

ESRA Bretagne
Mémoire de fin d'études

Louis LEFOL MUSIL

De la Conception à l'Immersion :
La Carte Postale Sonore Binaurale

ISTS 3 Audiovisuel 2024

<https://louis-lefolmusil.com/memoire/memoire.html>

Sous la direction de

Léa CHEVRIER

Remerciements

J'aimerais tout d'abord remercier Léa CHEVRIER qui m'a accompagné pour la supervision de mon mémoire.

Pour mes 3 années d'études, j'aimerais remercier l'ESRA Bretagne ainsi que tous les professeurs qui m'ont apporté toutes ces connaissances.

Pour la réalisation de mon projet, j'aimerais remercier la ville de Reims, Nicolas CHAUVET pour la visite de la cave de champagne ainsi que le stade Delaune et Mohamed.

Je souhaite aussi remercier mes parents qui m'ont toujours soutenu que ce soit dans ce mémoire mais plus généralement dans toutes mes études.

Merci à Raphaël JANOUEIX pour la relecture et j'aimerais de manière globale, remercier toutes les personnes qui, de près ou de loin, m'ont permis de faire ce projet et ce mémoire.

Sommaire

Remerciements -----	2
Sommaire -----	3
Introduction-----	5
1) Historique et Fonctionnement -----	6
1. Le Binaural-----	6
1. Théorisation et création -----	6
2. Fonctionnement -----	7
2. La carte postale sonore -----	8
2) Création de programmes binauraux -----	10
1. Aspect technique -----	10
1. Prises binaurales natives -----	10
2. Synthèse binaurale -----	12
2. Aspect esthétique-----	14
3. Création personnelle-----	15
1. Sujet-----	15
2. Production-----	15
3. Réalisation -----	16
4. Difficultés rencontrées -----	18
3) Étude de cas-----	19
1. Création personnelle – Carte Postale de la ville de Reims -----	21
2. Espaces Sonores – Kuthodaw Paya Mandalay-----	23
3. La Grande-Motte - Carte Postale Sonore (Été)-----	26
4. ARTE Radio - <i>Mingalabar</i> -----	28
Conclusion -----	31
Glossaire -----	32

Webographie	33
Discographie	33
Bibliographie	33

Introduction

J'ai toujours aimé les cartes postales. C'est un moyen facile de partager un souvenir de vacances. Mais une image ne suffit pas toujours pour revivre un moment. Le son, au même titre que la vidéo, peut plus facilement nous replonger dans un lieu.

Le binaural quant à lui, permet, au moyen d'un simple casque, d'obtenir une immersion totale. Ce principe me fascine autant qu'il m'interroge. Il permet d'offrir une qualité d'espace qui rivalise avec des systèmes multicanaux coûteux.

La carte postale sonore, et plus précisément la carte postale sonore binaurale, répond à cette question en permettant une expérience d'écoute immersive pour l'auditeur. Ce procédé permet non seulement d'immerger l'auditeur dans un paysage sonore, mais aussi d'offrir une nouvelle manière de voyager sans quitter son domicile. Il s'agit également d'un programme facile d'accès pour les personnes malvoyantes ou immobilisées qui leur permet d'explorer le monde par le biais de l'ouïe.

Ce mémoire explore la conception et la réalisation de cartes postales sonores binaurales. Dans un premier temps, il s'attardera sur les fondements théoriques et techniques du son binaural ainsi que sur l'évolution de la carte postale sonore. Ensuite, il analysera les méthodes de création d'un contenu binaural en abordant les aspects techniques, esthétiques et pratiques, avec la création d'une carte postale sonore de la ville de Reims. Enfin, l'étude de cas de ma production, ainsi que de plusieurs cartes postales sonores existantes viendra enrichir la réflexion en proposant une analyse critique et comparative.

Ainsi, à travers ce travail, nous chercherons à répondre à la question suivante :

Comment se construit une carte postale sonore binaurale ?

1) Historique et Fonctionnement

1. Le Binaural

1. *Théorisation et création*

La première apparition du concept de son « binaural » remonte à 1839 par l'allemand Heinrich Wilhelm Dove. Il a mis en évidence que lorsque deux sons, avec une fréquence légèrement différente, sont diffusés séparément dans chaque oreille, le cerveau perçoit alors un battement dont la fréquence est égale à la différence entre les deux¹.

Bien que cela soit une expérimentation sur le son « binaural », Dove a utilisé ce phénomène afin d'enrichir la musique, afin d'obtenir des battements supplémentaires. C'est pour cela que le terme de « battements binauraux » est resté pour parler de ce genre de son binaural.

En 1881, l'ingénieur français Clément Ader expérimente le son « binaural » grâce au Théâtrophone (fig.1), un appareil qui permettait d'écouter les représentations de l'opéra de Paris à distance².

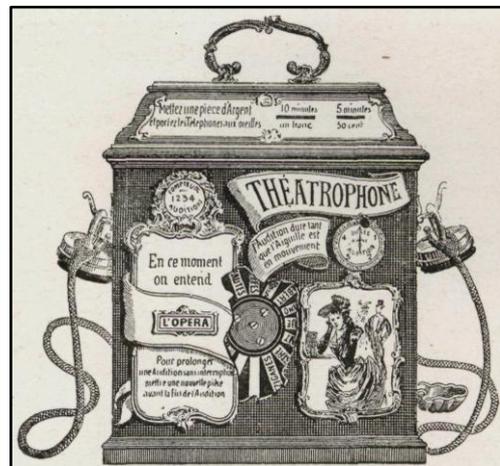


Figure 1 : Théâtrophone

Ce n'est qu'à partir de 1933 que l'entreprise AT&T (*American Telephone and Telegraph Company*) présente au « *Chicago World Fair* » la première tête binaurale « Oscar ». Ce sera finalement en 1972 que Neumann sortira sa tête KU-80, qui deviendra par la suite la KU-81 et la KU-100 encore utilisées aujourd'hui³.

Aujourd'hui, le « battement binaural » ne sert que dans la méditation et le « binaural » de Ader ne signifiait que du son qui n'était pas monophonique (binaural → « deux oreilles », deux canaux).

Seule la définition du son binaural en tant que technique de spatialisation sonore n'est restée et ce mémoire a pour but de traiter de cela.

¹ 1839 – *Expérimentation des battements binauraux*, Heinrich Wilhelm Dove.

