

# Le son pour l'image À l'ILOI

Éléments de connaissances

Michel Geiss

## Introduction

La gestion du son dans l'univers du visuel est un art à part entière. C'est aussi un vaste domaine, depuis la captation jusqu'au rendu d'un projet terminé, en passant par les choix et créations des musiques et des effets sonores.

Comme avec toutes les formes d'art, la connaissance technique, la maîtrise des outils, sont de nature à permettre d'atteindre de façon optimale un objectif souhaité.

Les techniques sonores sont complexes et diverses. En dehors des établissements d'enseignement, elles sont pratiquées de façon professionnelle par des spécialistes qui en ont fait leur activité principale. Cependant, les métiers de l'image nécessitent aussi une base de connaissances du son qui accompagnera les projets.

Peu d'ouvrages récents ont été publiés en français sur ce sujet. Ce document s'adresse à la fois aux enseignants de tous les cursus de l'ILOI et à leurs stagiaires. En effet, le son est présent dans tous les domaines liés à l'image animée, depuis le jeu vidéo jusqu'au cinéma. Il s'agit ici d'un « état des lieux » en 2023.

Deux fondamentaux sont à prendre en compte.

D'une part le fait que l'audition humaine a ses propres constantes, maintenant bien connues et non modifiables. Il est essentiel de les connaître, de même que les caractéristiques physiques du son, elles aussi faisant partie des éléments qui échappent à notre intervention. En connaître les principes de base permet de comprendre plus rapidement les effets que par l'expérience seule.

D'autre part, la technologie exploite ces principes fondamentaux du son. Et on le sait, les matériels, les logiciels, sont en évolution permanente et entraînent des modifications progressives des modes et méthodes de travail. L'arrivée de l'intelligence artificielle commence à concerner l'ensemble de la sphère audiovisuelle et va entraîner un bouleversement des outils et des activités. Alors, même si ce travail d'« inventaire » ne fait état que du contexte de 2023, il servira à mieux comprendre les implications de l'apport des futures technologies.

En principe, les stagiaires de l'ILOI ne se destinent pas aux métiers du son. Pourtant, sans qu'il s'agisse de former des spécialistes de ce domaine, l'ILOI s'efforce de leur apporter les éléments essentiels qui vont leur permettre de s'intégrer au mieux dans une équipe lors de leur arrivée dans le milieu professionnel de l'image, ou simplement d'acquérir la capacité d'autonomie dans des projets complets.

Tout en prenant en compte la grande importance des travaux pratiques à l'ILOI, l'objectif de ce document est de contribuer aux efforts des formateurs de l'établissement pour transmettre leurs connaissances et leur expérience au service de la réussite des études.

# Connaissance du son

## Qu'est-ce que le son ?

- Le son est une vibration de l'air audible par les êtres humains.
  - Dans le vide, sans air, il ne peut y avoir de son.
  - L'air est constitué de molécules : azote, oxygène, vapeur d'eau
  - Nous ne pouvons pas entendre tous les mouvements de l'air, mais seulement ceux auxquels notre audition est sensible.

## La pression atmosphérique

- Elle représente le poids de la masse d'air. Elle se mesure avec un baromètre.
- Elle est plus élevée à basse altitude qu'à haute altitude.
- Elle varie très lentement, en fonction des mouvements des masses d'air.
- Elle se mesure avec un baromètre ; quand il est enregistreur, il permet d'observer ses variations sur une certaine durée
- Le son est l'effet de variations très rapides de la pression atmosphérique au niveau des oreilles
- On donne le nom d'oscillations aux variations de la pression atmosphérique autour de sa valeur de base. C'est ce qui produit le son



*Figure 1 Un baromètre enregistreur. La variation de la pression atmosphérique dessine une courbe. L'équivalent en audio est beaucoup plus rapide. On l'appelle la forme d'onde.*

## Oscillations et sons

- Une oscillation mécanique produit un son quand elle fait vibrer l'air suffisamment rapidement. C'est, entre autres, le rôle d'un haut-parleur ou d'un casque audio grâce à leur membrane.



- Ci-dessous un équivalent d'une oscillation mécanique (mais ne produisant pas de son à cause de sa lenteur) :

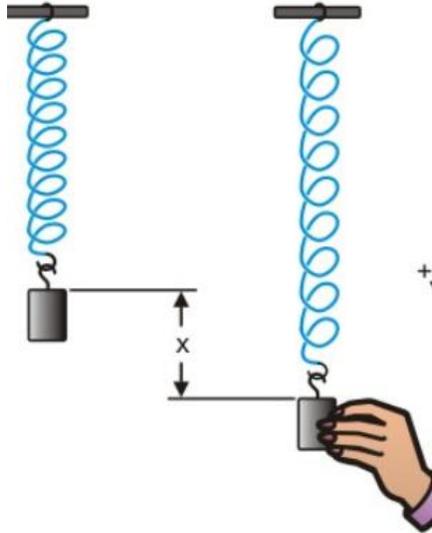


Figure 2 Une oscillation mécanique

### Amplitude, période

À gauche de la figure 2, le ressort est au repos (pas de vibration). À droite, on étend le ressort. Quand on le relâche, il commence à osciller, puis s'amortit :

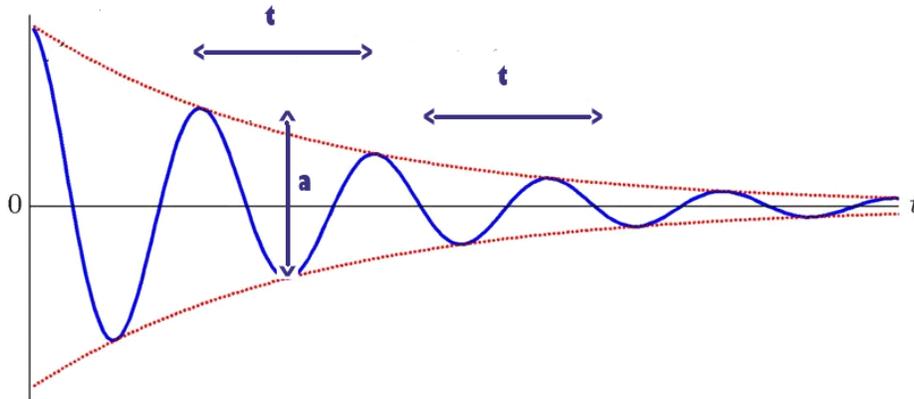


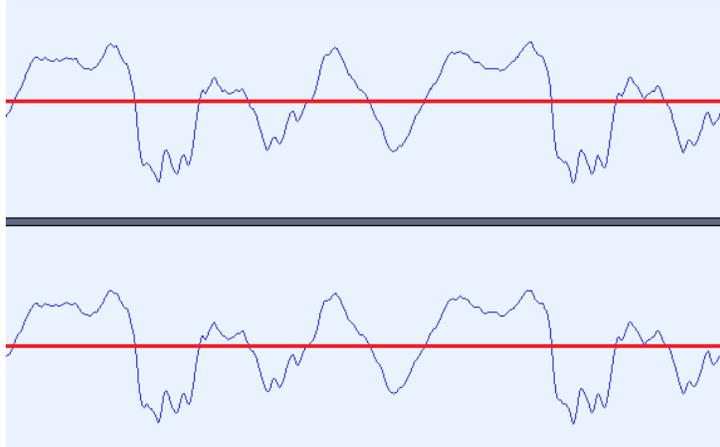
Figure 3 Amplitude, période

Cette figure représente l'**amplitude**  $a$  de l'oscillation (axe vertical) en fonction du temps  $t$  (axe horizontal). L'amplitude est donc l'importance de la variation, son intensité. Dans la figure ci-dessus elle diminue progressivement.

Cette oscillation est dite « périodique » parce que sa forme se répète identiquement sur la durée. La **période** est la durée qui sépare deux moments identiques de l'amplitude (les deux flèches «  $t$  »).

## Forme d'onde

Cette représentation des variations de l'amplitude en fonction du temps s'appelle la **forme d'onde** (l'onde étant ici une variation mécanique d'un objet, l'onde acoustique étant une variation de la pression atmosphérique autour de sa valeur moyenne).



*Figure 4 Une forme d'onde sonore stéréophonique ; le canal gauche est représenté dans la partie supérieure et le canal droit au-dessous ; les lignes horizontales en rouge représentent la valeur moyenne autour de laquelle les variations se produisent*

La connaissance de la forme d'onde dans son principe peut avoir une utilité pratique. En effet, un certain nombre de logiciels d'édition audio permettent de modifier une forme d'onde en la redessinant partiellement. Et modifier une forme d'onde a pour effet de modifier le son. Pour cela, il est utile de savoir que les sons graves ont une évolution plus lente, donc une courbe plus étalée que les sons aigus dont les évolutions sont plus serrées. Par exemple, quand on entend un « clic » parasite, un son aigu très bref, on peut le supprimer en redessinant la section de la forme d'onde correspondante comme ci-dessous :



*Figure 5 On voit dans la partie droite la rupture de la forme d'onde créée par le clic*

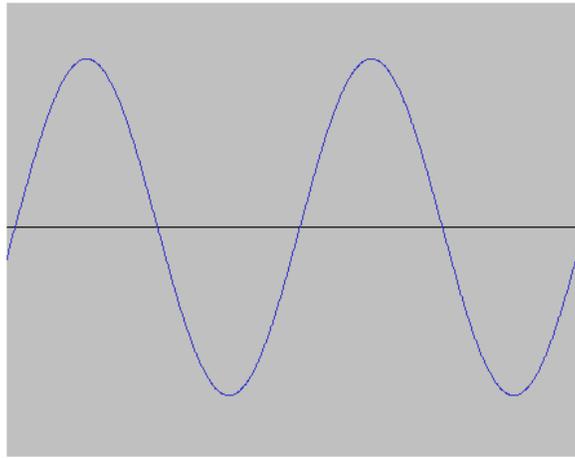


Figure 6 La forme d'onde après avoir redessiné la partie avec le clic

## La vitesse du son

- Le son se propage dans l'air à 340 m/s au niveau de la mer à 15°C (mais 4 fois plus vite dans l'eau)
- Elle baisse en fonction de l'altitude
- Elle augmente en fonction de la température : 349 m/s à 30°C au niveau de la mer
- Ces notions de vitesse sont utiles pour comprendre beaucoup de phénomènes acoustiques

## Fréquence et longueur d'onde

- La **longueur d'onde** est la distance parcourue par le son dans l'air par une période de l'oscillation. (exemple : à 340 m/s, avec une période  $1/100^{\text{ème}}$  de seconde un son parcourt  $340 \text{ m}/100 = 3,40 \text{ m}$ ).  
La longueur d'onde de ce son est donc de 3,40 m.
- La fréquence est le nombre d'oscillations par seconde, indiquée en **Hertz** (symbole **Hz**) ou **kilohertz** (symbole **kHz**)

## Effets de la vitesse du son

- À 1 m de distance, le son met 3 ms (millisecondes) pour nous parvenir
- À 100 m, il met 300 ms (1/3 s)
- Les molécules d'air des vibrations sonores rebondissent sur les obstacles et produisent deux phénomènes fréquents et essentiels dans la prise de son. Le son réfléchi est renvoyé avec un délai dépendant de la distance de la source sonore.

**L'écho** : répétition d'un son renvoyé par un obstacle, après une durée dépendant de sa distance (aller/retour). Il arrive que l'écho se répète plusieurs fois en disparaissant progressivement.

**La réverbération** : répétitions multiples et amorties du son sur toutes les surfaces d'une pièce - sa durée dépend du volume de la pièce et des matériaux qui la composent

## Sons, infrasons, ultrasons

- Les limites de l'audition humaine sont de 20 vibrations de l'air par seconde à 20 000 par seconde
- Au-dessous de 20 par seconde (20 Hertz ou Hz) on les appelle infra-sons
- Au-dessus de 20 000 par seconde (20 000 Hertz ou 20 kHz) on les appelle ultra-sons
- Les sons à basse fréquence sont qualifiés de **sons graves**
- Les sons à fréquences élevées sont qualifiés de **sons aigus**

## Spectre sonore et timbre

L'énergie des sons se répartit en fonction des différentes vibrations qui le composent. En effet, un son est rarement pur. Il se compose presque toujours d'un mélange de multiples fréquences dont le volume varie en permanence et qu'on entend ensemble. Cet ensemble complexe de vibrations compose la matière sonore.

**Le spectre sonore est la représentation des énergies de ce que contient le son, en termes de répartition entre les graves et les aigus.**

- Il peut s'observer avec un **analyseur de spectre**
- Il peut se modifier avec un **filtre** ou un **égaliseur**
- Dans la réalité, le spectre sonore évolue très rapidement, en même temps que les variations du son. Mais on peut en observer une moyenne dans certaines zones de fréquences. Des dominantes peuvent parfois indiquer des problèmes à corriger.

Le **timbre** est le nom qu'on donne à la sensation acoustique produite par le spectre sonore. Chaque instrument acoustique a son propre timbre.

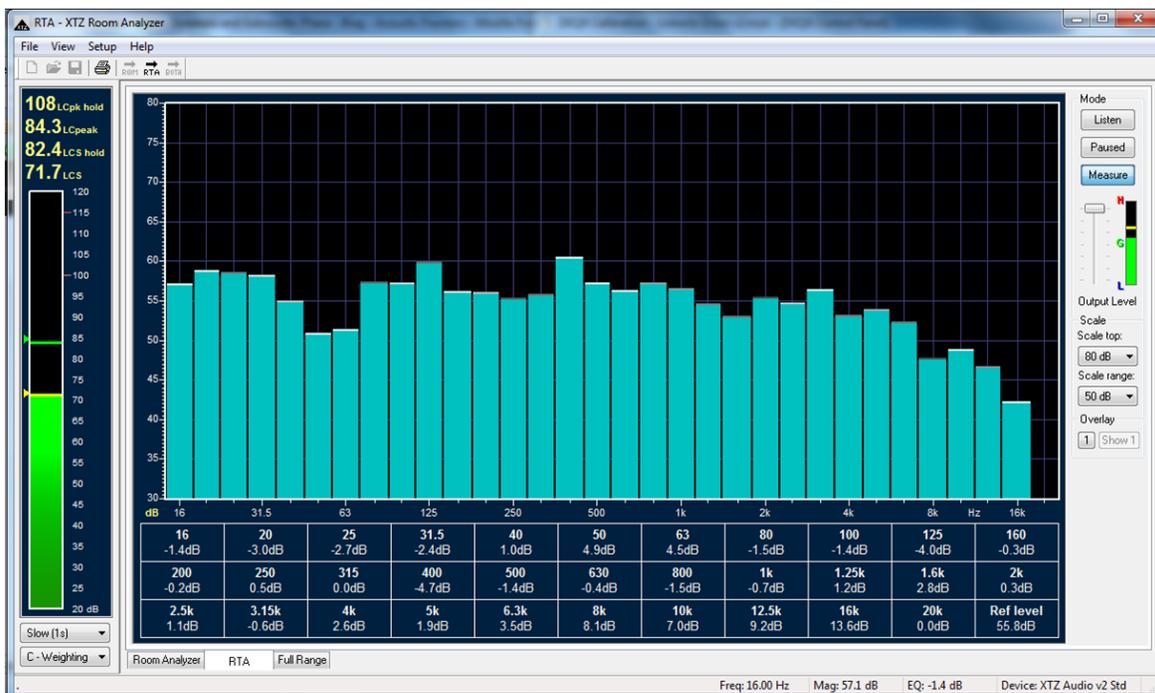


Figure 7 L'écran d'un analyseur de spectre. De gauche à droite, les sons graves jusqu'aux sons aigus.

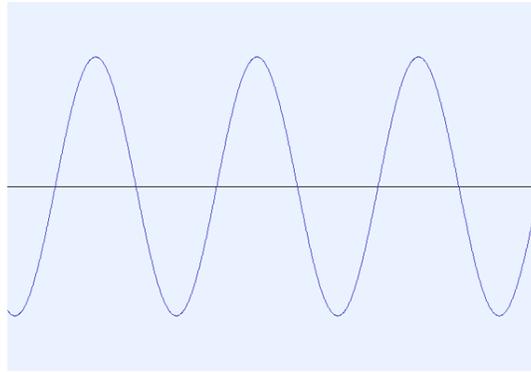
## La propagation du son

- Les sons graves ont une plus grande portée que les sons aigus qui se dispersent plus rapidement
- L'énergie des fréquences élevées subit une perte en relation avec la distance
- Les aigus sont essentiels pour la perception de la « présence » des sons

## Le signal sonore

- Les vibrations acoustiques peuvent être transformées en vibrations électriques équivalentes par une captation
- On appelle **signal** ces vibrations électriques
- Ce signal peut être enregistré, restitué et traité : amplifié, atténué, compressé, expansé, corrigé en fréquences.
- On produit ce signal avec des **microphones** (les **micros**) ou des capteurs piezo-électriques
- Dans l'autre sens, les vibrations électriques peuvent être transformées en sons grâce aux haut-parleurs et aux casques

*Le son le plus pur est celui produit par un signal dont la forme d'onde est sinusoïdale*



*Figure 8 Une forme d'onde sinusoïdale*

## Le bruit

- C'est le son le plus complexe qui puisse exister. Il n'a pas de fréquence particulière mais il est le résultat d'un mélange de toutes les fréquences, depuis les graves, jusqu'aux aigus.
- C'est le son qu'on entend près d'une cascade. À très faible volume le bruit (**noise** en anglais) est un élément perturbateur présent dans les circuits électroniques (le bruit de fond).
- Dans la pratique de l'audio, le bruit de fond est un bruit électronique qu'on doit essayer d'éviter, notamment en choisissant du matériel à faible bruit, en particulier les préamplis, ou en enregistrant avec des volumes suffisants.
- C'est le rapport **S/N (Signal to Noise) en dB** qui définit le bruit de fond dans les matériels électroniques. Il est mentionné dans leurs caractéristiques. Plus ce rapport est élevé, moins le bruit de fond est présent.



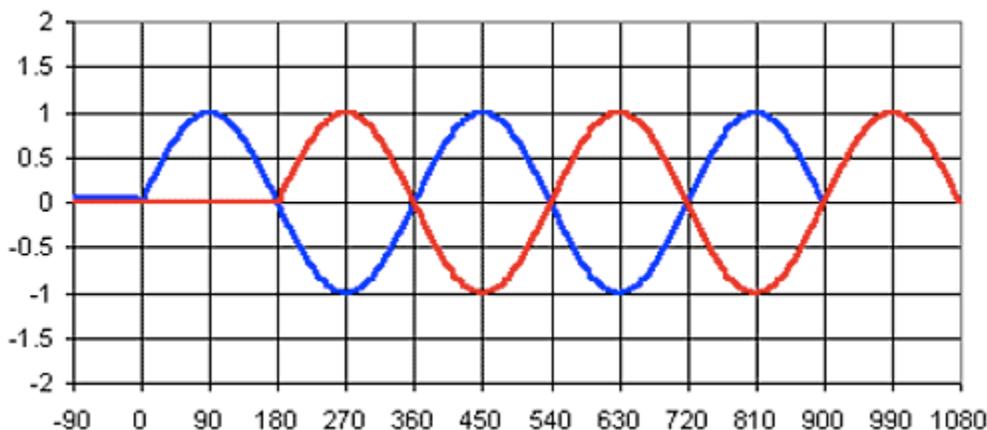
*Figure 9 Une forme d'onde de bruit*

Ce bruit de fond est l'un des problèmes des enregistrements, surtout avec des sources à faible volume. Depuis longtemps on fait la guerre au bruit de fond, entre autres avec les anciens procédés de réduction de bruit de Dolby ou dbx pour la lecture des pistes sonores optiques ou magnétiques. L'arrivée du numérique a pratiquement éliminé le problème, mais d'autres sources de bruit d'origine électronique subsistent.

Au contraire, le bruit de fond du lieu correspondant à l'ambiance d'un tournage est un élément utile à la crédibilité des prises. Que ce soit en intérieur ou en extérieur, il est fortement conseillé de l'enregistrer seul dans les mêmes conditions que celles du tournage. Il pourra être très utile en postproduction pour rendre plus naturels des raccords d'éléments enregistrés différemment.

## La phase

- C'est le décalage de temps entre deux signaux.
- « En phase » veut dire sans décalage. Les deux signaux se renforcent.
- « Hors phase » signifie que les deux signaux sont en opposition et s'annulent



*Figure 10 Les deux signaux en bleu et rouge sont hors phase : l'un augmente pendant que l'autre diminue ; on note une différence de phase de une demi longueur d'onde entre les deux courbes*

Entre « Hors phase » et « En phase », de multiples déphasages se produisent. La différence de phase concerne donc la présence de deux signaux. En audio, les deux situations ci-dessus de deux signaux identiques simultanés ne se trouvent quasiment jamais. Les signaux ont des fréquences qui varient en permanence. Les déphasages aussi. De plus, de multiples fréquences composent le son global. Beaucoup d'entre elles sont décalées les unes par rapport aux autres. Or, de très petits décalages entre deux sons mélangés produisent une modification du timbre du son.

Par exemple, quand le son provenant d'une paire d'enceintes est perçu directement et aussi réfléchi par une console, le rendu sonore des enceintes est en partie modifié. Leur emplacement par rapport à l'opérateur doit donc être pris en compte pour éviter les erreurs de jugement.

Par ailleurs, entre deux micros proches, en fonction de leur espacement, le son résultant du mixage des deux micros sera peut être modifié à cause du décalage temporel du son arrivant sur l'un des micros par rapport à l'autre. Dans ce cas, certaines composantes du son s'annulent. On appelle **déphasage** ce décalage temporel. Il s'exprime en degrés, comme on le voit sur l'axe horizontal de la figure ci-dessus. En cas de doute, il est conseillé de couper l'un des deux micros et de constater un éventuel changement dans le timbre du son. Auquel cas on changera la position de l'un des micros.

