUERT Jens, *Spatial Hearing*, MIT Press, 1997

DESHAYS Daniel, *Pour une écriture du son*, Paris, éditions Klincksieck, coll. 50 questions, 2008

HUGONNET & WALDER, *Théorie et pratique de la prise de son stéréophonique*, Paris, Eyrolles, 1995

MALHAM Dave, *Spatial Heading Mechanisms and Sound Reproduction*, University of York, England, 1998

MILLOT Laurent, *Traitement du signal audiovisuel : Applications avec Pure Data*, Paris, Dunod, 2008

PARENTHOEN Yann, *Propos d'un tailleur de son*. Editions phonurgia nova, 2002

PERIAUX Bergame, OHL Jean-luc, THEVENOT Patrick, *Le son multicanal*, Paris, Edition Dunod, 2015

Littéraire

BRENNAN J.H., *La Horde des démons*, Paris, Folio Junior, 1984

JACKSON Steve, *Les Démons des profondeurs*, Paris, Gallimard, coll. « Défis fantastiques », 1986

JOMUNSI Neil, *Menu Cthulhu*, Paris, Editions numériques Walrus, 2013

OULIPO, *La littérature potentielle*, Paris, Gallimard, coll. Folio Essais, 1973

ROSENTHAL Pierre, *Le Prisonnier*, Paris, Le Livre de Poche, coll. « Histoires à jouer: Missions spéciales », 1988

SAGOT Gildas, *Jeux de rôle : tout savoir sur les jeux de rôle et les livres dont vous êtes le héros*, Paris, Gallimard, 1986

Immersion et transmédia

CAILLER, DENIS, SAPIEGA (Sous la dir. de), *Histoire du Transmédia*, Paris, L'Harmattan, coll. Champs Visuels, 2014

JENKINS Henry, *La culture de la convergence*, Paris, Armand-Colin, 2013

ROSE Frank, *The Art of Immersion. How the Digital Generation is Remaking Hollywood, Madison Avenue, and the Way We Tell Stories*, New York, Norton and company, 2011

Relatif au théorie du jeu

EDWARD Ron, « The right to Dream » in *The Forge*, 2003

EDWARD Ron, « Narrativism: Story Now » in *The Forge*, 2004

EDWARD Ron, « The whole model - this is it » in *The Forge*, 2003

Articles, revues et conférences

Scientifique

BEGAULT, WENZEL, ANDERSON, « Direct comparison of the impact of head tracking, reverberation, and individualized head-related transfer functions on the spatial perception of a virtual speech source », *JAES*, 2001

CHENG, COREY, WAKEFIELD, GREGORY, « Introduction to Head-Related Transfer Functions (HRTFs): Representations of HRTFs in Time, Frequency, and Space », *JAES*, 2001

GERZON Michael, « Ambisonics in multichannel broadcast and video », *JAES*, 1984

MENZIES, AL-AKAIDI, « Nearfield binaural synthesis and ambisonics », *JAES*, 2007

NOISTERNIG, MUSIL, SONTACCHI, HOLDRICH, « 3D binaural sound reproduction using a virtual ambisonic approach. » In *Virtual Environments, Human- Computer Interfaces and Measurement Systems*, Symposium international, 2003

NOISTERNIG Markus [sous la dir.] , « A 3D Ambisonic Based Binaural Sound Reproduction System » , *AES 24th International Conference on Multichannel Audio*

TISSEAU Jacques, *Réalité virtuelle et complexité*, Manifeste scientifique du CERV, 2004

ZENOUDA Hervé, « Son, interaction, immersion et effets de présence » in *Journées Neptune : Environnements immersifs*, Colloque à Toulon, 17 avril 2012

Littéraire

AARSETH Espen, « The joy of reading role-playing games », *The Guardian*, 19 juin 2015

BOUCHARDON Serge, « Le récit littéraire interactif : une valeur heuristique », *C&L*, Janvier 2008

BOUYALA-FROSIN Bernadette, « Les livres dont vous êtes les héros », *Actes de Lecture* n°21, Mars 1988