² France Musique – *Le théâtrophone, le streaming classique avant l'heure !*

³ SoundGirls – *Recording and Binaural Audio*

2. Fonctionnement

Le binaural se base sur trois principes de différences interaurales¹. Il y a d'abord la Différence Interaurale de Temps ou ITD (*Interaural Time Difference*). L'ITD est le délai que met un son à arriver à une oreille, puis à l'autre. Le délai peut durer de 10 μ s à 600 μ s (0.6ms) selon l'azimut (fig.2) de la source sonore par rapport aux oreilles. Les fréquences les plus utiles sont principalement inférieures à 1.5 kHz car au-delà, la longueur d'onde devient plus petite que la distance interaurale ($\frac{34000}{1500} \approx 22cm$).

Ensuite, l'IPD (ou *Interaural Phase Difference*) est la différence de phase entre les deux oreilles. L'IPD est liée à l'ITD.

Finalement, l'ILD (ou *Interaural Level Difference*) est la différence d'intensité du signal sonore entre les deux oreilles. Cette différence est efficace à partir de 1.5-2 kHz et il peut y avoir une différence maximum de 20 dB SPL pour des fréquences supérieures à 5 kHz en azimut de $\pm 90^\circ$.

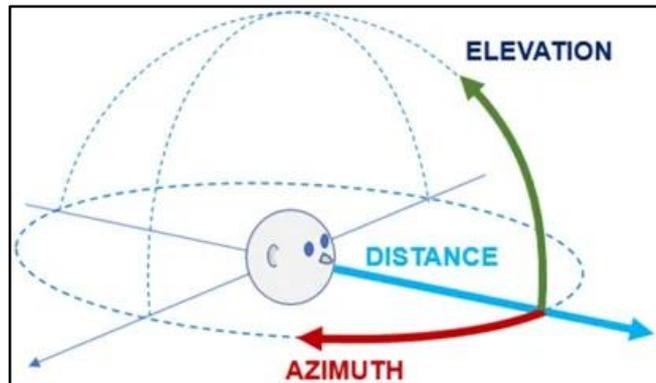


Figure 2 : Nom des axes

En plus de ces trois principes physiques, il existe des paramètres physiologiques. La forme de la tête, des oreilles, des épaules sont des éléments qui changent la réflexion des ondes sonores à l'entrée des oreilles. Cela induit donc des changements d'intensité, de fréquences dans le spectre qui aident à la localisation des sons. De plus, les micromouvements de la tête permettent de plus facilement détecter l'origine d'un son.

Tous ces paramètres, à l'exception des mouvements de tête, se retrouvent dans une fonction de transfert relative à la tête ou HRTF (*Head Related Transfer Function*). Ces HRTF sont à la fois utiles pour la captation du son, mais aussi pour sa restitution ; elles permettent d'adapter la perception du binaural pour chaque personne, mais cela permet également de pouvoir écouter des programmes multicanaux au casque grâce au *downmix*.

¹ Frontiersin – Auditory localization: a comprehensive practical review

Il existe 3 méthodes afin d'acquérir ces HRTF. La première consiste à utiliser une tête binaurale placée dans une chambre anéchoïque et d'émettre une série plus ou moins importantes d'impulsions sonores tout autour de la tête. La seconde méthode est similaire à la précédente mais consiste à placer un humain avec des microphones intra-auriculaires à la place de la tête. Cette approche est toutefois moins facile à standardiser étant donné que chaque personne a un HRTF naturel différent. Enfin, la troisième méthode consiste à faire des modélisations 3D basées sur des mesures anatomiques ou des scans. Certaines entreprises comme Embody, Sony ou encore Apple permettent donc de créer des HRTF « sur mesure » avec une simple vidéo d'une personne.

Bien que les HRTF soient assez efficaces, il reste tout de même certaines zones d'incertitude de localisation du son¹. Sur l'axe azimut, l'incertitude est d'environ 7 ° en frontal, 11 ° en arrière et 20 ° en latéral. De plus, l'élévation entre 20 et 40 ° est elle aussi difficile à localiser.

2. La carte postale sonore

Les premières traces d'une carte postale sonore remontent au début du XXe siècle. Le Phonopostal² est donc breveté en 1905 par Brocherioux, Marotte et Tochon. Le principe était d'inscrire un message sonore sur la carte postale appelée « sonorine ». Bien que l'idée ait été novatrice pour l'époque, le besoin d'avoir un appareil spécifique pour lire la carte a contribué à l'échec commercial du Phonopostal.

C'est à partir des années 1920 que certaines cartes postales intégrant des disques vinyles miniatures furent produites. Les disques permettaient de pouvoir écouter de la musique ou des messages vocaux enregistrés directement dessus. Ce type de carte postale a été popularisé dans les années 1950 par Phonoscope.

¹ Bergame Périaux – Jean-Luc Ohl – Patrick Thévenot - *Le son multicanal - De la production à la diffusion du son 5.1, 3D et binaural.* (p.125)

² Musée de la Poste – *Phonopostal - Phonographe*

Aujourd'hui, grâce au streaming, la carte postale a su devenir un format narratif immersif. France Culture ou ARTE Radio proposent des cartes postales sonores permettant de voyager le temps du programme.

Il est difficile d'estimer la part de marché que représentent les cartes postales sonores car aucune donnée n'existe. Cependant, leur popularité croissante est évidente à la vue de la place qui leur est donnée sur les plateformes de streaming. Ce format plaît notamment pour son immersion, ce qui induit un engagement auprès des auditeurs.

2) Création de programmes binauraux

1. Aspect technique

1. *Prises binaurales natives*

Afin de réaliser des prises natives en binaural, il existe plusieurs façons de procéder selon le temps et le budget à disposition.

Si le temps et le budget le permettent, il est intéressant d'utiliser une tête binaurale commerciale. Ce dispositif permet d'obtenir les meilleurs résultats binauraux et il s'agit du système de référence en termes d'enregistrement professionnel. Ces têtes, comme par exemple la Neumann KU-100 (fig.3), sont basées sur une moyenne mondiale de forme de têtes et d'oreilles humaines. Elles permettent donc une meilleure standardisation du signal pour chaque individu. En plus de la tête, il est aussi possible de rajouter un torse afin de prendre en compte l'effet « *torso-related* » qui comprend notamment les diffractions du son induites par le cou et les épaules.



Figure 3: Neumann - KU-100

Les microphones sont omnidirectionnels et placés au fond du conduit auditif artificiel, à l'endroit où se trouve le tympan humain.