Autre phénomène, quand une surface lisse se trouve entre la source sonore et la capture par un micro, celle-ci est le résultat du mélange entre le son direct et le son réfléchi par effet de miroir. La légère différence de timing entre les deux donne un déphasage qui modifie le timbre du son direct par leur mélange. Et puisque les sons contiennent de multiples fréquences, là aussi, celles qui correspondent à des oppositions de phases s'annulent et modifient le son résultant.

C'est la raison pour laquelle le problème du déphasage entre son direct et son réfléchi mélangés peut affecter les prises de sons.

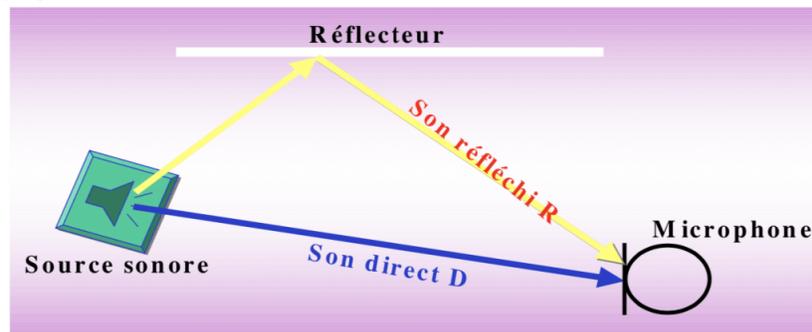


Figure 11 Le micro capte le mélange du son direct avec le son réfléchi

Prenons un exemple. À 340 m/s la différence de timing entre deux micros distants de 34 cm est de 1 ms. Ces 34 cm correspondent à la longueur d'onde d'un signal à 1 kHz. À cette fréquence, les deux signaux sont « en phase » et leur amplitude est augmentée quand on les mélange. À cette même distance, à 2 kHz, le décalage est d'une demi-période et les 2 signaux s'annulent pour leur contenu à 2 kHz. Par conséquent, dans la figure ci-dessus, si le trajet indirect du son par la paroi réfléchissante parcourt en tout 34 cm les fréquences de 2 kHz s'annulent en partie (le son reçu par le micro est modifié).

Dans le travail sur le son, les déphasages s'observent avec un **phasemètre**.

## L'effet Doppler

- Il s'agit de la variation de fréquence d'une source sonore en mouvement quand elle est entendue par un observateur fixe. Quand elles s'approchent, les ondes sonores sont compressées dans l'air, donc leur fréquence est plus haute. Quand elles s'éloignent, elles sont étirées et leur fréquence baisse. C'est par exemple l'effet qu'on entend quand un jet ou un véhicule passent à proximité. La hauteur du son d'origine est d'abord plus grande, devient celle d'origine en passant à proximité, puis diminue en s'éloignant.
- L'effet Doppler peut se comprendre par le fait qu'avec un son en mouvement à la vitesse de 340 m/s, un cycle du signal se trouve raccourci par son rapprochement, puis rallongé par son éloignement.

Par exemple, si on filme le mouvement des vagues en bord de mer avec un drone, quand il est en position fixe, elles se répètent avec un certain rythme. Si on met le drone en mouvement vers le large, le même rythme des vagues apparaît en accéléré sur le film. Au contraire, en faisant revenir le drone vers le bord de mer, on observe un mouvement des vagues ralenti.

Dans le cas de la perception de l'effet Doppler, au lieu du drone, c'est le jet ou le véhicule qui se déplacent, et le film est l'effet sur notre écoute.

C'est le mouvement relatif entre une source sonore et la position d'écoute qui produit cet effet de changement de fréquence.

Un exemple d'effet Doppler que les musiciens connaissent : celui de la cabine Leslie. La rotation d'un haut-parleur tournant provoque des éloignements et rapprochements réguliers de la source sonore. Le changement de la hauteur perçue correspond au vibrato perçu et dont on modifie la vitesse en agissant sur celle du haut-parleur.

## L'octave

- C'est la différence de hauteur entre 2 sons dont l'un a une fréquence double de l'autre : exemple, l'octave supérieure d'un son à 440 Hz est un son à 880 Hz

## La distorsion

Il s'agit de la déformation d'un son par rapport à l'original. La distorsion la plus répandue est la distorsion harmonique. C'est celle qui rajoute des harmoniques supplémentaires au son de base. Elle peut se produire lors de la capture et donc être présente sur le fichier enregistré à cause d'un excès de volume, ce qui est à éviter absolument. Mais les traitements ultérieurs peuvent aussi la générer. Même une faible distorsion peut être repérée par des oreilles entraînées. On peut l'éviter en étant attentif aux volumes d'enregistrement.

La distorsion peut se produire à cause de système d'écoute, ce qui est moins grave, par exemple par excès du réglage de volume sur des moniteurs de faible puissance. Dans ce cas, il suffit de diminuer ce volume pour rétablir un son normal.

## La perception des sons

Nous percevons les sons par un ensemble de processus. Le son pénètre dans nos oreilles. Il produit des signaux nerveux envoyés au cerveau qui les analyse et les interprète. Il ne s'agit pas d'un processus technique, mais d'un processus biologique et psychologique. En tant que tel il est influençable par notre psychisme. On ne peut donc pas comparer notre audition avec une captation neutre par micro. On n'entend pas le son avec nos oreilles, mais avec notre cerveau.

### La perception du volume sonore

Notre perception du volume est de deux sortes : le **volume absolu** et le **volume relatif**.

Le premier concerne l'échelle de pression acoustique absolue entre le minimum perçu et le maximum tolérable.

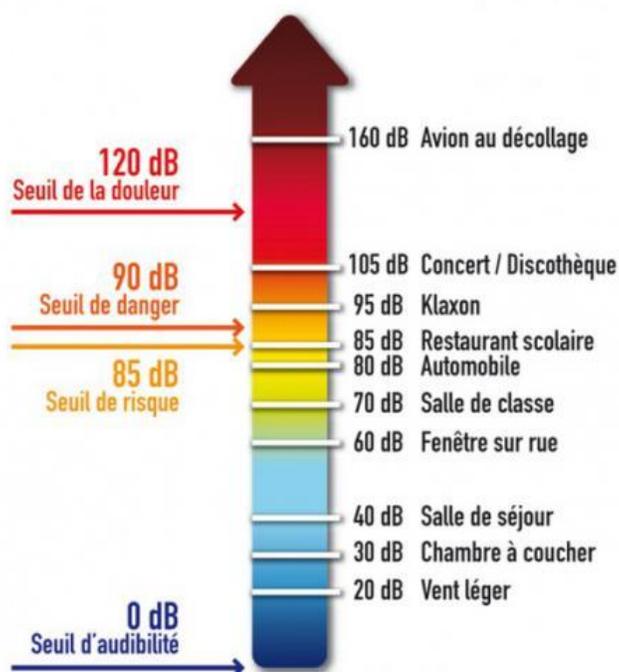


Figure 12 L'échelle des volumes perçus. La valeur de 0 dB est celle qui représente le volume acoustique à partir duquel on peut entendre le son.

Les différents volumes de cette échelle s'expriment en dBA, les décibels acoustiques. Le volume relatif correspond à la perception de la différence de volume entre un son et un autre. Lors de l'écoute, c'est ce qui correspond au « relief » sonore qu'on appelle la **dynamique**. Cet écart entre deux volumes s'exprime sous forme de rapport en **dB**.

Le volume acoustique se mesure avec un **sonomètre**.



Figure 13 Le sonomètre numérique PEAKMETER PM6708

## La perception des fréquences en fonction du volume

- Depuis longtemps il a été constaté qu'à faible puissance sonore, on perçoit mal les sons graves et les sons aigus, mais qu'on est plus sensible aux fréquences moyennes (appelées « médiums »), celles qui, entre autres, correspondent aux sons de la voix
- Plus on augmente le volume sonore, mieux on perçoit les sons graves et aigus
- Ce phénomène propre à l'être humain est représenté par la courbe d'isophonie suivante (aussi appelée courbe de Fletcher-Munson)

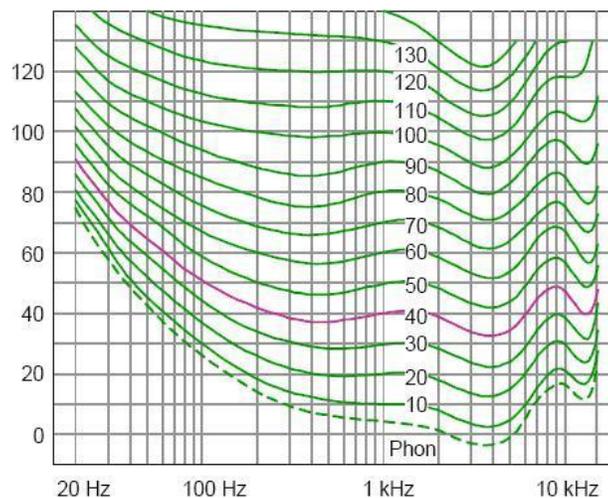


Figure 14 La courbe d'isophonie. En vertical, le volume sonore, en horizontal les fréquences. Les chiffres centraux indiquent pour la courbe correspondante la pression sonore nécessaire pour obtenir la même sensation de puissance

À cause de ce changement de perception des fréquences en rapport avec le volume et pour éviter les erreurs de jugement, on doit s'efforcer de travailler aussi longtemps de possible à volume constant, surtout au moment du mixage.

## La perception en fonction du timbre

Entre deux sons à même pression sonore, celui dont le contenu en fréquences est le plus riche est perçu avec plus d'intensité que l'autre. En d'autres termes, entre deux sons de même énergie acoustique mais de contenus en fréquences différents, celui qui contient plus d'aigus est perçu avec plus d'intensité. Cette notion est très importante lors du travail sur les sons avec l'égaliseur (voir ci-dessous).

## La perception en fonction de la distance

- Plus on s'éloigne d'une source sonore, moins on perçoit les aigus qu'elle contient
- Or, les aigus contribuent à la sensation de « présence » de la source
- **Cette notion est fondamentale en captation sonore, où cette sensation de présence est souvent indispensable**

## La perception en fonction de la direction

- Nous pouvons déterminer la provenance des sons sur le plan horizontal et vertical : avant, arrière, gauche, droite, dessus, dessous
- Notre cerveau interprète un ensemble de facteurs et les traduit en informations directionnelles
- L'un des facteurs concerne la distance entre nos deux oreilles : à la vitesse de 340 m/s et un espacement de 20 cm entre deux oreilles, un son arrivant totalement à gauche ou à droite met environ 6 ms (millièmes de seconde) pour parvenir à l'autre oreille. C'est ce qu'on appelle la différence inter-aurale (ou Inter Aural Difference).
- Les autres facteurs de notre perception spatiale comprennent la forme de la tête, sa position par rapport au buste, la forme des oreilles
- La perception de la direction est dépendante des fréquences sonores : les fréquences aiguës sont les plus directives, contrairement aux fréquences graves qui le sont très peu. En effet, la longueur d'onde d'un son à fréquence de 10 kHz est de 3,4 cm, très inférieure à l'intervalle entre les deux oreilles, alors qu'à 100 Hz elle est de 3,40 m, en dehors de nos capacités de différenciation d'une oreille à l'autre.

# L'écoute des sons

- Ce qu'on entend sans y prêter attention et ce qu'on écoute réellement correspondent à deux attitudes différentes
- Travailler sur le son nécessite une écoute attentive de tous les sons, agréables ou désagréables, et un jugement de leurs qualités et leurs défauts.
- L'écoute est un ensemble de dispositions qui nécessitent concentration et discrimination

## Le son qu'on entend

- Les sons les plus faibles sont masqués par les plus forts.
- En extérieur, le son est amorti. En intérieur il subit des réflexions sur les surfaces : murs, sols et plafonds

- On entend les sons qui proviennent de l'avant, de l'arrière, du dessus. Nous sommes immergés dans le son.
- En intérieur on entend un mélange entre le son direct et le son réfléchi renvoyé par les surfaces  
**Plus la source est éloignée, plus le rapport volume des réflexions/volume de la source augmente (sauf en extérieur où le son est amorti)**
- **Le son qu'on entend est aussi un mélange entre celui des sources sonores dominantes et des bruits divers; lors des captations; plus on est proche de la source, plus le rapport son utile/sons indésirables sera élevé (donc moins ces derniers seront gênants). À prendre en compte lors des enregistrements :**
  - **En intérieur on doit éviter de capter à des distances excessives**
  - **En toutes circonstances on tente d'éviter les bruits parasites**
- Le son qu'on entend de notre propre voix est le mélange du son direct avec les résonances des os de la tête. C'est pourquoi les autres nous entendent différemment et c'est aussi ce qui explique pourquoi on perçoit différemment un enregistrement de notre voix.

## **La fatigue auditive**

- Elle se produit progressivement lors des longues séances de travail
- Elle est dépendante du volume sonore : plus il est intense, plus la fatigue est rapide
- Une exposition à un volume excessif peut conduire à des pertes auditives irréversibles.

## **L'accoutumance**

- Notre audition s'habitue rapidement à l'écoute répétée des sons et accepte des erreurs de jugement qu'une écoute extérieure sortie du contexte pourrait révéler.

# **La reproduction du son : haut-parleurs et casques**

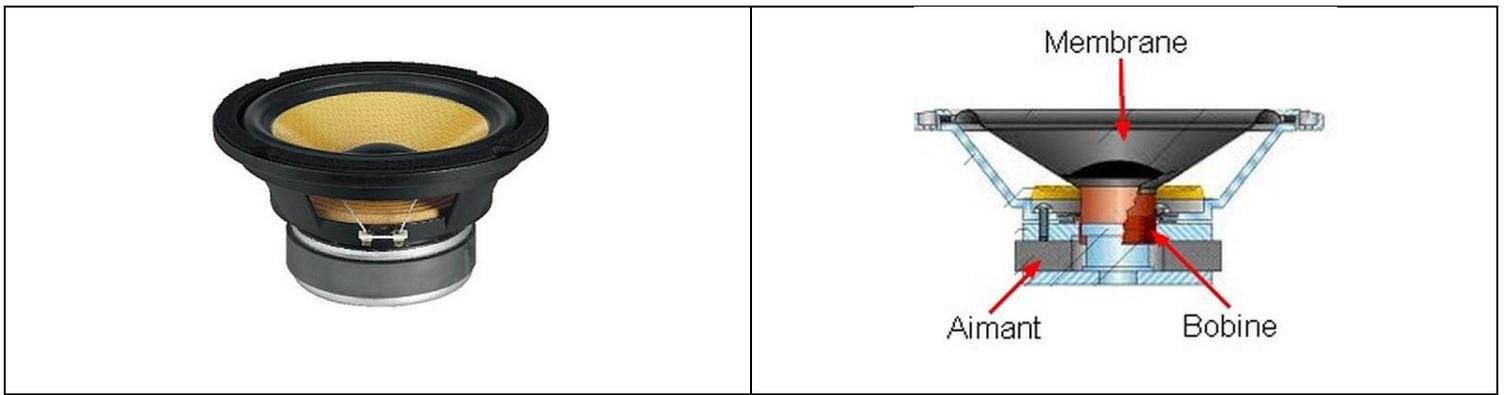
## **Les haut-parleurs**

La reproduction du son nécessite de faire vibrer l'air, souvent par l'intermédiaire de haut-parleurs dont les membranes, aussi appelées diaphragmes, suivent les oscillations électriques qu'on leur soumet. Le principe est de produire le mouvement d'une bobine mobile couplée au diaphragme dans un champ magnétique produit par un aimant. Quand la bobine est parcourue par un courant variable, elle se met à vibrer au rythme du courant, et le diaphragme fait vibrer l'air et produit le son. Pour résumer, un haut-parleur transforme l'énergie électrique en énergie sonore.

Il faut prendre en compte le fait qu'aucun haut-parleur n'est parfait. Chacun a ses propres défauts, notamment sous forme de résonances et de distorsion.

Un seul haut-parleur reproduit mal l'ensemble des sons entre les graves et les aigus. Les membranes de grand diamètre ne reproduisent pas correctement les sons aigus, et à l'inverse celles de petit diamètre reproduisent mal les sons graves.

Pour combiner le meilleur des deux, on a construit des « moniteurs », qu'on appelle aussi des « enceintes », ce dernier mot étant plus utilisé en Hi-Fi.



*Figure 15 La structure du haut-parleur : on voit l'aimant et la bobine mobile à laquelle la membrane est fixée. Autour de la membrane, une suspension souple la maintient en position centrée.*

Les moniteurs sont des caisses rigides dans lesquelles on fixe les haut-parleurs. Ces moniteurs comportent deux haut-parleurs (un pour les graves et un pour les aigus) ou plus, par exemple un troisième pour les médiums. En spécialisant chaque haut-parleur pour ses capacités de reproduction, on peut obtenir un résultat optimum, l'objectif étant de restituer le plus fidèlement possible l'ensemble des fréquences contenues dans les sons qu'on écoute.

## Les casques

Les casques sont des systèmes d'écoute individuelle. Ils sont constitués de petits haut-parleurs dans chaque oreillette. Le principe est le même que celui des haut-parleurs ci-dessus, à part leur taille. Très proches du tympan ils nécessitent peu de puissance pour produire un volume acoustique important. Leur faible dimension et la légèreté de leur membrane leur permet d'avoir une très bonne reproduction des sons depuis les graves jusqu'aux aigus. Leur fidélité dépend de leur construction et leur qualité de reproduction dépend des modèles et de leur prix. Ils ont chacun leurs caractéristiques sonores. Il existe des casques fermés qui isolent des sons extérieurs, et des casques semi-ouverts qui les laissent partiellement passer. Les casques Bluetooth ne sont pas destinés aux professionnels du son.



*Figure 16 Les détails de la construction d'un casque. On voit le petit haut-parleur à droite.*

# Pourquoi s'intéresser au son à l'ILOI ?

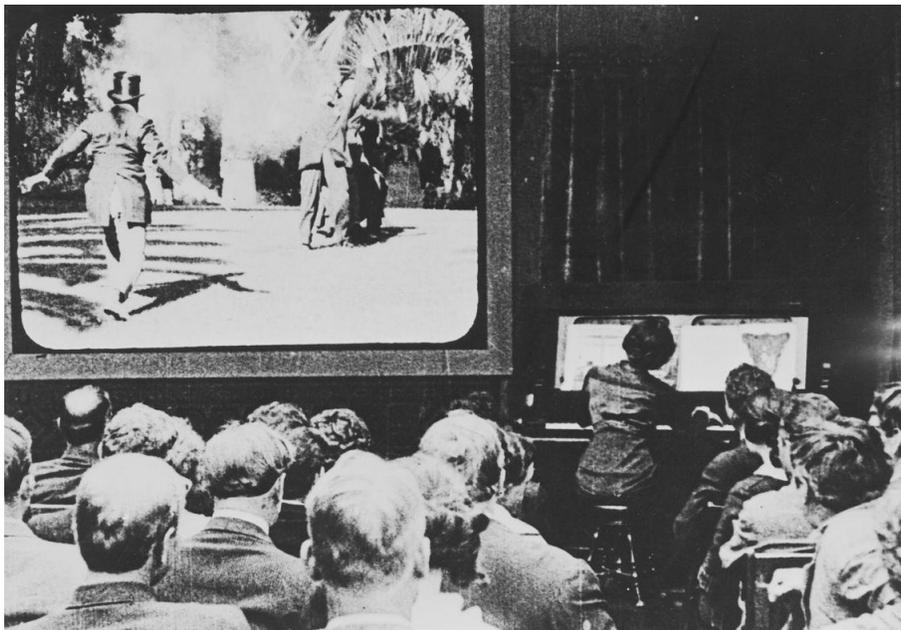
- Le son est indispensable dans la plupart des projets audiovisuels
- Il doit être de qualité
- Pour une bonne qualité il nécessite des connaissances
- De plus en plus, on doit s'en occuper soi-même
- Un son bien géré renforce le visuel et donne un profil professionnel au projet, même avec un visuel imparfait.
- Un son de mauvaise qualité et mal géré diminue l'impact du visuel et donne un profil amateur au projet, même avec un visuel correct
- Un projet terminé sera vu et entendu parmi d'autres et sera inévitablement comparé avec eux.

## L'importance du son dans les réalisations audio-visuelles

Le son est omniprésent dans l'environnement naturel quotidien. Le son qu'on entend ne demande pas une attention soutenue. Or, la majorité des projets audio-visuels sont destinés à une audience captive, dont l'attention est focalisée sur le projet diffusé.

C'est dans ce contexte que le son prend une grande importance. C'est lui qui permet d'augmenter la concentration d'une audience sur le visuel.

Même à l'époque des films muets, on a estimé que le son était nécessaire. Sans autre possibilité à l'époque, on a fait jouer des pianistes ou des organistes pendant les projections.



*Figure 17 À droite, une pianiste de cinéma*

Par la suite, le son a pris une importance de plus en plus grande, que ce soit pour renforcer le visuel, pour lui donner plus d'impact, ou pour donner un style particulier à la production.

La musique joue aussi un rôle qui peut-être essentiel. Certains réalisateurs ont fait preuve d'une grande originalité, en utilisant des musiques inattendues dans certains contextes, comme dans *2001 l'odyssée de l'espace* de Stanley Kubrick, où les musiques de Johann et Richard Strauss apportaient une dimension majestueuse à la vision des vaisseaux spatiaux.

Les dialogues occupent naturellement une place prépondérante par le rôle des acteurs dans le scénario. Leur compréhension doit être optimale.

Par ailleurs, les sons ambiants sont aussi indispensables, puisqu'ils correspondent au réalisme et la crédibilité des scènes.