CLEMENT Jean, « La cyberlittérature entre jeu littéraire et jeu vidéo », *Formules n°10*, Paris, Agnès Vienot, juin 2006

MENELDUR, « Interview de Paul Mason » in *La Bibliothèque des aventuriers*, 2009

SOUCHIER Emmanuel, « Les Livres dont vous êtes le héros », *Trousse-Livres*, n°58

Mémoire ou thèse

Relatif au son

ANGLIONIN Louis, *De nouveaux outils pour le suivi des mouvements de la tête*, Mémoire de fin d'étude de Louis-Lumière, 2013

AUSSAL Matthieu, *Méthodes numériques pour la spatialisation sonore, de la simulation à la synthèse binaurale*, Thèse de l'Ecole Polytechnique, 2014

CHEVRIER Léa, *Expérimentation des techniques binaurales appliquées au documentaire radiophonique*, Mémoire de fin d'étude de Louis-Lumière, 2015

DIENER Elliot, TOULON Arnaud, *Réalisation d'un jeu audio*, Mémoire de fin d'étude de Louis-Lumière, 2015

LLAVE Adrien, *Simulation et modélisation de l'éloignement d'une source*, Mémoire de fin d'étude de Louis-Lumière, 2015

MALLO Eléonore, *Etude pour la création d'une bande dessinée sonorisée*, Mémoire de fin d'étude de Louis-Lumière, 2015

MATAHRI Hanna, *Adaptation d'une pièce de jazz en ambisonique*, Mémoire de fin d'étude de Louis-Lumière, 2014

SALAUN Stéphanie, *Formats Audio 3D Hiérarchiques : Caractérisation objective et perceptive des systèmes Ambisonics d'ordres supérieurs*, Thèse de l'INSA Lyon, 2009

SALMON François, *Monitoring au casque d'un outil de spatialisation sonore*, Rapport de stage de Louis-Lumière, 2015,

Littéraire

BOUCHARDON Serge, *Le récit littéraire interactif. Narrativité et interactivité*, thèse soutenue le 7 décembre 2005 à l'Université de Technologie de Compiègne

CAUVIN Aurélie, *La Littérature Hypertextuelle, analyse et typologie*, Maitrise de Lettres Modernes 2001, Université de Cergy Pontoise

MUNZEL Andreas, *Le livre audio en France et en Allemagne : Motifs, freins et contextes*, Mémoire d'économie, Strasbourg Business School, 2013

Sites internet

Audigame, plate-forme de développement et de diffusion d'aventures interactives

<http://audigame.fr>

Audiopub, site de l'Audio Publishers Association américaine, diffusant les études des ventes de livre audio : http://audiopub.org/press/Sales_Survey_APA_2015_Final.pdf

Projet BiLi, projet de recherche collaboratif sur l'écoute binaural

<http://www.bili-project.org/le-projet/>

Content is ART, web-journal culturel. Propose des analyses sur les livres dont vous êtes le héros : <http://contentisart.com/fort-mcmoney-le-courage-la-chance-et-la-destinee/>

Core Sound, Fabricant de matériel audio :

<http://www.core-sound.com/TetraMic/1.php>

Darkshire, plate forme d'échange sur le jeu de rôle. Propose des analyses sur des théories du jeu : <http://www.darkshire.net/~jhkim/rpg/theory/threefold/>

Dolby, acteur majeur dans la technologie audio. Propose deux nouveaux systèmes de diffusions domestiques *immersifs*:

<http://www.dolby.com/us/en/guide/speaker-setup/index.html>

Embrace Cinema Gear, fabricant de matériel audiovisuel indien

<http://www.embracecinema.com/gear/product-view.php?slug=brahma-in-zoom>

Flux, éditeur du SPAT

http://www.fluxhome.com/products/plug_ins/ircam_spat-v3

Github, plate-forme de diffusion de code source. Héberge l'ensemble des informations nécessaires afin de réaliser le head-tracker utilisé dans ce mémoire ainsi que des objets externes Max.