Ces têtes sont utilisées dans des contextes variés pouvant aller de l'enregistrement de concert classique à la captation de paysages sonores pour de la fiction, en passant par la recherche en psychoacoustique et la réalité virtuelle.

Ces systèmes ont des avantages notables comme le réalisme de l'enregistrement notamment dans la perception de la hauteur et de la distance mais également le fait que l'enregistrement est utilisable directement en binaural. Cela évite le traitement qui peut être long et moins immersif lié à la synthèse.

Cependant, comme tout système, il n'est pas universel mais de plus, cela reste du matériel assez coûteux (de l'ordre de plusieurs milliers d'euros juste pour la tête) et assez peu discret pour des captations en public.

C'est donc pour cela que depuis les années 2000, la fabrication artisanale de têtes binaurales s'est démocratisée. Cela permet de réduire le coût et d'avoir une tête pour quelques centaines d'euros. Ces systèmes sont souvent développés à des fins pédagogiques, artistiques ou de recherches expérimentales, mais ils peuvent aussi être utilisés dans le cadre d'enregistrements professionnels au budget limité.

La structure crânienne doit imiter la forme et la densité d'un crâne humain. Il est donc possible d'utiliser du polystyrène, du bois ou de la résine.

Le pavillon auriculaire est la partie la plus sensible car c'est à cet endroit que le son sera principalement capté. Le plus souvent, les pavillons seront réalisés en silicone.

Enfin, concernant les microphones, il est possible d'utiliser des Panasonic WM-61A afin de réduire les coûts tout en gardant un son avec peu de bruit. Ils peuvent être remplacés par des DPA 4060 afin d'avoir une bien meilleure qualité sonore, tout en gardant le prix de la tête proche du millier d'euros.

Ce système est intéressant d'un point de vue budgétaire, mais il permet également de pouvoir expérimenter de façon plus facile l'acoustique liée à la densité et au pavillon des individus. Ces têtes dites DIY (*Do It Yourself* – Faites le vous-même) sont généralement utilisées sur le terrain car elles sont plus légères que des têtes commerciales. Elles sont aussi beaucoup utilisées dans la musique expérimentale et dans l'ASMR ou le podcast, notamment pour leur prix.

Il existe une dernière méthode pour réaliser des enregistrements binauraux. Cela consiste à placer des microphones directement dans les pavillons auditifs d'individus (fig.4). Ce système permet un enregistrement à la première personne plus intuitif car l'individu devient en quelque sorte le microphone. De plus, elle lui permet une restitution optimale au casque car il s'agit de son HRTF.



Figure 4: Earsight binaural

La légèreté et la discrétion du système se prêtent parfaitement à des usages mobiles en situations réelles ainsi qu'aux mouvements de la tête. Cependant, les microphones intra-

auriculaires sont très sensibles aux bruits du corps et de vêtements, et la restitution au casque peut poser problème pour certaines personnes car le HRTF peut ne pas leur convenir.

2. Synthèse binaurale

La synthèse binaurale désigne l'ensemble des techniques numériques qui permettent de simuler une écoute immersive au casque. Il permet de transformer des sons monophoniques ou multicanaux en sons binauraux. Afin de stocker toutes ces données, il est possible d'utiliser le format SOFA (*Spatially Oriented Format for Acoustics*). Il s'agit d'un standard ouvert développé par l'AES (*Audio Engineering Society*) et normalisé sous la référence AES69-2015. Il est principalement utilisé dans le mixage d'objets des programmes.

Le SOFA permet de stocker plusieurs catégories de données. Il y a d'abord les mesures HRTF d'un individu ou d'une tête, les positions des sources et des capteurs, c'est-à-dire leur azimut, élévation et distance, et enfin les métadonnées liées au fichier telles que l'identité du sujet, sa morphologie, le protocole de mesure utilisé, etc... Enfin, le fichier stocke le système de coordonnées utilisé (SOFA/IRCAM, SOFA/CIPIC). Cette liste de données stockées n'est pas exhaustive et c'est pour cela qu'il n'est pas rare d'avoir des fichiers SOFA faisant plusieurs centaines de Mégaoctets selon la qualité et la précision des réponses impulsionnelles réalisées.

Il existe de nombreuses bases de données d'HRTF au format SOFA avec chacune leur particularités¹. Il y a par exemple la base de données « CIPIC », réalisée par UC Davis, qui compte 45 sujets aux morphologies diverses. La « Listen » de l'IRCAM compte 51 sujets avec des mesures réalisées en chambre anéchoïques ou encore la « FABIAN » de TU Berlin qui ne compte qu'un sujet, mais dont la résolution spatiale est fine. L'utilisation de certaines bases de données dépend donc du résultat recherché.

Le principal avantage à l'utilisation du SOFA est son interopérabilité. Il est possible d'utiliser le même fichier dans des moteurs audio, des DAW (*Digital Audio Workstation*) ou des SDK (*Software Development Kit*) de réalité virtuelle. De plus, le stockage de données est extensible ce qui signifie qu'il est possible de faire évoluer un SOFA, tout en gardant les

¹ Sofaconventions.org

données précédentes, ce qui permet par extension une personnalisation plus facile pour l'utilisateur.

Toutes les données des fichiers SOFA peuvent par la suite être utilisées dans des plugins binauraux et notamment les réverbérations. Bien que ce système offre un réalisme maximal, la faiblesse de cette technique est la limite de l'espace. Le résultat est cohérent spatialement avec la réponse impulsionnelle mais dépend du point de mesure. Dans le cadre d'une fiction ou d'une carte postale sonore binaurale, ce n'est pas une grande contrainte, mais quand il s'agit du jeu vidéo, cela devient tout de suite plus compliqué.

C'est donc pour cela qu'il est possible de créer des réverbérations binaurales algorithmiques. Elles ont l'avantage d'être extrêmement flexibles car chaque paramètre peut être changé à la volée, mais elles peuvent manquer de réalisme et de localisation si le moteur est mal calibré.

Il existe de nombreux plugins et moteurs audio prenant en charge le binaural de synthèse comme le Spat Revolution de l'IRCAM, la HOARverb de Flux ou encore la DearVR Pro de Dear Reality.