## Historique du son au cinéma

### Les grands films qui ont changé radicalement le travail sur les sons

- **Star Wars** : création de multiples effets sonores et le début de la spécialité de **Sound Designer**, dont l'importance capitale est maintenant acquise, particulièrement dans le domaine du jeu.
- **Apocalypse Now**, film avec 6 pistes séparées, et la reconnaissance du statut de Sound Designer dans les crédits. Le format son utilisé servira ensuite de base pour les systèmes 5.1 du cinéma et du DVD.
- **Le retour du Jedi** inaugure la certification THX pour la qualité image et son des cinémas qui y adhèrent.
- **Batman Returns** utilisait le **Dolby Digital** pour la 1ère fois, en 92
- **Jurassic Park** en 93 marque le début du format **DTS**
- **Last Action Hero** lance l'utilisation du format **SDDS**
- Le début du format **Dolby Atmos** se produit avec le film d'animation *Brave*

Autres systèmes multicanaux : ceux des salles **Imax** et **Omnimax**.

Le développement des technologies sonores a suivi de nombreuses étapes, d'abord avec la téléphonie, puis avec le Gramophone, l'enregistrement (mécanique ou magnétique) et le disque. Le premier microphone aurait été conçu en 1875.

Un certain nombre de projets ambitieux de cinéma, tirant parti des développements précédents ont été le point de départ d'avancées majeures dans le domaine du son pour les images, essentiellement en cinéma.

### Les débuts

- **En 1925**, Thomas Edison disait :  
« Mon projet était de synchroniser la caméra et le phonographe, de telle façon que les sons soient enregistrés en même temps que les images, et reproduire les deux en harmonie. Nous avons les premières versions des images parlantes dans notre laboratoire il y a 30 ans ».

En fait, Edison avait filmé une personne qui éternuait et avait enregistré le son sur l'un de ses disques sonores. Le film et le son avaient ensuite été rejoués ensemble, et constituaient la toute première réalisation audio-visuelle jamais réalisée. La durée n'était que de quelques secondes, mais l'idée était la base de tous les développements futurs.

- **En 1922**, deux inventeurs allemands ont mis au point le procédé Tri-Ergon. Le son était enregistré sur une piste optique directement sur le film, permettant une synchro parfaite son-image. Le principe : un éclairage modulé servait de base pour une impression sur la pellicule. À la relecture, une cellule photo-électrique retraduisait les variations lumineuses en signaux électriques audio. Ce principe a été utilisé jusqu'à l'arrivée du son numérique sur pellicule.
- **En 1926** la Fox Film a développé un autre système de son sur film : le Movietone, pour produire des bobines d'infos.

Au début des années 20, les laboratoires Bell ont mis au point un système de son sur disques 33 tours, le *Vitaphone*.

- **En 1927**, Warner Brothers produisait « The Jazz Singer » avec quelques paroles synchronisées. Continuant sa progression, Warner Brothers lançait en 1928 le premier film avec dialogues, « Lights of New-York ».

## Les développements ultérieurs

Puis, « Fantasia » de Walt Disney, sorti en **1940**, un film basé sur des animations accompagnées par une musique symphonique, a mis en œuvre l'enregistrement et le mixage multipistes. L'enregistrement a pris 42 jours et a nécessité 150 km de film magnétique. Une piste de clics a été rajoutée par la suite pour servir de guide pour les équipes d'animation. Pour la 1ère fois, un système à 3 haut-parleurs était mis en œuvre à l'avant : Left, Center et Right, et 3 autres à l'arrière.

Fantasia a réellement été le film le plus important dans l'histoire des innovations du son pour le film.

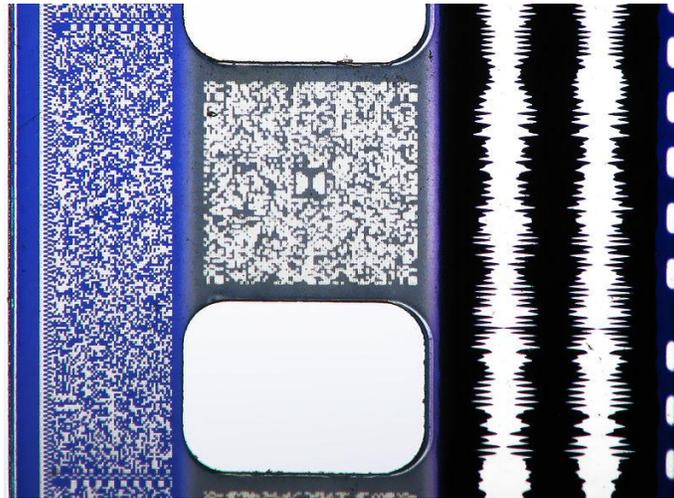
En même temps, la télévision commençait à capturer une audience domestique, Pour conserver un maximum de public, l'industrie du cinéma souhaitait réagir et augmenter l'intérêt des salles de cinéma. De nouveaux standards sont apparus, en 35mm et 70mm, dont le Cinerama (**1952-1962**), très coûteux. Il s'agissait d'une technique de prises de vues par trois caméras synchronisées (fonctionnant avec des pellicules 35mm et projetées sur un écran « extra large » et courbe d'une ouverture de 146°). Le mot Cinerama est une contraction de cinéma et de panorama.

Ensuite, le Cinemascope (**1953-1967**), procédé utilisant l'**anamorphose** pour compacter l'image dans la largeur sur un film 35mm lors des prises de vue, avec un rapport ½, et la décompacter à la projection pour obtenir une image panoramique.

## Le son lui-même a évolué en parallèle avec l'image

Différents standards pour le son en cinéma sont apparus, avec une augmentation progressive de la qualité :

- Dolby Stéréo (**depuis 76**)
- Dolby 6 pistes (L, ML, C, MR, R, S), Dolby Baby-Boom, autre format 6 pistes
- Dolby Split Surround, autre 6 pistes avec canal de graves
- Dolby Digital (**à partir de 92**), début du son numérique au cinéma en format 5.1. en compression AC3, ensuite le DTS (**depuis 93**) en 5.1 aussi, avec un CD spécial à lecture synchronisée avec le film avec compression spéciale, le SDDS (Sony Dynamic Digital Sound) de Sony Columbia Pictures, (**depuis 93**), en 7.1.
- Dolby Atmos, format actuel (**depuis 2012**)



*Figure 18 Les pistes sonores sur une pellicule 35 mm. À gauche la piste SDDS, au centre, la piste Dolby Digital entre les perforations, à droite la piste stéréo analogique avec ses formes d'onde, et à droite la piste de synchronisation du DTS.*

## Le cinéma numérique

- Il a débuté à la fin des années 90
- Il concerne la production, la distribution et la projection
- La pellicule a disparu et a été remplacée par des supports numériques
- Le standard image est 2k (2 048 x 1 080 pixels) ou 4K (4 096 x 2 160 pixels)
- Tournage, montage, postproduction son et image sont en numérique
- La distribution : anciennement par livraison par véhicules des films pellicule sur bobines, puis pour les films numériques par internet, par canaux satellites, ou par envoi de supports physiques (disques durs, disques Blu-Ray)
- Les projecteurs numériques Laser sont utilisés (Barco, Christie)
- L'audio numérique est de très haute qualité : quasiment pas de bruit de fond ni de distorsion, très grande dynamique
- Fréquence d'images de 24 ou 48 par seconde
- Audio jusqu'à 16 canaux WAV 48 ou 96 kHz 24 bits non compressés
- Immersion sonore complète par ajout de la verticalité du son : procédés ATMOS, DTS-X et Auro-3D



*Figure 19 Le projecteur numérique laser 4k Barco SP4K-20 20 000 Lumens*

## L'écoute du Dolby Atmos

- Au cinéma
- À domicile (installation de plusieurs haut-parleurs, ou barres de son)
- Au casque (la version binaurale)



*Figure 20 Une salle Dolby Atmos*



*Figure 21 Une barre de son Dolby Atmos, l'Ambeo Soundbar Max de Sennheiser*

Même si le Dolby Atmos s'est largement répandu les autres standards numériques coexistent avec lui : le DTS-X et l'Auro-3D.

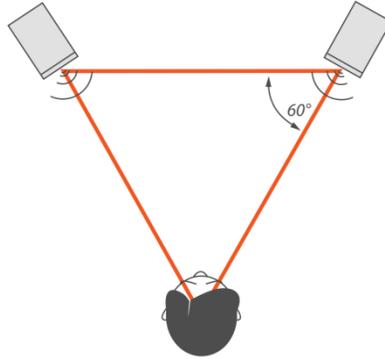
## L'audio immersif

### Les grands principes : Mono/Stéréo/Multiphonie

Dans ses débuts, l'audio était enregistré et écouté en un seul canal. Même avec plusieurs haut-parleurs, le son pouvait provenir de multiples directions mais avec un contenu strictement similaire. C'était la monophonie. Elle ne subsiste que sur les téléphones.

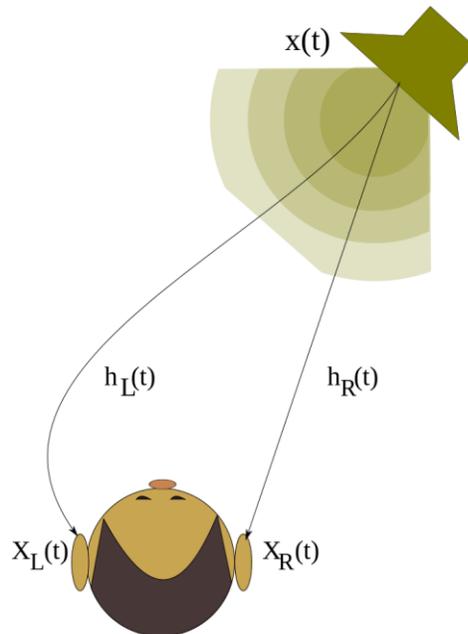
Ensuite, vers la fin des années 50, on a commencé à produire des albums de musique pour lesquels un canal est attribué à la gauche, et l'autre à la droite. C'était la stéréophonie et le début de la spatialisation dont l'objet est de reproduire le son dans un espace plus ou moins large, dans lequel on peut localiser la provenance des sons. En stéréophonie, les sons proviennent de la gauche et de la droite de sources différentes. Les deux canaux s'entendent en se recouvrant partiellement, créant une sorte de panorama

sonore. Dans une position d'écoute standard, en triangle équilatéral, on entend une reproduction équilibrée espacée sur 60 degrés. Le mélange acoustique des sources gauche et droite permet d'entendre un canal central virtuel quand deux éléments sonores sont diffusés simultanément sur les deux enceintes avec le même volume.



**Figure 22 La position d'écoute optimale en stéréo**

Comme vu précédemment, la sensation de provenance du son en stéréo dépend du temps de son arrivée entre les deux oreilles. À 340 m/s le son met 0,6 ms pour aller d'une oreille à l'autre. S'il arrive plus tôt sur l'oreille de gauche, nous l'entendons à gauche. Plus tôt à droite et nous l'entendons à droite. S'il arrive en même temps aux deux oreilles, nous l'entendons au centre. C'est cette perception des différences de temps d'arrivée aux deux oreilles des sons des deux enceintes qui produit la sensation d'espace. Sans la perception à gauche d'une partie du canal droit et inversement par le mélange dans l'air, l'écoute au casque ne respecte pas le champ stéréo du triangle équilatéral créé en studio.



**Figure 23 Ici le son arrive plus vite sur l'oreille de droite : il est perçu plus à droite qu'à gauche**

La stéréophonie est restée dominante jusqu'à aujourd'hui, aussi bien sur les CD, la radio FM, la télévision et le streaming en ligne.

Le principe de la production en stéréo inclut la captation par micros et le mixage bidimensionnel réalisé sur une paire d'enceintes. Lors du mixage, une source monophonique peut être positionnée entre la gauche et la droite grâce à un réglage appelé « panoramique » (en anglais « Pan Pot » ou simplement « Pan ») en affectant sa répartition plus ou moins à gauche ou à droite.

Un mouvement récent vers une spatialisation plus étendue s'est répandu et développé. La multiphonie ou « son surround » désigne une diffusion spatialisée qui entoure l'auditeur. Les sons reproduits parviennent de l'avant et de l'arrière et en hauteur. On parle d'**immersion sonore**.

En résumé :

- La monophonie concerne la production et la reproduction du son avec un seul canal
- La stéréophonie concerne la production et la reproduction du son avec deux canaux, gauche et droite
- La multiphonie concerne la production et la reproduction du son avec des canaux multiples, horizontalement et verticalement

## Historique du son immersif

- Tout a commencé avec la création par la société Dolby d'un principe capable de restituer 4 sources à partir des 2 pistes optiques d'un film 35 mm, le Dolby Stereo. Grâce à un principe d'encodage réalisé en studio, en plus de la stéréo, on produisait un canal « Surround » à partir du mélange des pistes gauche et droite, le **Lt Rt** (Left total Right total). Ce dernier était envoyé en même temps sur les haut-parleurs de gauche et droite de l'arrière. Bien qu'imparfait, cet enveloppement sonore produisait un meilleur relief que la monophonie.
- Le Dolby Stereo étant compatible avec les projecteurs monophoniques précédents a été largement adopté par les exploitants, à partir de 1975.
- En 1977 le Dolby Stereo est popularisé par le film Star Wars utilisant ce standard.
- L'Omnimax, apparu en 1992, a été l'un des précurseurs de l'immersion sonore
- Puis les formats multicanaux (5.1, 7.1) sont apparus avec une séparation complète des canaux grâce aux pistes numériques sur la pellicule 35 mm
- Les pistes Dolby Stéréo sont restées présentes sur les pellicules 35 mm après l'arrivée des pistes optiques numériques. Elles servaient de secours en cas de défaillance de lecture du numérique. Elles ont disparu lors de l'arrivée du remplacement de la pellicule par le cinéma numérique.

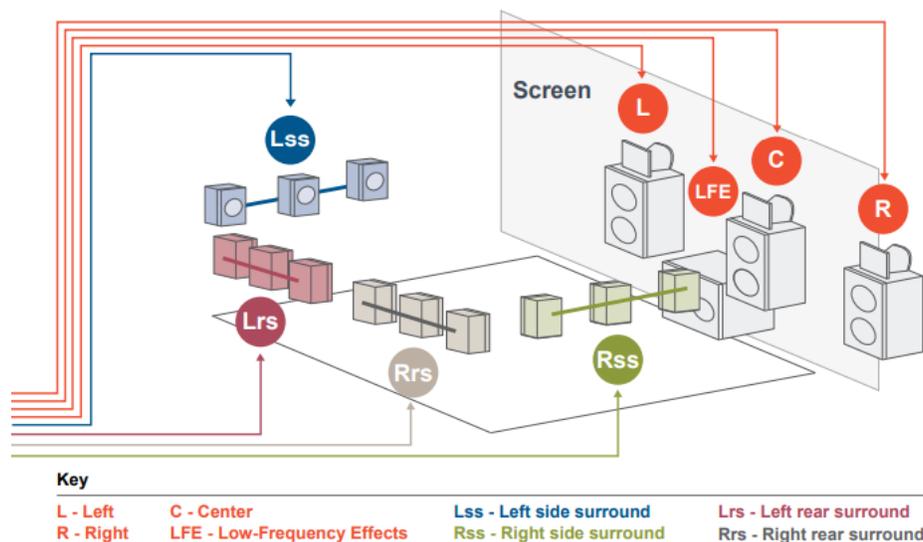


Figure 24 L'implantation d'un ensemble d'enceintes d'une installation 7.1

- Suite à l'arrivée du cinéma numérique, à la dimension horizontale de la diffusion, on a ensuite ajouté le son vertical, pour une immersion encore plus grande dans le son, permettant la perception de hauteur.
- Pour y parvenir, on a augmenté le nombre de canaux. De nouveaux standards à haut-parleurs ont été produits : le Dolby Atmos, l'Auro-3D et le DTS:X

Entre autres, l'évolution majeure du Dolby Atmos est d'une part l'ajout de canaux séparés diffusés en hauteur par des enceintes dispersées dans la salle, et d'autre part le concept d' « objets » sonores (voir plus loin), avec un effet de spatialisation est encore plus réaliste.



## Le son immersif sur haut-parleurs

Avant l'arrivée des formats « Surround », le son au cinéma était donc monophonique et ne comprenait qu'un seul canal, même si plusieurs haut-parleurs étaient utilisés.

L'immersion sonore concerne l'illusion de se situer dans un environnement sonore spatial complet, suggérant la réalité de l'audition humaine dans une ambiance naturelle.

C'est donc avec l'arrivée de la stéréophonie qu'on a commencé à enregistrer et à reproduire une simulation de spatialisation, qui s'est étendue à un espace environnant multidirectionnel.

Dans les cinémas, depuis l'arrivée du Dolby Stereo, on a essayé de simuler notre habitude d'entendre le son dans toutes les directions, et plus seulement à l'avant comme la stéréo s'y limite.

Il existe 3 principes de diffusion spatialisée :

- **Le principe de canaux** : Chaque source est affectée horizontalement à une enceinte (les 6 ou 8 canaux du 5.1 et du 7.1 par exemple) ; c'est le principe immersif qui a prévalu au cinéma jusqu'à récemment où chaque source correspond à chaque enceinte de façon fixe. L'enregistrement des pistes audio peut être constitué de sources monophoniques, stéréophoniques ou multiphoniques. Le mixage final peut combiner les différentes sources de façon statique ou en mouvement, grâce au principe de panoramique.
- **Le principe d'objets** : il est aussi destiné à une diffusion par des enceintes et correspond à leur disposition. Chaque source est affectée à un canal qui contient le son et ses métadonnées de volume et de positionnement avant, arrière et au-dessus. On ajoute donc la dimension verticale à la dimension horizontale, et d'autres canaux. Le principe est très flexible, étant adaptable à toutes les configurations d'installations et reste compatible avec celles existantes. C'est ce choix qu'ont adopté Dolby avec l'Atmos et ses concurrents le DTS-X et l'Auro-3D. Ces sociétés proposent des outils spécifiques pour gérer les objets dans un mixage complet. Celui-ci conserve la compatibilité avec les formats précédents par l'utilisation de « beds » (des ensembles de pistes spatialisées, en standard 5.1 ou 7.2). Avec ce principe, on réalise la position des sons dans l'espace par la gestion des objets. La spatialisation est donc contenue dans un volume demi-sphérique : horizontal et au-dessus.
- **L'ambisonie**, est un principe très différent, non pas basé sur des sources fixes ou des canaux, mais des « scènes » qui représentent une sensation multidimensionnelle. Le champ sonore d'une sphère virtuelle est une représentation du son créée par une combinaison de sources sonores,

appelée **Format-B**, plutôt que par des sources individuelles de haut-parleurs. La scène comprend toutes les informations des objets sonores virtuels : direction, position dans l'espace. Elle peut être modifiée de différentes manières, en temps réel de façon interactive : c'est l'intérêt de l'ambisonie en particulier en réalité virtuelle. Dans ce cas, le capteur de position d'un casque peut associer les mouvements de la tête à la provenance du son. C'est un principe particulièrement flexible qui permet de nombreuses manipulations sur le son perçu.

- L'ambisonie repose sur un principe mathématique complexe, appelé « harmoniques sphériques », d'une précision de localisation d'autant plus grande qu'elle est d'« ordre » élevé. Le 1<sup>er</sup> ordre permet la prise en compte des dimensions horizontale et verticale, représentées par combinaison de 4 sources virtuelles. Avec une ambisonie du second ordre on dispose de 8 sources, donc avec une meilleure précision de localisation, etc.
- La gestion de l'ambisonie se réalise avec des logiciels disposant de cette fonction. **Reaper** de Cockos est l'une des plateformes logicielles à l'avoir intégré dans sa dernière version. Mais de nombreux développeurs ont créé des logiciels autour du principe de l'ambisonie, sous forme de plug-ins ou de logiciels stand-alone.
- Le procédé est très utilisé dans le domaine du jeu vidéo, en partie pour son interactivité.
- L'ambisonie peut s'écouter soit sur un ensemble de haut-parleurs, soit sur un casque stéréo standard ou des écouteurs à bande passante correcte, par encodage de la source et décodage à la lecture, grâce au principe de la HRTF.
- On trouve deux formats pour l'ambisonie : le « Format-A » et le « Format-B » le 1<sup>er</sup> pouvant être transformé en Format-B.
- **Le Format-A n'est pas un format d'écoute. C'est un format intermédiaire d'enregistrement en 4 pistes, destiné à être converti en Format-B.**



Figure 26 La gestion de l'ambisonie dans Reaper de Cockos (Format-B) avec l'Ambisonic Toolkit

Le format sphérique de l'ambisonie permet de gérer les informations en provenance du dessus et du dessous.

## Le son immersif au casque

Grâce à l'étude de notre système de perception, on a développé des procédés de spatialisation pour les casques stéréo. Tout est basé sur le principe binaural, terme général qui veut dire l'écoute par les deux oreilles. Dans son utilisation générale, le terme « binaural » concerne le moyen par lequel notre perception permet la localisation multidimensionnelle des sons.

Aujourd'hui, le mot binaural se réfère à l'ensemble des procédés d'immersion sonore sur casque stéréo.

Les technologies binaurales englobent des méthodes d'enregistrement, de traitement, de synthèse et de reproduction du son qui sont spécifiquement conçues pour produire les propriétés de localisation tridimensionnelle lors du couplage direct diffuseur/oreille..

L'enregistrement binaural et sa reproduction, ainsi que la synthèse binaurale font appel à différents procédés.

Pour reproduire sur une écoute au casque la sensation du son provenant de différentes directions, on dispose de deux options.

- La restitution d'un enregistrement direct par une tête artificielle dotée d'un micro de précision dans chaque oreille. C'est le moyen le plus simple pour reproduire la sensation d'immersion sonore. Il a pourtant ses limites : le coût du matériel, son encombrement et son poids.
- La captation par une combinaison de 4 capsules (voir ci-dessous) en application du principe de l'ambisonie et grâce à l'exploitation de la HRTF. On peut obtenir de bons résultats et une certaine flexibilité lors du décodage des pistes enregistrées.