<https://github.com/ptrbrtz/razor-9dof-ahrs/wiki/Tutorial>

https://github.com/mccartyn/Arduino_Dev/blob/master/libraries/SimpleMessageSystem/MaxMsp%20Example/asciimessage2max.mxb

Harpex, éditeur de plug-in de spatialisation

<http://harpex.net/about.html>

Holophone, fabricant de solutions microphoniques multicanales.

<http://www.holophone.com>

Indie-RPG, site hébergeant le forum *The Forge*, très actif dans la conception des théories du jeu : <http://www.indie-rpgs.com/forge/index.php>

Ircam, centre de recherche musicale et acoustique français

www.ircam.fr/eac.html?tx_ircam_pi4%5BshowUid%5D=1&ext=4

ISO, organisation de normalisation internationale :

http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=62835

Kickstarter, site de crowdfunding :

<https://www.kickstarter.com/projects/1569945514/brahma-affordable-ambisonics-microphone>

Les numériques, site d'actualités sur les nouvelles technologies : lesnumeriques.com

Linux Audio Conference : <http://lac.linuxaudio.org/2013/papers/51.pdf>

Matthias Kronlachner, créateur de la suite Ambix :

<http://www.matthiaskronlachner.com/?p=2015>

MH Acoustics, fabricant du microphone EigenMike : <http://www.mhacoustics.com/products>

MPEG-H Alliance, site présentant les avancées concernant cette norme :

<http://www.mpeghaa.com/presentations/MPEG-H%20Fraunhofer%20HPA%20February%202015.pdf>

Maison des sciences de l'homme. Porte le projet HOA Library

<http://www.mshparisnord.fr/hoalibrary/ambisonie/les-champs-sonores/>

Multiphonie, Blog de Jean-Marc Duchenne, musicien electro-acoustique

<http://multiphonie.free.fr/plugins4.htm>

Noise Makers, éditeur de plug-in de spatialisation binaurale

<http://www.noisemakers.fr/binauralizer/>

Numerama, site d'actualité numérique :

<http://www.numerama.com/pop-culture/127741-jouiez-a-livre-dont-etes-heros-twitter.html>

Page 42, Blog de Neil Jomunsi : <http://page42.org/a-propos/>

Place to Go, People to Be, site d'information sur le jeu de rôle

<http://ptgptb.fr/theorie-des-niveaux-d-immersion-multiples/>

Intel, fabricant de microprocesseur :

<https://software.intel.com/en-us/articles/realsense-overview>

Sparkfun, fabricant de carte électronique : sparkfun.com

Spheric, Blog du monteur son et mixeur de film en réalité virtuelle Gael Nicolas

<http://www.spheric.fr>

Syndicat national de l'édition, discute des enjeux du livre audio :

www.sne.fr/wp-content/uploads/2015/06/Fiche-SNE_livre-audio-juin-2015.pdf

Surround Library, sonothèque et documentation sur les technologies multicanales :

<http://www.surround-library.com/white-papers>

TSL, fabricant des microphones SoundField

<http://www.tslproducts.com/soundfield-type/soundfield-microphones/>

Walrus, éditeur de Menu Cthulhu :

<http://www.walrus-books.com/menu-cthulhu-une-nouvelle-dont-vous-etes-le-heros/>

3D SoundLabs, fabricant du casque Sound One

www.3dsoundlabs.com/?lang=fr#section_4

Discographie

The Barber Shop, démonstration binaurale par QSound

<https://www.youtube.com/watch?v=IUDTlvagjJA>

Thérèse Raquin, le livre audio réalisé en binaural natif par RoadBook

www.roadbookcollections.fr/content/10-3d-son

NouvOson, plateforme de diffusion des créations binaurales et multicanales de Radio France

<http://nouvoson.radiofrance.fr>

Ludographie

A blind Legend, Jeu audio sur smartphone en binaural

<http://www.ablindlegend.com>

Deep Sea, jeu audio créé par Robin Arnott en 2006, centrée sur la privation sensorielle