Figure 5: DearVR Pro

Tous ces plugins sont basés sur les mêmes techniques de spatialisation. Il y a la décorrélation spatiale qui ajoute de légères différences interaurales dans le champ diffus ; la compensation de rotation de la tête (YAW) qui fait évoluer les réflexions en fonction de la rotation de la tête (uniquement dans les moteurs audios dynamique) ou encore le filtrage basé

sur la distance qui applique une coloration spectrale à la réverbération selon la distance, généralement en appliquant une perte dans les aigus.

Ces plugins peuvent travailler sur des canaux discrets ou sur des objets. Chaque méthode a ses avantages et inconvénients. Pour les canaux discrets, la puissance de calcul pour la synthèse ainsi que pour la restitution du signal sera moins importante que pour les objets. Cependant, travailler avec des objets permet de s'adapter à n'importe quels systèmes d'écoute car les sources sonores ne sont plus liées aux canaux mais à l'espace. Ils permettent également de pouvoir faire du binaural dynamique, c'est-à-dire que le son bouge autour de l'auditeur en même temps qu'il tourne la tête.

Dans le cadre d'un programme binaural non dynamique, qui comporte donc deux canaux fixes, le choix des canaux discrets semble ainsi évident.

La synthèse et la prise native ont chacune leurs avantages et inconvénients mais dans le cadre de la fidélité et de l'immersion, la prise native sera en règle générale plus intéressante que la synthèse. La chercheuse Elizabeth M. Wenzel a notamment démontré dans une étude de 1993 que seulement 25% des sujets localisent correctement les sons en synthèse binaurale¹. L'intérêt de la synthèse est donc principalement lié au jeu vidéo, là où la distance importe réellement.

2. Aspect esthétique

Concernant l'esthétisme d'un programme binaural, il est assez difficile de pouvoir traiter le signal en postproduction. Par exemple, l'utilisation d'un égalisateur altère le spectre fréquentiel, ce qui détruit le filtrage en peigne naturel lié à la spatialisation. Les prises binaurales deviennent donc stéréophoniques.

Bien que le choix des microphones puisse avoir un lien avec le budget et le temps accordé, il joue un rôle plus classique de couleur sonore. Le choix des réverbérations, des bases de données utilisées joue sur la définition du signal et donc sur l'émotion recherchée. Il serait peut-être intéressant d'utiliser un SOFA moins défini dans le cadre d'une fiction sonore horrifique afin de perdre plus facilement l'auditeur, afin de le plonger au cœur de la détresse des personnages.

¹ Wenzel (1993) - *Localization using nonindividualized head-related transfer functions.*

Tout cela fait partie des choix liés à une production, mais il faut également faire attention à ne pas vouloir non plus « abuser » de l'utilisation du binaural. Il se peut qu'un programme stéréophonique soit plus en lien avec le sujet traité que du binaural. Si l'immersion ne s'y prête pas, il est important de ne pas forcément imposer le binaural. Ce dernier étant encore nouveau pour les auditeurs amateurs, imposer le binaural là où il n'a pas de réel intérêt pourrait les repousser à écouter des programmes binauraux complexes, ne voyant pas la différence avec un programme stéréophonique habituel.

3. Création personnelle

1. *Sujet*

Cherchant, au travers de ce mémoire, à comprendre comment se créer une carte postale sonore binaurale, il me paraissait assez évident de tenter de créer moi-même un tel programme afin de plus facilement comprendre la difficulté et les enjeux liés au binaural.

Étant originaire de Reims, le sujet que j'ai choisi m'est apparu comme une évidence : réaliser une carte postale sonore de la ville de Reims. Pour ce faire, je me suis rendu à la Cathédrale de Reims, au Stade Delaune et enfin dans une cave de champagne. Ces lieux représentent selon moi des aspects à la fois touristiques, artistiques et culturels de la ville.

2. *Production*

Avant toute chose, il était important pour moi de déterminer l'importance du binaural dans cette production. Afin de retranscrire au mieux les lieux que j'ai choisi, il était nécessaire d'avoir un format immersif pour pouvoir profiter au maximum de la grandeur des espaces, notamment la cathédrale et le stade Delaune.

Lors de la phase de préproduction, j'ai donc commencé par réaliser un premier plan de ce que je souhaitais incorporer à ma carte postale. Pour la cathédrale, j'ai donc imaginé l'entrée dans celle-ci, le bruit des cloches et des chants grégoriens ainsi que les bruits des pas de personnes se déplaçant à l'intérieur. Ensuite, changement de lieu, direction une cave de champagne avec le bruit des pas en lien, ainsi que le bruit des bouteilles manipulées par le

caviste. Finalement, après être sorti de la cave, direction le stade Delaune lors d'un match afin de vivre au plus près l'ambiance.

Concernant la liste du matériel utilisée, j'ai opté pour un Zoom F6 et une paire de Earsight Binaural (fig.6). Ce choix se base sur 2 critères : le premier économique et le deuxième praticité / discrétion. Cela m'évite de devoir investir ou louer du matériel coûteux, et cela me



Figure 6: Matériel de tournage

permet également de pouvoir réaliser des mouvements de déplacement plus facilement. Le principal inconvénient de ce système est le *monitoring*. Il était impossible de pouvoir enregistrer et écouter en même temps, ce qui m'a demandé plusieurs prises test afin de comprendre comment fonctionne et sonne le matériel simplement en regardant les vumètres. Le second inconvénient est évidemment le choix du HRTF. Le risque de ce genre de pratique est le manque de standardisation pour le plus grand nombre d'auditeurs. Enfin, afin de pallier au problème de saturation, toutes mes prises ont été réalisées en 32 bits flottants – 48 kHz.

Enfin, vis-à-vis des autorisations et lieux de tournage, certaines contraintes se sont présentées à moi notamment pour le stade Delaune. Il a donc fallu pour cela que je contacte le responsable sécurité afin de lui parler de mon projet, ce à quoi il m'a autorisé l'accès.

3. Réalisation

Après avoir réalisé tous mes enregistrements, j'ai choisi de partir sur REAPER pour le montage et le mixage. Je l'ai choisi car il est léger, modulable et suffisant pour ce que demande le binaural en tant que tel.

J'ai donc d'abord commencé par importer mes pistes, puis je les ai rangées par catégories (fig.7). J'ai ensuite tout réécouté, puis trié ce qui me paraissait du plus au moins exploitable. J'ai également utilisé les marqueurs d'objets afin d'annoter les événements importants de mes enregistrements comme des bruits de cloches, de bouteilles ou encore de buts marqués.

