

École Nationale Supérieure Louis Lumière
Mémoire de fin d'études
Son 2014

Lucie BOURELY

Le mixage d'une prise de son globale

Analyse des procédés d'optimisation
de remise en phase des microphones d'« appoint »

Sous la direction de

Jean CHATAURET
Thomas VINGTRINIER

Rapporteurs

Mireille FAURE
Eric URBAIN

SAINT-DENIS, mai 2014

Résumé

La prise de son de musique acoustique demande une certaine méthodologie. On pratique à l'heure actuelle la méthode dite de « prise de son globale » : on place un couple de microphones stéréophoniques qui capte l'ensemble de notre scène acoustique puis on place des microphones d'« appoint » au plus près de chaque source présente dans cette scène acoustique. Cette multi-microphonie entraîne une double captation des sources sonores. Les signaux obtenus seront corrélés, mais non complètement identiques : déphasage de temps et d'intensité. Lorsqu'au mixage, on en vient à mélanger ces signaux, le signal d'« appoint » sera en avance dans le temps par rapport aux signaux du couple ce qui peut atténuer ou renforcer certaines fréquences, donc perturber l'image musicale souhaitée. Pour pallier à ce décalage, l'usage est de retarder les signaux d'« appoint ». Dans ce mémoire, nous étudierons l'effet qu'entraîne ce retard dans la pratique du mixage. De plus, avec les mixages réalisés, il sera étudié l'influence du traitement, par retard temporel au mixage, sur l'écoute audiophile et musicale de l'enregistrement.

Mots Clés

- Prise de son
- Prise de son globale
- Stéréophonie
- Retard temporel
- Jugement de goût musical
- Optimisation de la phase
- Microphones d'« appoint »
- Ecoute audiophile
- Ecoute musicale

Abstract

The recording of acoustic music needs a certain methodology. What is nowadays usually used is a method called : “global sound recording”. A stereophonic couple of microphones is installed in the middle of the acoustic stage : it will record all the sounds contained in the studied acoustic stage. Very often some additional microphones are installed as close as possible to every present source in this acoustic stage. This "multi-microphonia" brings a double captation of the sound sources. The obtained signals will be correlated but not exactly similar : time dephasing. When at mixing step, all the signals are mixed together, the additional closer signal will be early in time compared with signals of the couple. To make up for this time gap, the usual way is to delay the additional closer signals. In this report, we will study the effect induced by this delay and its treatment in the practice of mixing. Furthermore, with the realized recordings and mixings, will be studied the influence of the treatment of this delay effect on the audiophile and musical listening of the recording.

Key Words

- Sound recording
- Global sound recording
- Stereo
- Time delay
- Musical taste evaluation
- Phasing optimisation
- Additional closer microphones
- Audiophile listening
- Musical listening

Remerciements

Jean CHATAURET
Thomas VINGTRINIER
Virginie LEFEBVRE
Olivier MONTAGNON
Hanna MATAHRI
Leo ROSSY-OTH
Jean WAGNER
Neven LESAGE
Louis GALLIOT
Cippora LENGLET
Clémentine RICHARD
Maholy SAHOLIARILIVA
Lazare HERSON-MACAREL
L'Orchestre Philharmonique du COGE
Pierre BATISTE
Thierry ALADENISE
Claire et Sylvestre DENIS
Véronique DOUCOT
Vincent DUCROT
Guillaume HEBERT
François GELIN
Lucas DERODE
Nicolas RANDRIAMARO
Paul-Aurélien BONNICHON
Erwan CORVELER
François HELLER
Geraud BOURELY

Table des matières

Résumé	2
Abstract	3
Remerciements	4
Table des matières	5
Introduction	7
1 Prise de son globale	9
1.1 Production/Reproduction musicale	9
1.2 Diffusion et perception du son acoustique	11
1.2.1 Timbre d'une source sonore	11
1.2.2 Effet de masque	12
1.2.3 Perception de localisation azimutale	13
1.2.4 Perception de profondeur	14
1.2.5 Filtrage en peigne	15
1.3 Méthodologie de la prise de son globale	21
1.3.1 Image stéréophonique	21
1.3.2 Différents systèmes d'enregistrement stéréophonique	23
1.3.3 Défauts de la stéréophonie	28
1.3.4 Définition technique de l'actuelle prise de son globale	30
1.4 Méthodes de prises de son utilisées pour l'expérience	33
1.4.1 Contrebasse seule	33
1.4.2 Quintette à vent accompagné d'une contrebasse	35
1.4.3 Orchestre symphonique	36
2 Mixage	38
2.1 Goût musical	38
2.1.1 Révolutions musicales	38
2.1.2 Évolution du son dans la musique baroque	40
2.2 Cas judicieux de l'introduction d'un retard temporel	42
2.3 Outils de mixage	44
2.3.1 Identification du problème de phase	44
2.3.2 Unités du retard entre deux signaux corrélés	48
2.3.3 Méthodes de calcul de retard entre « appoint » et couple	49
2.3.4 Outils d'introduction de retard	51
2.4 Mixages de l'expérience	56
2.4.1 Choix des outils	56
2.4.2 Mixage de la contrebasse	58
2.4.3 Mixage du quintette et contrebasse	59
2.4.4 Mixage de l'orchestre symphonique	60
2.4.5 Mesures selon la norme Loudness EBU R128	62
3 Enquête de perception	63
3.1 Sujets de l'expérience	63
3.1.1 Critères d'écoute	63
3.1.2 Description du public interrogé	64
3.2 Méthode d'enquête	65
3.3 Analyse des résultats	67
3.3.1 Analyse de la distinction de différences	67
3.3.2 Analyse du vocabulaire employé	68
3.3.3 Jugement de goût	69
Conclusion	71

Bibliographie	72
Livres	72
Cours	72
Thèses et mémoires	72
Articles	72
Ressources internet	73
Table des illustrations	76
Annexes	77
A Courbes de perception	77
B Documentation des microphones utilisés	78
C Documentation du matériel utilisé	85
D Formules mathématiques	88
D.1 La transformée de Fourier	88
D.2 La cross-corrélation	88

Introduction

Aux débuts de la prise de son stéréophonique musicale, on a généralement défini la prise de son globale par l'utilisation d'un dispositif unique à deux microphones (couple) ou à trois microphones (arbre Decca). Cette méthode de prise de son avait pour référence la restitution d'une « réalité » qui était celle de la « meilleure place possible » dans la salle de concert. Certaines conventions sonores en résultent, comme la recherche d'un réalisme, d'une fidélité à une scène acoustique.

Aujourd'hui l'acception du terme « prise de son musicale globale » s'est élargie : on peut considérer qu'il désigne la captation simultanée de plusieurs sources sonores dans un même espace acoustique.

Dans ce contexte, si l'on a recours à la multi-microphonie, alors les signaux produits par les différents microphones sont inévitablement corrélés, puisque chaque microphone capte plusieurs des sources sonores. En d'autres termes, les différents signaux électriques produits comportent des caractéristiques temporelles et fréquentielles communes. Ces signaux sont pourtant différents, les microphones n'étant ni identiques ni coïncidents. La sommation de ces différents signaux réalisée lors du mixage a pour conséquence des annulations ou augmentations de l'amplitude de modulation de certaines fréquences, altérant le signal original capté par chaque microphone.

On peut donc anticiper des distorsions de timbre, et des distorsions d'image, conséquences du filtrage en peigne dû à l'addition des signaux « d'appoint » au signal « global ». Est-ce ce qu'entendent des sujets non professionnels ?

La question que je souhaite étudier ici est celle de la correction des décalages temporels inopportuns des signaux corrélés, décalages causés par la distance physique entre des micros captant simultanément la même source.

Cet effet est-il perceptible pour le public qu'il vise ? Il l'est déjà pour l'ingénieur du son qui réalise ce mixage. Et quelle satisfaction apporte-t-il à l'auditeur ? De quelle nature sont les différences perçues ? Ont-elles trait à la modification de l'image sonore ? (profondeur, relief...) À la qualité des timbres ? À un ressenti plus personnel, et propre à chacun ? Nous évaluerons la validité des problématiques évoquées par des tests et des entretiens ouverts. Il est donc logique de supposer que d'autres problématiques ou d'autres formulations surgiront, ce dont on ne peut que se réjouir.

Nous étudierons dans quelle mesure un public concerné, par la captation et la reproduction musicale, est sensible à la variation de la qualité sonore entraînée par les différents usages de la remise en phase. Il s'agira d'évaluer comment ces phénomènes et les solutions proposées sont perçus et évalués.

La finalité de ce mémoire est la conception d'une méthode d'utilisation des outils de retard temporel pour une optimisation dans la post-production musicale après une prise de son globale. L'objet de notre étude portera sur la musique classique enregistrée et s'adressera à son public dédié à la fois amateur de musique et dans une certaine mesure, audiophile.

Nous avons évoqué plus haut l'évolution des méthodes de prises de son globales. Elles sont rendues possibles par l'évolution technologique. (console de mixage, enregistrement multi-piste, apparition des retards temporels avec la technologie numérique, des nouveaux greffons logiciels (plugins) qui permettent aujourd'hui d'optimiser la cross-corrélation de signaux complexes.) On doit en parallèle considérer l'évolution du goût en matière de reproduction sonore lui-même lié à l'évolution des conditions d'écoute, à la diversification des médias et à l'augmentation de l'écoute du disque¹ par rapport à la diminution de la fréquentation des concerts.

1. L'écoute du disque comprend toutes les pratiques d'écoute musicale à partir d'un système de reproduction sonore. Je n'entend pas parler encore de l'écoute religieuse d'un disque. Mais de toutes les personnes qui écoutent un disque chez elles, en y focalisant toute leur attention ou pas.

1 Prise de son globale

1.1 Production/Reproduction musicale

Dans son livre, *Bien entendu, itinéraire d'un audiophile*, Francis Ibre nous parle des intervenants d'une production musicale :

« La reproduction musicale doit être appréhendée comme une activité artistique. Elle est en ce sens, d'un bout à l'autre du processus, depuis l'évènement sonore réel - le concert - jusqu'à sa représentation virtuelle - le disque écouté - le moyen d'expression de la créativité de chacun des hommes qui interviennent dans ce processus complexe. Tout d'abord le compositeur, puis l'interprète, le facteur d'instruments s'expriment à travers l'œuvre musicale. Puis l'ingénieur du son, l'architecte en tant qu'acousticien de la salle, le technicien du mixage, apportent leur représentation, sous la forme d'un enregistrement. Ensuite le constructeur d'appareils haute-fidélité, l'installateur, le fournisseur d'accessoires, l'architecte de la salle d'écoute donnent leur interprétation de cette représentation. Enfin l'auditeur fait l'expérience de la perception sonore, au travers de l'image sensorielle qui lui est transmise. » [6]

Proposons ici un découpage un peu différent de ces intervenants, orienté vers leur implication dans la production du disque. On peut séparer trois groupes : un premier groupe indépendant de cette production dans la réalisation de leur tâche, les seconds situés au cœur de la fabrication physique du disque et les troisièmes impliqués dans la réception de ce produit. On comprend que l'activité du troisième groupe dépend de celle du second groupe elle même dépendante de celle du premier. C'est l'équation de cette chaîne humaine que l'on va décrire ici.

Le compositeur, le luthier, le musicien et l'acousticien font partie de la chaîne d'une création immédiate du son. Pour eux la musique a une notion de temporalité qui est déterminée par le geste humain. Pour donner lieu à l'existence de la musique ils interagissent : les prérogatives techniques de la construction d'une salle (son architecture, ses matériaux), puis dans cet espace donné, la manière de jouer d'un instrument, construit d'une certaine facture et interprétant une certaine partition écrite, idée transcrite du compositeur. Les interférences entre ces quatre personnes (qui ne se sont pas forcément rencontrées ou concertées) donnent une identité particulière au rendu final, audible lors d'une représentation unique : le concert, car même d'un soir à l'autre, si l'idée musicale du compositeur reste la même, par exemple dans le cas d'un compositeur décédé, le public variera, ce qui modifie le travail acoustique, l'instrument changera, bougera, de part différents faits d'humidité, de température, de pression, et le musicien ne sera pas exactement dans le même état d'esprit, car il aura déjà donné un concert la veille. La

musique de concert est unique, elle a une marque précise dans l'espace temps, à cause de l'omniprésence humaine.

L'ingénieur du son, le directeur artistique, le concepteur de matériel d'enregistrement et de matériel Hi-Fi, le graveur de disque, le monteur, et le mixeur sont des parties de la chaîne de reproduction du son. Ils sont à la recherche de l'exactitude, de la perfection sonore reproductible de manière identique à l'infini. Ils veulent faire entendre à d'autres, dans un espace temps différent, ce qu'ils entendent lors du concert. Ils ont une volonté de création sonore universelle, pour un rendu différent des premiers intervenants : le disque. La reproductibilité transparente du concert peut être un idéal, mais il n'est pas le seul. La beauté d'une sonorité inaudible par nos oreilles seules peut en être un autre. Le choix du matériel technique de l'ingénieur du son se fait dans cette connaissance de l'idéal recherché. Ses outils pour créer un son ne sont plus seulement un espace, un instrument et un musicien (comme c'est le cas pour un compositeur) mais également toute une machinerie de microphones, de câbles, d'amplificateurs, de convertisseurs, etc. Alors, tout comme le musicien recherche un instrument parfait et une salle parfaite, l'ingénieur du son cherche le matériel qui convient le mieux pour recréer un son déjà existant mais d'une nouvelle manière, sur un nouveau support. Le son gravé sur le disque est un compromis entre la volonté de tous les intervenants de cette chaîne de création sonore et des réalités techniques.

L'audiophile, l'architecte de son auditorium et le concepteur de son matériel Hi-Fi² sont les derniers intervenants dans la création sonore du disque. En effet, en fonction des choix techniques de l'audiophile, reflet de son goût esthétique, le son qui lui parviendra sera différent. L'universalité du support que constitue le disque est à ce stade niée. Le son reproduit à partir d'une même gravure de disque va varier en fonction des conditions d'écoute de ce dernier. L'audiophile a donc une responsabilité sur ce qu'il écoute, et son expérience auditive sera tout aussi unique que l'interprétation concertante du musicien, cela même s'il écoute un seul disque, son attention auditive variant pendant toute la durée de l'écoute.

2. retransmettant le son : lecteur, amplificateur, enceintes, etc

1.2 Diffusion et perception du son acoustique

Avant de questionner les méthodes d'enregistrement, comprenons la diffusion d'une onde sonore de sa source jusqu'à son espace de diffusion. En effet, ce sont ces phénomènes physiques qui justifient le système de prise de son au regard du résultat sonore désiré. Le choix de la multi-microphonie se comprend d'abord par le phénomène physique qu'est l'onde sonore. Dans une production de disque, avec des sources acoustiques, on cherche à reconstituer un son proche de la réalité, souvent qualifié de naturel. En même temps, un nouveau son est créé, dans une production du disque, à cause de l'électronique et mécanique de captation (microphones, amplificateur), stockage (convertisseur, enregistreur, disque), restitution et diffusion (hauts-parleurs).

1.2.1 Timbre d'une source sonore

Le timbre permet à notre oreille et à notre cerveau de dissocier et distinguer entre elles les sources sonores. Ce timbre prend en compte toutes les caractéristiques sonores produites par la source et ne peut se résumer à seulement quelques critères. Cependant, pour notre étude, ramenons le « au spectre du son, c'est à dire à l'ensemble constitué par le fondamental et la série des partielles ou harmoniques. Le timbre est la sensation élaborée par le cerveau en réponse à cette analyse spectrale. » [3]. L'oreille humaine discrimine les timbres grâce au spectre fréquentiel et harmonique d'une source variant dans le temps : l'évolution temporelle de l'intensité de toutes les fréquences et de leurs harmoniques créées par une source sonore. Carl Stumpf montra, en 1930, que dans cette évolution la portion d'attaque des sons avait une grande importance dans cette discrimination des timbres ainsi que la décroissance du son. Cette faculté de discriminer les sources sonores se fait dans un laps de temps très court : « il est courant de dire que l'ouïe humaine a besoin d'au moins 50ms avant de reconnaître les différents éléments qui caractérisent le timbre d'un son. » [3]. Notre mémoire associe ensuite ces caractéristiques physiques à un souvenir auditif et visuel d'une source sonore. Ainsi, ce n'est qu'après avoir écouté une première fois une certaine source et l'avoir nommée que notre cerveau est capable de distinguer cette source des autres.

On comprend par cette courte définition du système complexe qu'est le timbre, que les filtrages en peigne risquant d'apparaître dans un dispositif d'enregistrement à deux microphones dissociés spatialement vont altérer et modifier le timbre des sources. Ces modifications auront pour conséquence un problème de distinction des instruments, ce qui est gênant dans un processus d'écoute musicale. Cette distorsion de timbre va jouer sur le jugement de goût de nos auditeurs, et c'est l'une de nos constatations lors des écoutes perceptives que l'on étudiera dans la troisième partie de ce mémoire. Cela est conforté par l'analyse faite dans *Le livre des techniques du son* de Daniel Mercier qui nous indique : « Le deuxième aspect du timbre concerne notre capacité à apprécier, pour

des sons provenant d'une même catégorie de sources, les fines variations possibles du mode de production, de la durée ou du contenu spectral. Nous dirons que tel piano a un son moelleux et rond ou encore que tel trombone est plus éclatant qu'un autre. » [3]

Le travail de prise de son prend une ampleur différente après cette analyse. En effet, suivant notre système d'enregistrement et notre méthode pour rendre ce travail le meilleur possible, selon un certain jugement de goût, il va toucher à l'interprétation même du musicien. D'un point de vue perceptif, le résultat ne sera pas perçu comme des erreurs de prise de son pour un public non averti, mais plutôt comme certaines erreurs d'interprétation du musicien. Nous verrons cela plus en profondeur dans la troisième partie de ce mémoire, mais c'est un résultat que l'on peut déjà anticiper. De plus, le lieu dans lequel joue un instrument imprime une marque liée au timbre : « Nous n'avons pas la même perception du timbre du même instrument en des lieux différents. » [3] Ce facteur va changer la manière dont on souhaite capter le son d'un instrument, comme nous le verrons dans la répartition entre champ direct et champ diffus (cf. partie 1.2.4, page 14). Cette composante du lieu aura des répercussions sur les facultés de discrimination des timbres par nos sujets interrogés ensuite.

1.2.2 Effet de masque

Lorsque la scène sonore est constituée de plusieurs sources, le résultat global donne lieu à une nouvelle perception par rapport à celle que l'on obtiendrait dans le cas de la superposition exacte de ces sources distinctes. Pour la musique acoustique : celle dite classique ou pour le jazz, plusieurs sources sonores différentes et distinctes produisent des ondes de manière simultanée dans le temps. Si les deux sources sont exactement identiques, on perçoit une augmentation de 3dB du niveau sonore global. En revanche, si ces sources ont des réponses fréquentielles très proches, ou des timbres proches, le résultat peut être la difficulté de percevoir l'une de ces deux sources. Celle dont l'intensité sera la plus forte prendra le pas sur la seconde source, jusqu'à ce que notre oreille se soit reconstituée un nouveau timbre ou fasse abstraction de la source la plus faible. Ce phénomène est l'effet de masque. Denis Mercier nous donne la définition suivante : « L'effet de masque est interprété comme un déplacement du seuil de l'audition provoqué par le son le plus fort et dépend de l'écart de fréquence et de niveau entre deux sons » [3]

Les musiciens, le compositeur et le chef d'orchestre sont conscients de ce phénomène, et souvent l'écriture musicale justifie la superposition de plusieurs timbres, ou bien l'évite par des jeux de réponses, ou de solos successifs. Mais lors d'un concert, notre vision nous aide à distinguer les sources, ce que l'on appelle l'effet « Cocktail Party » : « capacité d'un auditeur à détecter une information utile dans un environnement très bruyant, qui se manifeste dans des conditions naturelles d'écoute (foule, concert, théâtre) alors qu'il ne peut se manifester en écoute *via*

un système de restitution sonore (enceinte, casque) »[3]. L'absence de la vue dans la restitution de la scène acoustique va mener le preneur de son à devoir dissocier les sources masquées et masquantes, s'il y en a. Il prendra souvent la précaution de capter le son au plus proche de la source masquée, grâce à un microphone d'« appoint ». Ensuite à l'étape du mixage, le mixeur tentera de démasquer les sources grâce à ces différents signaux captés lors de la prise de son.

1.2.3 Perception de localisation azimuthale

Il existe également un moyen de démasquer certaines sources, lorsque les différences d'intensité sont faibles, en utilisant notre faculté à localiser de façon azimuthale les sources. Notre système auditif est constitué de deux capteurs : les oreilles, comme tous les systèmes sensoriels humains. Cette constitution permet à notre système de perception de distinguer la localisation d'une source sonore dans l'espace. Nous sommes capables de localiser une source selon un angle plus ou moins grand. Cet angle sera plus précis si une source se situe sur la droite perpendiculaire à celle formée par nos deux oreilles, et beaucoup plus grand si une source se situe sur la droite formée par nos deux oreilles (notre localisation frontale est plus précise).

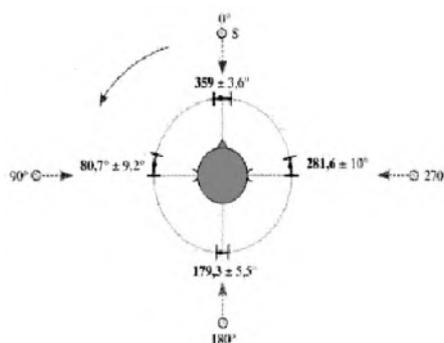


FIGURE 1 – Différentes zones d'incertitude de localisation sur le plan horizontal. D'après Haustein et Stimmer en 1970.

Les positions réelles sont dans le schéma indiquées par des flèches. Et les sources utilisées pour l'expérience sont des salves de bruit blanc de 100ms.

Cette localisation dépend de la différence de temps Δt et de la différence d'intensité Δi qu'ont les ondes sonores quand elles parviennent à chacune de nos oreilles, selon « Lord Rayleigh, en 1877 et 1907, puis Von Békésy en 1930. »[3]. Tout d'abord il y a la différence temporelle binaurale, ou encore effet de précedence (effet Haas). Tous ces termes techniques nous parlent du même effet : qu'un son parvenant à notre oreille droite, par exemple, au moins 1ms avant d'arriver à notre oreille gauche paraîtra venir du côté de notre tête où le signal audio lui est parvenu en premier. Ensuite la différence d'intensité binaurale, qui est la différence d'intensité

de l'onde sonore parvenant à l'une de nos oreilles par rapport à l'autre, due à la perte d'intensité de l'onde sonore dans sa dispersion dans l'air. Ces localisations naturelles sont réutilisées par le système de prise de son stéréophonique (cf. 1.2.3, page 13).