Table des illustrations

- P.21 : *Figure 1. Les frontières du récit littéraire interactif*
http://www.persee.fr/doc/colan_0336-1500_2008_num_155_1_5376
- P.24 : *Figure 2. Représentation du Threefold Model par Irina Rempt*
<https://groups.google.com/forum/#!topic/rec.games.frp.advocacy/-5KmJkNUNkc%5B1-25%5D>
- P.27 : *Figure 3. Artwork du livre Menu Cthulhu, propriété des éditions Walrus*
<http://www.walrus-books.com/menu-cthulhu-une-nouvelle-dont-vous-etes-le-heros/>
- P.31 : *Figure 4. Immersion, interaction et présence selon Jacques Tisseau*
<http://www.cerv.fr/wp-content/uploads/2012/07/manifesteFR.pdf>
- P.39 : *Figure 5. Affiche d'époque faisant la promotion du Théâtrophone*
<http://www.lesnumeriques.com/casque-audio/audio-binaural-experience-extremement-realiste-a2141.html>
- P.40 : *Figure 6. Schéma de présentation de la tête mécanique Oscar d'AT&T*
<http://www.lesnumeriques.com/casque-audio/audio-binaural-experience-extremement-realiste-a2141.html>
- P.41 : *Figure 7. Schéma représentant la différence interaurale de temps (ITD)*
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.442.5500&rep=rep1&type=pdf>
- P.42 : *Figure 8. Représentation des cônes de confusion relatifs aux ILD et ITD*
http://music.miami.edu/programs/mue/Research/jwest/Chap_2/Chap_2_Spatial_Hearing.html
- P.43 : *Figure 9. Schéma représentant l'acquisition et l'utilisation des HRTFs*
(www.fks.oeaw.ac.at)
- P.48 : *Figure 10. Vision d'ensemble du SPAT, transcodé en plug-in par Flux*
http://www.fluxhome.com/products/plug_ins/ircam_spat-v3
- P.50 : *Figure 11. Plug-in Binauralizer* : <http://www.noisemakers.fr/binauralizer/>
- P.53 : *Figure 12. Représentation des composantes W,X,Y et Z*
<http://www.surround-library.com/white-papers>
- P.55 : *Figure 13. Représentation des harmoniques sphériques des ordres 0 à 3*
<http://www.mshparisnord.fr/hoalibrary/ambisonie/ambisonie-ordre-superieur>

- P.56 : *Figure 14. Principe d'encodage Ambisonique destiné à une restitution binaurale avec head-tracking* : http://iem.at/~noisternig/publications/Noisternig_2003d.pdf
- P.59 : *Figure 15. Plusieurs microphones Eigenmike*
<http://www.mhacoustics.com/products>
- P.61 : *Figure 16. Plug-in AMBI PAN de Noise Makers*
<http://www.noisemakers.fr/product/ambi-pan/>
- P.63 : *Figure 17. Le Plug-in B-format 2 Binaural de Noise Makers*
<http://www.noisemakers.fr/ambi-head/>
- P.65 : *Figure 18. Scène sonore polarisée et scène sonore équi-énergie*
Le son multicanal, Paris, Edition Dunod, 2015
- P.74 : *Figure 19. Illustration du Yaw, Pitch, Roll*
<https://software.intel.com/en-us/articles/realsense-overview>
- P.75 : *Figure 20. Carte électronique Razor IMU (sparkfun.com)*
- P.76 : *Figure 21. Carte FTDI (sparkfun.com)*
- P.76 : *Figure 22. Bluetooth Mate Silver et Batterie LiPo 850mAh/3,7V (sparkfun.com)*
- P.81 : *Figure 23. Classification des sources sonores en fonction de la technologie de reproduction utilisée*
- P.89 : *Figure 24. Workflow de diffusion du prototype*
- P.103 : *Figure 25. Graphique de présentation du MPEG-H 3D Audio de Fraunhofer IIS*
http://www.mpeg-haa.com/presentations/NeuendorfM_NAB2014_MPEGH_Status_v013.pdf
- P.104 : *Figure 26. Présentation du décodeur audio du MPEG-H, un synthétiseur binaural peut donc être rajouté en fin de chaîne de décodage*
http://www.mpeg-haa.com/presentations/NeuendorfM_NAB2014_MPEGH_Status_v013.pdf
- P.105 : *Figure 27. Exemple de configuration 7.1.4 de Dolby At Home*
<http://www.dolby.com/us/en/guide/speaker-setup/index.html>