Après cela, j'ai réalisé une phase de sélection et de montage de mes pistes. J'ai d'abord pris des ambiances du parvis de la cathédrale ainsi que des cloches et de l'ambiance intérieure. À cela j'ai rajouté de l'orgue, des bruits de portes et de pas.



Figure 7: Rushes tournages

La première partie se déroule donc dans la cathédrale selon le chemin suivant : Tout d'abord, l'auditeur arrive sur le parvis de la cathédrale puis entre par la façade occidentale. Il se retrouve alors dans la nef. De l'orgue joue plus loin, au niveau du transept et des gens marchent dans les bas-côtés ainsi que dans le déambulatoire.

Après cela, un bruit de porte intervient, ce qui nous envoie dans la charpente de la cathédrale. C'est un endroit bien plus calme bien que la structure ne soit pas en bois mais en béton.

Afin de passer à ma deuxième partie, j'ai utilisé des sons de cloches extérieurs et des pas dans un escalier afin de faire passer l'auditeur aux caves de champagne.

J'ai choisi de mettre différentes sonorités de bouteilles. Les plus aiguës sont des bouteilles classiques (0.75L) alors que les plus graves sont des magnums (1.5L).

Enfin, la dernière partie tourne autour du stade Delaune. J'ai donc récupéré l'ambiance du stade, ainsi qu'un moment où un but a été marqué.

Concernant les traitements, à part une légère égalisation sur certains éléments, il n'y a rien d'autre. Mes prises ayant naturellement une grande dynamique, mon programme a un LRA de 20 et un niveau de -24 LUFS (fig.8).

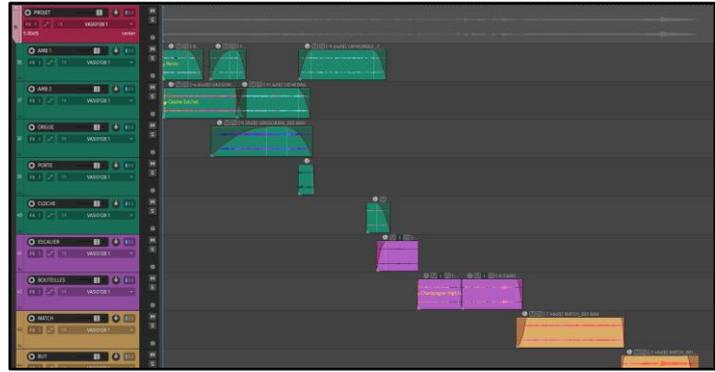


Figure 8: Montage Final

4. Difficultés rencontrées

L'une des principales difficultés de ce projet a été les bruits parasites liés aux personnes autour de moi. Il était évident que je ne pouvais pas demander aux gens de faire moins de bruits. Il a donc fallu choisir les meilleures prises, et malgré cela, il reste encore certains bruits que je n'ai pas réussi à retirer au nettoyage.

Un autre problème a été le fond d'air sur les ambiances dans la cathédrale. Il a donc fallu que je le corrige à l'égalisation, en baissant légèrement le bas de spectre afin de laisser de la place au reste, sans pour autant abîmer ma spatialisation.

3) Étude de cas

Après avoir créé et expérimenté une carte postale binaurale, il est intéressant d'analyser d'autres contenus du même genre afin de voir les similitudes et les différences afin de pouvoir cerner au mieux le problème.

Pour cela, il est nécessaire de mettre en place un barème précis et détaillé de ce qui peut être attendu d'une carte postale sonore binaurale.

D'abord, il est important de s'attarder sur la qualité de la captation :

Critère	Description	Questions à se poser
Clarté du signal	Propreté, absence de bruit, bonne intelligibilité	Y a-t-il des bruits parasites ? Les sons sont-ils nets ?
Dynamique sonore	Nuances, différences de niveaux	Y a-t-il une compression excessive ?
Bruit de fond	Faible, équilibré, ambiant, gênant	Est-ce que le bruit de fond nuit à l'écoute ?
Proximité des sources	Présence, distance apparente, perspective	Est-ce qu'un son est proche ou loin ?

Ensuite, il faut écouter la spatialisation du binaural :

Localisation latérale	Précise, stable, immersive	Les sons sont-ils correctement localisables de gauche à droite ?
Localisation verticale	Précise, stable, immersive	Les sons sont-ils correctement localisables de haut en bas ?
Distance apparente	Éloignement crédible, gradation des plans	Les sons sont-ils correctement localisables d'avant en arrière ?

Mouvements des sources	Trajectoires claires, naturelles, continues	Y a-t-il des mouvements ? Sont-ils fluides et perceptibles ?
Externalisation	Naturelle, stable	Les sons paraissent-ils venir de l'extérieur ?
Immersion globale	Enveloppement, sensation de présence	Se sent-on plongé dans une scène réelle ou fictive ?

Il y a ensuite l'intelligibilité narrative et la compréhension de la scène :

Lisibilité des plans sonores	Hiérarchie claire, bon mixage	Peut-on distinguer les différents plans ?
Cohérence spatiale de la scène	Réalisme ou crédibilité du placement des sources	La disposition spatiale a-t- elle du sens ?
Transitions	Fluidité	Les transitions sont-elles audibles ou justifiées ?
Effet narratif du son	Évocateur, informatif, immersif	Le son raconte-t-il quelque chose à lui seul ?

Après, il faut s'interroger sur la composition sonore et esthétique du programme :

Choix des sons	Pertinence, variété, originalité	Les sons choisis sont-ils significatifs ?
Rythme et temporalité	Cadence naturelle ou construite	Le montage sonore suit-il un rythme, une progression ?
Équilibre fréquentiel	Spectre harmonieux	Le spectre est-il bien distribué ?
Créativité et signature sonore	Esthétique originale, point de vue sonore	Ressent-on une intention artistique spécifique ?

Enfin, il y a la pertinence par rapport à l'objectif ou au sujet :

Cohérence avec le lieu ou le thème	Fidélité	La carte évoque-t-elle efficacement un lieu, une ambiance ?
Engagement sensoriel	Sensation, émotion, souvenir	La scène sonore suscite-t-elle une réaction chez l'auditeur ?
Lisibilité pour un auditeur extérieur	Accessibilité, universalité ou subjectivité maîtrisée	Un auditeur lambda peut-il comprendre le sujet ?