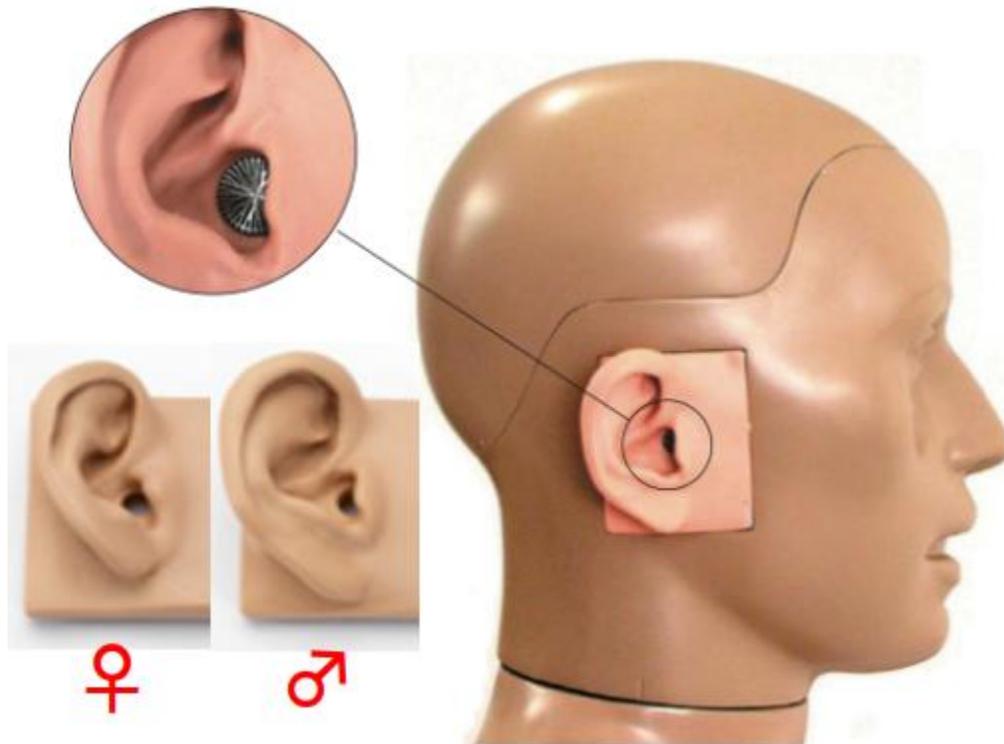
## Le principe de la HRTF

**HRTF** est l'abréviation de Head Related Transfer Function, ou Fonction de transfert en relation avec la tête. Il s'agit d'un système de filtrage issu de l'analyse du résultat d'une captation sonore par une tête artificielle. Celle-ci est disposée dans un ensemble de haut-parleurs qui l'entourent. Le résultat est analysé en fonction de la provenance du son dans toutes les directions : gauche, droite, avant, arrière, dessus, dessous et distance variable, le tout en fonction des fréquences.

Lors des études sur notre perception des sons, on a constaté qu'en dehors des différences de timing entre nos deux oreilles, un son qui leur parvient subit des modifications avant d'arriver à nos tympans. Ce phénomène concerne la forme de la tête et sa position par rapport aux oreilles, ainsi que de la forme des oreilles. Il se produit des réflexions, sources de déphasages, des résonances et des filtrages à certaines fréquences.

Notre sensation de la provenance des sons est liée à la différence de perception d'une oreille par rapport à l'autre qui comprend le décalage temporel, l'écart de volume et les réflexions sur la tête et le torse. L'analyse des sons environnants captés par la tête artificielle prend en compte ces différents facteurs.

On en déduit une modélisation mathématique, traduite sous forme logicielle, de l'écoute du son et de sa traduction physique. Chaque physiologie humaine pouvant avoir des différences plus ou moins importantes par rapport à une moyenne générale, certains logiciels permettent de choisir parmi des variantes de décodages les mieux adaptés à chacun, ou même de déterminer la fonction HRTF de chaque personne pour une sensation immersive optimisée.



*Figure 27 La tête artificielle du mannequin KEMAR, qui a servi à l'élaboration de la HRTF par l'Ircam, avec un micro placé dans chaque oreille. D'autres types de têtes artificielles sont utilisés pour d'autres analyses HRTF.*

Le HRTF englobe donc toutes les caractéristiques de perception de la provenance du son par l'oreille humaine : ses délais d'arrivée entre les deux oreilles, la différence de volume entre les deux, mais aussi le filtrage opéré par les oreilles et les phénomènes de phase qui se produisent par les réflexions sur la tête et sur le buste. Ces phénomènes très complexes combinés forment notre perception multidirectionnelle, bien au-delà de la stéréo.

À noter que grâce à un encodage basé sur le principe HRTF, chaque système de spatialisation sur haut-parleurs est adaptable pour une écoute au casque. Ce principe binaural part du fait que chaque oreille couplée à un écouteur reçoit uniquement le son qui lui est destiné, comme ce qu'on entend dans les conditions d'écoute naturelles.

C'est sur ce principe que repose le principe binaural du plug-in développé par l'Ircam sous le nom de HEar. En interne, dans le logiciel de mixage, par l'intermédiaire d'une matrice d'aiguillage, l'audio est dirigé vers les haut-parleurs virtuels qui simulent l'installation réelle.



Figure 28 La matrice de répartition du plug-in HEar de l'Ircam.

Le principe HRTF reste celui de la synthèse binaurale pour casques. Mais puisque chaque personne a sa propre physiologie, cette fonction de transfert est plus ou moins différente de l'un à l'autre. C'est une moyenne générale qui est utilisée dans les différents procédés actuels utilisant une modélisation unique pour tous. Mais on commence à proposer des méthodes de personnalisation pour optimiser le résultat individuellement. Par exemple, le plug-in Immerse de Steinberg propose de créer son propre profil HRTF grâce à un principe basé sur l'intelligence artificielle. On prend une photo de chaque oreille et le logiciel produit le profil correspondant.

## L'enregistrement binaural direct

Historiquement, les premiers enregistrements en binaural ont été réalisés avec une tête artificielle, comme la KU 101 de Neumann. Un micro est inséré dans chaque oreille et capte les sons avec les transformations qu'ils subissent avant de parvenir à l'intérieur, de façon similaire à l'oreille humaine. Les deux signaux captés conjointement constituent le signal binaural. À l'écoute de celui-ci avec un casque, on entend un son proche de celui de notre audition volumétrique.



Figure 29 La tête artificielle KU 101 de Neumann

## L'enregistrement binaural en Format-A

Il s'agit d'une toute autre méthode de captation immersive. On capte le son avec 4 micros cardioïdes orientés dans 4 directions. Les 4 signaux enregistrés constituent le Format-A.

Ce format brut est un groupe de 4 signaux audio, non directement écoutable. Pour ce faire, on utilise une application de transcodage qui transforme le Format-A en Format-B ambisonique en deux canaux, comme pour la stéréo.



*Figure 30 Le micro Ambeo VR de Sennheiser et ses 4 capsules*



*Figure 31 L'enregistreur ambisonique Zoom H3-VR et ses 4 capsules.*

## La synthèse binaurale

Grâce à l'exploitation du principe HRTF, à partir de logiciels dédiés, on peut créer des espaces ambisoniques artificiels, en positionnant ou en mettant en mouvement en 3D des sources virtuelles à volonté (voir par exemple la fenêtre de Reaper).

Avant d'aller plus loin, pour prendre un exemple simple, à partir d'un mixage en 5.1 ou 7.1 réalisé dans un environnement de haut-parleurs on peut réaliser par traitement numérique un fichier binaural en deux pistes. Dans ce cas, on n'aura bien sûr une transcription binaurale du fichier d'origine, donc pas de verticalité. Pour obtenir cette dernière, on a recours aux techniques issues de la HRTF.

Dans le sillage du Dolby Atmos au cinéma, considérant le potentiel commercial, Dolby a étendu le concept de l'écoute immersive sur haut-parleurs à celle au casque, le **Dolby Atmos Binaural** comportant la dimension verticale.

Par ailleurs, entre autres, Sony a développé le 360 Reality Audio, utilisable aussi bien sur casques que sur haut-parleurs.

La synthèse binaurale concerne particulièrement la création des jeux vidéo. Les fonctions immersives sont intégrées de façon native dans **Unreal Engine** et **Unity** par la prise en compte de l'ambisonie de 1<sup>er</sup> ordre. De même, les middlewares comme wWise d'Audiokinetic ou FMOD intègrent la gestion de l'ambisonie. Les applications qui intègrent l'interactivité directionnelle font partie de l'intense développement autour de l'immersion sonore au casque dans le jeu vidéo.

Mentionnons l'Ambisonic Toolkit, un ensemble de logiciels libres compatibles avec Reaper pour créer des projets ambisoniques imaginaires, *Binauralizer* de *Noisemakers*, et *Dear VR* issu des travaux de Sennheiser.

## Analogique ou numérique ?

L'audio a été analogique à ses débuts. Tous les matériels ont été construits sur ce principe de variations électriques continues: le courant électrique de l'audio varie constamment et peut prendre toutes les valeurs, entre 0 (absence de courant) et un maximum. C'est ainsi qu'ont été conçus micros, haut-parleurs, enregistreurs et reproducteurs, amplificateurs et équipements de traitement sonore, y compris les premiers synthétiseurs.

La différence fondamentale de l'audio numérique (on dit aussi « Digital », l'équivalent anglo-saxon) par rapport à l'audio analogique est d'ordre électrique. Alors qu'en analogique les variations sont progressives, en numérique, on n'a que deux valeurs de courant, une minimum, à laquelle on donne le chiffre « 0 » et une maximum, appelée « 1 ». Les zéros et les uns du numérique sont donc une représentation de ces deux états du courant. Exemple, on bascule un interrupteur, le courant passe, une lampe s'allume. On le bascule dans l'autre sens, le courant ne passe plus, la lampe d'éteint.

Depuis l'arrivée du numérique, l'audio analogique n'a pas disparu, mais il s'est raréfié. Il est maintenant toujours utilisé par une catégorie d'ingénieurs du son qui lui trouvent des qualités plus ou moins subtiles qu'ils apprécient. En parallèle, il reste aussi des adeptes de la pellicule argentique...

La très large adoption du numérique par les professionnels nous a incité à lui accorder une place bien plus importante que celle de l'analogique. Cependant, il subsiste par nécessité des liaisons analogiques, notamment pour les micros et les moniteurs.

## L'audio analogique

- C'est le domaine du **signal analogique**
- Il est composé de **variations électriques progressives et continues**, en correspondance directe avec celles du son.
- Exemple : un rampe d'accès dont la pente est progressive; on peut mesurer la hauteur par rapport au sol à n'importe quel endroit
- Le signal analogique s'enregistre et se traite directement par des **circuits analogiques**
- Il se transmet par des câbles ou par modulation/démodulation HF (Hautes Fréquences)
- Ses caractéristiques s'expriment en **dB (les décibels)** pour le volume audio électrique, en **Hz (Hertz)** pour la fréquence, et en **V ou mV (Volts ou millivolts)** pour le volume électrique
- Sa connectique et son câblage sont spécifiques (voir en fin de document).

L'interconnexion entre les différents équipements a fait appel à des câbles de deux sortes :

### **Le câble asymétrique.**

Il est composé d'un fil central, appelé « âme », entouré d'une tresse métallique, le « blindage », qui sert à bloquer les rayonnements électromagnétiques par sa mise à la masse (le point de référence électrique).

### **Le câble symétrique**

Il est composé de deux fils centraux pour le transport du signal, entourés d'une tresse métallique, le « blindage », qui sert aussi à bloquer les rayonnements électromagnétiques.



*Figure 32 À gauche un câble symétrique et à droite un câble asymétrique.*

Avec une liaison symétrique, on dispose d'une double protection contre les rayonnements électromagnétiques, d'une part grâce au blindage et d'autre part avec un principe astucieux de déphasage (voir plus haut ce terme). Sur l'un des deux câbles, le signal est inversé et donc transmis en opposition de phase par rapport à l'autre. À l'extrémité on fait la différence entre les deux signaux en les inversant à nouveau, ce qui permet de récupérer le signal d'origine. Mais les signaux parasites n'étant pas mis en hors-phase sont captés simultanément sur les deux câbles et arrivent en même temps à l'extrémité. Puisqu'on effectue une différence, ils sont annulés.

**Le câble de connexion aux enceintes non amplifiées** (maintenant essentiellement remplacées par les enceintes amplifiées) assure la connexion entre un amplificateur et une enceinte. Il n'a pas besoin

d'être protégé des interférences mais doit avoir une section suffisante pour assurer le mieux possible le passage de l'intensité nécessaire pour les haut-parleurs.

**Le câble de connexion aux enceintes amplifiées** est un câble audio blindé normal, en principe symétrique (mais souvent aussi en asymétrique), qui transmet un signal non amplifié, l'amplification étant interne aux enceintes.

Voir la partie « Connectique » plus loin.

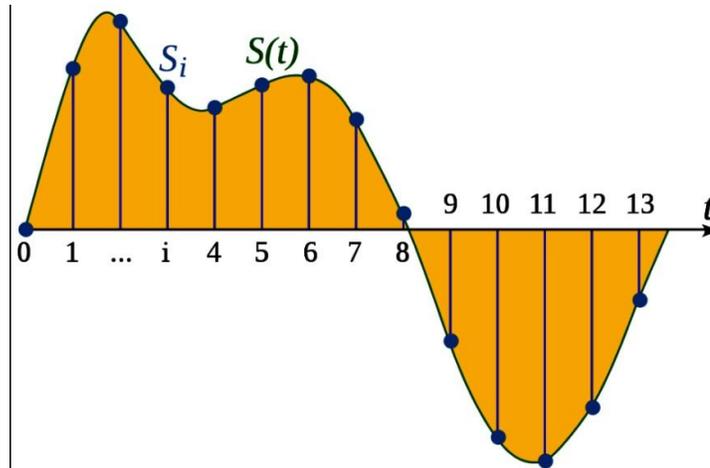
## L'audio numérique

En audio, le terme « numérique » concerne la représentation du son sous forme séries de nombres, ainsi que tout ce qui concerne sa gestion. Ces nombres sont issus d'une transformation des variations analogiques progressives du son en successions rapides d'alternances « tout ou rien » (les 0 ou 1) successives. En effet pour entrer dans le domaine du numérique, le signal analogique doit être converti. La captation en numérique d'un signal analogique s'effectue avec un **convertisseur analogique/numérique (A/N)**. À l'inverse, on transforme un signal numérique en analogique avec un **convertisseur numérique/analogique (N/A)**.

- C'est le domaine du **signal numérique**
- Il est composé de variations discontinues, résultat d'une mesure d'échantillons très rapprochés de l'amplitude instantanée du signal audio analogique, mesure exprimée sous forme de codes binaires composés uniquement des valeurs 0 et 1 : dans les circuits électroniques numériques quand le courant ne passe pas on lui donne la valeur 0, et quand il passe c'est la valeur 1.
- Exemple similaire à la notion d'échantillons numériques : un rampe d'accès dont la pente est en escaliers; on ne peut mesurer la hauteur par rapport au sol que sur chaque marche
- L'information numérique de base est constituée de deux états « tout ou rien » électriques qui constituent l'unité d'information numérique de base, qu'on appelle **le bit** qui a pour valeur 0 ou 1.
- Les bits sont souvent groupés par blocs de 8, les **octets**, appelés **Bytes** (lettre **B**) ou **kilobytes** (lettre **kB**) : par exemple 2 octets pour 16 bits (2x8 bits) ou 3 octets pour 24 bits (3x8 bits)
- Le **débit binaire** est le nombre de bits transmis par seconde quand ils sont transmis l'un après l'autre « en série ». Il s'exprime en bit/s ou b/s ou kb/s (kilobits par seconde), à ne pas confondre avec les B/s ou kB/s. Exemple : 8 **kb/s** correspondent à 1 **kB/s**
- Le signal numérique s'enregistre et se traite par des **circuits numériques**
- Il se transmet par des câbles filaires ou optiques ou par Internet, ou par HF (Wi-Fi, ou Bluetooth)
- Ses caractéristiques : la **fréquence d'échantillonnage** en **Kilohertz**, sa précision en volume en **nombre de bits** liée à son **relief dynamique**, son volume en **dBFS (decibels Full Scale)** avec pour valeur **maximum 0 dBFS**
- La fréquence d'échantillonnage définit la capacité à représenter correctement les fréquences aigues avec précision. Deux fréquences standard se sont imposées : 44 100 échantillons par seconde (**44,1 kHz**) pour le CD et 48 000 par seconde (**48 kHz**) pour l'image animée. Le théorème de Nyquist indique que pour représenter correctement un signal numérisé d'une certaine fréquence, on doit l'échantillonner à au moins le double de cette fréquence. Par exemple, pour correctement capturer un signal à 20 000 Hz, il faut au minimum l'échantillonner à 40 000 Hz. Au besoin, on peut facilement convertir une fréquence d'échantillonnage en une autre.
- Le relief dynamique dépend de la précision de la mesure des échantillons : plus on affecte de bits à cette mesure séquentielle, plus la précision est grande, plus les faibles niveaux seront correctement représentés. En 8 bits, on dispose de 256 valeurs pour représenter une amplitude,

en **16 bits** 65 536 valeurs (un équivalent d'une dynamique de 96 dB), ou en **24 bits** 16 777 216 valeurs (un équivalent d'une dynamique de 144 dB). Sauf pour le CD en 16 bits, le standard 24 bits s'est imposé dans l'audio numérique. Le **32 bits** est parfois utilisé. Il présente un avantage : grâce à sa dynamique extrêmement élevée, il permet de moins se soucier du réglage de volume lors des enregistrements.

- La numérisation des signaux électriques est réalisée par un convertisseur analogique/numérique et sa restitution audible s'obtient par un convertisseur numérique/analogique
- La connectique des signaux numériques et son câblage lui sont spécifiques



*Figure 33 Le processus de numérisation : les segments verticaux représentent les mesures successives rapprochées de l'amplitude*

## Les câbles du numérique

Les liaisons numériques s'effectuent par des câbles qui ressemblent à ceux de l'analogique : simple câble blindé en semi-professionnel (standard S/PDIF) ou paire de câbles blindés (standard AES/EBU). Mais les câbles pour le numérique ont une particularité : les très hautes fréquences qu'ils transportent nécessitent qu'ils répondent à un critère d'impédance spécifique. Les câbles audio risquent de poser des problèmes, surtout avec une longueur conséquente.

## La connectique du numérique

Là aussi, on trouve des similitudes avec les connecteurs de l'analogique, sauf pour les liaisons optiques (voir en fin de document).

## L'arrivée d'Internet et la réduction de débit

Pour le transport des fichiers audio, la diffusion du son sur les premiers réseaux à débit limité a déclenché le développement des formats de compression de débit. Il s'agit de procédés de **compression psychoacoustique** (à ne pas confondre avec la compression de dynamique), basés sur le principe de masquage des sons faibles par les sons forts par l'audition humaine. Le signal à transmettre est plus simple et le résultat permet de diviser par 10 le débit d'origine, tout en conservant une qualité audio correcte.

Le **MPEG Moving Pictures Experts Group** définit les normes de compression et de diffusion de l'audio et de la vidéo :

- MPEG-1 Layer III (le **MP3**) : encodage mono ou stéréo à 32 kHz, 44,1, ou 48 kHz avec un débit de 32 à 448 kbps en fonction de la qualité, souvent en 44,1 kHz/128 kbps avec une qualité moyenne..
- MPEG-2 : compatible MPEG-1, ajoute le format 5.1
- Le MPEG4 est un format compacté « conteneur », qui inclut vidéo et audio

Autres formats de compression de débit audio : **AC3 (Dolby)**, **SDDS (Sony)**, **DTS**, **AAC (Apple)**, **WMA (Windows)**

Avant l'arrivée du tout numérique dans le cinéma, on a utilisé ces types d'encodage pour le son des films.

## Le stockage de l'audio analogique

- Il utilise la bande magnétique, un ruban de film plastique recouvert d'une fine couche magnétique.
- Les réglages du magnétophone sont délicats, en fonction des bandes. Un entretien régulier des magnétophones est nécessaire.
- Les bandes magnétiques ont évolué progressivement, depuis la bande stéréo ¼ fr pouce jusqu'à la bande multipistes 24 pistes 2 pouces comme celle ci-dessous.



*Figure 34 Une bande analogique multipistes*

*La bande magnétique ci-dessus de 2 pouces de large (env. 5 cm) de 4,3 kg, (coût plus de 300 Euros) de 26 cm de diamètre peut enregistrer 27 minutes d'audio sur 24 pistes, soit 6 h 45 mn de son monophonique (76 cm/s). L'équivalent, numérisé en haute qualité tient sur une clé USB de 64 GB.*

## Le stockage de l'audio numérique

- Après avoir utilisé la bande magnétique numérique (de 2 pistes à 32 pistes) on est passé d'abord aux disques durs, puis aux disques SSD, aux mémoires Flash (EEPROM), aux clés USB, aux disques optiques (CD, DVD), et plus récemment au Cloud.
- Il est essentiel de penser que **TOUT** support numérique peut devenir totalement inutilisable, soit en cas de défaut d'une bande magnétique numérique, soit en cas de panne de composants internes, que ce soit un disque dur tournant, un SSD, une clé USB ou une carte Flash. Les conséquences peuvent être catastrophiques, surtout quand de nouveaux enregistrements image ou son s'avèrent impossibles.

Il est donc impératif de réaliser dès que possible une sauvegarde image et son sur des supports conservés à un endroit différent de celui des originaux.

## L'audio dans l'ordinateur

- Un ordinateur comporte un convertisseur analogique/numérique et un convertisseur numérique/analogique :
- Il permet la captation des sources externes et la restitution en audio sur ses haut-parleurs ou au casque de toutes les informations qui transitent en numérique dans ses circuits internes : enregistrements, lecture, streaming, traitements, synthèse de sons

## Les interfaces audionumériques externes

Elles sont complémentaires à l'audio interne de l'ordinateur.