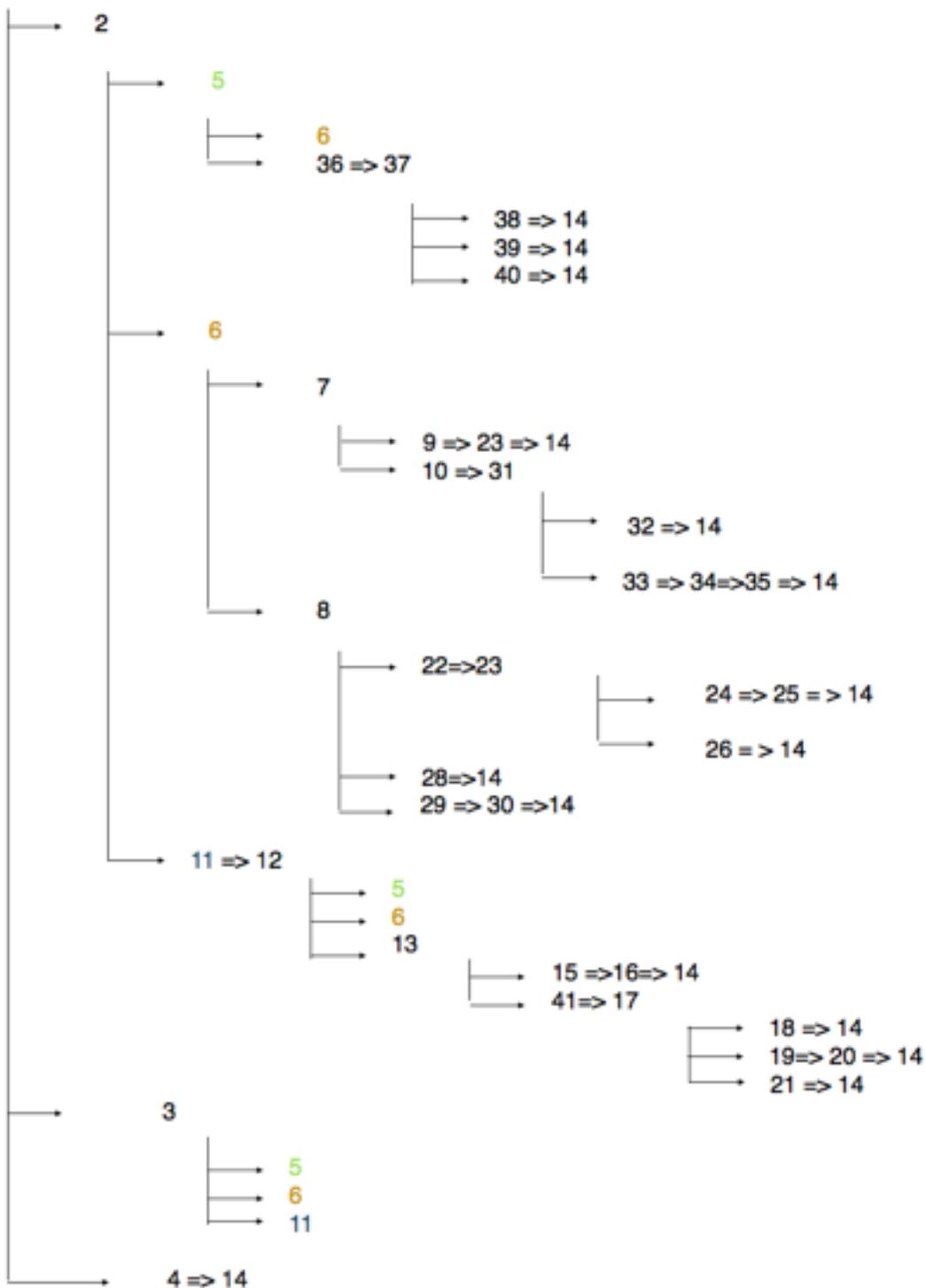
Annexes

Annexe 1 : Organigramme de *Menu Cthulhu*, le livre dont vous êtes le héros de Neil Jomunsi

Organigramme du livre:

(Le chapitre 14 étant le chapitre où le personnage est mort et doit recommencer l'aventure)