1. Création personnelle – Carte Postale de la ville de Reims¹

Après avoir détaillé le processus de création de cette carte, il est maintenant intéressant de l'analyser. Chaque point du protocole ci-dessus sera donc étudié afin d'établir les points positifs et négatifs de chaque carte. Ma création dure environ 4 minutes (fig.9).



Figure 9: ville de Reims

Concernant la qualité de captation, à part quelques bruits parasites dans certaines prises, elles sont globalement claires. Les sons sont nets et identifiables.

La dynamique est grande, peut-être un peu trop. Mais globalement, la différence entre les caves et le stade permet d'obtenir une variation intéressante.

Le bruit de fond est faible. Le vent a été atténué principalement dans le bas du spectre afin d'éclaircir l'écoute.

Pour la proximité des sources, certains sons sont plutôt proches et d'autres lointains. Cela donne du relief à l'écoute.

¹ Louis LEFOL MUSIL - Carte Postale de la ville de Reims

Pour ce qui est de la spatialisation, la localisation latérale est assez précise, tout comme la localisation avant / arrière.

La localisation verticale est plus subtile. Tout se joue sur la réverbération des lieux plus que réellement sur des éléments sonores verticaux.

Il y a peu de mouvements de sources sonores mais ils sont tous fluides et perceptibles. Ce point est certainement influencé par l'utilisation de mon HRTF étant donné que les prises sont natives.

L'externalisation est naturelle et le son semble venir de l'extérieur.

L'immersion globale est réussie. La scène semble réelle bien que les transitions entre les différents lieux soient audibles.

La scène est dans l'ensemble assez compréhensible. Les différents plans sonores ne se chevauchent pas et se complètent bien.

La spatialisation de la scène est réaliste. Les sources sont placées principalement en avant mais la réverbération naturelle augmente le réalisme.

Bien que les transitions soient audibles, elles restent fluides. Elles ne choquent pas l'auditeur dans son écoute, et elles permettent de facilement comprendre le changement de lieu au travers des changements d'ambiances.

Le son a pour effet narratif le voyage. Le but est de faire se balader l'auditeur au travers des lieux importants de la ville.

Concernant l'esthétique du programme, le choix des sons semble assez pertinent avec l'effet recherché. Les sons sont originaux, notamment avec la cave de champagne.

Le rythme est progressif du début à la fin. La carte commence dans la cathédrale pour se terminer dans le stade Delaune.

L'équilibre fréquentiel est assez bien respecté. Le bas du spectre n'est pas trop présent et les aiguës ne sont pas trop agressif pour les oreilles.

Pour ce projet, il n'y avait pas de réelle intention artistique autre que de partager les sonorités de la ville.

Vis-à-vis de l'objectif, les sons représentent fidèlement la ville de Reims. Que ce soit la cathédrale ou la cave de champagne, ce sont des sonorités typiques et iconiques de la ville.

Selon moi, l'engagement de l'auditeur au travers de l'écoute de ce programme est assez fort sur le début, puis réduit vers la fin. Cela est dû à l'universalité d'un stade de foot, à l'inverse d'une cave ou d'une cathédrale qui sont plus spécifiques.

Enfin, après l'avoir fait écouter à une personne qui n'est pas originaire de Reims, la compréhension du sujet est totalement abordable, même sans jamais avoir visité la ville.

Ce que je retiens de cette création est la difficulté liée aux bruits extérieurs à la scène que je souhaitais mettre en son. Malgré quelques bruits parasites, ce premier projet est pour moi une réussite, et me permet de imaginer de futurs axes d'améliorations autant technique qu'artistique.

2. Espaces Sonores – Kuthodaw Paya Mandalay¹

Comme nous l'apprend la description disponible sur Soundcloud, « le Kuthodaw Paya (fig.10) est l'un des temples immanquables de Mandalay ». Cette carte postale sonore a donc pour but de nous faire vivre un concert de gongs prenant place dans ce temple. Ce programme date de 2013 et dure 3 minutes.



Figure 10: Kuthodaw Paya

Pour commencer sur la qualité de la captation, le signal est extrêmement clair. Il est intelligible et n'a pas de bruits parasites à part à la fin où un léger bruit de craquements est audible. Ce bruit est d'ailleurs assez perturbant car à la première écoute, l'externalisation sur cet élément fonctionne assez bien.

La dynamique est agréable. Il y a une grande variation naturelle qui plonge l'auditeur en immersion.

¹ Soundcloud – Espaces Sonores - Kuthodaw Paya Mandalay

Le bruit de fond est faible voire quasiment nul à certains moments. Cela rend l'écoute agréable et immersive.

Pour finir sur la captation, la proximité des sources est assez bien gérée. Certains sons sont proches, d'autres plus lointains mais tout en gardant une cohérence spatiale.

Pour la spatialisation du binaural, la localisation latérale est vraiment précise. Chaque coup de gong peut être pointé du doigt sans pour autant les voir. Les voix les plus proches ont aussi cette précision, qui s'atténue plus elles s'éloignent.

Pour ce qui est de la localisation verticale, il est assez difficile d'entendre la précision, mais cela n'a sûrement pas à voir avec la qualité de la spatialisation mais simplement du fait qu'il y a peu de variation haut / bas dans ce programme.

La distance avant / arrière ressemble à la localisation latérale dans sa précision. Les gongs et les voix sont bien placés dans la distance, ce qui donne une sensation à la fois d'ouverture et d'externalisation.

Bien que les gongs soient fixes, les voix bougent. Pendant leur mouvement, leur trajectoire est claire et fluide. La sensation d'être sur place avec des gens qui passent autour de nous est réussie.

Comme dit rapidement ci-dessus, l'externalisation fonctionne bien, sauf peut-être pour les coups de gong très proches de l'auditeur. Les gongs semblent venir du casque et l'effet d'externalisation s'atténue.

Globalement, l'auditeur se sent rapidement plongé dans la scène, et même s'il n'a jamais expérimenté ce genre de son en vrai, la scène paraît réelle.

La compréhension de la scène est assez facile. En premier plan, il est possible d'entendre principalement des gongs, au deuxième plan, des gongs et des voix, et au troisième, l'ambiance du temple.

Au niveau de la cohérence spatiale, le réalisme fonctionne bien et la sensation de mouvement permet à l'auditeur de comprendre pourquoi certains gongs sont plus forts que d'autres et inversement.