1.2.4 Perception de profondeur

Notre système auditif est capable de distinguer les sources lointaines des sources proches, car les ondes sonores perdent en intensité, et se reflètent sur les parois du lieu dans lequel elles se diffusent, constituant deux champs : le champ direct et le champ diffus. Le champ direct, selon *Le livre des techniques du son*, est constitué « des ondes sonores parvenant directement, c'est-à-dire sans avoir subi de réflexions sur les parois » [3]. Dans une acoustique très réverbérée, ce champ direct se situe donc dans une zone très proche de la source sonore. D'un point de vue mathématique, ce champ direct, pour une source située assez loin des parois d'une salle, est représenté par l'énergie directe selon la formule :

$$E_d = \frac{WQ}{4\pi cr^2} \quad (1)$$

Cette énergie a pour unité le *Joule/metre*³, où W est la puissance acoustique de la source sonore, c la vitesse de dispersion des ondes sonores : $c = 340m/s$, r la distance source-auditeur et Q le facteur de directivité de la source. Cette décroissance en $\frac{1}{r^2}$ est très rapide. Ce champ direct se situe dans un espace très proche de la source sonore. On comprend donc que, pour le capter, le microphone doit se situer dans cette zone, d'où le rapprochement des microphones d'« appoint ».

Le champ diffus, quand à lui, est « l'ensemble des ondes parvenant à l'auditeur après avoir subi une ou plusieurs réflexions ». [3] Et il est mathématiquement exprimé par la formule :

$$E_r = \frac{4W}{cR} \quad (2)$$

Avec R la constante d'absorption de salle. Ce champ diffus dépend de la puissance acoustique de la source et des caractéristiques de la salle. Plus une source est puissante ou plus l'espace est réverbéré (avec un coefficient R faible) plus cette énergie sera grande. Le champ diffus va donc varier avec l'espace où a lieu l'enregistrement. Si ce champ diffus dans cet espace d'enregistrement est intéressant, le preneur de son cherchera à le capter par son dispositif de prise de son. En revanche, si il apporte des défauts (une réverbération trop grande par exemple, ou une décroissance peu linéaire), le preneur de son cherchera à le dissimuler.

Entre ces deux champs acoustiques, on définit une distance critique qui se situe à l'endroit où $E_d = E_r$. Lorsque l'on résout cette équation, on obtient alors : $r = \sqrt{\frac{RQ}{16\pi}}$. On constate que cette distance dépend de la directivité de la source sonore et du caractère réverbérant du lieu d'enregistrement. C'est souvent à cette distance que l'on cherchera à placer le couple stéréophonique, car elle correspond à un équilibre entre source sonore et environnement.

Si au cours du mixage, on augmente l'intensité du microphone d'« ap-point », et donc du champ direct, notre oreille percevra un rapprochement spatial vers la source. En effet, c'est dans ce champ direct que l'on trouvera de la précision dans les attaques et les phénomènes transitoires de la source, ainsi que la présence du musicien avec sa respiration et le bruit des doigts sur son instrument. Il s'agit de sons très faibles par rapport à ceux produits par l'instrument lui-même, qui perdent de l'intensité et disparaissent dans le champ diffus et sont parfois inaudibles même à la distance critique. En revanche, si au cours du mixage, on n'écoute que le couple de microphone placé à la distance critique ou dans le champ diffus, la source sonore, l'instrument, sera auréolée d'une acoustique qui permet de mettre une certaine distance entre l'auditeur et l'instrument. Cela crée une sensation d'espace. La source ponctuelle paraîtra moins précise, et parfois plus difficilement localisable. Cette source sera restituée dans un espace temps donné pouvant être jugé agréable. Et lorsque l'on se trouve dans le cas de plusieurs sources sonores, on trouvera dans le champ diffus une fusion des timbres et un gain d'homogénéité. Ces critères se retrouveront dans l'analyse faite par nos sujets non avertis en matière sonore, dans la troisième partie de ce mémoire.

Enfin, la présence de plusieurs microphones, certains dans le champ direct et d'autres dans le champ diffus entraîne une différence temporelle de l'onde sonore dans les deux signaux obtenus. Cette différence, pour une même onde sonore, se traduit par une modification de la phase d'une onde par rapport à une autre aux propriétés fréquentielles très proches, et donc certains filtrages en peigne.

1.2.5 Filtrage en peigne

Pour parler du filtrage en peigne, je voudrais redéfinir succinctement, d'un point de vue physique, le son, en citant une nouvelle fois Denis Mercier : « les sons, les bruits sont des propagations d'ondes de pression dans les milieux élastiques »[3]. Une onde se définit elle-même par sa fréquence, son amplitude et sa phase. Et la phase répond à l'équation suivante :

$$Phase(^{\circ}) = \frac{Distance(m) * Frequence(s) * 360^{\circ}}{340m/s} \quad (3)$$

La phase d'une onde est la situation instantanée de son cycle : ses nœuds (0° , 180° , 360°) ou ses ventres (90° , 270°) et ce qu'il y a entre les deux. Lorsque l'on capte une même onde à deux instants ou lieux différents, on observe deux signaux identiques en fréquence, proche en amplitude (si ces instants ou lieux sont proches) et de phase souvent différente (suivant la relation de ces instants ou distances avec la longueur d'onde de cette onde). Prenons comme exemple, un signal audio, sinusoïdal de fréquence 100Hz. Ce signal a pour longueur d'onde 3,4 m, ce qui signifie que si l'on suppose cette onde entretenue, et non décroissante en intensité, que l'on additionne ce même signal à la source et 3,4m plus loin : deux microphones à une distance de 3,4m, on obtiendra un signal deux fois

plus fort que le signal initial, ce qui correspond à une augmentation de niveau de 6dB. On prend l'hypothèse d'additionner des signaux de même intensité, de même fréquence et de phase identique.

$$x_1 = A.\sin(\alpha_1) \quad (4)$$

équation de l'onde à la distance d_1 de la source, avec A son intensité et α_1 est l'argument oscillant de l'onde, sa phase : $\alpha = f * 2\pi * t$, f représente la fréquence de l'onde (en Hertz) et t l'évolution temporelle (en secondes) ou encore $\alpha = \frac{\lambda}{c} * 2\pi * d$, λ représente la longueur d'onde de l'onde, c la célérité du son dans l'air (340 m/s) et d la distance de l'onde par rapport à son point d'origine. On a choisi ici cette réduction mathématique qui ne prends pas compte le fait que les ondes sonores sont décroissantes et non entretenues afin de simplifier les équations et le problème qui nous intéresse ici.

$$x_2 = A.\sin(\alpha_2) \quad (5)$$

equation de l'onde à une distance d_2 , avec A son intensité et α_2 son déphasage par rapport à la source. Si $d = d_2 - d_1 = 3,4m$ alors $\alpha_1 = \alpha_2$ et l'onde résultante x aura pour équation :

$$x = \sqrt{(A.\sin(\alpha_1) + A.\sin(\alpha_2))^2 + (A.\cos(\alpha_1) + A.\cos(\alpha_2))^2} \quad (6)$$

Or, $\alpha_1 = \alpha_2$:

$$x = 2A\sqrt{\sin^2(\alpha_1) + \cos^2(\alpha_1)} = 2A \quad (7)$$

On constate que pour des signaux en phase, donc captés à deux endroits séparés d'une distance égale à la longueur d'onde du signal, l'addition de ces deux nouveaux signaux se résume à l'augmentation de l'intensité sonore. On peut l'illustrer par le graphe suivant :

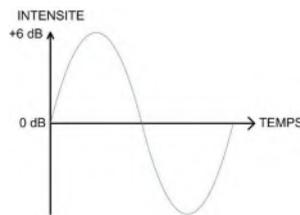


FIGURE 2 – Addition de signaux sinusoïdaux en phase

Réalisons le même calcul, pour deux captations réalisées à une distance d'un quart de cette longueur d'onde, dans le cas de ce signal de 100Hz, prenons $d=0,85$ m. Alors :

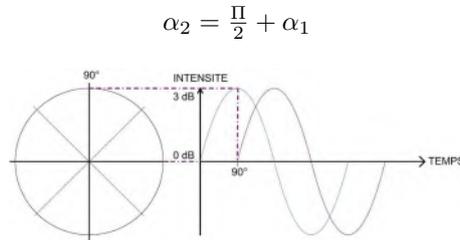


FIGURE 3 – Deux signaux sinusoïdaux déphasés de 90°

D'où :

$$x = \sqrt{(A.\sin(\alpha_1) + A.\sin(\alpha_2))^2 + (A.\cos(\alpha_1) + A.\cos(\alpha_2))^2} \quad (8)$$

$$x = \sqrt{(A.\sin(\alpha_1) + A.\sin(\frac{\pi}{2} + \alpha_1))^2 + (A.\cos(\alpha_1) + A.\cos(\frac{\pi}{2} + \alpha_1))^2} \quad (9)$$

$$x = \sqrt{2} * A \quad (10)$$

On a une augmentation d'intensité sonore de $\sqrt{2}$ fois l'une des ondes captées, c'est à dire de 3 dB. La résultante de cette addition est l'onde rouge dans le graphique suivant :

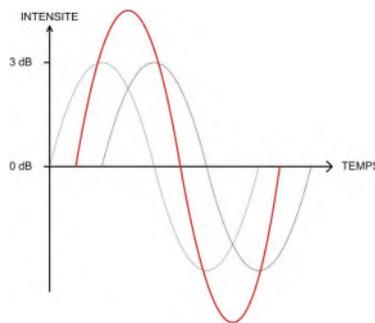


FIGURE 4 – Addition de signaux sinusoïdaux déphasés de 90°

Réalisons enfin la même opération pour deux signaux en opposition de phase, c'est-à-dire avec une distance $d=1,7m$, alors :

$$\alpha_2 = \Pi + \alpha_1$$

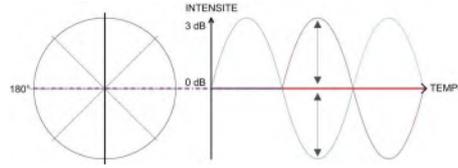


FIGURE 5 – Deux signaux sinusoïdaux en opposition de phase

D'où :

$$x = \sqrt{(A.\sin(\alpha_1) + A.\sin(\alpha_2))^2 + (A.\cos(\alpha_1) + A.\cos(\alpha_2))^2} \quad (11)$$

$$x = \sqrt{(A.\sin(\alpha_1) + A.\sin(\pi + \alpha_1))^2 + (A.\cos(\alpha_1) + A.\cos(\pi + \alpha_1))^2} \quad (12)$$

Or :

$$\sin(\pi + \alpha) = -\sin(\alpha) \text{ et } \cos(\pi + \alpha) = -\cos(\alpha)$$

On obtient le résultat :

$$x = 0 \quad (13)$$

Le signal résultant de cette addition est nul, les signaux s'annulent, et l'onde n'existe plus.

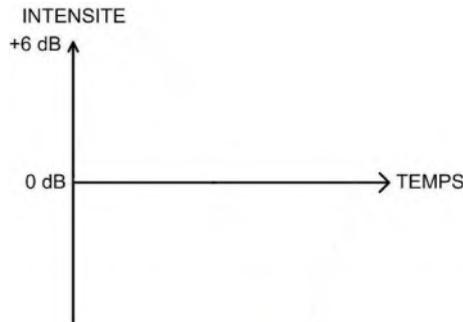


FIGURE 6 – Addition de deux signaux sinusoïdaux en opposition de phase, l'onde a disparu, il ne reste rien.

On a vu ainsi trois cas, exemples du résultat de l'addition d'une même onde captée deux fois à différentes distances. Pour chaque fréquence, il existera une distance où les signaux captés seront en opposition de phase, et une autre où les signaux seront en phase. On peut voir que dans le cas d'un signal audio complexe, avec beaucoup de fréquences, la double captation de cette onde à deux endroits séparés d'une distance donnée va entraîner l'annulation de certaines fréquences et l'augmentation d'autres lors de l'addition de ces deux signaux. C'est ce que l'on appelle **le filtrage en peigne**.

Prenons l'exemple d'un son contenant toute la bande de fréquence harmonique, un bruit blanc par exemple. On diffuse ce son par le biais d'une enceinte, et on l'enregistre à deux endroits de l'espace séparé d'une distance de 0,68m. Si on réalise un spectrogramme de la somme des deux signaux obtenus par cette opération.

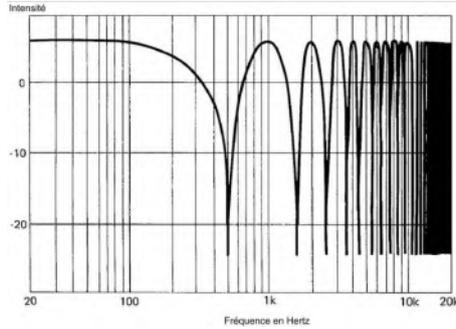


FIGURE 7 – Exemple de réponse en fréquence suite à un phénomène de filtrage en peigne

Comme nous l'avons vu au-dessus, les annulations ou additions vont être dépendantes de la distance entre les deux microphones. Cette distance est le double d'une longueur d'onde d'une certaine fréquence. Et c'est cette fréquence qui sera en opposition de phase entre les deux signaux enregistrés. Elle disparaîtra donc dans la somme. Dans le schéma ici, la première fréquence à s'annuler a pour valeur 500Hz.

On constate que d'autres fréquences s'annulent. Il s'agit des fréquences dont la distance représente le triple, le quadruple, un multiple entier de la longueur d'onde. Faisons le calcul alors dans le cas où :

$$\alpha_2 = n * \Pi + \alpha_1$$

Où n est un entier naturel non nul quelconque. Alors :

$$x = \sqrt{(A.\sin(\alpha_1) + A.\sin(\alpha_2))^2 + (A.\cos(\alpha_1) + A.\cos(\alpha_2))^2} \quad (14)$$

$$x = \sqrt{(A.\sin(\alpha_1) + A.\sin(n * \pi + \alpha_1))^2 + (A.\cos(\alpha_1) + A.\cos(n * \pi + \alpha_1))^2}$$

Or :

$$\sin(n * \pi + \alpha) = -\sin(\alpha) \text{ et } \cos(n * \pi + \alpha) = -\cos(\alpha)$$

On obtient le résultat :

$$x = 0 \quad (15)$$

On voit par le calcul, que 500 Hz ne sera pas la seule fréquence à s'annuler, il y aura aussi une annulation à 1000Hz, à 2000Hz, à 4000Hz etc. Ce sont toutes ces fréquences qui s'annulent et provoquent un filtrage multiple. Cette relation de multiple dans le domaine des fréquences est due à l'équation :

$$f = \frac{c}{\lambda} \quad (16)$$

où f représente la fréquence, c la célérité du son dans l'air, soit 340m/s et λ la longueur d'onde. Si l'on reprend notre première définition de la phase, en utilisant cette équation, on trouve le lien direct entre la distance de nos deux microphones et la longueur d'onde des ondes sonores enregistrées :

$$Phase = \frac{Distance}{\lambda} * 360^\circ \quad (17)$$

Pour remédier à ce phénomène, et rendre l'addition cohérente au moment du mixage, on introduit un retard au signal capté par le microphone placé au plus proche de la source. Ce retard doit être cohérent avec la distance que l'on a introduit entre les deux microphones. On appelle ce procédé : **la remise en phase** des signaux. On verra dans la deuxième partie, tous les moyens techniques et les choix esthétiques possibles dans l'introduction de ce retard.

1.3 Méthodologie de la prise de son globale

« Il n'est pas question, pour le preneur de son, d'occuper avec ses micros le « fauteuil idéal » saisissant la scène sous 50° . A cet emplacement, on mémoriserait surtout la belle impression musicale, l'image qu'on tentera d'approcher. »[10] La difficulté de l'enregistrement d'une scène de musique acoustique se résume bien ici dans le constat de Claude Bailbé, car notre système auditif n'est pas linéaire comme peut l'être un couple de microphones, nous réalisons une sélection auditive, pour percevoir l'information qui nous intéresse. Ce « rapprochement », dans l'enregistrement de la musique acoustique, est l'objectif premier du preneur de son. « Si c'est une symphonie d'un répertoire très classique, notre plus grande réussite, c'est de nous faire complètement oublier. Mais pour se faire oublier, il faut quelquefois beaucoup d'imagination... » nous dit Pierre Lavoix, grand ingénieur du son, dans le livre *Le Son des Musiques* de François Delalande. La grande invention des années 60 pour se rapprocher, oublier la prise de son fut la stéréophonie. Enfin, on percevait un espace, une localisation azimutale des sources, et on avait de plus en plus la sensation d'« être » au concert.

Pour atteindre cet objectif de la transparence sonore entre les musiciens et les auditeurs, les preneurs de son ont mis au point une technique particulière : la prise de son globale. Pour comprendre d'où vient cette méthode, je vais interroger d'abord le problème de l'écoute stéréophonique, son histoire, puis comprendre les limites de l'enregistrement par un couple seul, et enfin les raisons de l'apparition de la solution des microphones d'« appoint »

1.3.1 Image stéréophonique

Le mot « stéréophonie » est apparu au début des années 30 en France : « la Stéréophonie sera réalisée dans la reproduction des sons et elle sera pour l'oreille ce que les yeux attendent du cinéma en relief » disait Georges-Clément Levy dans son article intitulé *L'avis d'un ingénieur sur le « Sonore et Parlant »* publié au journal *l'Hebdo-film* du 19 avril 1930. Il y parle d'une nouvelle méthode, la « Stéréophonie », utilisée pour remplacer l'orchestre de fosse dans les salles de cinéma, du relief rendu par la présence de ces deux enceintes au lieu d'une unique ainsi que de la distinction des timbres rendue par ce système de reproduction sonore. Et ce n'est pourtant qu'en « 1931 que A.D. Blumlein de la toute jeune firme EMI dépose un brevet pour une technique d'enregistrement stéréophonique. »[7] Mais le passage de la théorie à une production industrielle a cependant pris du temps. « Un grand pas en avant est effectué vers 1936, lorsqu'on parvient, cette fois, à graver les deux signaux sonores différents dans un seul et même sillon. Mais on ne disposera pas d'un résultat satisfaisant tant qu'on aura pour support les matériaux couramment utilisés pour la fabrication des 78 tours. Il n'y aura pas de stéréo « correcte » tant qu'on n'utilisera pas le vinyle. »[7]

Dans le milieu musical, la stéréophonie connaîtra son essor en 1954, avec l'apparition des premiers microsillons stéréophoniques : le 33 tours par minute. Ce nouveau support audio permet l'écoute sans interruption de

plages durant 35 minutes. Les labels profiteront de cette nouvelle opportunité pour enregistrer les grandes symphonies, jusque là assujetties au changement de face au bout de 5 minutes (pour les disques de 30 cm). Ce sera l'âge d'or du disque analogique 33 tours et de l'enregistrement stéréophonique. Tous les supports audio qui suivront seront stéréophoniques. Daniel Lesueur relève cependant qu' « il est curieux de constater que la stéréo aura bien longtemps un petit côté « aristocratique ». Le procédé reste longtemps strictement confiné aux 33 tours. »[7]

Le système stéréophonique est le suivant : on enregistre une scène musicale grâce à deux microphones placés judicieusement. On reproduit le signal enregistré par chaque microphone dans l'enceinte qui lui correspond : le signal du microphone situé à droite dans l'enceinte droite, et le signal du microphone situé à gauche dans l'enceinte gauche. « Le but à atteindre est de tenir compte, puis de reproduire, les différences perçues d'une oreille à l'autre, en raison de leur emplacement légèrement différent par rapport à la source d'émission sonore. »[7] Avec cette méthode, on a réalisé une grande quantité de productions discographiques et on a mis en place des méthodes, des pratiques et des traditions de production. On peut lui donner la définition suivante, la stéréophonie est « La restitution d'une portion cadrée et limitée de l'espace acoustique à trois dimensions, anamorphée sur deux dimensions ». [8] (Francis Wargnier) ou encore « Un dispositif, des méthodes, des pratiques et des traditions de productions d'« images sonores » par deux flux sonores corrélés. » (Jean Chatauret)[8].

« Une reproduction stéréophonique s'obtient à l'aide de deux enceintes acoustiques constituant, avec la place de l'auditeur, les sommets d'un triangle équilatéral. Les deux enceintes se comportent alors comme deux sources sonores placées dans ce qu'il est convenu d'appeler « le cône de vigilance de l'oreille ». Ainsi les vibrations acoustiques arrivent aux oreilles droite et gauche de l'auditeur dans des conditions similaires à celles d'une écoute naturelle. L'oreille droite percevra d'abord le signal de l'enceinte de droite, puis celui de l'enceinte de gauche avec un certain retard et une légère atténuation. [...] Lorsque les deux signaux sont corrélés, les sources semblent alors provenir de la zone délimitée par les deux enceintes acoustiques. C'est le cas pour des signaux monophoniques dirigés artificiellement mais aussi des signaux stéréophoniques pour lesquels la corrélation existe dès la prise de son. »[3]

Il existe, à la reproduction, plusieurs sources sonores : les haut-parleurs, dits sources « réelles », et les sources « virtuelles » générées par la corrélation des signaux enregistrés puis diffusés dans les haut-parleurs. L'auditeur va percevoir des sources enregistrées entre les deux haut-parleurs, comme par exemple, les instruments d'un quatuor à cordes. L'auditeur perçoit toutes les caractéristiques physiques du son d'un violon, d'un alto et d'un violoncelle sans que pour autant aucun de ces instruments ne soient présents avec lui dans son espace d'écoute. On percevra également des sources dites « fantômes », ce sont les sources « virtuelles » indésirables, créées par le système de reproduction, soit dans l'espace de

diffusion (d'audition, d'écoute...), soit directement dans l'espace d'enregistrement. Un signal stéréophonique peut s'écouter dans un espace d'écoute donné, à l'aide de deux enceintes placées de façon judicieuse et réfléchi, ou bien à l'aide d'un casque audio, c'est-à-dire un haut-parleur placé très proche de chacune des deux oreilles, de sorte que le signal entendu est indépendant de tout espace d'écoute. Cette dernière méthode d'écoute sera peu traitée ici, car elle est peu utilisée par le public interrogé. De plus, le système stéréophonique utilisé et réfléchi utilise l'inter-modulation entre les canaux droit et gauche, pour créer une image stéréophonique, la plus proche possible de notre perception binaurale de l'univers sonore.

1.3.2 Différents systèmes d'enregistrement stéréophonique

Nous allons à présent évoquer les principaux systèmes microphoniques mis en place jusqu'à ce jour pour réaliser un **couple stéréo**. Par ce terme, je veux évoquer l'association de deux microphones à l'image de l'association de deux haut-parleurs. Je fais cette description car l'introduction de microphones d'« appoint » se justifie, entre autres, par les distorsions parfois non voulues de ces systèmes de prise de son. Historiquement, chaque système correspondait pendant longtemps à l'identité de certains très grands labels.

Le couple AB

Le couple AB est composé de deux microphones omnidirectionnels placés à au moins 40 cm de distance, et orienté avec un angle variable (souvent entre 0° et 120°), de préférence en direction de la source sonore car ces microphones sont directifs dans les hautes fréquences. Les choix de distance et de d'angle permettent de choisir l'angle de prise de son³. Le signal du microphone situé à gauche de la source sonore sera diffusé par l'enceinte gauche du dispositif stéréophonique et réciproquement, le signal du microphone situé à droite par l'enceinte droite.