- Leur qualité audio est meilleure que celle de l'ordinateur
- Elles ont une ou plusieurs entrées pour des micros de studio
- Certaines permettent d'enregistrer plus de 2 canaux simultanément
- Elles se connectent à l'ordinateur en USB ou Thunderbolt
- Elles ont souvent une connexion numérique.

## Les objectifs du son pour les images

- Renforcer le visuel et son impact
- Augmenter la crédibilité des images
- Mettre en condition le spectateur
- Augmenter la concentration de l'audience
- Donner un style défini à une production
- On utilise souvent l'expression symbolique « Habillage sonore », qui montre que le son peut être considéré comme une sorte de costume adapté à chaque production
- La musique joue un rôle très important par son choix, son originalité et son adaptation à un scénario. Elle est en relation avec les symboles culturels liés à la musique et nos codes mentaux.
- Mettre en condition lors de changements de scènes par l'ambiance sonore (tension, suspense, joie, tristesse)
- Servir de lien entre des scènes et produire une sensation de continuité pour des scènes tournées dans des conditions très différentes
- Maintenir l'attention dans les scènes sans action

## La mise en condition du public

Le son a un pouvoir sur l'esprit, que les réalisateurs mettent à profit pour conduire le public dans des états émotionnels déterminés, lui provoquer des émotions plus ou moins fortes, et aussi dans le but de garder les esprits en éveil et éviter la monotonie. Un scénario bien écrit est conçu pour comporter non seulement des parties calmes en alternance avec d'autres plus ou moins intenses, mais aussi pour créer des ruptures images et sons lors du montage.

Par exemple, une scène dans une ambiance feutrée d'intérieur peut passer en cut sur une autre très bruyante et provoquer un choc émotionnel. Le réalisateur dispose donc de ce moyen pour conditionner le spectateur et manipuler son état d'esprit pour maintenir son intérêt.

Le son peut contribuer aux sensations de calme, d'intimité, ou d'excitation, de suspense, d'impact, en étant au service du scénario. Quand ils sont continus, les sons aigus créent généralement du suspense et de la tension, alors que ceux qui sont graves suggèrent la dignité, la solennité, la dramatisation.

Par ailleurs, les sons forts sont imposants, intenses et menaçants, alors que les sons faibles sont perçus comme délicats, subtils, légers, hésitants.

Le tempo rapide d'une musique provoque une tension plus grande (dans une séquence d'action par exemple), alors que les tempos lents évoquent le calme, la sérénité, ou la tristesse.

Les effets sonores peuvent être utilisés pour transmettre un message par suggestion, sans nécessiter une illustration visuelle. Exemple, un coup de feu seul peut suffire pour comprendre ce qui s'est passé, sans l'image correspondante.

Les réalisateurs talentueux osent utiliser des accompagnements sonores sans relation avec les images et jouer sur les contrastes ainsi obtenus.

## Réalisme et précision du son

Deux critères se côtoient constamment en matière de son pour l'image : le réalisme du son et sa précision.

**On doit « croire » à ce qu'on voit, être dans le récit. C'est le principe de base dans le rôle du son en lien avec les images.** Contrairement à notre expérience du vécu, où le son et l'image sont indissociables et font partie de notre perception globale, vue et audition ensemble, la vision d'images sur un écran est une expérience subjective. On est dans l'illusion, dans le virtuel, mais les capacités d'imagination de notre esprit nous entraînent dans une sorte d'illusion de réalité.

Lors de la prise de son, on aura à l'esprit le besoin de vérité sonore dans les situations de tournage. Pour autant, en conditions réelles, en fonction des difficultés rencontrées, on pourra être amené à trouver des compromis entre vérité et qualité. On conçoit par exemple que la prise de son de comédiens sur le terrain est l'un des meilleurs moyens pour obtenir cette crédibilité, puisque son et image sont en parfaite correspondance.

La crédibilité subjective en face d'une image sonorisée n'étant que partielle, elle nécessite une participation augmentée de l'audition, pour renforcer la crédibilité des images.

Pour obtenir ce résultat, on cherche à obtenir la meilleure présence possible des éléments sonore. Dans notre vie, nous sommes plus attentifs aux sons proches. Or, ceux-ci ont pour caractéristique globale d'avoir un contenu en fréquences plus large que celui des sons distants. On appelle « présence » cette sensation de proximité des sons. Pour créer ou renforcer cette sensation de présence, on dispose de plusieurs moyens, plus ou moins compatibles avec le réalisme souhaité.

# La réalité sonore augmentée

## Qu'est-ce que la réalité sonore ?

C'est l'ensemble des sons qu'on entend dans la vie quotidienne : tous les sons quels qu'ils soient et d'où qu'ils viennent.

Dans tout projet visuel où le son est présent, le public n'entendra jamais une réalité sonore, mais plutôt une simulation de la réalité. Il faut bien comprendre que même des sons réels, quand ils sont perçus au travers d'un système de reproduction sont modifiés par celui-ci, qui ne peut restituer fidèlement le véritable son réel. Sans oublier le fait que les sons diffusés avec un film sont passés par de multiples captures et traitements ! On entendra donc toujours autre chose que la réalité. C'est notre cerveau qui nous permet de penser qu'on est dans le réel.

Ce contexte d'écoute artificiel incite à concevoir le son comme un élément créatif, dépassant la réalité et doté d'une dimension fictive propre à servir le projet.

À partir de cette constatation on peut considérer que ce n'est pas la recherche d'une réalité sonore illusoire qui importe le plus, mais plutôt celle des moyens pour obtenir l'efficacité de l'impact des sons lorsqu'ils sont reproduits dans les conditions d'écoute prévues : téléphone, ordinateur, télévision, casque stéréo, cinéma.

En tant que réflexion globale sur l'usage du son en audiovisuel on gardera donc à l'esprit que le son à l'image n'a pas nécessairement pour but de respecter fidèlement la réalité, sauf dans les reportages ou les documentaires. Quand on est dans le cadre d'une histoire, d'une fiction, on emmène le spectateur en dehors de la réalité qu'elle soit sonore ou visuelle. De plus, tout procédé sonore qui permet d'augmenter le vécu du spectateur et l'impact des images favorise est à envisager.

## Augmenter la réalité

Par conséquent, même si le storyboard nécessite un environnement sonore très proche de ce qu'on entend dans la réalité, cette réalité peut avantageusement être « augmentée » avec des sons hyperréalistes s'ils peuvent amplifier le propos visuel. En effet, selon les conditions, les sons captés sur le terrain peuvent être assez peu efficaces, et quand ils sont insuffisamment présents ils doivent être utilement remplacés par d'autres.

Il s'agit aussi d'augmenter le volume des parties faibles des dialogues en se rapprochant ainsi virtuellement de la source sonore, soit lors de la prise de son, soit ultérieurement lors de la postproduction en gérant les volumes lors du mixage.

On a donc déduit qu'au lieu de chercher à coller à la réalité, celle du son naturel tel que nous le percevons en condition réelle, on utilise le son pour augmenter l'intérêt ou l'impact des images, quitte à prendre de grandes libertés par rapport à cette réalité. Il s'agit d'aider le cerveau à croire à ce qu'il voit. Par exemple, un personnage qui chute sur le sol sera accompagné par un bruit sourd qui augmentera la sensation de réalité de cette chute, même si cette chute ne produit réellement qu'un faible bruit.

Une partie du travail de création d'effets sonores hyperréalistes peut être confiée à un bruiteur, qui recrée des sons plus évocateurs et plus présents que les sons enregistrés lors du tournage, par une capture à proximité des accessoires de sa collection, en studio.

Par exemple, à cause des contraintes de placement des micros, la captation du son des pas d'un comédien enregistrés lors d'un tournage n'a qu'un faible intérêt sonore et peu de présence. Dans ce cas, les bruiteurs les recréent en marchant sur place dans un studio sur différentes surfaces. On connaît les exemples typiques du son des bruits de pas dans la neige obtenus en piétinant sur de la fécula de pommes de terre, ou le son des sabots de cheval avec des noix de coco, ou encore le bruit du feu avec de la cellophane froissée.

En dehors du remplacement des sons du tournage, l'ajout des bruits qui correspondent aux images est indispensable quand les dialogues sont remplacés en studio.

Lors de l'enregistrement des bruitages, la distance de captation jouera un rôle important dans la crédibilité et l'utilité des sons enregistrés : trop près et l'effet sera caricatural, trop loin et il manquera de précision et sera affecté par l'acoustique environnante. Une distance de 40 à 80 cm entre le micro et la source sonore sera une base de travail.



*Figure 35 Un bruiteur dans son studio*

Pour la capture des bruitages, on utilise souvent un micro hyper-cardioïde similaire à ceux qu'on place sur une perche habituellement. Sa directivité permet d'être très focalisé sur la source sonore et d'éviter au maximum de capter les réflexions de la pièce. Mais rien n'empêche d'utiliser un modèle cardioïde, surtout dans une acoustique neutre.

Cette production en studio des sons du tournage est appelée **Foley** aux Etats-Unis, du nom de celui qui a été à l'initiative de ce travail, James Foley.

L'arrivée du sampler (en français échantillonneur), machine ou logiciel permettant de capter des sons en mémoire (les échantillons, appelés « samples ») et de les déclencher par un clavier MIDI a d'abord facilité le placement d'effets sonores en synchro dans un film. La lecture instantanée de sons préenregistrés a permis l'utilisation de banques d'effets sonores sous issues de catalogues. Ensuite, l'intégration directe de fichiers audio dans les logiciels audio/vidéo a rendu facile leur placement et leur calage précis par rapport aux images dans la timeline (DaVinci Resolve 15 Fairlight, ProTools Avid, Nuendo Steinberg, Logic Pro, Reaper Cockos). Les lecteurs logiciels de samples (Kontakt de Native

Instruments, UVI Engine) peuvent aussi être utilisés pour déclencher les sons de leurs bibliothèques par des événements MIDI de pistes destinées à cette fonction.



Figure 36 La fenêtre principale du lecteur d'échantillons Kontakt 6 de Native Instruments qui comporte des fonctions de transformations des sons.

Des logiciels apparaissent aussi pour produire des effets sonores, comme celui-ci-dessous. Le métier de bruiteur ne fera peut-être plus partie que de l'histoire du cinéma dans l'avenir...



Figure 37 Spécialisé dans la re-création des bruits de pas, EDWARD ULTIMATE utilise le lecteur de samples Kontakt de Native Instruments pour les déclencher en synchro avec l'image

Comme dans d'autres domaines, de nouveaux moyens informatiques se développent. On voit arriver des logiciels qui permettent de créer des effets sonores par synthèse. Un exemple : Tsugi Studio propose un ensemble de logiciels qui sont conçus pour générer toutes sortes d'effets et leurs mouvements dans l'espace. Les effets sonores sont produits par des algorithmes composés d'« objets » qu'on peut assembler à volonté dans la version de base, une sorte de synthétiseur d'effets. Ce mode logiciel de création sonore, appelé « Audio Procédural » ne fait pas appel à des échantillons d'audio, mais plutôt à des modèles mathématiques. Par conséquent, au lieu de disposer d'une matière

sonore figée, on peut rapidement générer les effets qui correspondent aux mouvements de l'image, en manipulant divers paramètres en temps réel.

Le travail de sound design et de bruitages évolue...

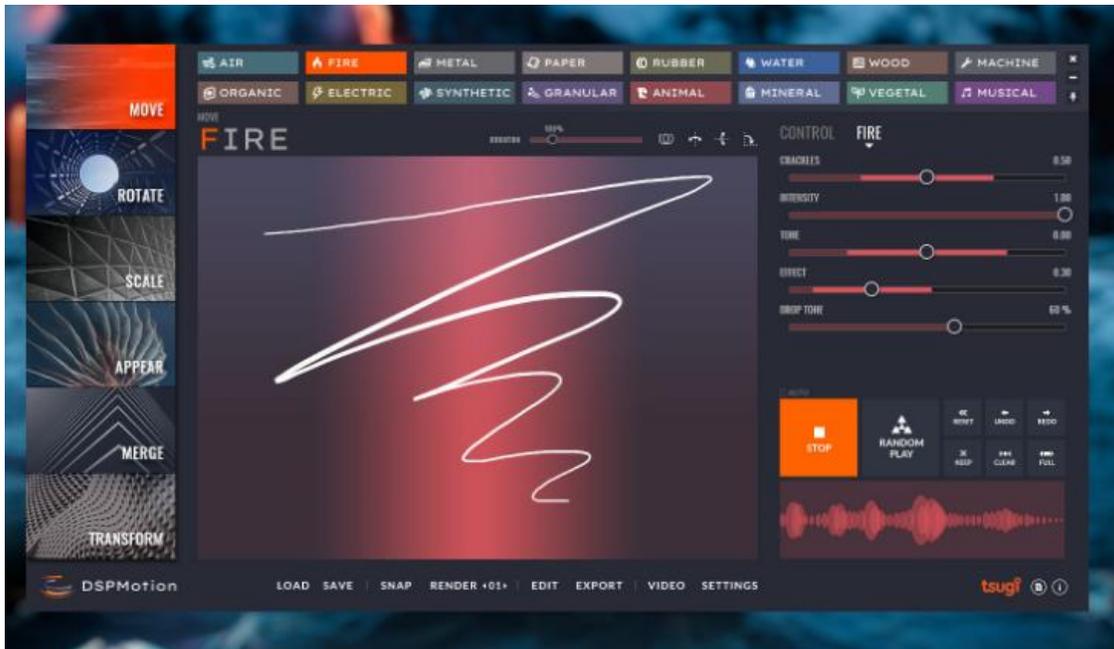


Figure 38 L'une des pages de DSP Motion de Tsugi. Les menus permettent de choisir différentes catégories d'effets et la fenêtre centrale de dessiner leur mouvement, en synchro avec la vidéo.

## Le rôle de la musique

**La musique joue un rôle très important ou même essentiel dans les projets visuels. Sa puissance sur les émotions est reconnue depuis longtemps. De plus, elle permet de créer des atmosphères ou des ambiances. Elle ne doit pas être considérée comme accessoire à une production, mais plutôt comme son complément indispensable.**

À ce titre, il est fortement conseillé aux stagiaires de l'ILOI de s'en préoccuper, bien avant la finalisation de leur projet, à l'étape de la pré-production (voir plus loin).

**La musique a le pouvoir de donner un sens aux images ou de l'augmenter et d'influencer leur compréhension. La perception d'une scène peut changer radicalement en fonction de la musique.**

Elle est capable de suggérer ou augmenter les sentiments d'action, de dramatisation, de tension, de suspense, de terreur, de violence, de romantisme de bonheur, etc.

Elle peut servir à la création de sensations de calme, d'intimité, d'excitation, d'énergie, de suspense, de tension, d'impact, au service du scénario, grâce à certains paramètres, les ingrédients qui constituent ses effets sur l'esprit : le rythme, l'intensité, l'harmonie ou la dysharmonie, le style, les instruments utilisés, le calage des éléments musicaux avec les images, l'originalité.

La musique n'est pas nécessairement toujours mélodique. Dans une scène à suspense on peut simplement utiliser une rythmique de percussions.

Lors du mixage, on décide du volume de la musique. Elle peut être présente, tout en restant discrète pour ne pas nuire à la perception de l'essentiel, ou bien au contraire puissante par moments.

Le choix des musiques ou des options de composition nécessite un dialogue entre le compositeur et le réalisateur : la musique est au service du scénario.

La musique composée en fonction des images nécessite un talent particulier et une large culture musicale. Détaillons les paramètres musicaux en fonction de leur pouvoir psychologique.

## **Les pouvoirs de la musique sur l'esprit**

### **L'harmonie**

Les musiques harmonieuses sont directement acceptées par notre esprit. Elles jouent leur rôle sans questionnement inconscient.

### **La dysharmonie**

Les musiques disharmonieuses sont dérangeantes. Elles produisent un sentiment de malaise, notre cerveau ayant des difficultés à les comprendre.

### **La tonalité**

Les tonalités majeures sont associées à la joie, la puissance, l'extraversion, alors que les tonalités mineures, sont associées à la tristesse, la beauté, la faiblesse, l'introversion.

### **La hauteur du son**

Produit une tension avec les sons aigus continus

Suggère dignité, solennité, dramatisation, danger avec les sons graves

### **Le volume du son**

Lié à la délicatesse, la subtilité, la légèreté, l'hésitation avec les sons de faibles volumes

Les sons forts sont imposants, intenses, menaçants

### **Le tempo (BPM)**

Un tempo lent produit des sensations de calme, sérénité, tristesse, ampleur

Un tempo rapide évoque action, tension, énergie

### **Les sons répétés rapidement**

Ils donnent une sensation de suspense

### **Les sons espacés**

Donnent de l'anxiété, de l'angoisse

### **Les ponctuations sonores**

C'est le point d'exclamation sonore. Placé après un mot d'un dialogue, il renforce son importance émotionnelle.

Quand il a acquis une bonne connaissance du scénario, le compositeur de musique à l'image a intégré tous ces éléments inconsciemment. Il a déjà une idée de l'esprit dans lequel il va travailler, en fonction des images.

## **Le pouvoir de suggestion de la musique :**

Un exemple : un plan large : une maison isolée dans la campagne.

Que se passe-t-il à l'intérieur ? Une fête de famille ? Une scène d'amour ? Un crime ? Un dîner calme ?

Pendant l'image sur ce plan, la musique choisie a le pouvoir de diriger notre imagination et nous mettre en condition pour créer le sentiment requis.

Le choix de la musique permet de mettre le spectateur en condition pour le préparer à la scène suivante. Une musique romantique va faire penser à une ambiance heureuse, alors qu'une autre, dramatique va suggérer un drame. C'est l'un des pouvoirs de la musique.

Plus généralement, il peut être parfois largement préférable de suggérer par la musique des sentiments ou des émotions que de recourir aux dialogues : on fait appel au ressenti du spectateur, il participe, il est impliqué.

Quand on a recours à un compositeur pour illustrer la musique d'un film, celle-ci prend en compte les événements marquants du montage pour les renforcer.

Le pouvoir de la musique sur les esprits est tel que le choix des musiques doit être en rapport avec les sensations qu'elle procure et non pas avec des préférences personnelles.

## Les grands principes d'utilisation de la musique

- La musique doit être placée au bon moment, arriver et repartir là où c'est nécessaire
- Il faut savoir où l'utiliser et où elle doit laisser la place au silence
- L'arrêt instantané de la musique à l'arrivée d'un dialogue peut procurer un intense effet émotionnel, ou simplement produire un relief utile pour donner de la vie au projet
- De même, le changement de musique en rapport avec un changement de plan, par exemple en cut sur un plan large, renforce le relief de la réalisation
- La musique doit rester discrète par rapport aux dialogues et aux effets sonores
- Sauf si elle est seule, une musique très intense est généralement à éviter, puisqu'elle devra être fortement atténuée pour ne pas nuire aux voix et aux sons d'ambiance
- Une sensibilité à la musique est nécessaire pour faire les bons choix. Il s'agit d'un art qui nécessite non seulement le ressenti des effets de tous types de musiques sur l'esprit, mais aussi une culture musicale très large de tous styles de musiques, de films ou autres.
- Le choix des musiques et leur emplacement incombe au réalisateur
- Dans la construction d'un film, des « temp tracks » ou « temporary tracks » sont parfois placées à certains moments par un réalisateur ; il s'agit de musiques existantes qu'il a choisies pour servir de guide au compositeur ou même à la sélection ultérieure de musiques de librairies. Les motivations des réalisateurs ne sont pas toujours de la plus grande élégance. Quelquefois, on recherche un compositeur capable de prendre comme exemple une musique dont les droits d'utilisation seraient hors budget. On lui demande alors une musique similaire, dans le même style mais légèrement différente pour qu'elle échappe aux droits relatifs à l'original.

## Les options pour la musique d'un projet visuel

Plusieurs choix s'offrent au réalisateur d'un projet. Ils dépendent de plusieurs facteurs : son niveau d'exigence, son désir d'originalité, sa culture musicale et... le budget disponible.

**Pour enrichir un film d'une bande sonore musicale, un producteur de cinéma ou de télévision dispose de deux solutions : commander une musique originale à un compositeur ou choisir dans des musiques existantes, chansons ou musiques publiées, ou musiques de production.**

Cependant, suivant la dimension d'un projet, avec un goût musical suffisant et peu de budget, un réalisateur peut composer sa propre musique avec des logiciels comme GarageBand (Apple) sur Macintosh, ou Music Maker (Magix) sur Windows. Avec ces logiciels, une composition se réalise par assemblage de « briques sonores » (aussi appelées « boucles » ou « loops »), qui sont des éléments rythmiques ou harmoniques. Les droits d'utilisation de ces boucles sont inclus. On peut ajouter des

éléments joués avec un clavier sur les instruments virtuels inclus ou rajoutés en plug-ins, ou même enregistrer par micro des musiciens avec leur instrument.