Dans cette carte, il n'y a aucune transition, ou du moins, elles ne sont pas audibles. Il y a simplement un « *fade in* » et un « *fade out* » assez longs pour entrer et sortir de la scène.

En tant que tel, le son ne raconte pas une histoire explicite. Voir la scène aurait potentiellement aidé à s'imaginer le scénario ou l'histoire que veulent nous raconter les gongs. Le son seul n'en reste pas moins intéressant et immersif, ce qui permet de laisser l'auditeur s'imaginer lui-même une histoire.

La composition sonore est agréable à écouter. Les sons choisis sont pertinents compte tenu du sujet. Leur signification est importante étant donné le lieu dans lequel se passe la carte postale sonore.

Le rythme naturel des gongs permet de fluidifier l'écoute, ce qui donne une cadence naturelle assez progressive. L'entrée est lente, le rythme accélère vers le milieu puis finit lentement, comme l'entrée.

Le spectre sonore est harmonieux. Il n'y a pas de fréquences qui viennent déranger l'écoute. Peut-être qu'une balance de volume plus approfondie des différents coups de gongs pourrait être réalisée.

Pour ce qui est de la créativité sonore, il ne semble pas y avoir une réelle intention artistique outre le fait de vouloir partager la sonorité de ces gongs. Ceci n'est pas un problème en soi puisque la carte postale sonore a pour but de faire voyager, à l'inverse de la fiction sonore qui cherche à créer quelque chose d'unique.

Enfin, pour ce qui est du rapport au sujet, la carte évoque de manière plutôt efficace ce lieu. Bien que n'y étant jamais allé, les sonorités sont fidèles à ce que la description laisse imaginer.

Pour moi, l'engagement sensoriel est réussi. Cette carte me rappelle des souvenirs de voyages et cela crée donc une envie d'en écouter plus.

Finalement, il paraît évident de penser qu'un auditeur lambda puisse se laisser emporter par le sujet et comprendre de quoi traite la carte même sans avoir lu la description au préalable.

Pour conclure sur cette carte postale sonore, elle est selon moi proche de l'excellence en termes de qualité de captation, de spatialisation, de narration, de composition et de pertinence. Dès la première écoute, il est difficile de ne pas voir ces 3 minutes passer, et donc de chercher à vouloir en écouter plus.

3. La Grande-Motte - Carte Postale Sonore (Été)¹

Cette carte postale sonore est une carte interactive disponible sur le site de la Grande-Motte (fig.11). Elle met en son la ville au travers des différentes saisons de l'année. Chaque saison comporte 5 enregistrements différents. C'est donc pour cela que dans cette analyse, seul l'été sera analysé. L'ensemble des enregistrements fait un peu plus de 4 minutes et les pistes sont au format 16 bits – 44.1 kHz.



Figure 11: La Grande-Motte

La qualité de captation est trop juste. Le signal manque beaucoup de clarté. Cela est dû à la quantification qui n'est pas suffisamment importante. Cela induit un manque de précision dans les aiguës, et donc un manque de précision dans le binaural.

Concernant la dynamique, elle est totalement suffisante compte tenu des ambiances sonores.

En bruit de fond, le vent et le souffle de quantification sont tout de même présents. Cela nuit à l'externalisation. Le programme en lui-même est fort, mais l'air qu'il y a entre l'oreille et les sources sonores est trop importants.

Pour la localisation latérale, elle est assez précise. Chaque son est facilement identifiable dans l'espace.

Alors que pour la localisation verticale, comme pour la carte précédente, il est difficile d'avoir cet effet de hauteur. Ce n'est pas un problème lié à la carte en elle-même mais plutôt aux sons de la carte.

Pour la distance avant / arrière, cela fonctionne bien, mais les sources sont pour la plupart toutes trop loin.

¹ La Grande-Motte – Carte Postale Sonore

Il y a peu de mouvement dans cette carte, mais le peu de son qui bouge est facilement traçable.

L'externalisation est pour moi le point faible de cette carte. Étant donné la qualité et le bruit de fond, il est difficile d'externaliser. J'ai toujours l'impression d'avoir un casque sur les oreilles et d'écouter un programme, je ne me retrouve pas transporté en dehors de chez moi.

Malgré cela, l'immersion est assez bien réussie. La scène semble réelle, bien que l'externalisation ne fonctionne pas trop.

La scène est compréhensible grâce à la mémoire. En tant que tel, les plans sonores sont particuliers. Le premier plan ne semble pas exister, et le deuxième et troisième plan semblent mélangés. Il est assez difficile de différencier les plans.

La cohérence spatiale est réaliste car les prises sont des captures sonores de moments réel.

Il n'y a pas de transitions entre les différents morceaux de carte postale. Il faut que l'auditeur clique lui-même sur l'endroit de la ville qui l'intéresse. Ce système est intéressant et permet de pouvoir imaginer un système de déclenchement automatique des sons lors d'une visite de la ville.

Comme pour la carte précédente, le son seul ne cherche pas à raconter une histoire différente ou supplémentaire aux sons qu'il est possible d'entendre. Cette carte se fait simplement le messager sonore de la ville.

Au niveau de l'esthétique, bien que le choix des sons soit en rapport avec la ville, il y a certains problèmes qui sortent l'auditeur de la carte.

Il n'y a pas de rythme car il s'agit simplement de prises sans montage.

Et concernant l'équilibre fréquentiel, le manque d'aiguës donne l'impression que le programme a un problème technique.

Pour ce qui est de la créativité sonore, l'idée de mettre la ville en son dans des lieux différents est intéressante.

L'objectif de cette carte était de mettre en son la ville de la Grande-Motte. Les lieux évoqués sont assez fidèles à la ville.

L'engagement de l'auditeur peut être altéré par la qualité mais il reste selon moi important car les sons utilisés parlent au plus grand nombre.

Pour conclure sur cette carte, bien que les lieux choisis soient intéressants, la réalisation dessert le sujet en sortant l'auditeur du moment. Afin d'augmenter l'immersion et l'engagement, un réenregistrement en meilleure qualité permettrait de corriger ce problème.

4. ARTE Radio - *Mingalabar*¹

Ce programme est proposé par ARTE Radio. Il a été créé il y a 11 ans et dure une quinzaine de minutes. Cette carte postale plonge l'auditeur au cœur de la Birmanie (fig.12). Elle a pour but de faire découvrir les sonorités du train, des temples, de la vie de tous les jours.