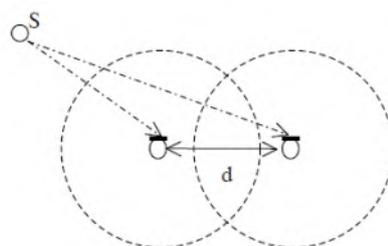


FIGURE 8 – Couple AB

L'image stéréophonique est créée par la différence de temps de captation entre les deux microphones pour une même source sonore. En effet, un

3. L'angle de prise de son est l'angle selon lequel on souhaite restituer la scène acoustique enregistrée. Le système de restitution reproduira dans son angle de 60° tout ce qui sera contenu par cet angle de prise de son. D'où l'anamorphisme du système, si cet angle de prise de son est différent de 60° .

même front d'onde atteindra l'un des deux microphone avant l'autre, créant une précédance, recrée par le système d'écoute stéréophonique, à laquelle l'oreille est très sensible. Ce système pose un problème de compatibilité monophonique, dû aux filtrages en peigne lors de l'addition des signaux droite et gauche. De plus, il y a une plus grande distorsion frontale pour ce système, que ceux seulement en Δi , avec un centre de l'image plus flou que les extrémités droite et gauche à cause de la distance entre les deux microphones. Chacun reproduira quelque chose de plus précis sur les extrémités qu'au centre qui est construit par un ensemble d'information capté par les deux microphones et donc des filtrages.

Le couple XY

Le couple XY est un microphone composé de deux capsules cardiodes coïncidentes placées avec un angle de 90° . Cet angle peut varier, modifiant l'angle de prise de son. Le signal de la capsule orientée vers la gauche de la source sonore sera diffusé par l'enceinte gauche du dispositif stéréophonique et le signal de la capsule orientée vers la droite par l'enceinte droite.

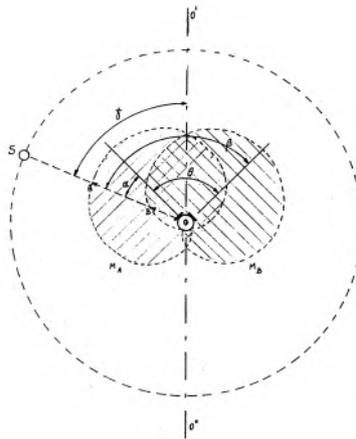


FIGURE 9 – Couple XY

La stéréo est créée par la différence d'intensité du signal capté par chacune des capsules. On nomme cet effet la stéréo d'intensité. Elle permet une plus grande précision de la localisation des sources dans l'image stéréophonique. La nature des capsules utilisées diminue le repliement de la scène arrière dans l'image utile, c'est-à-dire que les sources sonores situées à l'arrière du couple sont atténuées. Mais le couple XY devant être composé de microphones cardiodes, donc à gradient de pression, est peu linéaire dans son rendu en fréquence. Il en résulte que le timbre des instruments est moins bien rendu. Les couples XY souffrent surtout d'une perte dans les basses fréquences, ce qui limite leur utilisation dans certaines circonstances, ou favorise leur utilisation dans d'autres.

Le couple ORTF

Le couple ORTF est constitué de deux microphones à capsules cardioïdes séparés d'une distance de 17cm et d'un angle de 110° .

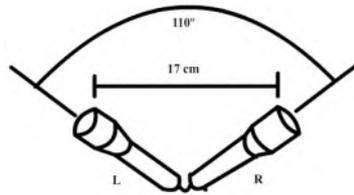


FIGURE 10 – Le couple ORTF

L'union d'une stéréophonie temporelle, l'écartement des microphones, et d'une stéréophonie d'intensité, la nature des capsules microphoniques et leur angulation, permet une plus grande cohérence de l'image, en diminuant les distorsions frontales et de perspective apportées par des systèmes seulement temporels ou d'intensité. Ce système a été mis en place par André Lamartine dans les locaux de l'ORTF, d'où il tient son nom. Il offre un angle de prise de son de 90° .

L'arbre Decca

L'arbre Decca est un système déjà multi-microphonique, et plus uniquement stéréophonique. Il est constitué de trois microphones, à capsules omnidirectionnelles, placés selon des angles de 90° à une distance d'environ 2 mètres.

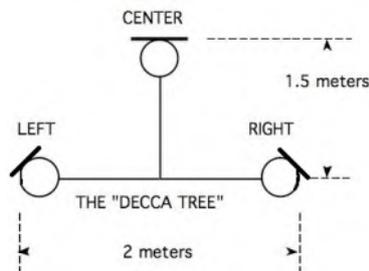


FIGURE 11 – L'arbre Decca

Ce système, appelé arbre car composé de plus de deux micros, est l'union d'un grand couple AB associé à une base monophonique. Il permet de rattraper les distorsions frontales présentes dans les couples AB classiques écartés de plus de deux mètres. Il fut très utilisé par le label anglais Decca, auquel il doit son nom, ainsi que par le label Mercury aux Etats-Unis, selon Pierre Lavoix. [5]

Le couple Blumlein ou Stéréosonic

Le Stéréosonic est l'association de deux capsules bidirectionnelles coïncidentes, orientées avec un certain angle physique qui permet la variation de l'angle de prise de son. Cette variation va de 70° pour un angle physique de 90° , à 140° pour un angle physique de 30° . La capsule orientée vers la gauche de l'image vient fournir le signal du canal gauche de notre système d'écoute stéréophonique, et la capsule orientée vers la droite de l'image celui du canal droit de ce système.

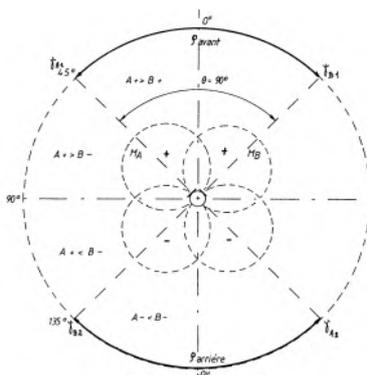


FIGURE 12 – Stéréosonic

L'avantage de ce système de captation est l'enregistrement « omnidirectionnel », à 360° , avec le repliement inversé sur la rampe stéréophonique, de l'image arrière, qui permet une distinction entre l'arrière en opposition de phase et l'avant en phase.

Le couple MS

Le couple MS est un microphone de deux capsules coïncidentes : une cardioïde (Mono ; M) dirigée vers la source sonore et une bidirectionnelle (Stéréo ou Latérale ; S) dirigée à la perpendiculaire de cette dite source afin de capter les sources sur les côtés. Pour la restitution de leur signaux enregistrés dans les deux canaux stéréophonique, on réalise un matricage :

$$\text{canalgauche} : V_A = a * V_M + b * V_S \quad (18)$$

$$\text{canalgauche} : V_A = a * V_M - b * V_S \quad (19)$$

V_A est la tension électrique du signal audio envoyée au haut-parleur

V_M est la tension électrique du signal audio créé par la capsule cardioïde (Mono ; M)

V_S est la tension électrique du signal audio créé par la capsule bidirectionnelle (Stéréo ; S)

a est le coefficient de présence monophonique et b le coefficient de présence stéréophonique.

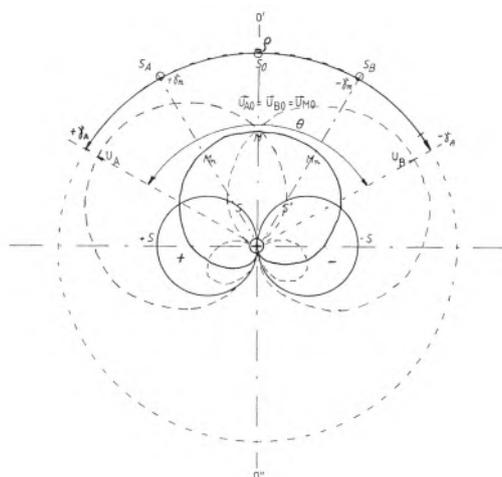


FIGURE 13 – Couple MS

L'utilisation de ces deux capsules permet un réglage de la largeur stéréophonique au mixage, suivant la valeur des coefficients a et b . Si $a = 1$ et $b = 0$, le résultat sera une restitution monophonique de la scène sonore. À l'inverse si $a = 0$ et $b = 1$, le résultat sera une image stéréophonique⁴ extra-large, avec un écrasement de l'espace sonore sur les côtés et l'annulation de l'image centrale. Le passage de l'extrémité gauche de l'image à celle de droite sera trop rapide. L'utilisation d'une capsule bidirectionnelle entraîne aussi un autre problème : la perception du hors phase entre les deux enceintes, avec la sensation que la stéréophonie dépasse les enceintes. Les sources sont alors difficilement localisables, le son provenant d'un espace différent de l'image définie contenue dans les 60° créés par les deux haut-parleurs. Ce système, selon Pierre Lavoix correspond au son Deutsche Grammophon des années 60, il permet une compatibilité idéale entre stéréophonie et monophonie, ce qui est un atout à l'époque.

Il existe d'autres systèmes que nous n'évoquerons pas ici, car on ne les retrouve pas dans les usages de la prise de son « globale ». Conformément aux usages actuels, en France, les prises de son de la partie pratique ont été réalisées grâce un couple AB, de 50 cm, les capsules pointées vers le sol, et un couple ORTF, les capsules pointées vers l'orchestre. Ces choix ont été réfléchis en fonction des sources sonores et du lieu acoustique de leur enregistrement, selon des critères esthétiques semblables à ceux de la majorité des productions actuelles dans leur domaine musical.

4. Il ne s'agit pas d'une « vraie » stéréophonie, mais d'une image reconstituée à partir d'un signal monophonique.

1.3.3 Défauts de la stéréophonie

Dans un article tiré de *L'audiophile* de juin 1991, Claude Bailblé [10] nous décrit les aberrations de la projection stéréophonique. Elles paraissent ici très pertinentes dans le sens où le plus souvent, ce sont ces aberrations qui justifient l'utilisation massive de microphones d'« appoint ».

« En se rapprochant, le couple microphonique favorise les instruments peu distants, tandis que les instruments plus obliques (plus lointains, donc) paraîtront trop diffus, trop flous, en raison du niveau moindre (en onde directe), du niveau augmenté (en ondes indirectes). C'est la *distorsion de présence*. » [10] On a vu plus haut que le champ diffus d'une source est, à une certaine distance, prépondérant sur son champ direct. Lorsque les sources sonores sont nombreuses, et donc étalées dans l'espace, le couple microphonique accentuera les distances entre les sources et entre les sources et le couple. En effet, il restituera les sources qui lui sont proches encore plus proches, car le champ direct est prépondérant dans le signal enregistré. Tandis que les sources plus distantes paraîtront plus lointaines encore, car le champ diffus est prépondérant dans le signal enregistré. Cette répartition entre champ direct et champ diffus n'étant pas linéaire suivant le lieu d'enregistrement, comme il est dit plus haut, il devient donc nécessaire d'enregistrer le champ direct de ces sources « lointaines » par le biais de microphones placés plus près des sources. En mélangeant ce nouveau signal à l'enregistrement stéréophonique, on corrige le problème de *distorsion de présence* en « rapprochant » certaines sources. C'est pourquoi le mixeur a à sa disposition tous les outils pour rééquilibrer avec sens les sources enregistrées.

« En outre, le premier rang des pupitres sera « vu » au premier plan, tandis que les derniers rangs seront captés sous un angle plus petit, accompagnés d'une réverbération plus présente. C'est la *distorsion de perspective* » [10] Lorsque les sources se déplacent sur un premier rang, devant le couple stéréophonique, le passage de l'extrémité gauche de l'image à celle de droite est plus rapide que s'il est fait à une certaine distance. Cela s'explique par l'ouverture plus ou moins grande de l'angle de prise de son. Les sources, dans l'image stéréophonique occupent une portion d'angle plus ou moins importante, on vient de le voir. Cette imprécision entraîne, en plus d'une distorsion de localisation, perspective, certains effets de masquage. En effet, pour deux sources que l'on localisera dans un même angle, ayant donc une même position relative dans l'image, si de plus spectralement ces sources sont proches, la source la plus forte en intensité sonore prendra perceptiblement le pas sur l'autre. Il s'agit d'un effet de masque, que l'on a vu plus haut.

La présence de microphones d'« appoint » devient alors nécessaire. Du fait de leur position, plus proche des sources, ils permettent d'éviter tous ces effets de masquage acoustique, et dans une scène sonore complexe,

d'enregistrer des sources sonores distinctes dans certaines limites. Leur signal, souvent monophonique, mixé avec l'outil de Pan Pot (raccourci anglais de « panoramic potentiometer », c'est-à-dire potentiomètre panoramique), qui est un déplacement linéaire ou logarithmique du gain d'un canal par rapport à un autre, permet de replacer ces sources dans l'image stéréophonique. On peut ainsi démasquer des sources. En introduisant un son de manière plus forte dans l'une des deux enceintes, cette source sonore paraîtra venir de cette enceinte. (cf. partie 1.2.3, page 13) Ces différences de perception gauche/droite répondent à une loi mesurée, entre autre par Henri Mertens ou encore par Gert Simonson. Les lois de Pan Pot correspondent à l'effet perceptif que l'on appelle la différence de perception naturelle de différence d'intensité : « *interaural level differences* ». Ces critères de panoramique sont très utilisés dans l'intégration de microphones d'« appoint ». On peut ainsi reconstruire par un effet d'intensité des sources sonores reproduites la scène stéréophonique.

De la même manière, il existe une différence de perception naturelle de différence temporelle des sources sonore : « *interaural time differences* »(cf. partie 1.2.3, page 13). A l'étape du mixage, on utilise cet effet en introduisant un retard différent de la transmission du signal audio entre les deux enceintes. Par exemple, si le même signal audio est transmis par l'enceinte droite de notre dispositif stéréophonique d'écoute 1ms avant sa transmission par l'enceinte gauche, ce son nous semblera venir d'une source située à droite de l'image stéréophonique⁵. Cette utilisation du retard temporel pour la localisation frontale stéréophonique est cependant moins précise que la localisation par phénomène d'intensité, d'où une utilisation limitée du retard temporel à cet effet.

C'est pourquoi, dans les mixages réalisés dans la partie expérimentale de ce mémoire, le Pan Pot d'intensité sera utilisé, pour les effets de localisation dans l'image stéréophonique. Le retard temporel est utilisé pour une cohérence du signal audio transmis et écouté. En effet, quand on mélange deux signaux audio d'une même source sonore, mais à deux endroits différents, on en revient à mélanger le champ diffus d'une source sonore avec son champ direct. Les deux microphones étant à deux endroits différents de l'espace, ils captent des ondes sonores similaires, notamment en fréquence, mais différents en termes de phase. L'addition de ces signaux risque donc l'annulation ou le doublage de certaines fréquences. Cet effet touchera forcément au spectre de la source sonore, et donc au timbre de celle-ci, perdant tout le côté naturel que l'on aura tenté de garder au cours de la prise de son par l'utilisation de microphones peu colorés spectralement et d'un système stéréophonique.

C'est ce problème qui nous intéresse dans ce mémoire. Pour permettre cette addition de la manière la plus cohérente possible, on introduit un retard sur le signal que l'on appellera « delay » du mot anglais qui signifie retard temporel.

5. Lorsque l'auditeur est placé à équidistance des deux enceintes.

1.3.4 Définition technique de l'actuelle prise de son globale

La production classique avec prise de son globale est une réalisation où l'on positionne un couple de microphone dit principal, qui construit l'image stéréophonique, auquel on ajoute des signaux enregistrés grâce à des microphones dit d'« appoint », situés plus près des instruments. C'est l'association des signaux corrélés captés par cette microphonie qui produit la prise de son. Nous avons vu au-dessus les phénomènes acoustiques pouvant justifier l'introduction d'une multi-microphonie dans l'enregistrement d'une scène acoustique complexe et composée de plusieurs sources. Cette justification correspond ou bien à des choix relevant du goût, ou bien à fournir au mixeur les éléments nécessaires pour résoudre les défauts du couple principal. Ces choix stylistiques ou pratiques s'expliquent par des réalités technologiques, culturelles et économiques, que je vais énumérer ici.

La technologie a beaucoup évolué, permettant de plus en plus facilement, et de manière mobile, l'enregistrement de nombreuses pistes audio. Ainsi, le multi-microphonie a réduit de beaucoup son coût économique, en peu de temps. Les premières consoles de la BBC dans les années 30, permettaient le mélange de 6 pistes. Puis les consoles à lampes, des années 60, ont permis le mélange de 6 à 12 pistes. La révolution a lieu dans les années 70-80 lorsque apparaissent les consoles à transistors puis celles à transistors à effets de champs qui permettent l'enregistrement multipiste jusqu'à 32 pistes. Si l'on met en parallèle l'histoire des supports de reproduction, on constate que les enregistrements stéréo qui apparaissent dans les années 60 ne bénéficient pas de la technologie du multi-piste. Les preneurs de son avaient donc une démarche différente avec le devoir de faire une balance à la prise de son, sans liberté de rééquilibrage au mixage comme à l'heure actuelle.

Dans la production de disque de musique acoustique, il existe deux types de réalisation : le disque dit « produit » où l'enregistrement a lieu dans un studio, sans public, et le disque dit « live », réalisé lors d'un concert. Le cas du disque « live » prend place dans une dynamique de restitution d'un événement. Les musiciens mettent l'accent sur l'aspect concertant de leur prestation, et le public veut pouvoir ressentir l'essence du concert tel qu'il l'a vécu, ou comme s'il avait été présent. Le preneur de son doit se faire le miroir de l'évènement pour produire un reflet le plus proche possible du concert. De plus, sa prise de son doit prendre en compte l'aspect visuel de sa présence, les microphones ne peuvent pas prendre une trop grande place et doivent rester discrets vis à vis des auditeurs présents au spectacle enregistré en « live ». La technologie le permettant, on utilise par exemple dans le cas de opéras, des petits microphones sans fil pour enregistrer la voix des chanteurs. Le temps de balance est court, et la présence du preneur de son sur le plateau est elle aussi réduite. Il utilise alors ces microphones d'« appoint » comme une sécurité par rapport au couple, devant restituer toute l'information musicale produite au cours du concert. Cette multiplication de l'information lui permet de réaliser un travail de rattrapage d'erreurs en post-production.

Le cas du disque « produit » est lui dans une dynamique de réflexion sonore en vue d'un unique résultat : le disque. Le preneur de son peut avoir une discussion sur le son avec les musiciens au cours de la séance de balance. L'utilisation ou non de la multi-microphonie est un choix commun au regard du résultat attendu. On se permet de choisir un lieu dont l'acoustique est intéressante et de prendre le temps d'une véritable balance. Et l'introduction d'« appoint » permet cette fois de former un son plus proche des instruments, où l'on entend toutes les impulsions transitoires du jeu des musiciens.

Cette recherche sonore s'explique entre autre par l'éducation musicale différente des musiciens actuels par rapport à ceux des années 60. En effet, ils ont grandi avec l'omniprésence de la musique actuelle dans notre environnement sonore. Et ce genre musical correspond à un goût pour la proximité des sources, les voix sont enregistrées de manière très proche, on perçoit toutes les impulsions de celles-ci. Il en va de même pour tous les autres instruments de ces styles musicaux enregistrés en ultra-proximité. La recherche d'un nouveau son pour la musique acoustique, en particulier dans le cas d'une production de disque, se rapproche donc naturellement de ce goût actuel pour l'ultra-proximité. Et pour obtenir ce résultat, le preneur de son ajoute des microphones d'« appoint ».

Notre univers sonore a beaucoup évolué depuis les années 60, car à présent, il est courant d'écouter de la musique dans toutes sortes d'environnements : dans la voiture, sur un baladeur, avec de petits autos-radios, ou encore sur nos ordinateurs. Tous ces systèmes de réception sont de qualité bien plus médiocre que les grandes enceintes d'autrefois. Le preneur de son doit prendre en compte ces nouveaux systèmes, et rendre le discours musical le plus lisible possible. L'enregistrement en proximité de sources acoustiques permet cette écriture lisible, en avantageant certaines sources lorsqu'elles interviennent dans le discours musical. Le discours sonore devient de plus en plus scolaire et prémaché pour l'auditeur.

« Les microphones d'« appoint » sont là pour assurer présence et définition, le couple constituant la base stéréophonique sera placé un peu plus en recul par rapport à la position qu'il aurait si on l'utilisait seul. Ainsi, on évitera les filtrages en peigne, en particulier avec les microphones d'« appoint » les plus proches. »[4] voilà ce que nous dit Denis Mercier de l'introduction des microphones d'« appoint ». Ce recul du couple peut être remis en question, car il n'est pas obligatoire, et souvent le placement du couple correspond à une place idéale dans le cas où il serait utilisé seul : on place le couple en l'écoutant seul d'abord. Et les microphones d'« appoint » sont ajoutés pour corriger les manques que l'on perçoit dans cette écoute du couple seul. Parfois, au mixage, certains microphones d'« appoint » ne seront pas utilisés, car leur présence ne représente qu'une sécurité.

Les filtrages en peigne dont parle Denis Mercier dans cette définition de la prise globale sont au cœur du sujet de ce mémoire. Car, si on cherche à les éviter au maximum, la corrélation des signaux enregistrés rends les filtrages en peigne inévitables. Le champ direct et le champ diffus

d'un même signal audio d'origine vont former deux nouveaux signaux audio corrélés. Ils auront certaines fréquences en commun, même si leur spectre général est différent. Leur addition provoquera l'annulation, la diminution ou l'augmentation (de 3dB) de ces fréquences communes. Or ce sont celles-ci qui vont définir le timbre d'un instrument.

« Le timbre peut être attribué à certaines structures physiques du son. Celles-ci se ramènent classiquement au spectre du son, c'est-à-dire l'ensemble constitué par le fondamental et la série des partielles ou harmoniques. Le timbre est la sensation élaborée par le cerveau en réponse à cette analyse spectrale. Parler de timbre n'a ainsi de sens que dans le cas de sons complexes et ce mot est principalement employé pour des sources sonores musicales. Une flûte et une clarinette sonnent différemment tout en émettant la même note, c'est-à-dire jouant la même hauteur.

Mais allier le timbre uniquement à la notion de fréquence fondamentale et série d'harmoniques ne suffit pas. Il faut tenir compte de l'instant précédant le régime établi (ou stationnaire) du son, c'est-à-dire les transitoires d'attaque. Il est en effet difficile à qui que ce soit de pouvoir faire la distinction entre une clarinette et une flûte à partir d'enregistrements dont on aura amputé artificiellement les attaques.

Certains facteurs dus notamment à l'environnement physique provoquent des colorations du timbre. [...] C'est la réverbération accompagnant l'extinction du son qui affecte ainsi la perception de la musique par le biais de multiples réflexions et de leur répartition dans l'espace. Toutes ces modifications ne nous empêcheront pas cependant de reconnaître le timbre de tel ou tel instrument ou de telle ou telle voix. En effet, le cerveau intègre ces différents paramètres et s'aide de la mémoire auditive pour décider de l'origine du son. »[3]

1.4 Méthodes de prises de son utilisées pour l'expérience

Pour ce mémoire, j'ai réalisé des prises de son en vue d'un mixage orienté sur l'utilisation du retard temporel entre un couple principal et des microphones d'« appoint », dans le cadre d'une reproduction d'une scène musicale acoustique. Pour cela, j'ai contacté des musiciens qui avaient besoin d'une prise de son, je me suis associée à des projets existants, pour me rapprocher d'une prise de son réaliste et actuelle. J'ai enregistré deux concerts : *Pierre et le loup* de Sergueï Prokofiev adapté pour un quintette à vent, une contrebasse et un récitant, ainsi que *La 5^{ème} symphonie* de Gustav Mahler interprétée par l'orchestre philharmonique du COGE (Chœur et Orchestre des Grandes Ecoles). Ce sont donc des productions qui se rapprochent d'un disque « live ». Ce choix a été fait pour des raisons pratiques, ce genre de prise de son, on l'a vu au-dessus, est le cas où l'utilisation de microphones d'« appoint » est essentiel pour le preneur de son, afin de disposer de toute l'information musicale avec un temps court d'installation.