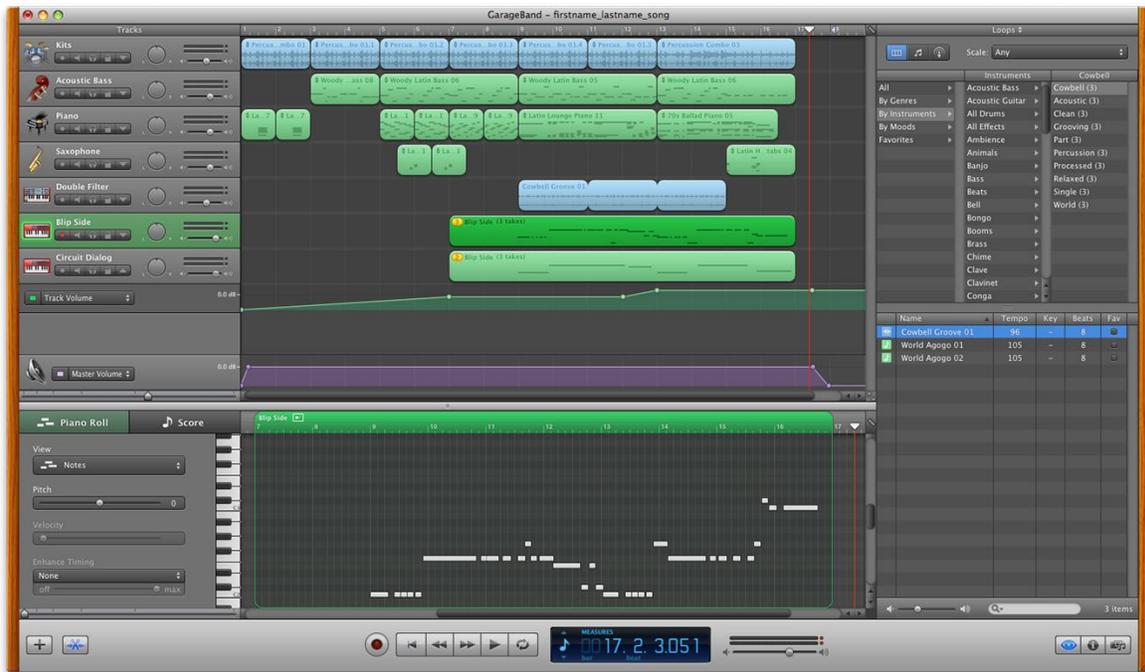


Figure 39 Le logiciel GarageBand d'Apple



Figure 40 Le logiciel Music Maker de Magix

Au sommet de l'échelle, avec un budget de long métrage conséquent, on se tournera plutôt vers un compositeur spécialisé dans la musique à l'image. Il travaillera en relation avec le réalisateur selon ses instructions. Le calage de la musique en fonction des événements du film et du montage renforceront son effet émotionnel, ou participera à son atmosphère générale. La musique pourra être enregistrée en studio et jouée par des musiciens avec leurs instruments, jusqu'à un orchestre symphonique pour les films à gros budget.

La commande d'une musique originale pour un film est un acte de confiance, mais n'empêche pas les discussions lors de la composition. Le dialogue réalisateur/compositeur est essentiel et doit contourner divers obstacles de communication.

Des difficultés de compréhension peuvent émerger, les termes des deux métiers étant très différents. Un réalisateur peut donc avoir des difficultés à exprimer ses souhaits. Et l'effet de la musique est tellement puissant qu'il peut prendre le pas sur celui du scénario et donner au réalisateur la sensation qu'il perd la main sur son film. Il sait que malgré sa confiance dans le compositeur il ne peut intervenir directement sur la composition, seulement sur des orientations.

Autre facteur d'inquiétude : le temps disponible pour la composition de la musique. C'est en effet au moment de la postproduction que le montage du film se termine et peut servir de base à la conception finale des sections musicales, surtout quand elles doivent être synchronisées avec les images. Et dans certaines circonstances ce temps peut manquer !

Entre ces deux options ultimes, musique autoproduite et musique composée spécialement, d'autres choix sont possibles, en fonction du budget disponible (celui qu'on aura bien voulu y consacrer !).

Un compositeur pourra utiliser des banques de sons diverses, y compris orchestrales, pour réaliser des compositions complètes en studio personnel. C'est très souvent le cas pour les téléfilms. Il pourra éventuellement faire appel à quelques interventions de solistes pour compléter son orchestration.

Ou bien on cherchera directement dans les catalogues de musiques celles qui conviendront le mieux, en s'acquittant des droits d'utilisation à régler.

Certains ont aujourd'hui recours à la place de marché en ligne Fiverr de travailleurs indépendants pour commander des musiques à des tarifs très bas.

## **Les droits de la musique dans l'audiovisuel**

Nous allons aborder ce sujet complexe sur deux plans : celui du cas général où tous les principes de base s'appliquent et celui moins contraignant de l'ILOI.

Les auteurs et les compositeurs disposent de diverses protections pour leurs créations, qui leur permettent d'en revendiquer la propriété et d'en tirer des revenus, qu'ils soient les principaux ou accessoires. Leurs droits sont très bien protégés en France. Il est indispensable d'en avoir conscience quand on envisage d'utiliser leur travail dans le cadre d'une production audiovisuelle.

Le strict encadrement de l'utilisation de musiques dans les projets où elles sont intégrées fait l'objet de contrats qui fixent les conditions d'utilisation et les rémunérations afférentes. Le producteur d'un audiovisuel est celui qui porte la responsabilité juridique de l'intégration d'une musique. Il peut éventuellement confier la gestion administrative complète des droits à un superviseur musical, qui peut aussi l'assister dans la recherche de compositeurs ou de musiques existantes.

Ce domaine des droits de la musique est complexe, mais il existe des méthodes pour une utilisation sereine dans le respect des droits et des créateurs et ainsi se préserver des problèmes.

Toutefois, il faut faire une distinction entre un projet interne à l'ILOI, et un autre, diffusé en public, d'autant plus s'il est monétisé, où toutes les demandes d'autorisations et précautions légales doivent être prises et tous les paiements de droits effectués.

Les titulaires de ces droits sur des musiques, auteurs, compositeurs, éditeurs ou ayants-droits, ne les réclameront certainement pas à des étudiants ou des stagiaires dans le cadre de leur scolarité où les risques juridiques sont inexistantes. D'ailleurs l' « exception pédagogique » au droit d'auteur est prévue dans la loi sur la propriété intellectuelle pour permettre, entre autres, l'utilisation d'extraits limités de musiques protégées par la loi (30 secondes maximum). Cette durée étant généralement insuffisante, on sera généralement amené à prendre les précautions qui sont l'objet de ce texte.

Cette section a pour objectif l'information des stagiaires concernant leur utilisation de musique dans leurs projets, ou même plus tard dans leurs activités extrascolaires ou professionnelles. Certains de ces éléments ne seront pas directement utiles dans l'établissement, mais nécessitent d'être mentionnés dans ce contexte.

## Le cas général

Dans le cadre général de l'utilisation pratique de musiques dans un projet d'images sonorisées, trois domaines générateurs de rémunérations pour les auteurs, les éditeurs ou les producteurs de musique sont à considérer : le droit d'utilisation, le droit d'exécution, et le droit de reproduction.

Et chacun de ces droits donnent lieu à des paiements en relation et qui sont en relation avec les utilisations de l'œuvre.

Trois types de musiques sont utilisables dans une production audiovisuelle : celles composées sur commande, celles du commerce (CD ou web), ou celles d'un catalogue de musique d'illustration.

Commençons par une notion de base. L'auteur d'une œuvre musicale est et restera propriétaire de sa création. S'il l'exploite, il est libre d'en fixer les conditions. Mais il peut aussi en céder l'exploitation sous contrat à des tiers qui de ce fait détiendront des droits.

Il existe deux grandes catégories de droits d'auteur pour la musique : le droit moral et le droit patrimonial (celui qui génère des rémunérations). Chacun de ces droits se divise en plusieurs.

**Le droit moral** est un droit du respect de l'œuvre et de son créateur. Il comprend :

- **Le droit à la divulgation** qui indique que l'auteur peut choisir de rendre public ou non son œuvre, par tous les procédés de son choix
- **Le droit à la paternité** implique que son nom et sa qualité (auteur, compositeur etc.) soient indiqués sur tous les supports (par exemple en streaming sur le web, sur un générique ou dans les crédits d'un jeu vidéo). Il peut choisir de le faire sous son nom ou sous un pseudonyme.
- **Le droit de retrait ou de repentir** : avec ce droit, l'auteur pourra retirer ou cesser l'exploitation d'une œuvre.
- **Le droit au respect de l'œuvre** qui permet à l'auteur d'autoriser ou d'interdire toute modification, transformation, coupure, changement de destination etc.

Le droit moral ne peut pas être cédé à un tiers. Il reste la propriété du créateur toute sa vie et se transmet aux ayant-droits à son décès.

**Le droit patrimonial d'auteur** est celui qui permet au créateur d'exploiter son travail selon ses souhaits.

Au plan du droit patrimonial, l'artiste dispose, à titre exclusif, du droit d'autoriser par écrit l'enregistrement de sa prestation, sa reproduction et sa communication au public, ainsi que de toute

utilisation séparée du son et de l'image de la prestation (par exemple un concert filmé ou une vidéo musicale) lorsqu'elle aura été fixée à la fois pour le son et l'image.

Même s'il détient ce droit, le compositeur peut déléguer à un éditeur l'exploitation de sa création.

Le droit patrimonial d'auteur comprend :

- **Le droit de représentation (le droit d'auteur), immatériel, en principe géré par la SACEM, indépendant du droit de propriété sur le support matériel de l'œuvre : il s'agit de la communication de l'œuvre au public par tout procédé, et notamment par l'exécution publique** lorsqu'elle est diffusée en radio, TV, cinémas, streaming etc. L'exécution publique est assortie d'un versement à la SACEM. Il est acquitté par le diffuseur (TV, radio, organisateur de festival, exploitant de cinéma). **Ce n'est donc pas le producteur de l'œuvre audiovisuelle qui l'aura à sa charge.**

**L'auteur-compositeur confie ses droits patrimoniaux à la SACEM quand il en est membre,** indépendamment du droit moral qu'il pourra toujours exercer (le droit moral est définitif et inaliénable). La SACEM établit le prix du montant collecté et reverse leur part aux bénéficiaires en fonction des relevés qu'on lui transmet. Elle dispose du monopole en France pour la collecte des droits patrimoniaux.



*Figure 41 La SACEM a un rôle de collecte et de répartition des droits des compositeurs membres.*

- **Le droit de reproduction,** autre droit d'auteur, géré par la SDRM (Société pour l'administration du Droit de Reproduction Mécanique) **concerne la fixation matérielle de l'œuvre (sa concrétisation) par tous procédés qui permettent de la communiquer au public d'une manière indirecte, comme pour le droit de représentation.**

Il s'agit de la duplication sur supports physiques de réalisations qui contiennent de la musique (DVD, CD-ROM, clés USB, jeux vidéo).

Lors d'une commande, l'autorisation de fabrication de supports physiques contenant de la musique est assujettie au pré-versement à la SDRM d'un montant proportionnel à la quantité fabriquée.

## La musique à l'ILOI

Les stagiaires de l'établissement réalisent des projets. Au sens juridique, ils sont des producteurs d'audiovisuels. En tant qu'œuvre de l'esprit, la production audiovisuelle est aussi l'objet de droits et d'obligations juridiques différents de ceux de la musique elle-même.

Pour ce qui concerne les différents droits de la musique, dans le principe, ceux qu'il faut connaître pour mieux comprendre les intérêts en jeu sont les suivants :

- **Le droit d'utilisation phonographique. C'est le droit du producteur de phonogrammes (celui qui investit dans la création et l'enregistrement) qui lui est accordé contre rémunération,** en contrepartie de son engagement financier : les frais d'enregistrement et de mixage, y compris les cachets des musiciens.

Le droit d'utilisation doit normalement faire l'objet d'un contrat spécifiant le nom du compositeur et ses coordonnées, le ou les morceaux utilisés, leur durée, le cadre de leur utilisation (cinéma,

télévision, jeu vidéo, publicité, festival, web), le ou les territoires de diffusion et l'intervalle de temps de l'utilisation. Le montant à payer pour ce droit (ou la gratuité), est associé à l'accord concernant les conditions de l'utilisation. Ces droits peuvent grandement varier suivant la notoriété des compositeurs.

Ceux qui créent et enregistrent leur musique (l'autoproduction) sont à la fois détenteurs du droit d'auteur et du droit du producteur.

#### - **Le droit de synchronisation**

Le terme « synchronisation », qu'on aurait pu appeler « droit d'intégration » (donc sans rapport avec un calage temporel), est relatif à l'utilisation de musiques existantes pour accompagner des images animées. Il concerne le pouvoir qu'ont les titulaires des droits sur l'œuvre et l'enregistrement sonore de négocier un montant pour l'utilisation d'une musique existante ou d'un extrait de cette musique dans un projet audiovisuel, qu'il s'agisse d'un long métrage, d'un documentaire, d'un film publicitaire ou aussi d'un jeu vidéo. Dans le cadre de l'utilisation de musiques pour les projets de l'ILOI, c'est ce droit de synchronisation qui est essentiel en pratique.

Les deux droits ci-dessus sont souvent gérés par l'éditeur de l'œuvre (si son auteur lui a confié son exploitation). L'utilisation d'une musique nécessite, dans le principe général, l'autorisation de l'éditeur, qui négocie avec l'utilisateur par contrat une redevance relative à ces deux droits, en fonction de la durée utilisée, du type d'exploitation et des territoires de diffusion. Suite au versement, l'éditeur donne son autorisation d'utilisation.

L'éditeur est propriétaire exclusif des droits d'exploitation de l'œuvre musicale. Il partage avec l'auteur des textes et le compositeur les rémunérations de la SACEM pour l'exploitation de l'œuvre. Le contrat d'édition est déposé à la SACEM sur lequel figure le partage des droits. L'éditeur et la SACEM versent les rémunérations aux auteurs et compositeurs selon ce partage.

L'éditeur peut permettre l'utilisation gratuite de musiques dans certaines conditions particulières, par exemple en établissement d'enseignement.

Le site de la SACEM permet de trouver l'éditeur par une recherche sur le titre de l'œuvre, mais une telle démarche a peu de chances d'aboutir à la gratuité, et encore moins pour les musiques très connues. Et il est quasiment impossible d'avoir une idée des droits d'utilisation au préalable.

Cela dit, les éditeurs de musique de catalogue ont des grilles de tarifs qui dépendent de l'utilisation des musiques pour le projet.

Quelques exemples (seulement indicatifs) de tarifs de droits d'utilisation dans un catalogue de musiques préenregistrées :

*Cinéma long-métrage : 450 €/mn*

*Téléfilm : 250 €/mn*

*Publicité (chaînes hertziennes France) : 1 500 € pour 1 mois*

Il est utile de rappeler que le droit d'exécution et le droit d'utilisation sont deux notions différentes. Par exemple, un compositeur membre de la SACEM qui réalise lui-même sa musique peut céder son droit d'utilisation, avec ou sans rémunération, mais reste titulaire de ses droits d'auteur qu'il conserve en totalité, ou qu'il peut céder partiellement à un éditeur.

## **Les musiques « libres de droits »**

Il faut aussi évoquer les musiques indiquées comme « libres de droits » de certains sites internet. On doit se méfier de cette appellation qui n'a pas de base légale. Il peut ne pas s'agir que du droit d'utilisation.

« Libre de droit » est la traduction de « Royalty Free », qui veut dire que l'utilisation des contenus n'est pas soumise à versement de royalties dépendantes des conditions de l'utilisation.

Les mentions « libres de droits » ou « gratuites » pour des musiques en téléchargement ne signifient pas que vous pouvez les utiliser librement en public. En effet, le droit d'utiliser une musique, c'est-à-dire le droit de s'en servir, ne veut pas dire que sa diffusion est gratuite.

De plus, la mention « libre de droits » ne signifie pas que l'utilisation d'un contenu est exempte de toute condition ou précaution.

**Les musiques dites « libres de droits » payantes.** Cette expression peut paraître paradoxale. Elle veut dire qu'après avoir effectué un paiement pour une ou plusieurs musiques, on a acquis un **droit d'utilisation** (mais pas un droit de diffusion) par le fait de ce paiement. Il peut s'agir d'une solution très abordable et simple pour utiliser légalement des musiques qui peuvent être d'excellente qualité dans un projet commercial. Le compositeur (et l'auteur s'il y a lieu) doit être crédité.

**Les musiques dites « libres de droits » gratuites,** sous licence « Creative Commons », souvent de moins bonne qualité. De plus, elles sont soumises à certaines restrictions d'utilisation, selon les sites qui les proposent où 6 types d'attribution existent, ce qui rend leur choix complexe.

Généralement, cette catégorie ne permet pas l'utilisation des musiques dans un projet commercial (monétisé).

Par exemple, YouTube met à la disposition des créateurs de vidéos des musiques et effets sonores libres d'utilisation. À noter que la reconnaissance automatique des musiques sur YouTube peut restreindre ou empêcher la diffusion d'une vidéo, ou même bloquer un compte quand on ne dispose pas des droits afférents à une musique protégée. En effet, certains titulaires de droits d'auteur utilisent « Content ID », le système d'identification de contenu automatisé de YouTube pour identifier et gérer leurs contenus musicaux protégés par des droits d'auteur sur les vidéos.

En fait, même si un compositeur ne demande pas de paiement pour la partie production de sa musique, s'il est inscrit à la SACEM (ou autre organisme à l'étranger équivalent), il doit percevoir les revenus issus des diffusions. Cette précision est utile dans le cadre des diffusions des projets à l'extérieur de l'ILOI (par exemple dans les festivals) pour lesquels l'organisateur des projections devra adresser à la SACEM la liste des musiques, leurs durées, et les auteurs/compositeurs et éditeurs s'il y a lieu. Pour toute projection d'un projet créé à l'ILOI, on devra donc fournir un relevé détaillé des musiques utilisées avec toutes les informations en relation. C'est le rôle de l'organisateur de la diffusion de respecter la loi en déclarant les musiques à la SACEM.

Bien sûr, le stagiaire de l'ILOI qui compose lui-même la musique de son projet audiovisuel est propriétaire de ses enregistrements et peut les utiliser librement (à condition de ne pas se servir d'extraits de musiques du commerce sous forme de samples dans ses compositions). S'il est membre de la SACEM, il a tout de même intérêt à les déposer et à déclarer leur utilisation. S'il n'a pas d'éditeur, il perçoit la totalité des droits d'auteur, lors de la diffusion.

### **Les musiques « libres de droits » sous licence Creative Commons**

Creative Commons est une organisation sans but lucratif pour la gestion de licences textes, images sons et musique. Pour cette dernière, il existe 6 catégories de licences, de la plus étendue à la plus restrictive, dont on doit prendre connaissance pour en **respecter les conditions d'utilisation**. Par exemple, la mention « CC No Derivative Works » ne permet pas l'utilisation sur une vidéo.

Librairies musicales sous licence Creative Commons :

AudioFarm

BeatPick

Cash Music

ccMixer

Epidemic Sound

Free Music Archive

Internet Archive's Netlabels Collection

Jamendo

Magnatune

Opsound

Podsafe Audio

SectionZ

Simuze

## Les musiques de catalogue payantes

Les musiques préenregistrées, dites « musiques de catalogue », « musiques de production » ou « musiques de librairie » ont un avantage : le réalisateur d'un projet qui les utilise dispose de versions terminées et il sait à quoi s'en tenir, plutôt que d'avoir à expliquer à un compositeur ce qu'il souhaite. La « dépense musique » est simple à gérer. Le filtrage des musiques dans les catalogues en ligne par mots clés permet une sélection ciblée avec pré-écoute.

Les sites professionnels qui diffusent ces musiques les organisent par catégories, types d'ambiance, et par styles musicaux avec leur tempo. Ces critères permettent de cibler les choix et de les affiner.

Liste partielle de musiques de catalogue payantes :

- 615 Music
- Audiojungle
- Audio Network
- DeWolfe Music
- Envato Elements
- Killer Tracks
- Los Angeles Post Music
- Manhattan Production Music
- Megatrax Production Music
- Omni Music
- Sound Ideas Music
- Sony Production Music
- SoundDogs
- UPPM Production Music
- Warner Chappell Production Music

**On doit retenir qu'en dehors des établissements d'enseignement, une utilisation illégale de musique dans un projet commercial expose à des risques juridiques, en fonction des infractions. Les sanctions peuvent être lourdes, jusqu'à l'interdiction de projection d'un film en salles ou une procédure judiciaire en contrefaçon.**

## En conclusion

Dans le cadre général, l'utilisation de musiques existantes dans un projet audiovisuel, appelée « synchronisation » nécessite l'autorisation de l'éditeur de l'œuvre (quand il en existe un), qui peut accorder l'autorisation en contrepartie d'une rémunération ou sans frais dans certains cas. L'absence d'éditeur est plus complexe : on doit faire une demande au compositeur et au producteur.

Le cadre de la musique de catalogue est l'un des plus simples. Une autorisation de l'éditeur est suffisante, puisqu'il regroupe de nombreuses musiques et gère les droits en relation.

**Quoi qu'il en soit, dans tout projet interne à l'ILOI, il est fortement recommandé de mentionner au générique le titre des musiques, les auteurs, compositeurs et interprètes, ainsi que les éditeurs.** L'absence de ces informations constitue une contrefaçon.

De plus, même à l'ILOI, en vue des utilisations extérieures, le principe est aussi de créer attentivement une liste des durées utilisées pour chaque morceau et sa position dans la time-line, début et fin. C'est ce qu'on appelle la cue-sheet. Ce travail sera bien plus facile à réaliser dès la réalisation que plus tard, lors d'une nécessité de déclaration pour une diffusion publique.

Par ailleurs, on ne doit pas attendre la fin d'une réalisation visuelle pour décider des moyens par lesquels on obtiendra les musiques appropriées. Une fois le script établi, il faut savoir lesquelles on va intégrer. Ces choix faits en amont permettent d'obtenir le cas échéant les autorisations dans les temps, en tenant compte des éventuels délais d'attente. De même, si on fait appel à un compositeur, il est nécessaire de préparer son engagement largement en amont du montage final.

## Les différents types de projets audio-visuels dans lesquels le son est présent

### - Le film vidéo

Il est destiné à la diffusion sur les chaînes hertziennes ou sur le câble, ou encore en projection locale (conférences, salons, congrès, entreprises, familles, etc...), ou encore sur le web

- **Le film** principalement diffusé en salles de projection pour le public, et décliné en DVD pour l'utilisation domestique.

### - Le dessin animé ou le film d'animation

- **Le jeu vidéo**, où au contraire de du film vidéo ou en salle, à visualisation linéaire, là, c'est l'interactivité qui joue un rôle essentiel et nécessite un accès séquentiel aux images et aux sons associés. Le jeu vidéo nécessite un travail complet sur le son, qui fait appel au sound designer, à l'ingénieur du son pour les prises de sons réels (voix) et les mixages, et à un sound programmer qui lui va intégrer les différents éléments sonores dans l'algorithme du jeu, en prenant en compte les limitations de la machine de destination (PlayStation, Xbox 360, Wii, PC ou Mac). Dans ce contexte, les effets sonores renforcent la sensation d'interactivité (action/réaction).