Figure 12: Birmanie

La qualité de cette carte est superbe. Le signal est clair, sans aucun bruits parasites. Chaque son est net et facilement identifiable.

La dynamique du programme est assez grande. Certains passages sont plus calmes alors que d'autres plongent l'auditeur dans un bruit presque constant.

Le bruit de fond est quasiment nul. Même lors des moments les plus calmes, aucun souffle de quantifications ne vient déranger l'auditeur dans son écoute.

Les sources sonores ont, entre elles, une belle perspective. Et qu'un son soit proche ou loin, il reste tout aussi défini qu'un autre, ce qui aide à la compréhension de la scène.

Comme pour les autres cartes postales, la localisation latérale est précise. L'immersion est telle, que l'envie de tourner la tête vers le son est présente.

Pour ce qui est de la localisation verticale, il n'y a pas réellement d'éléments sonores qui jouent sur la hauteur à part bien sûr la réverbération naturelle des lieux.

Pour la localisation avant / arrière, la majorité des sources sont devant mais certaines sources qui tournent autour de l'auditeur passent bien derrière.

¹ Soundcloud – ARTE Radio - *Mingalabar*

En parlant de mouvements, ils sont tous fluides, et certains sont d'ailleurs surprenants (pas de façon négative). Ils permettent de jouer avec l'habitude que prend l'auditeur à écouter le binaural, ce qui permet de se « re-rendre » compte du côté immersif.

L'externalisation de la carte est donc impressionnante. Il est facile d'oublier le casque sur les oreilles et de se laisser transporter par les sonorités de la Birmanie.

L'immersion globale est également intéressante. La scène semble réelle bien que certaines transitions transportent l'auditeur dans une sorte de fiction.

La scène est totalement compréhensible. Les 3 plans sonores sont facilement identifiables et permettent d'obtenir un beau relief.

La cohérence spatiale est réaliste. Le placement des sources sonores est tout à fait crédible vis-à-vis de ce qu'il serait possible de vraiment entendre. Seules les transitions ne sont pas cohérentes, mais cela n'influe pas sur la compréhension de la scène.

Pour les transitions, elles sont audibles et dénotent du côté réaliste de la carte. Cependant, elles apportent une touche originale qui est totalement justifiée par les sonorités de gongs et d'eau.

Le son seul semble vouloir raconter une autre histoire. Il y aurait donc un double sens à trouver dans cette carte. Le premier lors de la simple écoute de la carte. Le second lors de l'écoute des différentes transitions et sonorités.

Concernant l'esthétique du programme, le choix des sons est à la fois pertinent, varié et original. Chaque son est à sa place, sans pour autant prendre le dessus sur les autres.

Le rythme est progressif. Le programme commence de manière assez forte, pour ralentir vers le milieu, et regagne en puissance pour la fin de la carte.

Le spectre fréquentiel est harmonieux et il joue avec les différentes sonorités afin d'explorer complètement tout le spectre.

Pour la créativité, il est assez facile de remarquer une esthétique assez originale. Cette carte n'est pas juste un enregistrement fixe de plusieurs lieux, mais réellement une histoire qui est racontée à l'auditeur.

Le rapport avec le sujet est complètement respecté. La fidélité des sons évoque efficacement les différents lieux qui sont proposés.

Grâce à la qualité de captation et aux idées esthétiques, l'engagement auprès des auditeurs est total.

La lisibilité pour un auditeur extérieur est donc facile. Cette carte est universelle et le sujet est compris dans les premières secondes d'écoute.

Encore une fois, ARTE Radio propose un programme d'une très belle qualité, autant sur le plan technique qu'esthétique. Les 15 minutes de ce programme semblent passer en quelques secondes, tant l'engagement dans ce programme est important.

Conclusion

Au travers de ce travail, j'ai appris comment la construction d'une carte postale sonore binaurale pouvait se faire.

Il est d'abord important de bien définir le cadre de son projet, ce qu'il est possible de faire ou non selon son temps et son budget. Il est donc nécessaire de connaître le matériel existant ainsi que les outils de post-productions disponibles au bon déroulement du programme.

De plus, il a été intéressant d'écouter d'autres cartes afin de se rendre compte des choses utiles à réutiliser dans de futures productions.

Ce qui ressort cependant de ce format est que la partie intention artistique est difficile à traiter. Réussir à proposer un autre sens dans une carte postale autre que celui du simple voyage dans un lieu reste compliqué, mais pas impossible.

C'est pour cela qu'il serait intéressant de s'attarder sur la différence qu'il peut y avoir en termes d'esthétisme entre une carte postale et une fiction sonore binaurale.

Glossaire

- Théâtrophone : Appareil mit au point par Clément Ader
- ITD (Interaural Time Difference) : Différence interaural de temps
- IPD (Interaural Phase Difference) : Différence interaural de phase
- ILD (Interaural Level Difference) : Différence interaural d'intensité
- HRTF (Head-Related Transfer Function) : Fonction de transfert relative à la tête
- Chambre anéchoïque : Lieu qui ne provoque pas d'écho dont les parois absorbent les ondes sonores et électromagnétiques.
- Externalisation : Capacité d'un son écouté au casque à donner l'impression qu'il vient de l'extérieur de la tête.
- Format SOFA (Spatially Oriented Format for Acoustics) : Fichier permettant de stocker diverses données liées à la spatialisation.
- LRA (Loudness Range) : Mesure l'écart dynamique dans un programme audio
- LUFS (Loudness Units relative to Full Scale) : Unité qui mesure le volume perçu d'un fichier audio, de manière standardisée.

Webographie

[Wikipédia – Battement binaural](#)

[France Musique – Le théâtrophone, le streaming classique avant l’heure !](#)

[SoundGirls – Recording and Binaural Audio](#)

[Frontiersin – Auditory localization: a comprehensive practical review](#)

[Musée de la Poste – Phonopostal - Phonographe](#)

Discographie

[Louis LEFOL MUSIL - Carte Postale de la ville de Reims](#)

[Espaces Sonores – Kuthodaw Paya Mandalay](#)

[La Grande-Motte – Carte Postale Sonore \(Été\)](#)

[ARTE Radio - Mingalabar](#)

Bibliographie

Bergame Périaux – Jean-Luc Ohl – Patrick Thévenot - *Le son multicanal - De la production à la diffusion du son 5.1, 3D et binaural.*

Wenzel - *Localization using nonindividualized head-related transfer functions.*