J'ai réalisé trois prises de son distinctes afin de différencier certains problèmes. On l'a vu, pour chaque enregistrement, les sources sonores et les lieux varient, faisant varier les dispositifs d'enregistrement. Par exemple, plus il y a de sources sonores, plus il y aura de microphones d'« appoint ». D'autre part, plus l'espace sonore est mat, plus la distance critique, où l'on place le couple principal, se trouve loin des sources.

1.4.1 Contrebasse seule

La prise de son de la contrebasse m'a semblé essentielle dans mon étude, car le filtrage en peigne est très sensible, et audible pour des sources fréquemment graves. Elle a eu lieu à la fin de la séance de raccord réalisée à la suite de l'enregistrement de deux concerts du *Pierre et le loup* pour quintette à vent. Les microphones étaient placés pour l'ensemble :

1. -un couple de microphones situé à l'avant de la scène, un couple de microphones 4006 de chez DPA (cf. l'annexe B, page 78) situé à une distance de 3,8 m.
2. - des microphones d'« appoint » placés devant chaque instrument : la contrebasse avait elle aussi son microphone d'« appoint », un TLM 103 de chez Neumann (cf. l'annexe B, page 78).

La distance entre le couple et l'« appoint » (3,8m), si on l'assimile à une longueur d'onde, correspond à une onde d'une fréquence de 88Hz, soit dans l'échelle logarithmique des hauteurs de note (avec comme référence un diapason à 440Hz) à un fa deux gammes en dessous du la⁴⁶. On peut s'attendre à ce que cette note disparaisse lorsque l'on additionnera le signal de l'« appoint » avec ceux du couple, parce que l'on se trouve à la distance entre le couple et le microphone d'« appoint » où la fréquence de 88Hz (tout comme la fréquence 176Hz, 352Hz, 704Hz...) est en opposition de phase.

6. voir tableau des correspondances entre fréquences et hauteur dans les Annexes



FIGURE 14 – Microphone d'« appoint » de la Contrebasse.

Le choix des microphones et de leur position, on l'a vu au-dessus, est réalisé pour que chaque signal écouté de manière isolée, paraisse bien. Ensuite, on réalise le mélange qui doit aussi être le mieux possible. On a donc choisi ces positions en écoutant seul le couple puis seul le microphone d'« appoint ». Ce sont les habitudes qui nous ont poussé à faire le choix de ces microphones :

- Le couple de microphones DPA a été choisi car ce sont des microphones omnidirectionnels, ici séparés d'une distance de 50cm afin d'obtenir une image stéréophonique construite sur un principe de différence de temps et pour leur grande linéarité en fréquence qui s'oriente vers une plus grande transparence et une couleur proche de la réalité. L'avantage est que l'on a plus d'ambiance de la salle, et peu de distorsion de présence. La contrebasse, située de plus au centre de l'image, subira peu la distorsion frontale à laquelle on peut s'attendre avec ce genre de système en particulier sur des fréquences graves.

- Le microphone TLM 103 de Neumann a été choisi car il s'agit d'un microphone à directivité cardioïde, afin de ne capter que la contrebasse, et à large membrane, idéal pour capter les instruments à fréquences graves. J'ai demandé au contrebassiste d'improviser un morceau de jazz mêlant pizz de contrebasse et percussion sur la table d'harmonie. Le résultat est une improvisation simple, parfois arythmique et sans mélodie détectable. Cela a un peu surpris le public ensuite interrogé, habitué à des pièces structurées, mais a eu le grand avantage de concentrer leur attention sur le son des différentes versions, en particulier le timbre des pizz et des sons percussifs, et de leur distance par rapport au système d'enregistrement. Cela a permis d'éviter le problème rencontré pour les autres extraits, à savoir, s'absorber, se concentrer seulement sur le jeu du musicien ou le

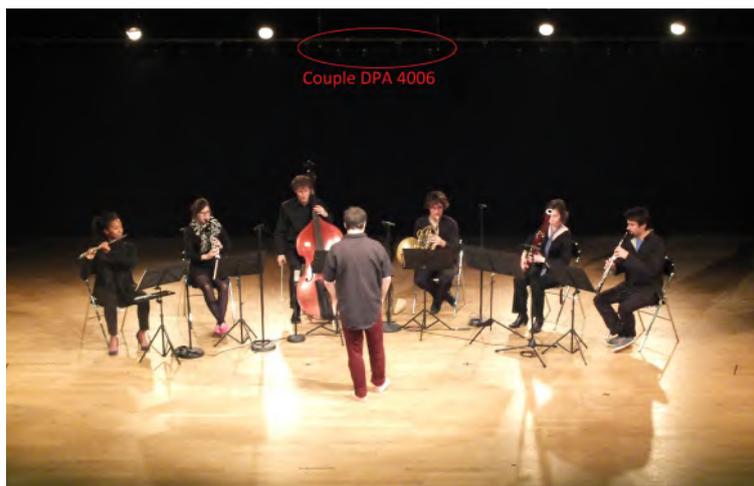
morceau interprété. Cela a eu enfin l'avantage de sortir mes sujets de leur zone de confort et d'expertise pour une concentration plus grande sur un sujet parfois nouveau pour eux : le son.

1.4.2 Quintette à vent accompagné d'une contrebasse

L'enregistrement du quintette à vent a pris place lors d'un concert de cet ensemble pour des classes de primaire. Cette prestation était filmée et enregistrée. J'ai aidé à l'installation de l'enregistrement sonore supervisée par Virginie Lefebvre. L'avantage de ce partenariat a été pour moi de pouvoir profiter d'un enregistrement réalisé avec un matériel professionnel et dans des conditions réelles. Sans intervention de ma part, la prise de son a été réalisée avec un couple de microphones et des « appoint » sur chaque instrument, comme c'est souvent le cas. La balance s'est faite pendant la répétition générale, qui s'est tenue avant le premier concert de la journée. Il y a eu ensuite deux concerts dans une même journée. J'ai pu ainsi faire un montage des meilleures prises entre les deux concerts et le raccord qui a suivi. Nous avons placé chaque microphone en écoutant chaque source de manière individuelle et le couple suspendu afin d'obtenir un résultat de l'ensemble du quintette et de l'ambiance de la salle où prenait place le public.

Intruments	Microphones ⁷	Angles (en degré)	Distance (en mètre)
Flute	Schoeps mk21	60	3,90
Clarinette	Neumann Km 184	30	3,85
Contrebasse	Neumann TLM 103	10	3,80
Cor	Neumann Km 184	10	3,80
Basson	Schoeps mk21	20	4,10
Hautbois	Neumann Km 184	60	3,90

Le comédien était quant à lui équipé d'un HF DPA 4060. Voici une photographie lors d'une répétition entre les deux concerts :



7. (cf. l'annexe B, page 78)

FIGURE 15 – Vue d’ensemble de la prise de son de *Pierre et le loup*.

Pour enregistrer, nous avons utilisé l’amplificateur et convertisseur Mics-tasy de chez RME (cf. l’annexe C, page 85), puis en ADAT nous sommes connectés à la Fireface 400, toujours de chez RME, (cf. l’annexe C, page 85), enfin en Firewire, nous étions connectés à un ordinateur où tournait un Pyramix.



FIGURE 16 – Installation en régie de l’enregistrement *Pierre et le loup*.

1.4.3 Orchestre symphonique

L’orchestre philharmonique du COGE m’a demandé d’enregistrer leurs concerts des 3 et 4 avril 2014 à l’église Saint Marcel. Il a eu l’amabilité de me laisser utiliser ces enregistrements pour mon mémoire. La seconde partie du concert était la 5^e symphonie de Malher dont j’ai extrait le début de l’Adagio pour orchestre à cordes et harpe ainsi que le début du 5^e mouvement afin de composer mes tests audio. Cet enregistrement a été pour moi très important dans ce mémoire, car le cas de la prise de son d’un orchestre symphonique se réalise avec un très grand nombre de microphones d’« appoint ». Ces microphones sont placés à des endroits très différents par rapport au couple. En comparaison, les microphones d’« appoint » placés pour le quintette paraissent être à équidistance du couple.

Le couple principal était composé de deux microphones mk5 de Schoeps (cf. l’annexe B, page 78) réglés sur les capsules cardioïdes. J’ai fait ce choix, car l’église Saint Marcel est un lieu très réverbérant, et réglé en omnidirectionnel, les microphones du couple perdaient beaucoup en précision. Les deux microphones étaient séparés d’une distance de 30 cm, sur une barre de couple adéquate, et d’un angle de 90°, dans une position d’un couple ORTF (cf. paragraphe 1.3.2, page 25). Le couple était placé à l’arrière du chef d’orchestre.

Les pupitres de cordes, placés en arc de cercle par rapport au chef d’orchestre, étaient donc relativement proche du couple. Devant chaque pupitre, près de chaque chef d’attaque, j’ai placé des microphones d’« appoint ». Pour les pupitres de contrebasse et de clarinette, j’ai fait la même chose. Un microphone était également placé en proximité de la harpe. Pour les pupitres des bois, j’ai fait le choix d’un autre couple de

microphones cardioïdes, placés aussi en position ORTF. Le but était de reprendre les attaques et la proximité des instruments parfois faible en intensité : les hautbois, les flûtes, les bassons. Les pupitres des cuivres et des percussions ont été laissés sans microphones d'« appoint », car ce sont des instruments à forte puissance sonore, et donc très perceptible dans le signal de notre couple principal.

Intruments	Microphones ⁸	Angles (en degré)	Distance (en mètre)
Violon 1	oktava MK012	50	2,1
Violon 2	oktava MK012	50	2,1
Alto	oktava MK012	20	2,9
Violoncelle	oktava MK319	20	2,9
Contrebasse	oktava MK319	60	5,77
Harpe	oktava MK012	30	6
Couple Bois	Sontronics STC-10	0	4,5
Clarinette	Sontronics STC-10	20	5,8

Tous les microphones étaient réglés selon leur directivité cardioïde. Et voici un schéma de la disposition de l'orchestre, elle correspond a une norme établie par le compositeur, Gustav Malher :

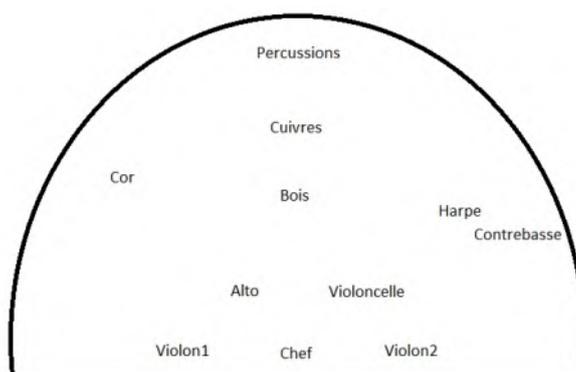


FIGURE 17 – Implantation de l'orchestre du *COGE* pour l'interprétation de la 5^{ème} *symphonie de Malher*.

Les sources étaient dispersées dans l'espace. Les sources les plus lointaines étaient des sources puissantes et directives, reculant la distance critique dans ce lieu très réverbérant. C'est pourquoi, j'ai fait le choix de ne pas leur assigner de microphones d'« appoint ». En revanche, il était important de venir au plus près des sources les plus faibles, ici les cordes, pour capter toutes les inflexions de leur jeu musical. C'est pourquoi ce sont ces sources que j'ai « appointé » dans ma prise de son.

Les enregistreurs utilisés étaient deux DR680 de chez Tascam, synchronisés (reliés en cascade), pour permettre aux deux enregistreurs de commencer et terminer leurs périodes d'enregistrement de façon simultanée.

8. (cf. l'annexe B, page 78)

2 Mixage

Le mixage correspond à la dernière étape dans un processus de production de disque. C'est le moment où l'on choisit le son à graver en regard des choix pris lors de la prise de son.

2.1 Goût musical

Dans la première partie du mémoire, j'ai évoqué les évolutions qui ont eu lieu pour la prise de son. Ces évolutions correspondent à des évolutions naturelles historiques du goût musical et sonore. Je souhaite évoquer ici comment et pourquoi ces évolutions ont eu lieu. Et ce, dans le but de comprendre le goût actuel pour certains mixages, et justifier non plus seulement d'un point de vue physique et mathématique, mais plutôt d'un point de vue sociologique et esthétique, les choix de mixage réalisés actuellement après une prise de son globale.

2.1.1 Révolutions musicales

Au cours des siècles, l'écoute et les pratiques musicales ont changé. François Delalande, dans son livre *Le son des musiques* [5] retient deux révolutions.

Tout d'abord celle de l'écriture musicale.

Jusqu'à l'apparition de l'écriture musicale, entre la fin du IV^e siècle et le IX^e siècle, la musique se transmettait par son interprétation, les générations mémorisaient la musique de leurs ancêtres et y apportaient forcément des modifications. On retrouve ce phénomène encore aujourd'hui dans les musiques dites « traditionnelles ». En Occident, l'écriture musicale coïncide avec une complexification harmonique de l'écriture. Il est difficile de dire quel phénomène a entraîné l'autre. La musique devient alors définitive sur plusieurs aspects : d'abord la hauteur, la rythmique, puis ensuite l'instrumentarium. Et enfin l'interprétation ? Sans aucun doute, la complexification de l'écriture musicale a amené de plus en plus de contraintes aux musiciens quant à leur interprétation. Mais l'écriture leur laisse toujours une plus ou moins grande part de liberté.

La seconde révolution, toujours selon Delalande, est celle de l'enregistrement. Soudain, un musicien ne fixe plus seulement des indications musicales, mais la musique elle-même. Et vient avec cela une nouvelle individualité, celle du son. Pour une même œuvre écrite, il existe parfois une multitude d'enregistrements. Pourquoi ? D'abord pour l'interprétation du musicien, qui est aussi variée qu'il existe de musiciens. Mais pas seulement, il existe aussi une nouvelle variété dans le « son ». L'enregistrement apporte une nouvelle technicité à la musique, le « son » qu'elle doit avoir. En effet, un « son pop » n'est pas le même qu'un « son jazz », même si parfois on retrouve les mêmes instruments. Ces derniers vont être « traités » par les machines d'enregistrement différemment en fonction de l'esthétique musicale que l'on cherche à créer. » [5]

Voilà un très bref aperçu de l'évolution de la pratique musicale, qu'en est-il de l'écoute ? Dans ce mémoire, c'est l'opinion de l'auditeur que l'on

cherche à connaître. Il me semble important de rappeler son histoire. Après la première « révolution » musicale, il fallait une certaine éducation pour « écouter » la musique d'une partition. En effet, seules les personnes ayant un certain niveau de « solfège » peuvent « écouter » la musique grâce à la lecture seule de la partition. Mozart dit-on écrivait ces œuvres sur une table sans l'aide d'un piano, et Beethoven en étant sourd. Mais il s'agit d'une écoute intérieure, qui ne passe pas par le biais mécanique des oreilles, et qui est seulement mentale. Jusqu'à l'invention de l'enregistrement, en 1877, il fallait donc des musiciens, personnes physiquement présentes, pour véritablement écouter et entendre de la musique, et cela, l'écriture ne l'a pas remplacé. Il faut aller au concert pour écouter et voir un musicien qui interprète l'œuvre écrite d'un compositeur. Ou bien lors d'un événement particulier : une célébration religieuse ou autre. L'expérience musicale correspond à un événement social particulier, qui fait titre d'exception, et qui se vit en groupe. L'enregistrement : la « seconde » révolution amène quelque chose de nouveau, d'inédit : la musique à la maison. D'abord par la radio, puis le 45 tours, le 78 tours, le disque numérique et enfin le lecteur mp3. Quelle révolution pour l'auditeur ! On ne l'oblige plus à se déplacer vers les musiciens ou la musique, à aller au concert, c'est le concert qui vient à lui. Enfin, il peut choisir sa musique parmi une multitude, et surtout il peut l'écouter seul. Plus besoin de supporter les autres au concert. La musique devient une expérience solitaire. Mais l'auditeur a soudain une nouvelle responsabilité : le son. Tandis qu'au concert, le public vient voir un produit fini, où son intervention sonore s'arrête à sa seule présence, l'écoute d'un enregistrement demande un matériel particulier dont la responsabilité incombe totalement à l'auditeur. On lui demande de choisir ses enceintes, son lecteur de disque, son amplificateur, son convertisseur, et puis ses câbles et enfin son lieu d'écoute. Si quelqu'un n'aime pas une musique, on peut lui reprocher son installation d'écoute. D'où la problématique de l'audiophile développée dans le livre de Francis Ibry, *Bien entendu, itinéraire d'un audiophile*, [6] qui explique la démarche d'écoute que doit avoir un audiophile avant de chercher des solutions physiques à son installation sonore.

La reproduction amène un nouvel outil dans le quotidien : le haut-parleur. Le son n'est plus créé par des instruments joués par des musiciens, mais par des haut-parleurs, qui à eux seuls sont capables de remplacer tout un orchestre. La compréhension de ce changement par les musiciens baroques est très perceptible dans leur manière d'appréhender la musique et leur interprétation. Ils ont été les précurseurs de notre goût actuel en matière de son pour la musique dite « classique ». Il y a chez eux une véritable recherche dans la compréhension et l'audibilité de la ligne musicale. C'est cette recherche qui justifie notre dispositif d'enregistrement, la prise de son globale.

Je vais donc évoquer et expliquer cette évolution.

2.1.2 Évolution du son dans la musique baroque

Pour la musique baroque, les évolutions techniques sont marquées par des esthétiques sonores différentes. Chaque nouvelle génération de musiciens a profité des passages sur des nouveaux supports pour apporter leur interprétation d'œuvres connues, comme les *Quatre saisons* de Vivaldi, car ils amenaient d'abord un nouveau son. Cette recherche se mène en parallèle à celle menée en musicologie pour se rapprocher de ce que les anciens pouvaient faire lorsque les œuvres ont été écrites.

« Le 78 tours captait les inflexions mélodiques, les phrasés, les rai-
deurs d'attaque, les nuances, les tempi, c'est à dire, tous les traits qui
définissaient traditionnellement l'interprétation. La nouvelle dynamique
du microsillon permet de saisir en plus tout ce qui est petit, à peine
audible, et se perdait dans le bruit de fond du 78 tours, à savoir : les
harmoniques qui ont peu d'énergie et concourent à la qualité du timbre,
les résonances (par exemple des cordes sympathiques), les petits bruits
qui donnent un caractère concret au son, tels que le frottement de l'ar-
chet, la respiration des musiciens, ou encore les bruits d'air, et surtout
l'espace, même en monophonie (la stéréo n'est utilisée sur disque qu'à
partir de 1957), puisqu'on peut enfin entendre clairement la réverbéra-
tion de l'édifice, ce qui impose de choisir des lieux pour leur acoustique
naturelle - d'ailleurs Harnoncourt a principalement enregistré avec un
couple suspendu et seulement un ou deux micros d'« appoint ». »[5]

« Les musiciens du Giardino Armonico sont les parfaits représentants
de la génération du numérique. Le groupe est fondé en 1985 [...] ; ils
n'auront enregistré qu'en numérique. Or ce qui frappe immédiatement
chez eux c'est le « son ». Bien sûr, ils héritent des recherches sur la
lutherie, qu'ils prolongent, des travaux musicologiques, qu'ils étudient,
des réflexions sur la rhétorique déthématisées chez Harnoncourt (1984
et 1985), mais ils héritent aussi du son de la musique « pop » de leur
enfance, ils savent ce qu'est la « pêche » d'une attaque, et bien sûr ils
maîtrisent la prise de son, l'équilibrage à la console et même le montage
sur ordinateur. »[5]

L'influence du renouveau baroque sur la musique occidentale et son
enregistrement est indéniable. On trouve deux problématiques dans ce
genre d'enregistrement. Les musiciens viennent avec un nouvel équilibre,
qui correspond au placement des musiciens dans un espace donné, et
parfois l'ingénieur du son peut se contenter de placer deux microphones
qui restituent cette réflexion, faite par le chef d'orchestre et l'ingénieur
du son. Un travail acoustique, la balance, réalisé en collaboration entre
ces deux acteurs du son. Ainsi en parle très bien Pierre Lavoix : « Pour
vous donner un exemple, les violons baroques ont moins de puissance
que les violons modernes. Mais comme ils vibrent très peu, lorsqu'ils
sont à l'unisson, brusquement ils donnent une puissance que n'ont pas
les instruments modernes (qui eux vibrent, donc l'énergie ne s'additionne
pas). Influencé par ces effets curieux devant les microphones, et par le

style des musiciens qui étaient eux aussi très différents (ils avaient un côté baba-cool), petit à petit, j'ai été obligé de changer mes habitudes. Moi aussi j'ai été tenté d'aller vers un peu plus d'« écologie » dans ma façon de travailler. De ne pas rechercher les effets. Donc j'ai diminué mes quantités de micros, considérablement. J'ai même tenté de renouer avec mes débuts, quand j'ai appris mon métier avec André Charlin ; je me suis retrouvé avec deux, quatre micros. Et alors, au lieu de déplacer les micros, et d'en rajouter, c'est les musiciens que l'on fait bouger... »(entretien de Jean-Marie Piel et Pierre Lavoix dans le Son des Musiques, p173)[5].

On voit que le style des musiciens change la méthodologie de travail. Ainsi ce changement ne vient pas seulement des progrès technologiques, mais aussi des mœurs, des attentes des musiciens. Et pourtant, dans les cas de Pierre Lavoix, l'équilibre instrumental existe lorsqu'il intervient. Ce n'est pas toujours le cas. Par exemple, l'apparition dans les basses continues d'instruments aux cordes pincées pose un nouveau problème : comment enregistrer le théorbe, la harpe, ou même parfois le clavecin, (instruments que l'on ne parvient pas toujours à entendre en concert) dans les grands forte de l'orchestre ? Et pourtant, il faut qu'ils apparaissent sur le disque, au même niveau que le pupitre de violons. Pourtant, on comprend facilement que les six violons ont une puissance d'intensité d'une autre dimension qu'un théorbe seul. Le microphone d'« appoint » pour cet instrument manquant de puissance devient indispensable. Et les autres instrumentistes en réclament aussi, ainsi peut grimper la surenchère de microphones. Tous les musiciens veulent leur microphone, dans la crainte que l'on ne les entende pas. Je ne dénonce pas ici une ignorance des musiciens, ou autre chose, seulement la nature humaine qui se veut ainsi. Chacun veut la même chose que son voisin, surtout lorsque dans l'écriture, les instruments ont une place égale.

L'introduction de microphones d'« appoint » vient également du fait que le couple ne suffit pas à distinguer toutes les sources, on l'a vu au dessus. L'équilibrage se fait alors au mixage, qui par cette présence d'« appoint » devient, comme dans les musiques actuelles, une véritable étape dans la création sonore. Les méthodologies ont beaucoup évolué à cette étape de la production du disque, pour les raisons que l'on a évoquées.

2.2 Cas judicieux de l'introduction d'un retard temporel

Avec l'habitude, le preneur de son connaît les cas où l'utilisation du retard temporel sur le signal d'« appoint » pourra être utile. Je souhaite évoquer ici quelques exemples concrets où le retard temporel est inutile, ou bien sert à autre chose que le mélange cohérent des signaux.