## La gestion du son dans les projets audio-visuels

Le son final des différents types de projets audio-visuels est le résultat de diverses sources :

- La prise de son live dans les documentaires, les news, certains films de fiction, les concerts en direct

- Le remplacement des dialogues (ADR, ou Automatic Dialogue Replacement)  
Contrairement aux news sur le terrain, pour optimiser la qualité des voix en cinéma, les prises de son du tournage sont souvent remplacées par des doublages en studio, notamment pour s'affranchir des bruits indésirables captés sur le terrain. Lors de cette opération, les comédiens se calent sur leurs voix d'origine avec un maximum de précision. Les versions en différentes langues sont aussi réalisées de cette manière.
- Les « voix off » des documentaires.
- Les effets sonores dont l'objectif est de renforcer l'impact des scènes, ou d'augmenter la sensation de réalisme.
- La musique, par son rythme, son intensité, son harmonie ou sa dysharmonie, son style, les instruments utilisés, est capable de créer ou d'augmenter les sentiments de tension, de suspense, de terreur, de violence, de romantisme de bonheur, etc.

# Les différentes phases de l'intégration du son dans les projets audio-visuels

## La préproduction

Il est essentiel de déterminer la place du son dans le projet, dès sa conception, lors de la lecture du script. Il s'agit de prévoir le type de sons à inclure et les moyens pour les obtenir : matériel de prise de son, effets sonores, musique. Une liste détaillée est nécessaire. On doit garder à l'esprit que le son final devra être constitué de multiples sources qui feront partie du résultat final, et pas seulement des captures des acteurs sur les lieux des tournages.

Un planning précis doit être établi entre le démarrage du projet et le rendu final.

On devra établir un budget pour ce poste (ou au moins une liste précise de ces différents éléments) : rémunérations des intervenants (free lance, intermittents, salariés), éventuellement location de matériel, achat de banques de sons (CD, DVD, disque dur) ou de sons à l'unité sur le web, musique.

Pour la musique, on devra se préoccuper du paiement des droits d'utilisation pour les musiques préenregistrées, dites « musiques de librairie » ou « musiques de catalogue », ou bien mandater un compositeur pour la création de musiques originales, quand le budget le permet. On se chargera aussi de la gestion des contrats avec les ayants-droits.

Dans les projets à gros budgets, un superviseur musical peut être mandaté par la production pour proposer des compositeurs adaptés au projet et gérer la partie administrative (contrats, déclarations, droits, facturations).

Lors de cette étape de la pré-production, on définit les conditions d'accès aux sites prévus pour les tournages (si nécessaire les autorisations d'accès), et on procède à la visite des lieux de tournage pour déterminer dans quelles conditions le son sera enregistré et prévoir les besoins techniques spécifiques.

**Tout ce qui peut éviter les problèmes de dernière minute et le stress doit être minutieusement préparé.** On prévoira tout les secours possibles, dont des batteries supplémentaires chargées, et bien sûr le matériel aura été vérifié et testé avant les déplacements sur le terrain.

# La production

C'est le moment des tournages images/son et celui où les enregistrements ont lieu sur le terrain, en live, par différents types de captures.

Si un compositeur a été mandaté pour le film, c'est aussi le moment où un contrat doit être établi et où il va commencer à travailler sur les directives du réalisateur.

La phase de production est aussi celle où on peut capter séparément tous les sons utiles au réalisme de la scène. Exemple : son de décollage d'un avion de ligne, bruit de passage d'une voiture, engins de travaux publics, ambiance de rue, de gare, d'aéroport, etc.). Ces prises réelles évitent le recours ultérieur aux banques d'effets sonores et les recherches fastidieuses qui s'ensuivent.

Les prises réelles ont aussi pour effet d'augmenter la crédibilité de la séquence, y compris par l'utilisation de sons non parfaitement enregistrés mais vraiment adaptés au contexte.

Lors de cette étape, on va accumuler tous les effets sonores nécessaires, avec plusieurs sources :

- Sons enregistrés sur le terrain
- Serveur de l'ILOI
- Création d'effets spéciaux
- Acquisition sur internet ou dans des collections spécifiques.

# La captation

La captation est l'ensemble des processus d'enregistrements par micros de tous les sons nécessaires à une production. On appelle aussi ces enregistrements des « prises de son », ou simplement « prises » (en anglais « takes »).

Elle met en œuvre différents types de micros et d'enregistreurs, selon les besoins spécifiques.

Elle peut concerner une source unique ou des sources multiples, suivant le contexte et les choix.

En majorité, les enregistrements ont lieu avec du matériel mobile en dehors d'un studio, en plein air ou en intérieur. D'autres enregistrements sont réalisés en studio : les doublages de voix, les voix-off et les bruitages d'un bruiteur.

## Les micros : caractéristiques

Il existe plusieurs catégories de micros, plus ou moins adaptés aux prises de sons selon leur nature. Et dans chaque catégorie les caractéristiques diffèrent. Les professionnels de la prise de son ont souvent leurs préférences.

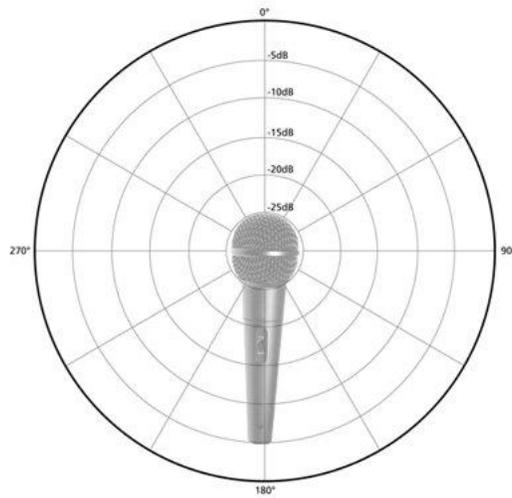
L'autre facteur dont vont dépendre les choix est le coût de chaque modèle. Entre deux micros d'une même catégorie, celui qui coûte dix fois plus cher que l'autre n'est pas dix fois meilleur. C'est pourquoi les choix se déterminent souvent par l'économie conjuguée avec l'efficacité.

### La directivité

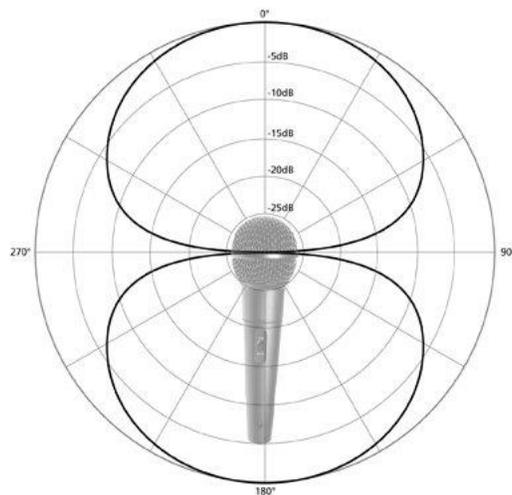
Selon leur type, les micros sont conçus pour capter le son de façon plus ou moins focalisée. On peut comparer cette focalisation avec celle des objectifs optiques : le grand angle « voit » une large partie du champ visé, alors que le téléobjectif se concentre sur une partie étroite, avec un angle étroit.

Cette focalisation du son est appelée « directivité », représentée sous forme de diagrammes qui représentent l'intensité de captation dans toutes les directions par rapport à l'axe du micro.

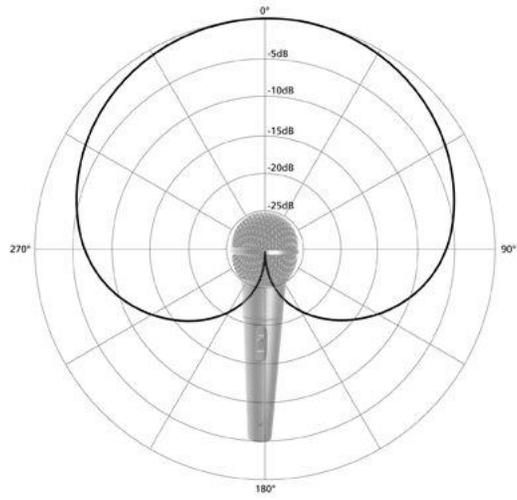
Les courbes des figures ci-dessous indiquent l'intensité de captation dans l'axe du micro en fonction de l'angle de provenance des sons.



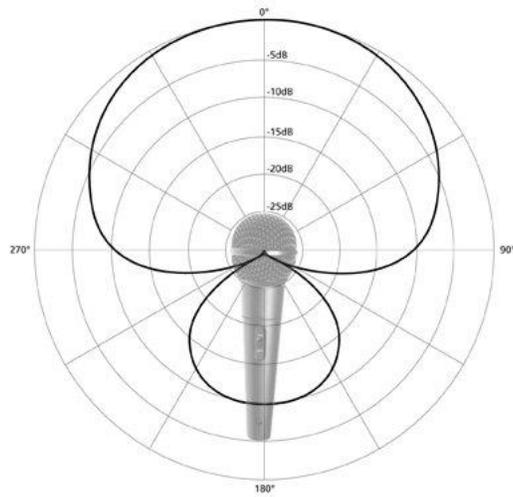
*Figure 42 Le micro omnidirectionnel*



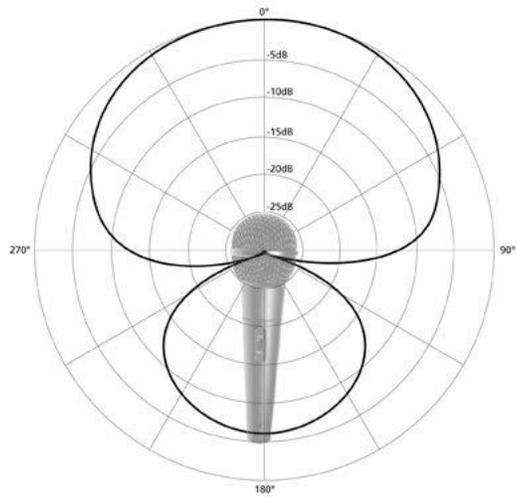
*Figure 43 Le micro bidirectionnel ou à courbe en 8*



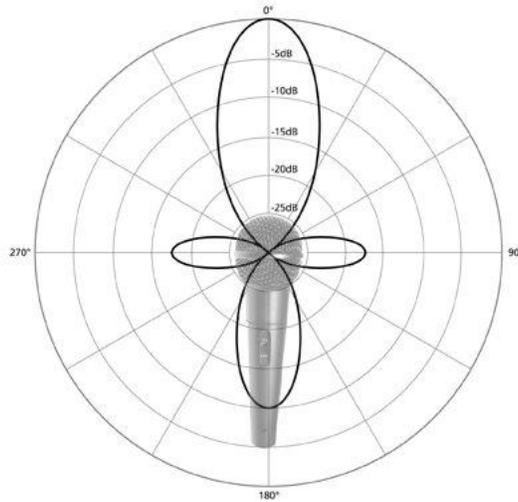
*Figure 44 Le micro cardioïde*



*Figure 45 Le micro supercardioïde*



*Figure 46 Le micro hypercardioïde*



*Figure 47 Le micro canon*

**En résumé, un micro peut être de différents types plus ou moins directionnels :**

- Omnidirectionnel (non directionnel ; capte le son dans toutes les directions)
- Bidirectionnel (directionnel à la fois vers l'avant et l'arrière)
- Unidirectionnel – cardioïde (directionnel vers l'avant)
- Supercardioïde (directionnel vers l'avant)
- Hypercardioïde (directionnel vers l'avant)
- Canon (directionnel vers l'avant)



*Figure 48 Le micro supercardioïde ME 66 Sennheiser à polarisation permanente sur support anti-vibrations.*



*Figure 49 Le micro cardioïde à condensateur TLM103MT Neumann*

Un autre type de micro est utilisé en extérieur, par exemple dans les documentaires sur les animaux, le micro parabolique. Sa parabole est un réflecteur qui exerce une focalisation acoustique et concentre les sons reçus dans son axe tout en rejetant ceux provenant des autres directions.



*Figure 50 Un micro parabolique*

## La technologie des micros

Différents principes techniques sont utilisés dans la construction des micros :

- **Le micro à condensateur** utilise un principe d'électrostatique (on dit souvent « micro statique »). Une très fine membrane chargée électriquement vibre en relation avec les sons, en face d'une plaque rigide. Les deux forment un condensateur. Cette vibration produit un très faible courant qui nécessite une préamplification. Pour produire la charge électrique on doit appliquer au micro une tension de polarisation (généralement 48V), appelée « alimentation fantôme ». Elle lui est transmise par le câble de connexion, lui-même connecté au matériel de captation : enregistreur, caméra ou mixette. Le micro à condensateur est très sensible. C'est aussi le plus fidèle. Il doit être protégé contre les effets du vent et monté sur un support anti-vibrations.
- Pour ce type de micro, on devra donc mettre en service l'alimentation fantôme par le switch correspondant sur le matériel auquel il est raccordé.

- **Le micro électret** est aussi un principe de micro électrostatique à condensateur, à la différence que sa polarisation est permanente et n'a pas besoin de celle fournie par l'alimentation fantôme. Mais cette dernière, ou des piles, sont tout de même utilisées pour alimenter un préampli interne qui augmente le volume électrique produit par la capsule (pour certains modèles c'est une pile interne qui joue ce rôle). Son principe le rend miniaturisable. C'est le type de micro de loin le plus répandu. On le trouve, en qualité basique, dans tous les téléphones et les ordinateurs portables. Mais il peut aussi être d'une excellente qualité et équiper les enregistreurs portables ou des micros de studio, ou encore les micros cravate.
- **Le micro dynamique** est basé sur le principe de la génération du courant créé par les mouvements d'une bobine mobile dans un champ magnétique produit par un aimant. Il n'a pas besoin d'être alimenté et ne doit pas être raccordé à une alimentation fantôme. C'est le micro le plus solide, mais aussi le moins sensible et le moins fidèle. On l'utilise à proximité, tenu à la main, pour des interviews ou en sonorisation.

**Parmi ces trois types, le micro à condensateur est le plus sensible. Le modèle dynamique est plutôt réservé aux interviews.**



*Figure 51 Le micro dynamique Shure SM-58*

## La capture audio en pratique

Aucun micro, aucun enregistreur n'est universel. Chacun des choix de captation répond à de nombreux facteurs, dont chacun a ses avantages et ses inconvénients.

### Le micro perche

C'est la combinaison d'une perche télescopique avec en principe un micro directionnel fixé à son extrémité. C'est le type de capture qui donne le son le plus naturel et le plus réaliste pour les dialogues sur le terrain. En extérieur on privilège un micro hypercardioïde qui focalise bien la capture, alors qu'en intérieur on préfère un cardioïde ou un supercardioïde. Pour la capture d'un son global, un micro omnidirectionnel peut aussi être utilisé.

Le micro supercardioïde étant très directif on l'oriente en biais plutôt qu'à la verticale pour mieux prendre en compte les mouvements des comédiens.

Le micro perche est généralement utilisé au-dessus de la source, dirigé de haut en bas. Mais quand le cadrage le permet, il peut parfois être utile de l'orienter de bas en haut.

En principe, par l'orientation du micro on essaie d'éviter les sons perturbants en provenance des sons gênants.



*Figure 52 La prise de son en biais par la perche*



*Figure 53 La position de la perche sur les épaules est parfois utilisée pour diminuer la fatigue lors des longues séances de captation*

Avantages de la capture par perche :

- Avec un micro de bonne qualité, le son capté sera d'une excellente qualité et très naturel
- On obtient des « plans sonores » en correspondance avec l'angle de vue
- Si la prise est de bonne qualité, on bénéficie du réalisme grâce à la capture de l'ambiance avec les voix
- Une capture correcte mais trop distante pourra éventuellement servir de piste témoin même si la qualité est insuffisante et qu'on est amené à réenregistrer les dialogues plus tard

Inconvénients :

- Les micros utilisés sont très sensibles au vent et aux manipulations de la perche
- On doit les équiper d'un kit anti-vent (coque et bonnette en fourrure appelée « chat-mort », ou « dead cat ») et d'une suspension mécanique
- La perche nécessite un opérateur dédié avec une expérience de ce type de capture ; il doit anticiper les interventions des comédiens pour diriger le micro vers ceux qui prennent la parole à tour de rôle.
- Le port de la perche crée une fatigue et nécessite des pauses de repos

- L'ombre projetée par l'ensemble perche/micro peut être visible sur l'image dans le cadre
- Quand les micros capturent des sons indésirables en excès ou sont mal positionnés, il est préférable d'utiliser les micros personnels comme solution de remplacement (micros cravate, dits « Lavalier »)
- Le micro hypercardioïde capte les sons graves environnants (sa directivité est mauvaise dans cette zone de fréquences)
- On aura souvent intérêt à mettre en service le switch coupe-bas quand il est présent



*Figure 54 Un kit grille anti-vent et la bonnette chat mort correspondante*

## **Le micro cravate (ou Lavalier)**

C'est un micro miniature qu'on porte sur soi. C'est généralement un micro omnidirectionnel, de type électret. Il est souvent construit de façon à favoriser les aigus pour compenser les pertes dues à sa position éloignée de la bouche et par le fait qu'il est souvent caché derrière un vêtement quand il doit être invisible.

Pour les interviews ou les discussions de plateau, même si on cache le câble, le micro peut être apparent. Mais généralement on préfère qu'il soit invisible.

En principe on le connecte à un émetteur fixé à l'arrière de la personne, associé à un récepteur à proximité de l'opérateur et connecté à un enregistreur ou une mixette mobile. Mais un enregistreur miniature fixé sur la personne peut remplacer le système émetteur/récepteur.

Par ailleurs, dans certains cas, par le fait qu'il est minuscule, on utilise un micro cravate pour capter le son des dialogues, par exemple dans une pièce trop petite pour une perche.



*Figure 55 Un micro cravate avec connectique pour smartphone*



#### Avantage du micro cravate

- Il permet un son de proximité, plus précis qu'un micro distant
- Il évite un excès de capture des bruits environnants
- Certains modèles se connectent sur un smartphone

#### Inconvénients

- Le son capté n'est pas naturel et peut nécessiter de recréer artificiellement l'ambiance sonore du lieu en postproduction
- Il ne permet pas les captures en relation avec le cadrage de l'image
- Il peut capter les bruits des frottements avec les vêtements
- Pour compenser le masquage sonore par les vêtements, il peut être nécessaire d'appliquer une augmentation des aigus en post-production

## Le micro caméra

Il s'agit du micro interne ou fixé sur une caméra ou un appareil photo DSLR.



*Figure 56 Le Rode VideoMic Pro installé sur l'appareil photo*

Le tournage avec un appareil photo est courant, son coût d'achat étant très inférieur à celui d'une caméra avec des objectifs similaires. En revanche, le son capturé par les micros omnidirectionnels intégrés est juste utilisable en tant que son « témoin ». Non seulement il est monophonique (ce qui peut éventuellement suffire), mais de plus, sa position à l'avant de l'appareil est sujette à son obstruction par les doigts ! Pour un bon résultat, il faut donc exclure son utilisation principale.

Première possibilité, utiliser un micro séparé supercardioïde fixé sur la griffe de l'appareil photo et branché sur sa prise micro. On obtient une bonne focalisation du son, mais on dépend de la qualité du préamplificateur interne, souvent affecté d'un bruit de fond électronique.

Une autre solution consiste à compléter le son enregistré par l'appareil photo par un son de meilleure qualité par une capture extérieure indépendante avec un enregistreur séparé avec un préampli de bonne qualité.

En principe, une caméra professionnelle comporte deux connecteurs pour le branchement de micros. Le but est de permettre à un seul opérateur de capturer image et son simultanément.

Le type de micro le plus souvent utilisé est généralement monophonique, du type micro canon compact, pour sa directivité qui permet une concentration du son dans l'axe de la caméra. Pour l'utilisation de ce type de micros, sur une caméra professionnelle on peut trouver un switch 48V pour leur alimentation. On peut aussi disposer de 2 entrées lignes et donc au total de 4 entrées audio séparées, réglables individuellement en volume. Un certain nombre de micros, notamment ceux utilisables sur un DSLR sont eux alimentés par des piles.

Avec une caméra professionnelle, le signal du micro est souvent amplifié par un préampli interne. Les réglages placés sur la caméra ajustent le volume de la captation, avec la visualisation sur une partie de l'écran de contrôle. Les micros avec préampli interne ont souvent un bruit de fond propre meilleur que celui des caméras et surtout que celui des DSLR.

Certains modèles de micros disposent d'un switch pour enclencher un limiteur qui évite les saturations lors des dépassements de volume imprévus. On peut aussi trouver des types de micros externes disposant de switches +10 dB ou -10 dB qui permettent d'adapter la prise de son au volume de la source. On a toujours intérêt à exploiter au maximum les possibilités d'amplification interne des micros pour éviter celle des caméras ou DSLRs qui ont plus de bruit de fond.

La qualité sonore des prises est surtout acceptable à une distance proche de la source. C'est la raison pour laquelle ce type de prise de son est surtout réservé aux interviews. La distance optimale pour une bonne qualité audio est d'environ 1m.

Qu'il soit fixé sur un appareil photo DSLR ou sur une caméra pour éviter sa capture de bruits mécaniques, le micro devra être monté sur un support anti-vibrations.

Une caméra grand public est équipée de micros incorporés. Si la qualité d'image peut être correcte, celle du son n'est acceptable que pour une utilisation personnelle. Pour un meilleur résultat, on a intérêt à doubler la capture sonore de la caméra par un enregistreur externe. Le son interne servira de référence pour être remplacé par l'autre avec recalage dans un logiciel de montage vidéo.