1



**Annexe 2 : Interface du jeu audio *DeepSea*.
Photo de Matthew Wegner**

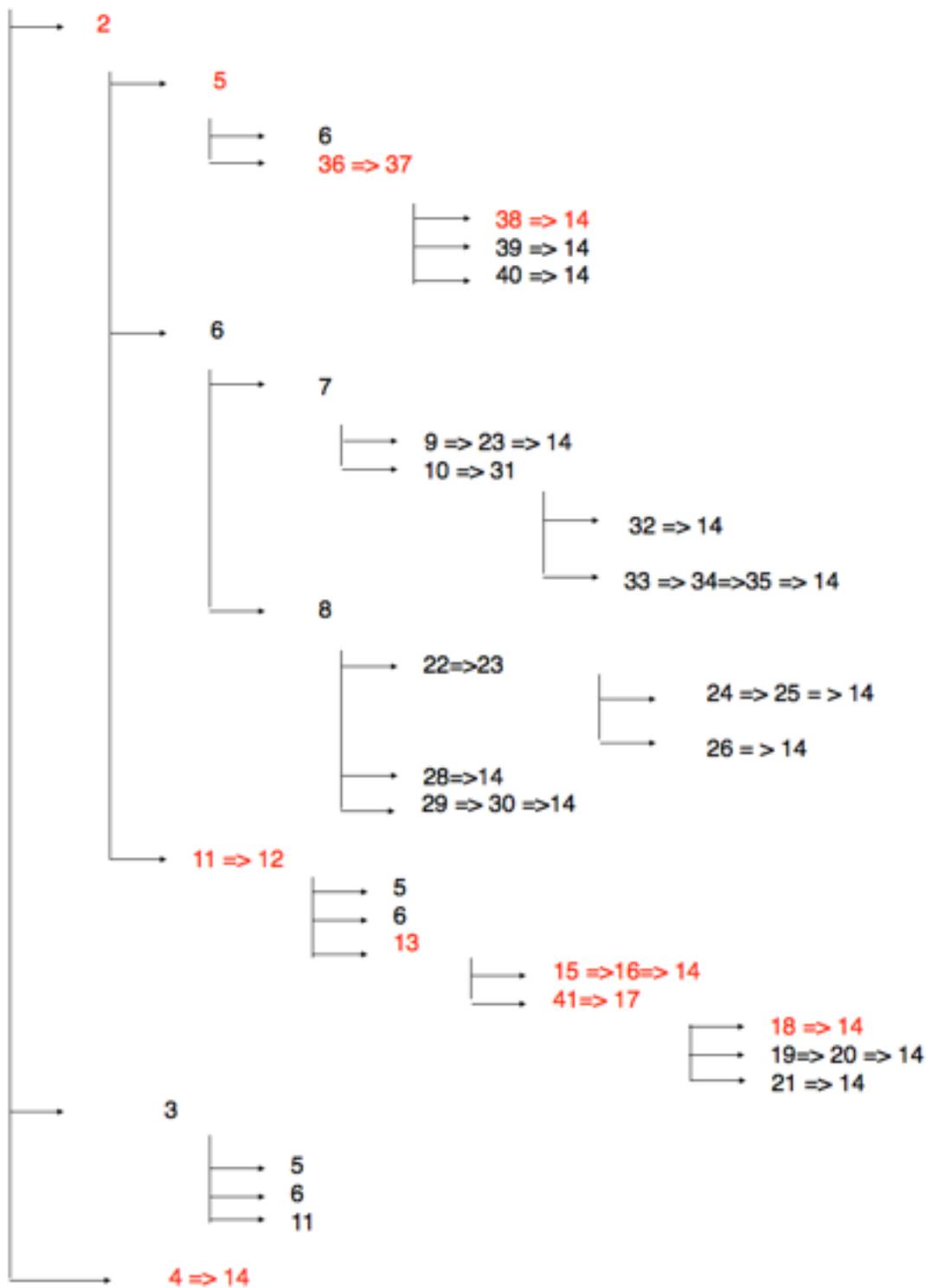


Annexe 3 : Mon organigramme de *Menu Cthulhu*, l'aventure sonore dont vous êtes le héros

Organigramme du livre:

(Les chapitres en rouge sont les chapitres que j'ai adapté au cours de ce mémoire)

1



Annexe 4 : Tableau récapitulatif des différentes techniques de spatialisation sonore

Périaux, Ohi, Thévenot, Le son multicanal, Dunod, 2015

Le tableau suivant présente une synthèse des techniques de spatialisation.

Techniques de spatialisation	Les signaux enregistrés	Manipulation de la scène encodée	Dispositif de restitution
Stéréophonie, « channel oriented » (2.0, 5.1, 7.1, 10.2, Auro-3D 13.1, 22.2, etc.)	généralement 1 signal par enceinte	limitée, scène sonore figée	fixe, lié au format
Ambisonie, « scene oriented » (Soundfield, HOA) ordre m, composantes	3D : $(m + 1)^2$ 2D : $2m + 1$	globale (rotation, distorsion de perspective)	adaptable, indépendant du format
WFS	objets + metadata de localisation	chaque source peut être traitée de façon indépendante	synthèse du front d'onde dispositif adaptable, indépendant du format (objets)
Objets, « object oriented » (Dolby Atmos, MDA, MPEG-H 3D, MPEG 4, MPEG 7...)	objets + metadata de localisation	chaque source peut être traitée de façon indépendante	adaptable, indépendant du format