- Dans le mixage d'une prise de son d'un orchestre symphonique, des microphones placés trop loin du couple stéréophonique⁹ peuvent souffrir de l'introduction d'un retard temporel. Les signaux que l'on tente de mélanger ne sont plus corrélés. Le couple capte le champ diffus de ces sources lointaines, tandis que les microphones d'« appoint », placés en proximité des sources capteront le champ direct, de proximité. Le retard que l'on va introduire n'aura aucun effet sur d'éventuels filtres en peigne qui ici, comme le disait Daniel Mercier (cf. partie 1.3.4 page 31) dans son approche de la prise de son globale, n'existent pas.

- Si l'instrument enregistré a un spectre complexe, par exemple un piano, un cor ou encore un basson, l'intérêt d'une remise en phase entre les signaux d'« appoint » et ceux du couple est plus réduite. Ces instruments ont des timbres très complexes et composés d'un grand nombre d'harmoniques. La disparition d'une ou plusieurs fréquences gêne peu notre oreille dans la détection de leur timbre.

- On peut jouer avec les filtrages en peigne, et les utiliser comme de véritables filtres, plus continus en phase que les égaliseurs de nos logiciels de traitement sonore. Par exemple, pour atténuer le « loup »¹⁰ d'un violoncelle ou d'une contrebasse. Un retard temporel correspondant à cette fréquence du « loup », et non plus au décalage entre les deux microphones, permet de réaliser un filtre fréquentiel sans aucune distorsion de phase sur le signal d'origine. Ce filtre va s'appliquer également aux partielles et harmoniques de la fréquence filtrée. Dans le cas particulier du « loup » cela est plutôt avantageux, car les instruments concernés par ce problème résonnent par sympathie. Le « loup » pris une octave au-dessus aura également une sonorité désagréable. Ces filtrages par retard temporel ne peuvent donc pas être utilisés dans tous les cas de figure.

- En l'absence de retard temporel, l'introduction du signal d'« appoint » correspond souvent, à l'écoute, à un assèchement de notre mixage. La réverbération naturelle de notre couple « disparaît », au profit des informations de proximité amenées par le signal de

9. au delà de 10m dans une salle de concert avec une réverbération moyenne, en présence d'un public.

10. note qui résonne plus que les autres dans la caisse de résonance de l'instrument. On le trouve plus souvent pour les instruments à cordes graves à fréquences graves.

l'« appoint ». Lorsque l'on retarde un signal d'« appoint » dans un mixage, on va avoir la possibilité d'augmenter son niveau sonore dans le mélange, en ne modifiant pas cette impression de réverbération. Les informations dues à la proximité seront mieux ajoutées à l'espace réverbéré créé par le couple principal. Pour faciliter le mélange entre signaux d'« appoint » et signaux du couple principal, on introduit souvent sur l'ensemble du mélange une réverbération synthétique. Il peut apparaître que nous n'ayons plus besoin d'introduire cette réverbération numérique supplémentaire à notre mélange, lorsque nos signaux d'« appoint » sont retardés de manière cohérente.

2.3 Outils de mixage

C'est au mixage que l'on utilisera les outils de remise en phase. J'ai analysé les outils à ma disposition avant de faire un choix que je vais expliquer. Ces différents mixages me permettront de réaliser une enquête de perception auditive auprès d'un public averti en matière discographique, mais peu averti en matière de prise de son. C'est pourquoi j'ai fait le choix de n'utiliser que des outils d'intégration spatiale dans l'image stéréophonique et de ne faire varier que trois paramètres : le PanPot d'intensité (fixe), le niveau sonore, et le retard des signaux d'« appoint », sujet de mon mémoire, pour optimiser mes mixages selon mon goût personnel, proche de ce que l'on trouve dans le commerce. En effet, lorsqu'une source est ajoutée par un microphone d'« appoint », il faut qu'elle apparaisse dans une localisation logique par rapport à l'image créée par le couple, d'où l'utilisation du PanPot d'intensité, et ensuite, lorsque cette source est retardée, son niveau dans le mixage global peut augmenter sans créer d'artefacts désagréables.

Le meilleur juge pour détecter ou résoudre les problèmes de phase, entraînés par l'enregistrement d'une même source par plusieurs microphones, reste toujours et encore nos oreilles et notre goût pour le son. Cependant, la réalité physique du problème étudié ici peut s'analyser avec des outils « scientifiques » proposés et disponibles grâce à la technologie numérique.

2.3.1 Identification du problème de phase

Au début d'un mixage en multi-microphonie dans le cas qui nous préoccupe ici, on écoute d'abord le couple seul. Celui-ci peut manquer de précision dans les attaques, dans les modes de jeux des musiciens, certains instruments peuvent être déséquilibrés par rapport à d'autres, etc. Tous les problèmes que l'on a évoqués plus haut peuvent apparaître. Pour tacher d'y remédier, on ouvre les microphones d'« appoint »¹¹ des sources qui nous semblent manquer dans le mixage. Mixés à un certain niveau, les signaux d'« appoint » seront mieux « distingués », plus audibles, reconnus, repérés. La source risque de paraître décorrelée si on augmente encore son niveau dans le mixage. Il faut trouver le niveau d'intensité idéal où la sensation lors du passage sans « appoint » et avec « appoint » soit « douce »¹². Même à ce niveau idéal, certaines fréquences dans le grave ou certaines transitoires peuvent encore manquer. Le son

11. Les signaux obtenus peuvent subir des problèmes de caractérisation des sources : lorsque les sources voisines de la source que l'on cherche à enregistrer sont plus fortes en intensité que la source voulue. Par exemple dans une scène acoustique avec de nombreux instruments, on va chercher à isoler dans le signal d'« appoint » ce qui nous intéresse, ce travail est souvent une égalisation fréquentielle, pas toujours linéaire en phase. Ce travail permet d'éviter l'influence trop grande de certaines fréquences ou certaines sources dans l'ajout d'un signal d'« appoint » dans le mixage. Par exemple, lorsque la contrebasse et la clarinette installées à proximité, dans le sextuor de *Pierre et le loup*, jouent en même temps, dans l'« appoint » de la clarinette toutes les fréquences en dessous de 200Hz, sont des apports de la contrebasse. Avant de l'ajouter dans le mixage, on peut faire le choix de retirer ces fréquences graves.

12. par douce, je veux dire : linéaire, sans discontinuité ni brutalité. Les microphones d'« appoint » doivent apporter de nouvelles informations sonores sans modifier le confort de l'écoute du couple seul.

paraît : « troué » ou encore « mou ». Une mystérieuse disparition de certaines fréquences graves, ou encore des pics d'intensité sur certaines autres fréquences non souhaités dans les médiums peuvent rendre le son métallique. Le mélange entre l'« appoint » et le couple entraîne un filtrage en peigne, on l'a vu dans la première partie, qui induit ces effets. Si l'on n'est pas sûr que ce soit vraiment l'effet d'un filtrage en peigne, on peut, de nos jours, faire appel à des analyseurs de spectres. On l'a vu, le problème que l'on cherche à résoudre est d'ordre fréquentiel, il faut donc analyser l'apport de chaque bande de fréquence dans le résultat de notre mélange. On peut vouloir faire appel à ces analyseurs de spectre pour comprendre exactement quelles fréquences sont absentes ou surabondantes. Ces analyseurs de spectre réaliseront des diagrammes de fréquences pouvant prendre plusieurs formes (cf. partie 2, page 16). Dans le commerce, on les trouve sous la forme de plugins, de greffons logiciels, déjà présents dans les logiciels de traitement sonore avec lesquels on travaille ou que l'on ajoute aux logiciels de traitement sonore. On est capable de distinguer le contenu fréquentiel¹³ d'un signal oscillant grâce à la transformée de Fourier discrète, car le signal analysé ici est numérisé donc discontinu (cf. l'annexe D, page 88). De cette information, on réalise des graphes pour rendre lisible l'information obtenue. Il existe plusieurs manières de représenter la répartition fréquentielle d'un signal, en prenant en compte le fait que cette répartition varie dans le temps. Les outils que nous avons à notre disposition sont capables de faire une analyse en temps réel. Les graphes que l'on va montrer ici sont en réalité évolutifs dans le temps, et correspondent donc à l'apport fréquentiel d'un signal audio donné à un instant donné. La représentation la plus simple est la suivante :

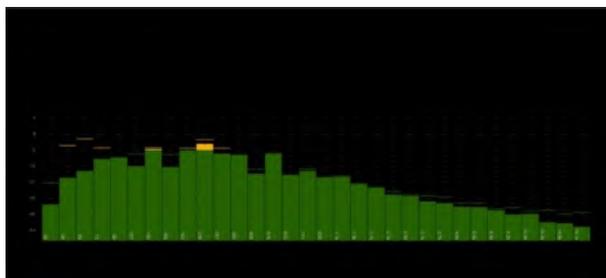


FIGURE 18 – Exemple d'un spectrogramme, ici par Samplitude Pro X.

On retrouve les bandes de fréquences en abscisses et l'intensité en ordonnées. Cette simple représentation a pour mérite d'être claire, mais elle est discontinue, en bande de fréquence.

Voici une représentation similaire mais linéaire en fréquence cette fois-ci :

¹³. ici on réalise la mesure de l'intensité de chaque fréquence dans un signal audio numérisé



FIGURE 19 – Exemple d'un spectrogramme, ici par Izotope

On retrouve les fréquences en abscisses et l'intensité en ordonnées. Mais cette fois, la représentation est graphique, continue en fréquences dans les abscisses. Mais cette représentation est évolutive dans les temps et ne permet qu'une lecture instantanée de l'apport fréquentiel du signal. Voici une représentation très proche, mais à trois dimensions : intensité, fréquence et temps.

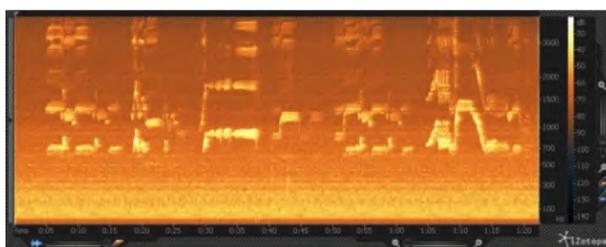


FIGURE 20 – Exemple d'un sonogramme, ici par Izotope

On trouve ici les fréquences en ordonnées, l'intensité en colorimétrie (plus le orange est clair plus l'intensité est grande) et le temps en ordonnées. Si l'on superpose cette représentation avec la « waveform ¹⁴ » du signal, on peut avoir ce genre de représentation :

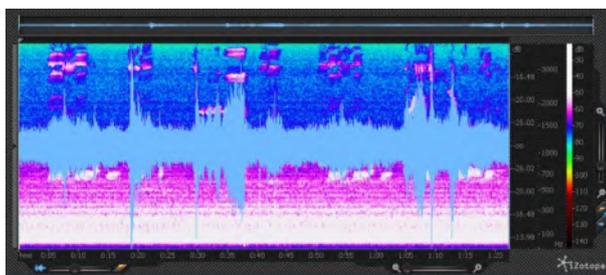


FIGURE 21 – Exemple d'un sonogramme , avec la waveform du signal, ici par Izotope

Ces représentations restent monophoniques. Il existe des outils pour mesurer la corrélation entre les apports fréquentiels des canaux droits et gauches.

14. Forme d'onde en anglais

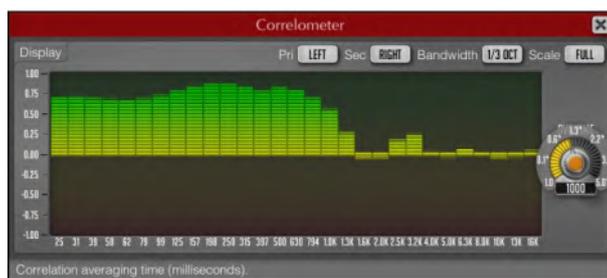


FIGURE 22 – Exemple d'un corrélateur-grame , ici par Voxengo

La représentation se fait de nouveau par bande de fréquences et en temps réel. L'information que l'on gagne par rapport aux autres représentations est que l'ordonnée représente la corrélation fréquentielle entre deux signaux (ici les canaux gauches et droits de notre mix stéréo). Si le graphe représente une barre au dessus de zéro (entre zéro et un), la bande de fréquence s'additionne (est en phase dans les deux signaux) en revanche si le graphe représente une barre en dessous du zéro (entre zéro et moins un), la bande de fréquence se soustraie (jusqu'à s'annuler quand la barre est à moins un) car les signaux sont hors phase. Si le graphe représente une barre à zéro cela signifie que le contenu fréquentiel, des deux signaux, à cette bande de fréquence est décorrélé. (cf. l'annexe D, page 88)

Le *Correlometer* du *PHA 979* de Voxengo est l'outil que j'ai utilisé, pour contrôler la corrélation des signaux par bande de fréquence et donc de potentielles modifications de timbre. (cf. partie 2.4, page 56)

Une autre compensation recherchée dans l'introduction des « appoints » peut être un rééquilibrage de l'espace stéréophonique. Par exemple si dans notre image, les sources situées à droite de l'image sont plus faibles que celles situées à gauche. Pour mesurer ces manques, il existe des outils d'analyses de niveaux de l'espace Stéréo comme le *Blue Cat's StereoScope Pro* :



FIGURE 23 – Exemple du *Blue Cat's StereoScope Pro*

La représentation de l'espace stéréophonique est ici linéaire. On peut en avoir une circulaire, comme dans le *StereoTools* de Flux :

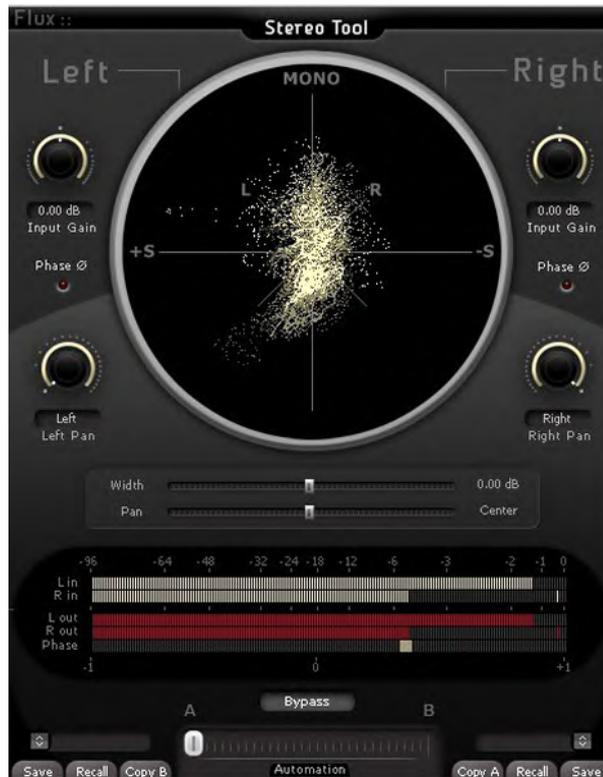


FIGURE 24 – Exemple du *StereoTools* de Flux

Avec ces outils, lorsque l'on ajoute les « appoints », « pannés », on peut mesurer l'influence en niveau de ces ajouts directement dans l'image stéréophonique.

2.3.2 Unités du retard entre deux signaux corrélés

On a vu dans le premier chapitre, que le déphasage de deux signaux correspondait à une distance physique dans le lieu de captation. La première unité de mesure de retard entre deux signaux corrélés est donc le **mètre**. Cette unité, comme l'on connaît la célérité du son dans l'air : $c = 340 \text{ m/s}$, peut être convertie en unité de temps : la **seconde**.

$$Temps(s) = \frac{Distance \text{ (en m)}}{c(m/s)} \quad (20)$$

La numérisation d'un signal se réalise par son échantillonnage dans une fréquence donnée, suivie par la quantification, sur un nombre défini de bits, de chaque échantillon du signal. Cet échantillonnage est une division temporelle du signal dans une unité beaucoup plus petite que la seconde, l'échantillon, le **sample** en anglais.¹⁵

$$Nombre\ des\ samples = \frac{Temps(s)}{Fréquence\ d'échantillonnage} \quad (21)$$

¹⁵. On choisit pour le reste du mémoire d'utiliser la nomination anglaise, car c'est ainsi que l'unité apparaît dans les outils utilisés.

Avec ces deux équations, on obtient la relation suivante :

$$Distance \text{ (en } m) = \frac{c(m/s)}{Distance(en m) * Fréquence \text{ d'échantillonnage}} \quad (22)$$

Cette dernière équation va nous permettre de faire le lien entre les différentes méthodes de calcul de retard entre les « appoints » et d'introduire des retards cohérents sur nos « appoints ». En effet, chaque outil fait le choix de l'unité la plus logique selon son raisonnement de fonctionnement.

2.3.3 Méthodes de calcul de retard entre « appoint » et couple

Le problème que l'on trouve est d'ordre fréquentiel. Pourquoi ne pas choisir de le résoudre par des égalisations sur le Master ? Ce n'est pas l'apport fréquentiel d'une seule source que l'on augmentera mais celui pour toutes les sources. Et faire cette égalisation sur le signal d'« appoint » ne modifiera que partiellement le problème du filtrage en peigne, en augmentant des fréquences qui ensuite dans l'addition seront annulées ou diminuées. Ce serait alors prendre le problème à l'envers, on le comprend. D'où l'utilisation qui doit être faite du retard temporel, cohérent avec la réalité acoustique et physique du problème que l'on rencontre. Avant de rajouter un retard temporel sur une quelconque piste, il faut que ce retard soit en cohérence avec la réalité acoustique à laquelle ce retard correspond. Pour cela, il y a des méthodes empiriques et des méthodes faisant appel à des logiciels de cross-corrélation. Nous allons en faire la liste ici :

Mesure physique

A la prise de son, le preneur de son peut directement mesurer, avec un mètre et un rapporteur, les distances et angles qui existent entre chacun des microphones. C'est une méthode simple et instinctive, mais très peu précise.

Mesure à la forme d'onde

Dans les logiciels de traitement sonore, le son est représenté par une forme d'onde. Il s'agit de la représentation graphique de l'amplitude de la pression d'air créée par l'onde sonore en ordonnée et le temps en abscisse. Si l'on réalise un son très court, comme un claquement de main, le premier front d'onde représenté dans tout logiciel de traitement de son est très facilement visible. Si on réalise ce son court et sec près d'un microphone d'« appoint », il devient facile de mesurer le temps que met cette onde pour parvenir à chaque microphone.

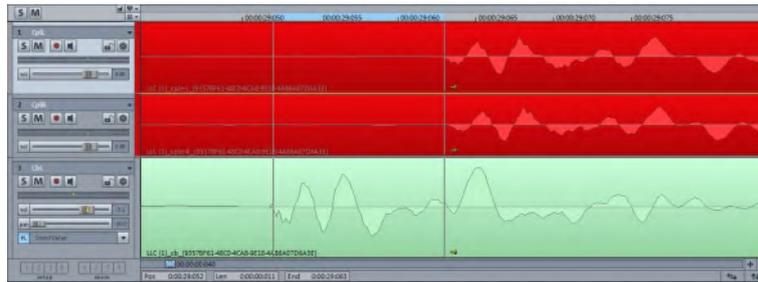


FIGURE 25 – Représentation graphique de trois signaux enregistrés par le biais d'un couple et d'un « appoint »

Dans cette représentation, les signaux du couple sont au-dessus et celui de l'« appoint » en dessous. Les pistes sont placées dans le temps, de sorte que l'on commence l'enregistrement en même temps sur les trois pistes. On constate que l'onde sonore est très proche entre les trois signaux, et parvient en premier au microphone d'« appoint ».

Mesure par cross-corrélation

On a vu dans le sous-chapitre sur l'identification des problèmes de phase, que la différence de phase entre deux signaux correspond à la corrélation entre ces deux signaux. Certains outils nous permettent de mesurer ces décalages de phase par un algorithme de cross-corrélation. On réalise des comparaisons du premier signal avec le second, qui lui est plus ou moins « décalé » dans le temps afin de trouver la valeur du retard qui permettrait une sommation en phase, et ce pour chaque bande de fréquence du signal. Comme les signaux analysés sont numérisés, donc discontinus, il suffit de faire la comparaison sur la valeur des échantillons et l'on obtient la valeur de l'écart entre les deux signaux. (cf. l'annexe D, page 88)

Certains logiciels utilisent cette méthode de calcul. Il est très intéressant de les utiliser pour avoir un résultat plus précis et cohérent que le résultat obtenu avec nos autres méthodes de calculs. Par exemple :

Auto-Align développé par Sound Radix ¹⁶

16. Comme j'ai utilisé ce logiciel comme outil de mesure, je le place dans ce sous-chapitre. J'ai conscience qu'il peut aussi s'utiliser comme un outil simple d'introduction de retard temporel.



FIGURE 26 – Exemple de deux signaux déphasés de 90° analysés par *Auto-Align* de Sound Radix

Dans l'exemple ci-dessus, on voit que les signaux sont déphasés de 90° . L'intensité des deux signaux comparés est affichée à gauche de la fenêtre. L'outil permet une visualisation graphique du déphasage, comme dans l'exemple, mais également du delay qui existe entre nos deux signaux. De plus cet outil affiche, en haut à droite, dans différentes unités (sample, msec, cm, pouces (unité anglo-saxonne de la distance)) du retard qu'il existe entre les deux signaux. Le logiciel propose d'introduire aussi directement ce résultat dans la piste concernée. Dans les mixages réalisés dans la partie pratique de ce mémoire, on choisira d'utiliser des outils plus simples pour introduire le retard calculé par *Auto-Align*. Il existe un autre logiciel, le *Track Align Pro* d'Audiowaves, qui réalise des opérations similaires mais avec beaucoup moins de visualisation sur leurs procédures¹⁷. Comme le fait *Auto-Align* de Sound Radix, le *Track Align Pro* d'Audiowaves introduit directement les retards qu'il calcule. Le fait de n'offrir aucune possibilité de réglage manuel m'a poussé à ne pas utiliser ces outils pour introduire des retards temporels dans mes mixages.

2.3.4 Outils d'introduction de retard

- Delays traditionnels

Dans tous les logiciels de traitement du son, sur les consoles numériques, il est possible d'introduire un retard temporel sur une piste audio. On

¹⁷. raison pour laquelle j'ai préféré utiliser *Auto-Align* au mixage de mes extraits

retarde le son d'une piste avant de l'envoyer dans le bus Master, c'est à dire notre écoute globale du mélange sonore. Ce retard peut s'exprimer en sample, en ms ou bien en pouce dans la majorité de ces logiciels. Il existe aussi différents greffons logiciels, plugins, qui remplissent cette fonction. Ces différents logiciels font office de « buffer »¹⁸ au signal. Ils le conservent quelques instants avant de le réintroduire dans le mélange tel quel. Aucune modification de phase, de fréquence ou de compression n'est réalisée. Ce processus simple a été rendu possible par le développement de l'informatique, où la fonction de mémoire préexiste dans toute la technologie numérique. Tout ordinateur fonctionne sur le principe de ces mémoires¹⁹ que l'on alimente et rappelle au moment voulu. Avant le numérique, des traitements analogiques, par électronique permettaient de retarder un signal audio par rapport à un autre. Ce retard analogique (électronique : à condensateur ou à transistor) était de très courte valeur et beaucoup plus difficilement manipulable et réglable que le retard numérique. Et si ce retard analogique était adéquat à l'utilisation que l'on veut en faire ici, le signal subissait une importante modification fréquentielle²⁰ à cause de la nature de l'électronique par lequel il devait passer pour être retardé. Ces « delays » ont surtout été utilisés en musique actuelle comme des effets. Ils permettent des effets d'échos ou de réverbération. Nous sommes très loin de la précision au sample près des outils actuels. De plus, la numérisation des signaux a permis une reproductibilité des signaux en grande quantité sans aucune modification de l'information.

Prenons l'exemple du delay proposés par Voxengo, utilisé dans les mixages de la partie pratique de ce mémoire :



18. « mémoire tampon » en anglais

19. mémoire vive RAM et mémoire morte ROM

20. coupe bande, l'utilisation de l'électronique entraîne une couleur sonore éloignée de notre approche transparente du son introduite grâce au numérique.

FIGURE 27 – Exemple d'utilisation du *Sound Delay* de Voxengo

Deux retards nous sont proposés dans ce logiciel, comme dans la majorité : au-dessus en ms, mètre ou « foot » qui est moins précis mais peut se calculer plus facilement par les mesures que l'on a vues au-dessus, ou bien en-dessous en sample, très précis, car on est au plus près de la précision du signal numérique, mais demande d'être mesuré par des outils plus complexes.

Dans ce mémoire, les retards de nos sources ont été mesurés en sample avec l'outil *Auto-Align* et introduit dans le mixage avec le *Sound Delay* de Voxengo.

Ces outils proposent d'introduire un unique retard temporel sur une piste audio quel que soit la spatialisation panoramique²¹ choisie. Mais il arrive que le retard temporel entre notre « appoint » et le microphone gauche de notre couple soit différent du retard temporel entre notre « appoint » et le microphone droit de notre couple²². Pour introduire ces différents retards temporels, on « ruse » : on duplique dans le mixeur les tranches de contrôle de nos pistes des « appoint »²³. Comme l'apport en intensité de notre piste est doublée, on retire 3dB sur notre piste audio²⁴. Chacune de ces tranches est « pannée » réciproquement complètement à gauche et à droite. On peut alors introduire un retard temporel différent sur chacune de ces pistes, retardant le signal de manière cohérente et différente à gauche et à droite de notre stéréophonie. La spatialisation en différence d'intensité se fait par la différence de niveau sur les nouveaux canaux gauche et droit de chacun de nos signaux d'« appoint ». On peut également contrôler ce paramètre (le niveau dans le mélange) en liant les pistes ou bien en envoyant ces deux tranches dans un bus stéréo (que l'on envoie ensuite dans le mélange final, stéréo).

C'est cette dernière méthode que j'ai utilisée dans les mixages de ma partie pratique. Cependant, pour éviter toutes ces manipulations, Merging Technologies a mis au point un greffon logiciel (plugin), le PanNoir.

- *PanNoir*, de Merging Technologies

Dans la dernière version du logiciel de traitement sonore de Merging Technologies, Pyramix 9, on trouve un nouvel outil d'introduction de retard temporel : le PanNoir. Cet outil est une commande de l'ingénieur du son Jean-Daniel Noir, d'où l'origine de son nom, pour faciliter l'intégration de retard temporel sur des signaux d'« appoint » mono ou stéréo par rapport à un signal principal stéréophonique. La société 44.1 a eu le gentillesse de me prêter une autorisation pour utiliser la dernière version dans Pyramix 9.

21. sous-entendu d'intensité

22. dans le cas des couples stéréophoniques à base de différences temporelles (cf. paragraphe 1.3.2, page 23, et paragraphe 1.3.2, page 25)

23. Dans un logiciel de traitement sonore, chaque piste de signal audio est dirigé par une tranche de console virtuelle (numérique)

24. Cette modification de niveau sonore peut se faire directement sur le signal et ne pas apparaître dans notre mixeur.

Le principe est de simplifier l'introduction d'un retard différent dans les canaux gauche et droit de notre stéréo. L'outil s'utilise comme un plugin que l'on place dans le mixeur sur la piste qu'il nous intéresse de retarder. Une fois mis en route, il se présente sous la forme suivante :



FIGURE 28 – Exemple d'utilisation du *PanNoir* de Pyramix

En haut de l'outil se trouve le « spatialisateur » qui permet de choisir l'angle²⁵, la distance, et la divergence de notre signal d'« appoint » par rapport au couple principal. Une représentation graphique²⁶ permet une visualisation de ce choix.

25. Cet angle permet un calcul de distance entre l'« appoint » et les deux signaux du couple et ne modifie pas l'intensité du signal dans les canaux gauche ou droit.

26. en double-cliquant sur le graphe, on peut l'agrandir pour gagner en précision dans le placement de notre source.

Plus bas dans l'outil, se trouve la « définition » de notre couple principal : choix de la directivité, de la distance et de l'angle entre les microphones. Au dessus, se trouve la visualisation de ce couple par une représentation de leur directivité sur une règle.

Entre ces deux outils se trouve une grille de « Damping »²⁷ qui permet un choix d'intensité de notre signal retardé. Enfin, tout en bas de l'outil, on trouve deux potentiomètres de gain : un pour chaque canal de la stéréo. Ces gains permettent de réaliser un PanPot d'intensité. Ils peuvent être réglés de manière indépendante, ou reliés²⁸entre eux suivant la règle classique du panoramique d'intensité.

Pour introduire un retard cohérent avec cet outil, l'unique possibilité est la mesure physique de l'écart des microphones à l'enregistrement. Il ne propose aucun algorithme complexe de cross-corrélation pour calculer le retard qui existe réellement entre nos signaux oscillants numérisés. C'est le plus grand défaut de cet outil, et c'est la raison pour laquelle je n'ai pas utilisé cet outil dans ma partie pratique.

L'avantage du PanNoir, est qu'il évite la duplication des pistes au cours du mixage si l'on souhaite introduire des retards différents dans les canaux droits et gauches de la stéréo. Cet outil fait office de PanPot d'intensité et d'introduction de retard de manière cohérente en fonction du lieu physique où l'on souhaite placer le signal d'« appoint » dans l'image stéréophonique.

27. « Atténuation » en anglais

28. bouton jaune entre les deux potentiomètres.

2.4 Mixages de l'expérience

Le mixage de manière générale fait appel à cinq outils différents dont pour certains nous avons peu parlé dans ce mémoire.

- le niveau d'intensité
- le PanPot, spatialisation en intensité
- l'égalisation fréquentielle
- la compression
- la réverbération

J'ai fait le choix de ne pas utiliser ces trois derniers outils, car ils entraînent une modification des caractéristiques des signaux d'« appoint » enregistrés lors de la prise de son. Dans ce mémoire, je souhaite montrer l'importance du retard temporel introduit sur les signaux d'« appoint ». A cette fin, il était important de modifier le moins possible les signaux retardés.

Ce retard temporel, delay, est le nouvel outil apporté par la technologie numérique. Ce sont les avantages et les limites apportées par cet outil, qui peut remplacer parfois, on l'a vu, une égalisation fréquentielle, ou une réverbération synthétique, qui seront éprouvées ici.

2.4.1 Choix des outils

J'ai justifié la majorité de mes choix dans les paragraphes précédents. Je vais rappeler les outils que j'ai utilisés avant d'évoquer la manière dont j'ai réalisé mes mixages en prévision de mes tests perceptifs.

Le logiciel de traitement sonore que j'ai décidé d'utiliser pour mixer mes enregistrements est Sequoia 12. J'ai fait ces mixages à l'aide de mon casque, K211 MkII de chez AKG. Pour être sûre de mes mixages, je suis allée vérifier les résultats deux fois sur des systèmes de monitoring avec enceintes. La première écoute m'a permis de détecter les défauts de mixage et la seconde de vérifier que j'avais réalisé les bonnes modifications²⁹.

Pour calculer les retards temporels que je souhaitais introduire, j'ai utilisé le logiciel *Auto-Align* de Sound Radix quand je l'ai pu. Ce logiciel ne peut calculer les retards nécessaires seulement si les microphones enregistrent une source unique. Dans les cas du quintette et de l'orchestre symphonique, où je n'avais pas d'enregistrement des sources prises individuellement, il m'a été impossible d'utiliser cette méthode pour mesurer les retards temporels que je devais introduire. J'ai donc réalisé des mesures à la « waveform » pour l'orchestre symphonique et j'ai fait confiance à mes mesures métriques pour le quintette à vent.

Pour introduire les retards temporels dans les pistes de mon mixeur, j'ai

²⁹. Celles-ci étaient des erreurs de niveaux de mes signaux d'« appoint » par rapport à l'importance de la source dans le mélange.

utilisé le plugin *Sound Delay* de Voxengo (cf. paragraphe 2.3.4, page 51). J'ai réalisé, dans mes différents mixages, la spatialisation d'intensité à l'aide de deux tranches de consoles, spatialisées à gauche et à droite, ayant des niveaux différents. (cf. paragraphe 2.3.4, page 53) Ce fut la première étape réalisée dans mon processus de mixage.

Pour mesurer la corrélation fréquentielle entre mes signaux droits et gauches, j'ai utilisé l'outil de mesure proposé par Voxengo dans le plugin *Correlometer* du *PHA 979* (cf. paragraphe 2.3.1, page 47).

De toutes les prises de son, j'ai sélectionné cinq extraits différents (après des montages, en accord avec les performances et la volonté des musiciens, grâce aux différentes prises faites lors des sessions d'enregistrement) :

- L'improvisation de jazz par la contrebasse seule

- Le thème du « loup » de *Pierre et le loup* par le cor, le basson et la clarinette du quintette.

- Le thème de l' « oiseau » de *Pierre et le loup* par la flûte, le hautbois, la clarinette et la contrebasse du quintette.

- Le début de l' « Adagio » de la 5^{ème} *symphonie* de Gustav Mahler par l'orchestre du philharmonique du COGE.

- Le début du « 5^{ème} mouvement » de la 5^{ème} *symphonie* de Gustav Mahler par l'orchestre du philharmonique du COGE.

Ces extraits ont des durées de 19 secondes à 2 minutes 30. Il m'a semblé important de limiter la durée des extraits car notre mémoire auditive est limitée. Pour pouvoir comparer différentes versions, il faut se souvenir de toutes. De tous mes extraits, j'ai réalisé deux ou trois mixages différents. Le premier mixage est réalisé avec des « appoint » sans aucun retard temporel. Leur niveau dans le mélange est donc moins important que dans les autres mixages.

Le second mixage est réalisé avec le même retard temporel introduit à gauche et à droite sur mes deux pistes d'« appoint ». Ce retard est une moyenne de mon calcul de retard entre les canaux gauche et droit. Avec ce nouveau retard, j'ai pu augmenter le niveau de mes signaux d'« appoint ».

Mon troisième et dernier mixage est réalisé avec des retards différents introduits à gauche et à droite sur mes pistes d'« appoint ». J'ai proposé à mes auditeurs ce dernier type de mixage uniquement pour l'extrait de la contrebasse seule, car les différences, lorsque je faisais cette opération pour les autres extraits étaient à la limite du perceptible.

2.4.2 Mixage de la contrebasse

L'extrait de la contrebasse seule dure 56 secondes. Il est composé de pizz et de percussions (réalisées sur la table d'harmonie de l'instrument). L'enregistrement, puis le mixage ont été réalisés à une fréquence d'échantillonnage de 48kHz et une quantification de 24 bits.

Le premier mixage a été réalisé en mélangeant les trois pistes, sans introduire de retard sur la piste d'« appoint » de la contrebasse. La contrebasse étant à 10° sur la gauche, je l'ai « panné » en conséquence. Ma volonté a été de favoriser le côté lointain de la contrebasse. L'« appoint » est donc à un niveau relativement faible, avec un fader à -9dB.

Nous sommes en présence d'une source unique, j'ai fait appel à *Auto-Align* de Sound Radix pour mesurer les retards que je souhaitais introduire au signal d'« appoint » mono. Voici le résultat obtenu, en samples :

gauche ³⁰	droite ³¹
536	547

Le second mixage a été réalisé en introduisant un retard moyen calculé à partir de mes résultats de mesure avec *Auto-Align*. C'est-à-dire un retard de 541 sample introduit dans les canaux gauche et droit de mon mélange. Avec ce retard, j'ai eu la possibilité d'augmenter de 3dB le niveau de mon signal d'« appoint », avec un fader à -7dB.



FIGURE 29 – Mixeur après l'introduction d'un retard de 541 samples dans les canaux gauche et droit du mélange de l'« appoint », dans l'extrait de la contrebasse.

Le troisième mixage a été réalisé en introduisant les deux retards mesurés par *Auto-Align*. Je n'ai pas pu augmenter le niveau de mon signal d'« appoint ». On mesure bien un gain de corrélation entre l'« appoint » et le couple, à gauche et à droite.

30. retard mesuré entre le signal d'« appoint » et le canal gauche du couple stéréo
31. retard mesuré entre le signal d'« appoint » et le canal droit du couple stéréo



FIGURE 30 – Mesure de la corrélation entre couple et « appoint » (ici à droite), après l'introduction d'un retard temporel dédié à gauche et à droite sur le signal d'« appoint », dans l'extrait de la contrebasse.

2.4.3 Mixage du quintette et contrebasse

Les extraits du « loup » et de l'« oiseau » de *Pierre et le loup* durent respectivement 19 et 27 secondes. L'enregistrement, puis le mixage ont été réalisés à une fréquence d'échantillonnage de 48kHz et une quantification de 24 bits. Les deux mixages ont été réalisés avec l'ouverture des microphones d'« appoint » des instruments jouant dans les extraits. Le premier mixage a été réalisé sans introduire de retard temporel aux signaux d'« appoint ». J'ai « panné » les signaux d'« appoint » en conséquence de leur positionnement dans l'image stéréophonique créée par le couple principal³².



FIGURE 31 – Mixage sans introduction de retards, de l'extrait du thème du « loup » de *Pierre et le loup*

Le second mixage a été réalisé en introduisant un retard moyen calculé à partir des mesures physiques réalisées au cours de l'enregistrement et

³². Selon leur implantation physique lors de l'enregistrement (cf. paragraphe 1.4.2, page 35)

des formules mathématiques adéquates³³. Voici les retards calculés et le gain du fader pour le mélange :

Source	Retard (en sample)	gain du fader (en dB)
Basson	550	-5
Cor	530	-11
Clarinette	535	-1
Flute	555	-9
Hautbois	555	-10
Contrebasse	541	-9



FIGURE 32 – Mixage avec l’introduction des retards, de l’extrait du thème de l’« oiseau » de *Pierre et le loup*

2.4.4 Mixage de l’orchestre symphonique

Les extraits de la 5^{ème} *symphonie* durent environ 1 minutes 30. L’enregistrement, puis le mixage ont été réalisés à une fréquence d’échantillonnage de 44.1kHz et une quantification de 24 bits. Les deux mixages ont été réalisés avec l’ouverture des microphones d’« appoint » des instruments jouant dans les extraits. C’est-à-dire la fermeture des signaux d’« appoint » dans l’extrait de l’*Adagio*, où seules les cordes jouent. Le premier mixage a été réalisé sans introduire de retard temporel à mes signaux d’« appoint ». J’ai « panné » les signaux d’« appoint » en conséquence de leur positionnement dans l’image stéréophonique créée par le couple principal³⁴. Le niveau sonore des signaux d’« appoint » était relativement faible.

33. formules exprimées dans la deuxième partie (cf. partie 2.3.2, page 48

34. Selon leur implantation physique lors de l’enregistrement (cf. paragraphe 1.4.3, page 48).



FIGURE 33 – Mesure de la corrélation entre le signal d’« appoint » des violons 1 et le canal gauche du couple, sans l’introduction d’un retard temporel sur les signaux d’« appoint », dans l’extrait du 5^e mouvement.

Le second mixage a été réalisé en introduisant un retard moyen mesuré à partir des informations relevées sur les « waveforms » (cf. paragraphe 2.3.3, page 49). Voici les retards mesurés et le gain du fader pour le mélange :

Source	Retard (en sample)	niveau du fader (en dB)
Violon 1	272	0
Violon 2	272	0
Alto	379	0
Violoncelle	379	-3
Contrebasse	749	-9
Harpe	778	-5
Couple Bois	584	0
Clarinette	758	-1

Avec l’introduction de ce retard, le niveau sonore de mes « appoints » a pu augmenter, et la corrélation entre les signaux du couple et les signaux d’« appoint » est devenue positive.



FIGURE 34 – Mesure de la corrélation entre le signal d’« appoint » des violons 1 et le canal gauche du couple, après l’introduction d’un retard temporel sur le signal d’« appoint », dans l’extrait de l’*Adagio*.

2.4.5 Mesures selon la norme Loudness EBU R128

Afin de connaître la cohérence en niveau de mes différents mixages, j’ai réalisé une mesure de mes mixages selon la norme Loudness EBU R128. Cette norme permet une mesure du niveau global d’un extrait audio, ainsi que la mesure de son niveau sonore le plus élevé (« peak » en anglais). Voici les résultats de ma mesure :

nom de l’extrait	niveau moyen (en dB _{LUFs})	« peak » (en dB _{FS})
Contrebasse sans retard	-32,6	-9,8
Contrebasse avec le même retard	-32,7	-7,1
Contrebasse avec des retards différents	- 32,4	-7,4
« loup » sans retard	-17,4	-2,3
« loup » avec le même retard	-17,5	-2,5
« oiseau » sans retard	-22,9	-7,5
« oiseau » avec le même retard	-23,1	-7,8
« Adagio » sans retard	-38,9	-19,9
« Adagio » avec le même retard	-38,2	-17,5
« 5 ^e mouvement » sans retard	-32,0	-11,9
« 5 ^e mouvement » avec le même retard	-31,7	-10,7

A la vue de ces résultats, on constate que l’ajout d’un retard, et l’augmentation du niveau sonore dans notre mixage n’a pas entraîné une augmentation importante du niveau sonore global de nos mixages. Ces mesures m’ont permis de régler le volume du système de restitution des sujets interrogés, afin de leur proposer une écoute cohérente, d’un extrait à l’autre.

3 Enquête de perception

Dans un article de l'*Audiophile* daté de 1994, Claude Bailblé nous dit : « Dans chaque canal sensoriel, il y a une composante « affective » et une composante « informative ». » C'est donc autant l'aspect affectif qu'informatif qui va m'intéresser dans cette étude, dans laquelle mes questions s'orienteront : vers le coté informatif de l'écoute, en comprenant ce que mes sujets entendent ; et le coté affectif, en demandant une préférence dans les extraits écoutés.

3.1 Sujets de l'expérience

Un audiophile est « un amateur de perceptions auditives, intéressé par les sensations qu'elles produisent, à la recherche des émotions qu'elles induisent » selon Francis Ibre. [6]

Ce comportement est souvent associé à l'écoute d'un répertoire « sérieux » : la musique classique ou la musique jazz. Ce choix personnel s'illustre dans les directions « marketing » de production musicale depuis le début de la stéréophonie. D'où le fait que « les premiers 33 tours stéréophoniques édités se cantonneront à des genres « respectables » : musique classique... ou lorsqu'on aborde le monde de la variété, des artistes comme Frank Sinatra. L'explication est d'ailleurs fort simple. Pour de strictes raisons économiques, peu de gens peuvent se permettre un coûteux équipement avec deux hauts-parleurs (une révolution!) »[7].

Cette réalité économique est toujours d'actualité. Les systèmes d'écoutes sont encore plus variés aujourd'hui qu'à l'époque des premiers disques stéréo. On écoute la musique dans sa voiture avec l'auto-radio, en travaillant grâce aux enceintes de nos ordinateurs, dans la rue grâce aux baladeurs et les petits écouteurs qui les accompagnent souvent, etc (cf. partie 1.3.4, page 31). A l'inverse de ces tendances, certaines sociétés de distribution, marquent leur différence par un accès à des supports de meilleure qualité. Ces nouveaux supports sont justifiés par la volonté de transparence du son. On recherche à faire oublier à l'auditeur le support de reproduction, et comme au début de l'enregistrement, l'auditeur veut être au concert auquel il n'assiste pas physiquement. Il y a donc deux approches de l'auditeur par rapport à l'audio. Le premier auditeur recherche le contenu musical lorsqu'il écoute un disque, il s'agit de l'amateur de musique. Le second recherche une sensation, une performance, il s'agit de l'audiophile. Au cours de mes entretiens, j'ai réalisé que l'amateur de musique est conscient des défauts de son système d'écoute. Par exemple, un sujet m'a indiqué que sa chaîne avait tendance à favoriser les fréquences aiguës.

3.1.1 Critères d'écoute

Dans ce même livre, Francis Ibre nous parle des différents degrés d'écoute qu'un auditeur doit avoir.

1. reconnaissance d'un air connu : ce premier degré d'écoute est le plus courant chez les auditeurs.

2. notion d'espace sonore
3. la matière sonore : « *des différents matériaux, de leur résonance propre, de leur grain, leur dureté, leur souplesse, ainsi que la taille de la source, de la place qu'elle occupe, de l'énergie qu'elle développe et qu'elle rayonne.* »(Francis Ibre) [6]
4. le mouvement : « *chargé d'intention, cause première de la vibration entendue. A la pointe, au sommet de ce niveau, se situerait la perception de l'intention pure : d'infimes détails, de petits bruits émergeant du silence, révèlent la présence silencieuse de l'artiste qui s'apprête à placer la première note* »(Francis Ibre) [6]

Les personnes que je suis allée interroger étaient toutes conscientes de ces critères. Et c'est dans cet ordre d'analyse que la majorité des auditeurs m'ont répondu. Leur analyse s'est porté d'abord sur le discours musical et sa compréhension. Puis, au regard de la nature des différences entre les versions de mixage que je leur demandais de comparer, leur analyse s'est orientée vers les critères d'« espace sonore », de « matière sonore » et de « mouvement » décrits par Francis Ibre. Tous ces amateurs que je suis allée interroger m'ont parlé, à la fin de l'entretien, du plaisir de l'écoute concentrée, qui permet l'accès à ces critères d'écoute.

3.1.2 Description du public interrogé

Le public que j'ai interrogé est hétérogène. Il est composé de personnes de tous âges, des deux genres, et qui pratiquent des métiers très différents. Cependant, mon choix s'est porté vers des personnes ayant une grande habitude de l'écoute musicale sur disque. Pour m'approcher de leurs écoute habituelle, j'ai réalisé les tests sur le propre système d'écoute de mes sujets. Le but de cette démarche est d'estimer d'un point de vue qualitatif l'influence de notre manipulation sonore³⁵ sur l'écoute d'un public averti. Comme je ne fais pas une étude statistique, mais une étude sur le goût, mes sujets ne sont pas contraints à réaliser le test d'écoute dans des conditions physiquement similaires. La similitude se trouve dans l'habitude et la connaissance qu'ils ont du système d'écoute sur lequel à lieu le test d'écoute. Cette étude a été réalisée au près de douze auditeurs différents, sur douze systèmes de restitution. Ces systèmes sont toujours placés dans le salon de mes sujets. Il est de différent ordre de qualité. Certaines personnes avaient des chaînes Hi-Fi standard, d'autres des systèmes très perfectionnés avec lecteur, amplificateur, et convertisseur séparés. Comme je fais ici l'analyse d'une écoute habituelle pour l'auditeur, j'ai considéré que ces différences de systèmes de reproduction étaient un atout dans le cadre de cette enquête. Lors de la production du disque, le mixeur considère que son travail doit pouvoir se traduire quelque soit le système de diffusion que possède son auditeur. J'ai soumis mes mixages à la même épreuve. L'homogénéité des réponses de mes sujets m'a prouvé que leur analyse ne dépendait pas de leur système d'écoute.