Binaural	2 signaux L, R	limitée, scène sonore figée	casque stéréo

Annexe 5 : Questionnaire distribuée à l'issue de la présentation publique du prototype

Nom / Prénom :
Professionnel du son // Non professionnel du son

Avis général:

Profil

Avez-vous l'habitude d'écouter du contenu binaural ? OUI NON

Avez-vous l'habitude de jouer à des jeux audio ? OUI NON

Avez-vous l'habitude de jouer aux jeux vidéos (plusieurs fois par semaine) ? OUI NON

Confort de l'aventure

Etiez-vous à l'aise avec ce casque sur la tête ? OUI NON

Avez-vous pris facilement en main l'ensemble des fonctionnalités ? OUI NON

Intérêt

Avez-vous eu la sensation « d'être » le héros ? OUI NON

Le récit vous a-t-il intéressé ? OUI NON

Avez-vous été intéressé(e) par les phases d'exploration ? OUI NON

Ce que vous avez préféré : - Les interactions avec l'iPad - L'exploration avec votre tête

Si vous deviez classer ces points d'intérêt de 1 à 3 (1 étant ce qui vous a semblé le plus intéressant)

- L'histoire
- L'interaction avec le contenu, grâce à l'iPad et au capteur
- La liberté de découvrir un espace sonore à 360 °

Auriez-vous souhaité que cette aventure soit plus longue ? OUI NON

Auriez-vous souhaité d'avantage d'interactions ? OUI NON

Seriez-vous prêt à payer pour ce genre de contenu ? OUI NON

Annexe 6 : Auditrice (*Amélie Hascoet*) de *Menu Cthulhu*

Photographie de Sofia Bendia