35. l'introduction d'un retard temporel sur les signaux d'« appoint ».

3.2 Méthode d'enquête

Dans cette enquête, je me suis déplacée chez mes sujets avec un CD-audio où j'avais gravé les extraits audio que je voulais leur faire écouter. Ces CD audio étaient gravés avec l'ordre suivant des extraits audio :

1. Contrebasse seule, extrait de jazz (version 1, 2, 3)
2. le trio cor, basson, clarinette, extrait du thème du « loup » de *Pierre et le loup* de Prokofieff (version 1, 2, 3)
3. le quatuor flute, clarinette, contrebasse, hautbois, extrait du thème de l'« oiseau » de *Pierre et le loup* de Prokofieff (version 1, 2, 3)
4. l'orchestre à cordes avec harpe, extrait de l'*Adagio* de la 5^e symphonie de Gustav Malher (version 1, 2)
5. l'orchestre symphonique, extrait du 5^e mouvement de la 5^e symphonie de Gustav Malher (version 1, 2)

J'ai choisi cet ordre, car il correspond à une augmentation graduelle du nombre de sources sonores. Le nombre de sources sonores, et la nature de ces sources modifient notre perception des éventuels filtres en peigne.

Pour chaque extrait, j'avais réalisé deux ou trois versions : avec ou sans retard temporel (différent ou similaire dans les canaux gauche et droit). Chaque extrait était la succession de ces différentes versions, placées sur différentes pistes pour permettre un accès plus facile à chaque version et à chaque extrait. Dans le cas des mixages des extraits de *Pierre et le loup*, j'ai répété le mixage avec l'introduction de retard. Cette répétition m'a permis de savoir si les distinctions entendues par mes sujets étaient dues à un phénomène psycho-acoustique ou bien réellement à la perception d'une différence. Je n'ai pas voulu faire cette répétition pour les extraits orchestraux, car les différences m'ont semblé plus perceptibles et que les extraits étaient plus long.

J'ai gravé différents disques où l'ordre des mixages variait afin de savoir si ce n'était pas l'ordre dans lequel on entendait les versions de mixage qui déterminait les remarques faites ensuite. En regardant les résultats, on peut conclure que cet ordre n'affecte en rien cette perception.

Pour chaque extrait, j'ai demandé à mon sujet d'écouter sans interruption les différentes versions que je lui proposais d'entendre. Après cette première écoute, je demandais à mon sujet les différences qu'il avait pu percevoir et à quelle version il accordait sa préférence. Ensuite, une seconde écoute des extraits débutait sous la direction de cet auditeur. Grâce à la télécommande de son système Hi-Fi, ou sur le lecteur de l'ordinateur, il avait accès aux différentes versions et aux différents passages qu'il souhaitait écouter. A la suite de cette écoute plus « libre », je reposais mes questions sur la nature des différences perçues. Pour orienter mon auditeur vers les problèmes qui me préoccupent dans ce mémoire, je terminais l'entretien en questionnant sa perception du timbre des instruments et de la spatialisation des sources (proches ou lointaines).

Mon étude s'est portée sur le jugement de goût de mon panel de sujet. Dans ce sens, il m'a semblé important de sortir de l'étude statistique. En effet, j'ai souhaité étudier les valeurs stylistiques de mes sujets, ce qu'ils appellent le beau en matière de « son ». Pour ne pas influencer leur critères de jugements, je ne leur ai pas donné à remplir un questionnaire à réponses multiples. A l'inverse, je leur ai posé la question simple à laquelle je voulais qu'ils répondent : « qu'entendez vous de différents entre les trois versions? Quelle versions préférez vous? ». Cela me permettait d'entamer une conversation plus large sur leur perception et attentes en matière de « son » sur un disque de musique classique ou jazz. J'ai pu aussi étudier les termes qu'ils attribuaient à certains effets qu'ils entendaient.

3.3 Analyse des résultats

3.3.1 Analyse de la distinction de différences

La première question que j'ai posé au cours de ces entretiens était la perception de la différence entre les extraits. Mes sujets ont perçus plus ou moins de différences entre les versions, suivant la nature des extraits.

Pour l'extrait de la contrebasse, la différence entre les trois versions était peu perceptible. Beaucoup de sujets n'entendaient pas les changements entre les mixages. Ce problème de distinction des versions se comprend par le caractère non mélodique de l'improvisation réalisée. Lors de la seconde écoute, avec une habitude prise pour le contenu musical, les sujets devenaient plus à même de percevoir des différences entre la version sans introduction de retard temporel sur les signaux d'« appoint »³⁶ avec les deux versions où un retard temporel sur les signaux d'« appoint » avait été introduit. La distinction entre la version où le retard introduit était similaire pour les canaux gauche et droit³⁷ et la version où le retard introduit était différent pour les canaux gauche et droit³⁸ était difficile à faire, voir impossible pour les sujets interrogés. Leur préférence s'est toujours portée vers les versions avec retard, par rapport à la version sans retard, après plusieurs écoutes.

Pour l'extrait du thème du « loup » la distinction entre les versions avec et sans retard n'est pas uniforme. Quatre auditeurs n'ont pas perçu de différences entre les trois versions proposées (deux versions avec retard et une sans retard) à la première écoute. La distinction entre les versions était perceptible surtout au début de l'extrait, selon tous mes sujets. La version avec retard permettait une plus grande distinction entre les trois instruments, tandis que la version sans retard était plus homogène d'un point de vue des timbres. Cette homogénéité a été analysée comme une prédominance du cor par rapport à la clarinette et au basson. Leurs timbres semblaient disparaître au profit de celui du cor. Aucun auditeur n'a perçu de différence spatiale entre ces versions.

Pour l'extrait du thème de l'« oiseau » les différences perçues ont varié en fonction des sujets. Les non-musiciens ne perçoivent que peu de différences, voire aucune. Les musiciens à instruments aiguës (flute, violon) ont perçu des différences de timbre et d'espace sur les instruments à

36. dans la suite de cette analyse, nous parlerons pour cette « version sans introduction de retard temporel sur les signaux d'« appoint » » de « version sans retard »

37. dans la suite de cette analyse, nous parlerons pour cette « version avec introduction d'un retard temporel similaire pour les canaux gauche et droit » de « version avec même retard »

38. dans la suite de cette analyse, nous parlerons pour cette « version avec introduction d'un retard temporel différent dans les canaux gauche et droit » de « version avec différents retards »

vents. Tandis que les musiciens à instruments graves (contrebasse, trombone) ou à percussions ont distingués des différences sur la contrebasse entre les versions. Curieusement, de manière générale, la version sans retard a été perçue plus lente en tempo que la version avec retard.

Pour l'extrait de l'*Adagio*, les différences perçues ont été immédiates pour la majorité des sujets. La version sans retard a été perçue lointaine, et plate dans les nuances. Tous les instruments ont paru être sur un plan sonore équivalent quelque soit leur rôle musical. La version avec retard a été perçue plus proche et dotée d'une grande image stéréophonique. La remarque la plus fréquente a été la perception d'un suivi de voix musical dans la version avec retard. Les parties solistes semblaient ressortir alternativement lors de leurs différentes interventions. (passage du thème des violoncelles aux violons, etc) De plus, la notion de pupitre de cordes semble apparaître dans la version avec retard, tandis que dans la version sans retard les auditeurs ont perçus le chef d'attaque de chaque pupitre plutôt que le pupitre entier. Cette impression a été entendue surtout pour le pupitre des violoncelles et des contrebasses. La harpe a été entendue lointaine dans la version sans retard.

Pour l'extrait du 5^e mouvement de la 5^e symphonie, les différences ont été moins grandes que pour l'*Adagio*. Elles sont plus centrées sur la perception du timbre des instruments. Les cordes paraissent plus rondes dans la version avec retard. Les fréquences graves des contrebasses et violoncelles sont plus précises dans la version avec retard. Les vents, eux paraissent sonner de manière plus aiguë dans la version avec retard. La sensation spatiale change peu entre les deux versions. J'ai eu souvent une critique, valable pour mes deux versions, d'un manque d'unité spatiale entre les cordes et les vents. Cela montre la limite de notre système d'intégration des signaux d'« appoint ».

3.3.2 Analyse du vocabulaire employé

Je vais présenter, et résumer dans un tableau les adjectifs qui ont été utilisés par les audiophiles et amateurs de musique interrogés pour qualifier leur perception sonore dans les extraits.

Nom de l'extrait	version sans retard	version avec retard
Contrebasse	lointaine ; son mat ; réverbéré ; son voilé	propre ; présence ; proximité ; grain dans les attaques ; percussif ; timbre plus clair ; chaud ; coloré ; agressif ; violent
Thème du « loup »	plat ; fondu niveau sonore faible	naturel ; distinction des instruments ; homogénéité ; vivant ; plus de présence harmonique ; plus de fréquences graves
Thème de l'« oiseau »	chirurgical ; équilibré distinction des instruments à vent plat musicalement	saturation dans les basses ; instruments à vents plus fondu ; lié ; évolution dans la nuance ; contrebasse distincte et claire
Extrait de l' <i>Adagio</i>	lointaine ; réverbéré pauvre ; plat ; écrasé	précision des timbres ; proximité ; homogénéité ; expressif ; gras espace idéal
Extrait du 5 ^e mouvement	trop fondu ; fouillis ; lointain pauvre ; plat ; écrasé	brillant ; homogénéité ; expressif ; équilibre ; gras

Dans cette analyse de vocabulaire, j'ai inclus les termes musicaux qui ont été évoqués. En effet, de manière générale, les auditeurs ont réalisé également une analyse musicale des extraits sonores. Les versions sans retard ont été perçues plus plates en nuances et en expression musicale. Tandis que les versions avec retard semblaient révéler à l'auditeur toutes les inflexions musicales créées par les musiciens. Pour tous les instruments à cordes, les versions avec retard faisaient apparaître l'utilisation par les musiciens du vibrato.

3.3.3 Jugement de goût

Tout les auditeurs, avec une écoute critique comparable, n'ont pas préféré les mêmes versions. Cependant, pour certains extraits il y avait peu d'ambiguïté. Pour clarifier l'analyse, j'ai comptabilisé le nombre de personnes qui ont préféré une version plutôt qu'une autre.

Nom de l'extrait	version sans retard	version avec retard
Contrebasse	0	12 ³⁹
Thème du « loup »	3	9
Thème du « oiseau »	3	7 ⁴⁰
Extrait de l' <i>Adagio</i>	0	12
Extrait du 5 ^e mouvement	1 ⁴¹	11

39. 4 sujets n'ont fait aucune différence entre les versions avec même retard ou avec une différence de retard ; 6 personnes ont préféré la version avec même retard et 4

Après la première écoute, les amateurs de musique ont fait souvent des commentaires liés à la musique. Ils percevaient des différences de notes⁴² ou de tempo. Les versions avec retard, quelque soit l'ordre de passage choisi entre les versions, ont parues être plus rapides pour sept sujets. Les différences d'expressivité ont été les plus nombreuses, il semblait à nos auditeurs que l'interprétation des musiciens changeait d'une version à l'autre. Tous, à la fin de l'entretien, lorsque je leur révélais la nature de la différence entre les versions qu'ils venaient d'entendre ont été surpris de l'influence du preneur de son et du mixeur sur la restitution du jeu des musiciens. En effet, lorsque certaines fréquences sont annulées, on perd l'information des micros-variations de jeu exécutées par le musicien. Il leur a semblé qu'avec mes mixages, j'avais participé aux intentions musicales des musiciens en amplifiant certaines inflexions du jeu musical.

personnes ont préféré la version avec une différence de retard

40. Deux sujets n'ont pas pu exprimer de préférences

41. La personne qui a préféré cette version a justifié son choix par un goût pour l'aplatissement des nuances dans cette version sans « appoint ».

42. seulement pour les extraits de la contrebasse.

Conclusion

Les tests auditifs ont montré que le public visé par la production de disque de musique acoustique percevait les effets introduits par le mélange cohérent des signaux du couple principal et des microphones d'« appoint ». Cette perception va au-delà des problématiques attendues : modification de l'espace sonore et de la qualité des timbres. Elle se traduit plus par une plus grande intelligibilité du discours musical et des inflexions musicales du jeu des musiciens (vibrato, précision des attaques). Cependant la perception de différences entre les mixages avec et sans retard entre les signaux du couple principal et des microphones d'« appoint » reste dépendante de la nature des sources. Pour les instruments à vent ces différences semblent être plus fines que pour les instruments à cordes.

La préoccupation de l'introduction d'un retard entre les signaux du couple principal et des microphones d'« appoint » paraît maintenant importante, selon la nature des sources à enregistrer. On a vu qu'aujourd'hui, les preneurs de son ont à leur disposition tous les outils nécessaires à la mesure et à l'introduction d'un retard temporel adéquat. Le choix que j'ai effectué ici a été guidé par une volonté de précision pour la mesure et de garder la maîtrise du réglage de la valeur des retards temporels introduits.

Nous pouvons maintenant définir une certaine méthode dans l'approche de ce problème d'introduction de retard temporel. Si le choix se porte sur une prise de son globale, le preneur de son doit analyser la nature des sources sonores et la composition spatiale de la scène acoustique à enregistrer avant d'envisager l'introduction de retards temporels. Cela lui fera gagner un temps précieux lors de la post-production. Enfin, pour les signaux d'« appoint » qu'il souhaite retarder, il a la possibilité, s'il détient un enregistrement de cette source seule dans la scène acoustique à enregistrer, d'utiliser *Auto-Align* pour mesurer le retard qu'il doit introduire afin d'éliminer la majorité des problèmes de phase. Il peut avec des outils tels que *Sound Delay*, utilisé dans ce mémoire, faire le choix de suivre cette mesure exacte, ou d'ajuster ce retard en suivant son écoute.

Bibliographie

Livres

- [1] COLLECTIF. *Dictionnaire encyclopédique du son*. Sous la dir. de Pierre-Louis de NANTEUIL. Dunod, 2008. ISBN : 978-2-10-005979-9.
- [2] Francis RUMSEY. *Spatial Audio*. Music Technology. Focal Press, 2001. ISBN : 0-240-51623-0.
- [3] COLLECTIF. *Le Livre des techniques du son. Tome 1 - Notions fondamentales*. Sous la dir. de Denis MERCIER. 4^e éd. Audio-Photo-Vidéo. Dunod, 2007. ISBN : 978-2-10-054919-1.
- [4] COLLECTIF. *Le Livre des techniques du son. Tome 3 - L'exploitation*. Sous la dir. de Denis MERCIER. 3^e éd. Audio-Photo-Vidéo. Dunod, 2007. ISBN : 978-2-10-049651-8.
- [5] François DELALANDE. *Le Son des Musiques. Entre technologie et esthétique*. Sous la dir. de Bibliothèque de recherche MUSICALE. 3^e éd. Buchet/Chastel, 2001. ISBN : 2-283-01850-1.
- [6] Francis IBRE. *Bien entendu. itinéraire d'un audiophile*. Elektor, 2005. ISBN : 286661-151-9.
- [7] Daniel LESUEUR. *L'histoire du disque et de l'enregistrement*. Carnot, 2004. ISBN : 2-84855-081-3.

Cours

- [8] Jean CHATAURET. « Cours sur la stéréophonie, première année à Louis Lumière. » ENS Louis Lumière, 2011.

Thèses et mémoires

- [9] Alice LEGROS. « Elaboration et reception d'un traitement sonore novateur d'une oeuvre musicale du repertoire occidental savant : pour une approche renouvelée de l'écoute et de l'enregistrement. » Sous la dir. de Isabelle VIAUD-DELMON. Mémoire de fin d'études. Conservatoire National Supérieure de Musique et de Danse de Paris ; Formation Supérieure aux Métiers du Son, 2012. URL : <http://alicelegros.tumblr.com/Memoirefsms2>.

Articles

- [10] Claude BAILBLÉ. « Acoustique musicale. La musique dans un triangle. » In : *L'Audiophile* (juin 1991). URL : <http://www.asrr.org/biblioteca/Revue%20Audiophile/index2.html>.

Ressources internet

- [11] DPA MICROPHONES. *Specifications du microphone 4006-TL*. Version 0.4. 2008-2013. URL : <http://www.dpamicrophones.com/en/products.aspx?c=item&category=191&item=24010#specifications>.
- [12] NEUMANN.BERLIN. *Specifications du microphone KM 184*. Version 0.4. 2014. URL : https://www.neumann.com/?lang=fr&id=current_microphones&cid=km180_data.
- [13] NEUMANN.BERLIN. *Specifications du microphone TLM 103*. Version 0.4. 2014. URL : https://www.neumann.com/?lang=fr&id=current_microphones&cid=tlm103_data.
- [14] OKTAVA. *Specifications du microphone mk012*. Version 0.4. 2014. URL : <http://www.oktava-online.com/mk012.htm#tech>.
- [15] OKTAVA. *Specifications du microphone mk319*. Version 0.4. 2014. URL : <http://www.oktava-online.com/mk319.htm#tech>.
- [16] RME. *Specifications de la Fireface 400*. Version 0.4. 2014. URL : http://www.rme-audio.de/en_products_fireface_400.php#5.
- [17] RME. *Specifications du convertisseur Micstasy*. Version 0.4. 2014. URL : http://www.rme-audio.de/en_products_micstasy.php.
- [18] SCHOEPS. *Specifications du microphone mk5*. Version 0.4. 2014. URL : <http://www.schoeps.de/fr/products/mk5/specs>.
- [19] TASCAM. *Specifications de l'enregistreur DR680*. Version 0.4. 2014. URL : http://www.lemicrophone.fr/produit.php?ref=DR680&id_rubrique=134&gclid=CJLq28SYx74CFQFwwodZUGATg.

Table des illustrations

1	Différentes zones d'incertitude de localisation sur le plan horizontal. D'après Haustein et Stimmer en 1970.	13
2	Addition de signaux sinusoïdaux en phase	16
3	Deux signaux sinusoïdaux déphasés de 90°	17
4	Addition de signaux sinusoïdaux déphasés de 90°	17
5	Deux signaux sinusoïdaux en opposition de phase	18
6	Addition de deux signaux sinusoïdaux en opposition de phase, l'onde a disparu, il ne reste rien.	18
7	Exemple de réponse en fréquence suite à un phénomène de filtrage en peigne	19
8	Couple AB	23
9	Couple XY	24
10	Le couple ORTF	25
11	L'arbre Decca	25
12	Stéréosonic	26
13	Couple MS	27
14	Microphone d'« appoint » de la Contrebasse.	34
15	Vue d'ensemble de la prise de son de <i>Pierre et le loup</i> . . .	35
16	Installation en régie de l'enregistrement <i>Pierre et le loup</i> . .	36
17	Implantation de l'orchestre du <i>COGE</i> pour l'interprétation de la 5 ^{ème} <i>symphonie de Malher</i>	37
18	Exemple d'un spectrogramme, ici par Samplitude Pro X. .	45
19	Exemple d'un spectrogramme, ici par Izotope	46
20	Exemple d'un sonogramme, ici par Izotope	46
21	Exemple d'un sonogramme , avec la waveform du signal, ici par Izotope	46
22	Exemple d'un corrélateur-grame , ici par Voxengo	47
23	Exemple du <i>Blue Cat's Stereoscope Pro</i>	47
24	Exemple du <i>StereoTools</i> de Flux	48
25	Représentation graphique de trois signaux enregistrés par le biais d'un couple et d'un « appoint »	50

26	Exemple de deux signaux déphasés de 90 ° analysés par <i>Auto-Align</i> de Sound Radix	51
27	Exemple d'utilisation du <i>Sound Delay</i> de Voxengo	52
28	Exemple d'utilisation du <i>PanNoir</i> de Pyramix	54
29	Mixeur après l'introduction d'un retard de 541 samples dans les canaux gauche et droit du mélange de l'« appoint », dans l'extrait de la contrebasse.	58
30	Mesure de la corrélation entre couple et « appoint » (ici à droite), après l'introduction d'un retard temporel dédié à gauche et à droite sur le signal d'« appoint », dans l'extrait de la contrebasse.	59
31	Mixage sans introduction de retards, de l'extrait du thème du « loup » de <i>Pierre et le loup</i>	59
32	Mixage avec l'introduction des retards, de l'extrait du thème de l'« oiseau » de <i>Pierre et le loup</i>	60
33	Mesure de la corrélation entre le signal d'« appoint » des violons 1 et le canal gauche du couple, sans l'introduction d'un retard temporel sur les signaux d'« appoint », dans l'extrait du 5 ^e mouvement.	61
34	Mesure de la corrélation entre le signal d'« appoint » des violons 1 et le canal gauche du couple, après l'introduction d'un retard temporel sur le signal d'« appoint », dans l'extrait de l' <i>Adagio</i>	61
35	Localisation des sources en fonction de la différence d'intensité entre les canaux gauche et droit d'un système de reproduction audio stéréophonique.	77
36	Localisation des sources en fonction de la différence de temps d'arrivée entre les canaux gauche et droit d'un système de reproduction audio stéréophonique.	77
37	Spécifications du microphone 4006-TL.	78
38	Réponse en phase du microphone 4006-TL.	78
39	Spécifications du microphone mk5.	79

40	Courbe de réponse en fréquence et diagramme polaire du microphone mk5 en position cardiode.	79
41	Spécifications du microphone mk21.	80
42	Courbe de réponse en fréquence et diagramme polaire du microphone mk21.	80
43	Spécifications du microphone KM 184.	81
44	Courbe de réponse en fréquence et diagramme polaire du microphone KM 184.	81
45	Spécifications du microphone TLM 103.	82
46	Courbe de réponse en fréquence et diagramme polaire du microphone TLM 103.	82
47	Spécifications du microphone mk012.	83
48	Courbe de réponse en fréquence et diagramme polaire du microphone mk012.	83
49	Spécifications du microphone mk012.	84
50	Courbe de réponse en fréquence du microphone mk319.	84
51	Image de la Fireface de RME.	85
52	Caractéristiques de la carte son Fireface 400.	85
53	Image du Micstasy de RME.	86
54	Caractéristiques du convertisseur Micstasy.	86
55	Image du DR680 de Tascam.	87
56	Caractéristiques de l'enregistreur DR680.	87

Annexes

A Courbes de perception

Des expériences ont été menées pour connaître précisément la localisation perçue en fonction de la différence de temps d'arrivée et la différence d'intensité. Voici les courbes représentant les résultats de ces expériences. Elles sont tirées de la publication de Wittek et Theile, présentées à la 112^e convention de l'AES.

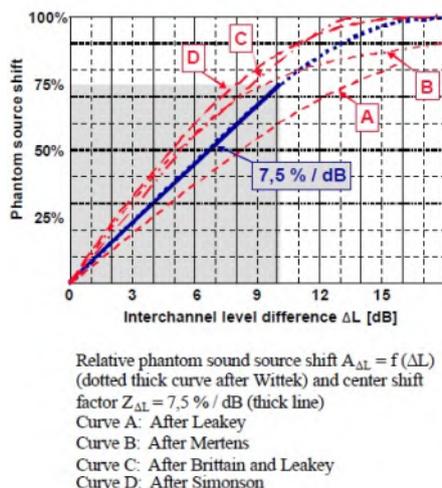


FIGURE 35 – Localisation des sources en fonction de la différence d'intensité entre les canaux gauche et droit d'un système de reproduction audio stéréophonique.

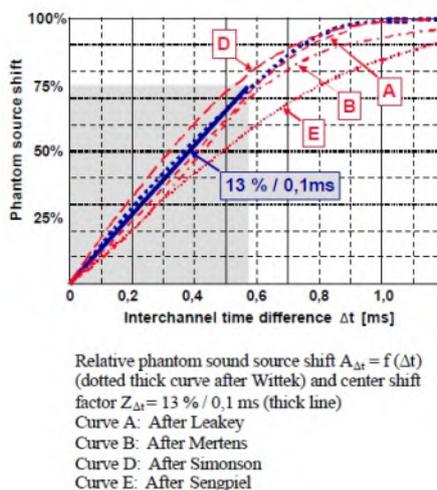


FIGURE 36 – Localisation des sources en fonction de la différence de temps d'arrivée entre les canaux gauche et droit d'un système de reproduction audio stéréophonique.

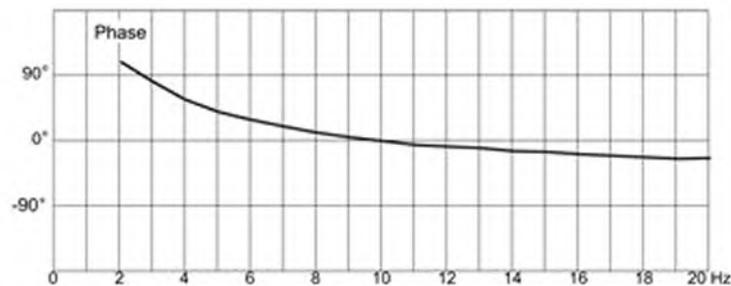
B Documentation des microphones utilisés

Le microphone 4006-TL omnidirectionnel de DPA : [11]

Specifications for 4006-TL Omni Microphone, P48, Transformerless

Directional characteristics:	Omnidirectional
Principle of operation:	Pressure
Cartridge type:	16 mm (0.63 in) pre-polarised condenser
Frequency range, ± 2 dB:	15 Hz - 20 kHz
Sensitivity, nominal, ± 2 dB:	35 mV/Pa; -29 dB re. 1 V/Pa
Equivalent noise level A-weighted:	Typ. 15 dB(A) re. 20 μ Pa (max. 17 dB(A))
Equiv. noise level ITU-R BS.468-4:	Typ. 27 dB (max. 29 dB)
S/N ratio, re. 1 kHz at 1 Pa (94 dB SPL):	79 dB
Total harmonic distortion (THD):	<0.5% THD up to 129 dB SPL peak, <1% THD up to 135 dB SPL peak
Dynamic range:	Typ. 120 dB
Max. SPL, peak before clipping:	143 dB
Output impedance:	<200 ohm
Cable drive capability:	Up to 100 m (328 ft)
Power supply:	Phantom P48
Current consumption:	2.60 mA
Connector:	3-pin XLR-M (Standard P48)
Weight:	150 g (5.29 oz)
Capsule diameter:	16 mm (0.63 in)
Microphone length:	165 mm (6.5 in)
Operating temperature range:	-40 to +70° C (-40 to +158° F)

FIGURE 37 – Spécifications du microphone 4006-TL.



On-axis phase response of DPA 4006-TL plotted using a linear frequency axis for evaluation of the phase response.

FIGURE 38 – Réponse en phase du microphone 4006-TL.

Le microphone mk5 de Schoeps : [18]

MK 5 - omnidirectionnelle	
Bande passante	20 Hz - 20 kHz
Sensibilité	10 mV/Pa
Niveau de bruit de fond acoustique équivalent (pondéré A)	14 dB-A
Niveau de bruit de fond acoustique équivalent (CCIR)	26 dB
Niveau de signal / bruit (pondéré A)	80 dB-A
Pression acoustique maximum à 0,5% DHT	133 dB-SPL

MK 5 - cardioïde	
Bande passante	40 Hz - 20 kHz
Sensibilité	13 mV/Pa
Niveau de bruit de fond acoustique équivalent (pondéré A)	15 dB-A
Niveau de bruit de fond acoustique équivalent (CCIR)	24 dB
Niveau de signal / bruit (pondéré A)	79 dB-A
Pression acoustique maximum à 0,5% DHT	132 dB-SPL

Caractéristiques mécaniques	
Longueur	27 mm
Diamètre	20 mm
Poids	25 g
Finition de surface	gris anti-reflet (g) ou nickel (ni)

Même si les tolérances de fabrication sont serrées, les capsules d'un même type peuvent être fournis appariées moyennant un supplément de prix.

FIGURE 39 – Spécifications du microphone mk5.

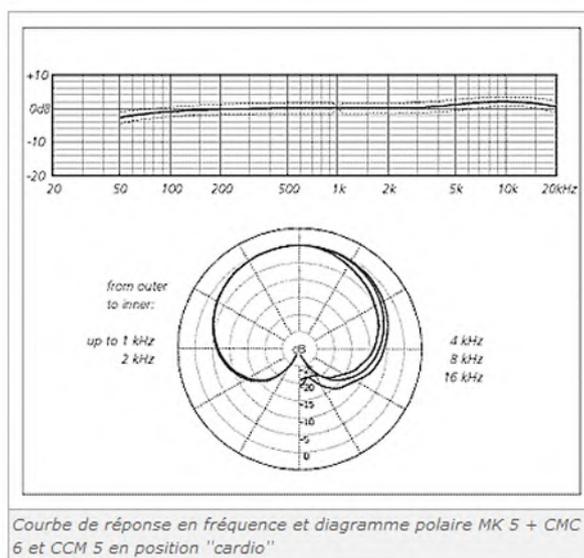


FIGURE 40 – Courbe de réponse en fréquence et diagramme polaire du microphone mk5 en position cardioïde.

Le microphone mk21 de Schoeps : [18]

MK 21	
Bande passante	30 Hz - 20 kHz
Sensibilité	13 mV/Pa
Niveau de bruit de fond acoustique équivalent (pondéré A)	14 dB-A
Niveau de bruit de fond acoustique équivalent (CCIR)	24 dB
Niveau de signal / bruit (pondéré A)	80 dB-A
Pression acoustique maximum à 0,5% DHT	132 dB-SPL
Longueur	22 mm
Diamètre	20 mm
Poids	17 g
Finition de surface	gris anti-reflet (g) ou nickel (ni)

Même si les tolérances de fabrication sont serrées, les capsules d'un même type peuvent être fournies appairées moyennant un supplément de prix.

FIGURE 41 – Spécifications du microphone mk21.

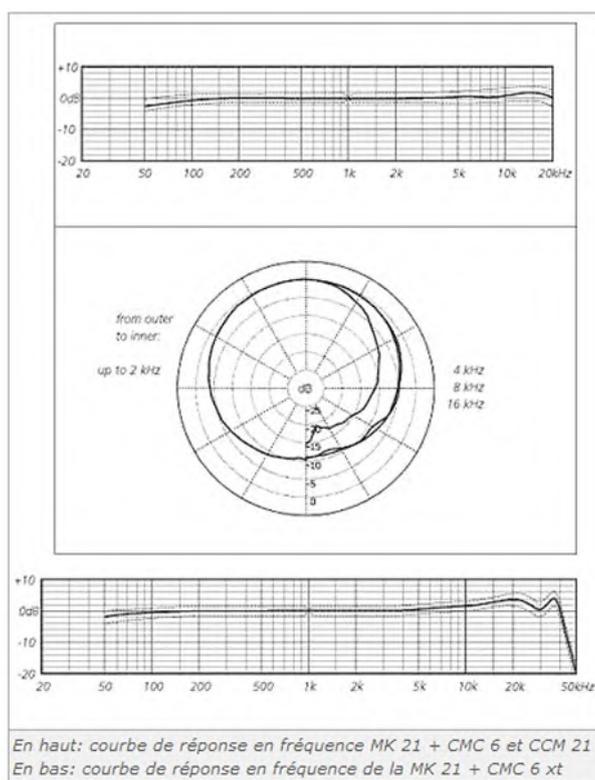


FIGURE 42 – Courbe de réponse en fréquence et diagramme polaire du microphone mk21.

Le microphone km 184 de Neumann : [12]

KM 183 / 184 / 185	
Principe de fonctionnement acoustique	Transducteur à pression/gradient de pression
Directivité	omnidirectionnelle/cardiode/hypercardioïde*
Réponse en fréquence	20 Hz à 20 kHz
Sensibilité à 1 kHz, charge 1 kOhm	12/15/10 mV/Pa ¹⁾
Impédance de sortie nominale	50 Ohms
Impédance de charge recommandée	1000 Ohms
Niveau de bruit équivalent, selon CCIR ²⁾	24/22/24 dB SPL ¹⁾
Niveau de bruit équivalent, pondéré A ²⁾	13/13/15 dB SPL (A) ¹⁾
Rapport Signal/Bruit, selon CCIR ²⁾ (au niveau SPL de 94 dB)	70/72/70 dB ¹⁾
Rapport signal-bruit, pondéré A ²⁾ (au niveau SPL de 94 dB)	81/81/79 dB ¹⁾
Niveau SPL maximal, pour THD = 0,5 % ³⁾	140/138/142 dB ¹⁾
Niveau de sortie maximal	+10 dBu
Alimentation fantôme	48 V ± 4 V
Intensité consommée	3.2 mA
Connecteur de sortie	XLR 3M
Poids	env. 80 g
Diamètre	22 mm
Longueur	107 mm

1) KM 183 / KM 184 / KM 185
 2) selon IEC 60268-1; Pondération CCIR selon CCIR 468-3, valeur de quasi-crête ; Pondération A selon IEC 61672-1, valeur efficace
 3) Le THD du préampli micro intégré est mesuré pour une tension d'entrée équivalente à la tension de sortie de la capsule pour le niveau SPL spécifié.

FIGURE 43 – Spécifications du microphone KM 184.

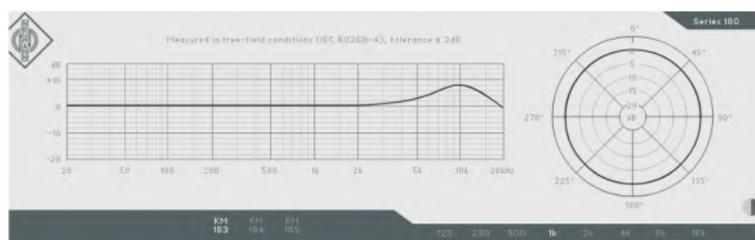


FIGURE 44 – Courbe de réponse en fréquence et diagramme polaire du microphone KM 184.

Le microphone TLM 103 de Neumann : [13]

Principe de fonctionnement acoustique	Transducteur à gradient de pression
Directivité	Cardioïde
Réponse en fréquence	20 Hz à 20 kHz
Sensibilité à 1 kHz, charge 1 kOhm	23 mV/Pa
Impédance de sortie nominale	50 Ohms
Impédance de charge conseillée	1000 Ohms
Niveau de bruit équivalent, selon CCIR ¹⁾	17,5 dB
Niveau de bruit équivalent, pondéré A ¹⁾	7 dB (A)
Rapport Signal/Bruit, selon CCIR ¹⁾ (au niveau SPL de 94 dB)	76,5 dB
Rapport signal-bruit, pondéré A ¹⁾ (au niveau SPL de 94 dB)	87 dB
Niveau SPL maximal, pour THD = 0,5% ²⁾	138 dB
Niveau de sortie maximal	+13 dBu
Tension d'alimentation fantôme	48 V ± 4 V
Intensité consommée	3 mA
Connecteur de sortie	XLR 3M
Poids	env. 450 g
Diamètre	60 mm
Longueur	132 mm

1) selon IEC 60268-1; Pondération CCIR selon CCIR 468-3, valeur de quasi-crête ; Pondération A selon IEC 61672-1, valeur efficace
2) Le THD du préampli micro intégré est mesuré pour une tension d'entrée équivalente à la tension de sortie de la capsule pour le niveau SPL spécifié.

FIGURE 45 – Spécifications du microphone TLM 103.



FIGURE 46 – Courbe de réponse en fréquence et diagramme polaire du microphone TLM 103.

Le microphone mk012 de Oktava : [14]

Technical specifications	
Mic type	Small diaphragm condenser
Polar Pattern	Cardioid, Hypercardioid, Omnidirectional
Freq. Response	20-20000Hz
Phantom voltage required	48 +/- 2V
Full impedance, module	less than 300 ohms
Weighted SPL (ref. DIN 45412)	18 dBA
Maximum SPL in 250-8000Hz range, (less than 0.5% THD)	130 db
Free field sensitivity at 1KHz	10 mV/Pa
Free field sensitivity roll off from 40Hz to 20KHz should not exceed	+/- 3 db
Weight with 3 capsules and pad	200 gr
Length, mm	128
Max diameter/width, mm	23
Accessories included	Mic holder
Accessories optional	Shock mount
The difference in free field sensitivity between 0° and 90° should be as follows:	
For omnidirectional capsule:	
in 40-1000 Hz range:	no more than 2 db
in 1-5 KHz range:	no more than 4 db
in 5-8 KHz range:	no more than 8 db
For cardioid capsule:	
in 250-8000 Hz range:	no less than 4 db
For hypercardioid capsule:	
in 250-5000 Hz range:	no less than 8 db
Average sensitivity difference between 0° and 180° for cardioid capsule in 63-12500 Hz range:	16 db

FIGURE 47 – Spécifications du microphone mk012.

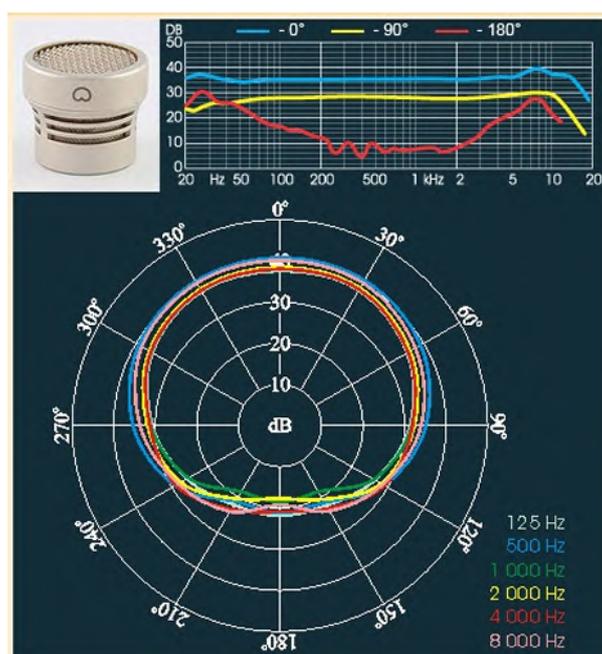


FIGURE 48 – Courbe de réponse en fréquence et diagramme polaire du microphone mk012.

Le microphone mk319 de Oktava : [15]

Technical specifications	
Mic type	Large diaphragm condenser
Polar Pattern	Cardioid
Freq. Response, Hz	20-18000
Output sensitivity, mV/Pa	13
Max SPL, dB for 0,5% THD	122
Self Noise (DIN), dBa	<14
Output impedance, ohm	200
Phase	Positive
Powering, V	48
Supply current, mA	8
Switches	-10dB / High-Pass Filter
Weight, g	490
Length, mm	205
Max diameter/width, mm	52
Temp. Range	-35degrees C to +45 degrees C
Relative Humidity	85% (+25 degrees C)
Accessories included	Mic holder
Accessories optional	Shock mount

A switch located at the left (High-Pass Filter) allows to change the microphone's cutoff frequency. This reduces low frequency interference directly at the input of the microphone amplifier. This setting also compensates for the unavoidable bass boost that occurs with all pressure gradient transducers when they are used at close distance (proximity effect).

The second switch attenuates the sensitivity by 10 dB. This option should be used in high SPL environments or by recording of loud instruments (for example saxophone).



FIGURE 49 – Spécifications du microphone mk012.

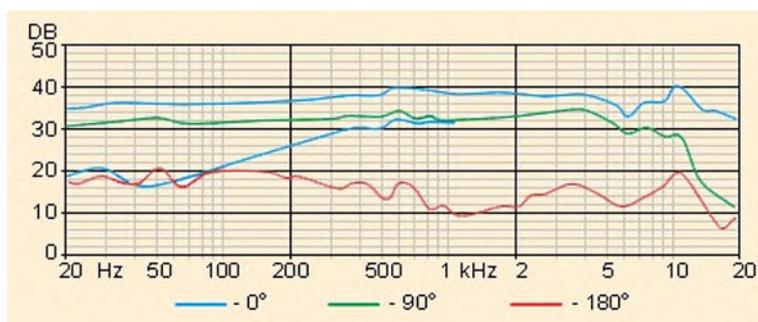


FIGURE 50 – Courbe de réponse en fréquence du microphone mk319.

C Documentation du matériel utilisé

Fireface 400 de RME : [16]



FIGURE 51 – Image de la Fireface de RME.

Input AD: 6 x 1/4" TRS (4 x Line, 2 x Line/Instrument), 2 x XLR/TRS Combo connector (2 x Mic/Line), all servo-balanced.

Output DA: 6 x 1/4" TRS, servo-balanced, DC-coupled signal path, 1 x 1/4" TRS unbalanced

Input Digital: 1 x ADAT optical or SPDIF optical, SPDIF coaxial (AES/EBU compatible)

Output Digital: 1 x ADAT optical or SPDIF optical, SPDIF coaxial (AES/EBU compatible)

MIDI: 2 x MIDI I/O via breakout cable (4 x 5-pin DIN jacks), for 32 channels low jitter hi-speed MIDI

Dynamic range AD: 110 dB RMS unweighted, 113 dBA

THD AD: < -100 dB (< 0.001 %)

THD+N AD: < -98 dB (< 0.0012 %)

Crosstalk AD: > 110 dB

Dynamic range DA: 110 dB RMS unweighted, 113 dBA (unmuted)

THD DA: -100 dB (0.001 %)

THD+N DA: -96 dB (0.0015 %)

Crosstalk DA: > 110 dB

Input/Output level for 0 dBFS @ Hi Gain: +19 dBu

Input/Output level for 0 dBFS @ +4 dBu: +13 dBu

Input/Output level for 0 dBFS @ -10 dBV: +2 dBV

Sample rate internally: 32, 44.1, 48, 64, 88.2 kHz, 96 kHz, 128, 176.4, 192 kHz

Sample rate externally: 28 kHz - 200 kHz

Frequency response AD/DA, -0.1 dB: 5 Hz - 20.4 kHz (sf 44.1 kHz)

Frequency response AD/DA, -0.5 dB: 1 Hz - 43.3 kHz (sf 96 kHz)

FIGURE 52 – Caractéristiques de la carte son Fireface 400.

Micstasy de RME : [17]



FIGURE 53 – Image du Micstasy de RME.

Frequency response -0.1 dB: 20 Hz - 100 kHz

EIN @ 30 dB Gain @ 150 Ohm: 122.3 dBu

EIN @ 40 dB Gain @ 150 Ohm: 126.5 dBu

EIN @ 50/60/70 dB Gain @ 150 Ohm: 127.4 dBu

THD+N @ 30 dB Gain: < -100 dB, < 0.001 %

Signal to Noise ratio AD (SNR) @ +30 dBu: 115.2 dB RMS unweighted, 118.5 dBA

Signal to Noise ratio AD (SNR) @ +21 dBu: 112.7 dB RMS unweighted, 116 dBA

Signal to Noise ratio AD (SNR) @ +13 dBu: 110 dB RMS unweighted, 113 dBA

Sync sources: Internal, AES, word, Option (MADI)

Sample frequencies: 44.1 kHz, 48 kHz, 88.2 kHz, 96 kHz, 176.4 kHz, 192 kHz, variable (external clocks)

Sample rate range: MADI: 32 - 192 kHz, word clock: 27 kHz - 200 kHz, AES: 28 kHz - 200 kHz

Jitter: Typical < 1 ns for internal, word clock, AES and MADI input

Jitter suppression: >30 dB (2.4 kHz)

Jitter sensitivity: all PLLs operate error-free even at 100 ns

Power supply: Internal switching mode PS, 100V - 240V AC, 60 Watt

Dimensions: (WxHxD) 483 x 88 x 200 mm

Warranty: 2 years

FIGURE 54 – Caractéristiques du convertisseur Micstasy.

Le DR680 de Tascam [19]



FIGURE 55 – Image du DR680 de Tascam.

Enregistrement/lecture

- ▶ Enregistrement mobile de haute qualité
- ▶ Dimensions compactes, utilisation facile
- ▶ Jusqu'à 8 pistes en enregistrement :
- ▶ 6 pistes + piste pré-mix stéréo interne, jusqu'en 96 kHz/24 bits
- ▶ 6 pistes via les entrées analogiques + 2 pistes via l'entrée numérique, jusqu'en 96 kHz/24 bits
- ▶ 2 pistes en 192 kHz/24 bits
- ▶ Support d'enregistrement carte SD/SDHC, fiable et facilement disponible
- ▶ Enregistrement et lecture aux formats Broadcast WAV (BWF) et MP3 (format MP3 avec firmware v1.10 ou ultérieur)
- ▶ Enregistrement multipiste MP3 (jusqu'à 4 pistes)
- ▶ Fréquences d'échantillonnage disponibles : 44,1 kHz, 48 kHz, 96 kHz, 192 kHz
- ▶ Résolution numérique : 16 bits/24 bits
- ▶ Buffer de pré-enregistrement : le fichier audio enregistré commence avant l'instant où vous avez effectivement appuyé sur la touche Record
- ▶ Fonction d'enregistrement automatique : l'appareil démarre et arrête automatiquement l'enregistrement en fonction du niveau d'entrée
- ▶ Gestion de marqueurs (fonction Locate)
- ▶ Fonctions de répétition de lecture

Entrées/sorties

- ▶ 6 entrées analogiques symétriques, niveau micro/ligne (4 connecteurs Combo XLR/jack TRS, 2 connecteurs jack TRS)
- ▶ Alimentation fantôme +48 V (activation via sélecteur mécanique, par paires d'entrées)
- ▶ Sélecteur de gain sur chaque entrée analogique (High/Low, commutation mécanique)
- ▶ Filtre passe-haut et limiteur commutables sur chaque entrée analogique
- ▶ 6 sorties ligne asymétriques (connecteurs RCA)
- ▶ Entrée et sortie numérique stéréo (format commutable, S/PDIF ou AES/EBU)
- ▶ Sortie casque avec réglage de niveau
- ▶ Port USB 2.0 haut débit, pour transfert des données vers PC
- ▶ Fonction Cascade, permettant d'utiliser simultanément plusieurs DR-680

FIGURE 56 – Caractéristiques de l'enregistreur DR680.

D Formules mathématiques

D.1 La transformée de Fourier

La définition mathématique de la transformée de Fourier discrète est la suivante : Pour un signal s de N échantillons :

$$S(k) = \sum_{n=0}^{N-1} s(n) \cdot e^{-2i\pi k \frac{n}{N}} \text{ pour } 0 \leq k \leq N$$

D.2 La cross-corrélation

En traitement du signal, la corrélation croisée (« cross-correlation » en anglais) est un algorithme qui permet la mesure de la similitude entre deux signaux.

Cet algorithme a comme écriture mathématique, dans un espace linéaire : Pour deux signaux A et B de même longueur, et soit C leur corrélation croisée :

$$C(t) = \int_{-\infty}^{\infty} A(\tau)B(t - \tau)d\tau$$