L'enregistrement en mode automatique (AGC ou « Automatic Gain Control ») est fortement déconseillé. La compression de volume enlève le naturel du relief sonore d'origine. La position manuelle est à privilégier.

Avantages du micro caméra :

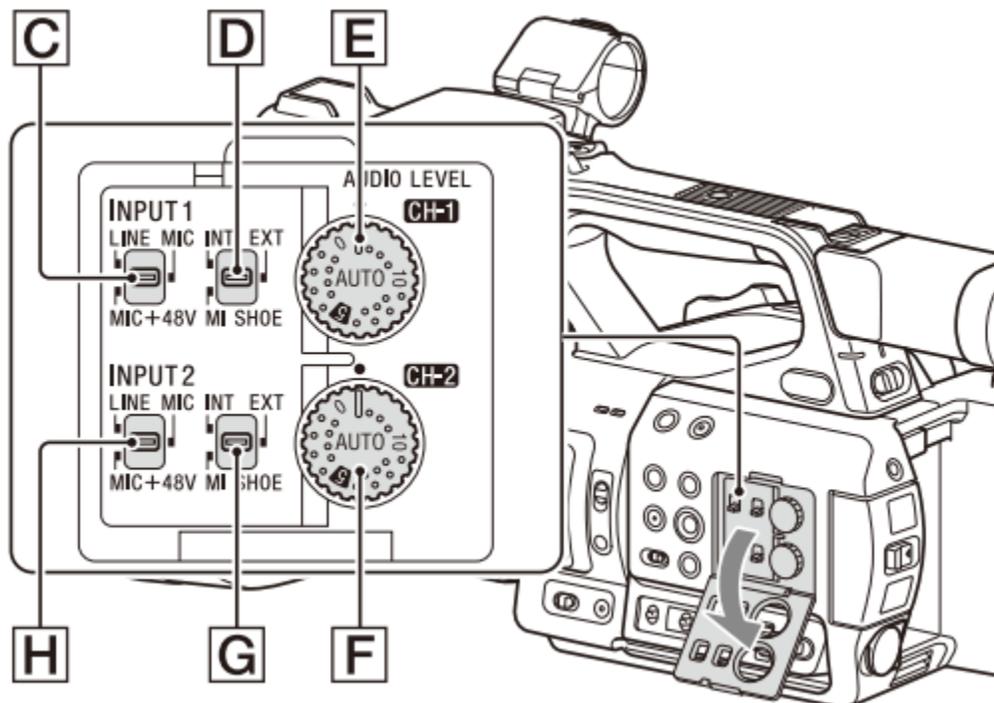
- Il est facile à mettre en œuvre, l'opérateur caméra est suffisant
- Le son capté est synchrone avec la vidéo
- Lorsque d'autres captations autonomes sont utilisées simultanément, le son de la caméra est utilisable pour le calage ultérieur des différentes sources. Dans ce cas, un clap sonore peut servir de référence.

Inconvénients :

- La distance entre la caméra et le sujet diminue la précision sonore
- Il peut capter les bruits mécaniques de certaines caméras
- La surveillance du volume enregistré et les réglages nécessitent une attention particulière



Figure 57 La connexion d'un micro sur une caméra professionnelle



## Le micro sans fil ou micro HF

En bref il s'agit de remplacer une liaison micro par câble par une liaison HF (Hautre Fréquence) avec un émetteur et un récepteur. Le terme « micro sans fil » ou « micro HF » ne concerne pas directement le micro, mais plutôt un ensemble micro-émetteur-récepteur permettant de capter le son à distance sans câble de liaison.

Se passer d'une liaison filaire présente divers avantages : la liberté de mouvement des acteurs, l'absence de câble visible dans le champ de capture et la facilité d'éviter la manipulation des câbles.

Dans cette catégorie on trouve surtout deux types de systèmes : le micro qu'on tient dans la main avec émetteur incorporé, ou l'association micro cravate/émetteur séparé.

Le premier type est approprié pour les interviews où on peut voir le micro sur l'image.

Le second concerne la capture par micro cravate avec émetteur attaché à la ceinture quand le micro doit être invisible.



*Figure 58 Les micros HF à émetteur incorporé en interview*

Comme dans la figure ci-dessus plusieurs micros HF peuvent être utilisés ensemble.



*Figure 59 Le micro cravate sans fil Sennheiser XSW2-ME2-A*

L'ensemble émetteur/récepteur forme une paire réglée sur une seule fréquence. Pour plusieurs captures simultanées autant de ces paires doivent être assemblées avec chacune sa fréquence propre.

Si se passer des câbles est une facilité, le principe a pourtant ses inconvénients.

Entre autres ce principe de captation par micro cravate ne permet pas de créer une perspective en relation avec le lieu, puisque le son capté est toujours proche de la source et toujours le même. De plus, le placement du micro et les bruits de frottements contre les vêtements sont souvent compliqués à gérer.

Dans un environnement urbain chargé en rayonnement électromagnétiques on peut subir des perturbations des transmissions sous forme d'interférences. Avec une liaison HF de faible qualité, des coupures de son sont possibles. Le principe « Diversity » offre une sécurité de réception en utilisant 2 antennes séparées pour basculer sur la meilleure réception en permanence.

Pour disposer de la meilleure qualité de transmission, une portée suffisante entre émetteur et récepteur doit être respectée. Certains modèles actuels (2023) sont utilisables jusqu'à 300 m, dans de bonnes conditions. Quoi qu'il en soit, on évitera d'atteindre les limites de portée pour éviter les coupures.

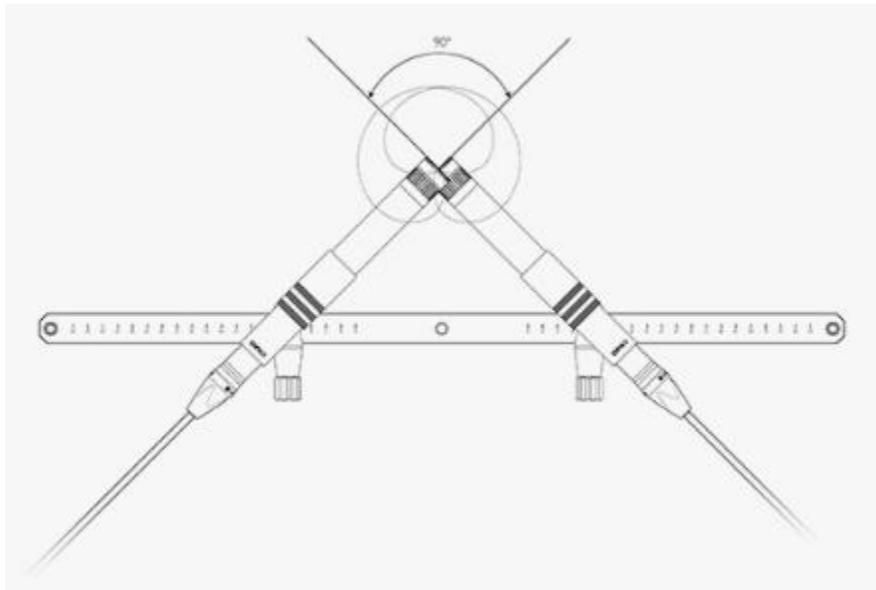
Les problèmes de portée peuvent aussi survenir quand l'antenne de l'émetteur n'est pas suffisamment dégagée, ou par la transpiration contre elle quand elle appuie sur la peau.

On prendra soin d'utiliser des piles neuves au début des sessions d'enregistrement.

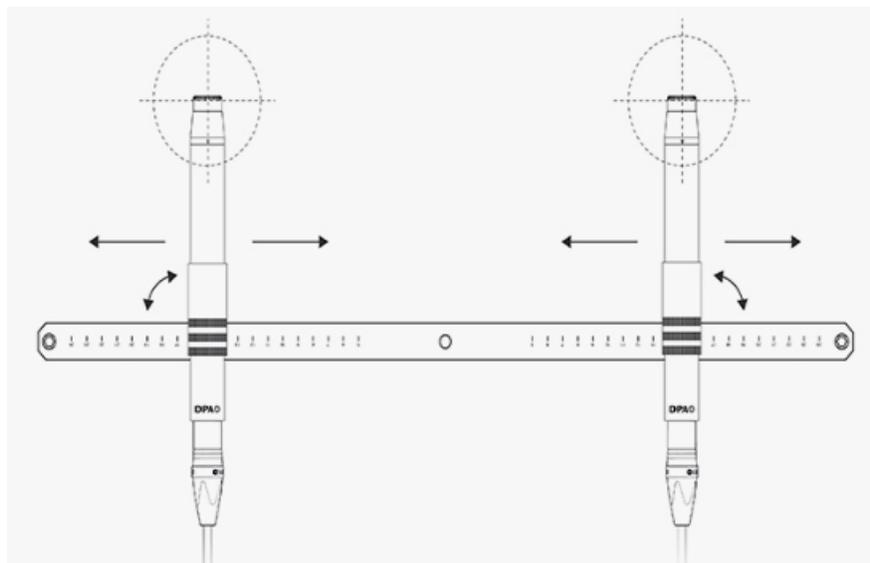
## **Le micro stéréo**

Il s'agit en réalité de l'association de deux micros directionnels, généralement orientés en angle de 90 degrés. C'est la configuration appelée **XY** qu'on appelle aussi paire coïncidente.

Autre disposition, celle où les 2 micros sont espacés en parallèle. C'est la configuration **AB**. On utilise des micros omnidirectionnels. Elle est moins directionnelle que la XY mais donne un plus grand sentiment d'espace.



*Figure 60 Deux micros en couple XY*



*Figure 61 Deux micros en couple AB*

Il existe aussi des micros stéréo intégrés dans lesquels deux capsules sont assemblées.

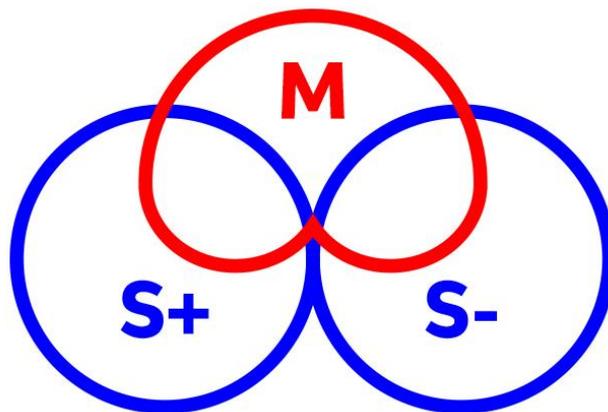


*Figure 62 Le micro stéréo Shure VP88*

## Les micros Mid/Side ou M/S

Ce principe consiste à combiner deux capsules pour obtenir une capture stéréo ajustable ultérieurement. Il s'agit donc d'un micro stéréo particulier. Une capsule cardioïde monophonique est dirigée vers l'avant et une autre à courbe en 8 est orientée à la fois à gauche et à droite.

Lors de l'utilisation, un système de décodage permet d'ajuster l'image stéréophonique, en dosant le rapport centre/gauche droite.



*Figure 63 Le principe de la captation Mid/Side. En rouge la zone de capture du micro cardioïde, en bleu, celle du micro à courbe en 8.*



*Figure 64 Le micro Mid/Side CMS-50 de Sanken*

## **Les micros surround**

Certains systèmes de capture intègrent un nombre plus ou moins grand de capsules directionnelles afin de capter le son dans de multiples directions. Ils sont peu répandus, mais leur intérêt se développe pour les productions sonores immersives dans le domaine du jeu vidéo ou pour la capture d'environnements immersifs.

## **Les supports antivibrations**

Les micros sont plus ou moins sensibles aux vibrations mécaniques, suivant les modèles. Celles-ci lui sont transmises par leur fixation sur un support, trépied, perche ou caméra. Elles produisent des bruits qui s'ajoutent à la captation et sont parfois très difficiles ou impossibles à supprimer ultérieurement. On a donc intérêt à les éviter autant que possible. L'un des moyens les plus efficaces pour les absorber est d'interposer une partie souple amortissante entre le micro et sa fixation.



*Figure 65 Un support antivibrations universel pour micro sur trépied*



*Figure 66 Un support antivibratoire pour fixation sur DSLR*

## **Les bonnettes anti-vent**

L'un des problèmes majeurs de la prise de son en extérieur est le bruit produit par le vent dans un micro. Quand ce problème se produit pendant une captation, la partie enregistrée peut être rendue inutilisable.

Pour se prémunir de cet inconvénient, on utilise une bonnette anti-vent. Le principe consiste à disperser l'arrivée d'air, tout en préservant au mieux les vibrations sonores à capter.

Il existe différents types de bonnettes, universelles ou spécifiques à certains micros, plus ou moins efficaces, adaptées aux différents corps de micros.

Celles qui arrêtent les flux d'air au maximum sont celles munies de longs poils. Les bonnettes en mousse sont bien moins efficaces.



*Figure 67 La bonnette Baby Ball Gag*



*Figure 68 La bonnette Softie de Rycote*



*Figure 69 La bonnette du H4N Zoom*



*Figure 70 Une bonnette pour micro sur DSLR*



*Figure 71 Une bonnette "Dead Cat"*



*Figure 72 Bonnettes en mousse*

## **Les enregistreurs**

### **L'enregistreur personnel**

De nombreux enregistreurs autonomes économiques à 2 canaux ou plus sont disponibles, pour la plupart dotés de micros stéréo intégrés de bonne qualité. Étant séparés de la caméra, ils peuvent être placés plus près de la source et capter un son plus précis. Un clap audio permet le calage ultérieur avec le son témoin de la caméra.

Certains de ces modèles ont des prises XLR pour le branchement de micros à condensateur, avec un switch pour la mise en service de l'alimentation fantôme interne.

Dans ce cas, on peut utiliser ce type d'enregistreur avec aussi bien un micro pour des voix off en intérieur ou un micro sur perche (avec recalage ultérieur en synchro avec le son de la caméra).



*Figure 73 Un enregistreur Zoom H6 fixé sur un DSLR*



*Figure 74 Le Portastudio X6 avec 2 prises XLR*

Avantages de l'enregistreur autonome :

- Il permet d'être placé à proximité de la source
- Sa qualité sonore peut être excellente
- Il est très pratique pour capturer des ambiances sonores sur le terrain
- Il peut être intéressant de l'utiliser à la place d'un système HF pour un micro cravate avec les modèles miniatures
- Il en existe sous forme adaptable sur la griffe d'un appareil photo

Inconvénients :

- Certains modèles en matière plastique captent les rayonnements électromagnétiques (smartphones, Wi-Fi, Bluetooth)
- Suivant les matériaux qui le composent, ils peuvent créer des bruits de manipulation
- Quand il est utilisé en remplacement d'un micro HF par un modèle attaché à la ceinture, le son capté ne pourra être contrôlé qu'une fois l'enregistrement terminé



*Figure 75 Le SR-R1 Saramonic avec micro cravate ; attaché à la ceinture, sa taille permet de le dissimuler sous les vêtements*

## Les enregistreurs professionnels

Les modèles professionnels ont des performances audio supérieures, notamment en matière de bruit de fond, de nombre de canaux et de robustesse.



*Figure 76 Le Scorpio de Sound Devices, Mixeur 32 canaux, enregistreur 36 pistes*



*Figure 77 Le Cantar-X3, un enregistreur numérique 24 pistes, très utilisé en prise de son TV et cinéma*

## Le doublage vocal

C'est la « post synchronisation ». Les anglo-saxons parlent d'ADR (Automatic Dialogue Recording). Dans un certain nombre de situations, l'enregistrement des dialogues est perturbé par les bruits environnants qui peuvent même les rendre inutilisables. Il peut s'agir des bruits de machineries ou autres, ou simplement un excès de bruits ambiants. Ou bien même on s'aperçoit au retour d'une série de prises en extérieur que dans une partie le micro était trop éloigné.

Une solution dans ce cas : faire venir plus tard les acteurs en studio et réenregistrer leurs dialogues. La prise de son a lieu dans un environnement neutre pour obtenir un son le plus naturel possible.

Les acteurs doivent être capables de réinterpréter leur texte avec les mêmes intonations, en se calant sur l'image projetée devant eux.

Cas particuliers, ceux des voix des dessins animés ou généralement ceux des personnages imaginaires. Il s'agit de créations vocales, aussi en relation avec le visuel, qui nécessitent un talent particulier.



*Figure 78 Une actrice en cours de doublage de son texte en studio.*

Un micro cardioïde ou hyper-cardioïde est placé à 50/60 cm en hauteur à l'avant.

Les développements des principes de l'intelligence artificielle rendent possible d'analyser ultérieurement un fragment de la voix des comédiens pour reconstituer leur dialogue par synthèse « text to speech ». On fournit le texte au logiciel qui le restitue avec la voix d'origine avec les intonations souhaitées. Les implications éthiques sont évidentes, mais néanmoins on peut s'attendre à voir ce procédé se répandre dans un proche futur.

## Les voix off

Il s'agit de toutes les voix sans correspondance directe avec l'image : les commentaires dans les documentaires ou les annonces publicitaires. Comme pour les doublages, on enregistre dans un endroit neutre acoustiquement (par exemple la cabine son de l'ILOI). Le son n'a pas besoin d'être aussi naturel que pour les doublages. On obtient un son chaleureux en étant proche du micro (entre 5 et 10 cm). Les graves de la voix sont alors augmentés par l'effet de proximité. Mais en étant trop proche, l'excès de graves peut rendre la compréhension du texte plus difficile et produire un son « embourbé ». On doit chercher le meilleur compromis. Un micro cardioïde à large membrane est généralement utilisé.

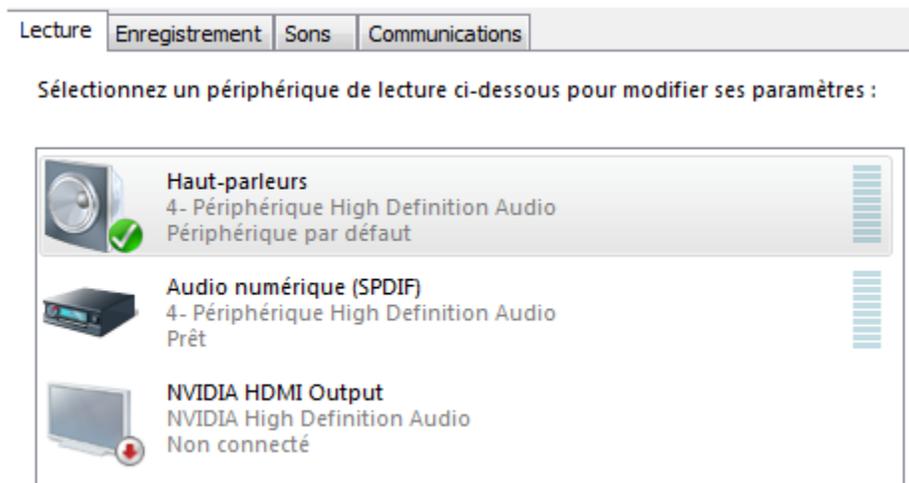
Là aussi, les développements de l'intelligence artificielle commencent à changer la donne. En 2023, on voit apparaître des annonces publicitaires destinées aux producteurs, proposant un choix de voix féminines ou masculines extrêmement réalistes qui prononcent un texte écrit, avec des choix d'intonations différentes. Une recherche sur Google avec « voice over » « AI » arrive par exemple sur « LOVO » ou d'autres sites qui proposent de nombreuses voix. Sur certains sites, pour tester on peut taper un texte et écouter le résultat. Même si ce qu'on entend est parfois un peu artificiel, il est évident que la qualité va beaucoup progresser dans les prochaines années.

## La capture des sons de l'ordinateur

En dehors de la capture par micros, d'autres formes d'acquisition audio peuvent s'avérer très utiles. Par exemple, tous les sons qui transitent par l'ordinateur peuvent être capturés avec une excellente qualité, en format wav ou mp3, notamment grâce au logiciel gratuit Audacity. Il faut cependant garder à l'esprit que l'utilisation publique de sons ou de musiques protégés par copyright captés de cette manière n'est pas légale.

Les paramètres du logiciel doivent être réglés de la façon suivante :

- **Hôte audio : Windows WASAPI**
- **Périphérique d'enregistrement : Haut-Parleurs – Périphérique High Definition Audio**
- **Canaux d'enregistrement (stéréo : 2, mono : 1)**



# Le contrôle des enregistrements

Tous les enregistrements doivent être contrôlés en permanence en cours de capture. Le moyen le plus sûr, est le contrôle au casque stéréo. On doit utiliser un casque fermé, étanche aux bruits extérieurs, pour la plus grande concentration possible.



L'un des meilleurs casques pour le contrôle des prises sur le terrain reste le MDR7506 Sony. Il isole bien de l'extérieur, il est léger, et sa grande clarté dans les aigus permet de distinguer le mieux possible les bruits parasites. Mais bien d'autres casques moins coûteux peuvent convenir, la condition principale étant qu'ils soient bien fermés pour la meilleure isolation possible. Par ailleurs, il est utile que leur port soit le plus confortable possible, surtout pour des séances de longue durée. Les charlottes pour casques peuvent être utiles, entre autres en cas de transpiration.

## Réflexions globales sur la prise de son

- Il est indispensable de faire autant que possible plusieurs captations image et son, surtout en cas de doute. Un problème sur l'une des prises pourra souvent être résolu en utilisant une autre.
- Par sécurité, il est fortement conseillé de réaliser au moins deux, ou même plusieurs captations simultanées, par exemple captation perche et micro cravate. Très utile en cas de problème constaté ultérieurement !
- Toujours avoir une piste témoin (p. ex. le micro camera) pour le calage d'autres prises.
- Pour re-crée au besoin l'univers sonore d'origine et éviter le silence total, enregistrer séparément l'ambiance locale (bruit de rue, ambiance intérieure...), éléments très utiles lors de réenregistrements des dialogues en studio en remplacement des originaux, ou dans des séquences voix-off

- Les évolutions technologiques des équipements d'enregistrement favorisent les captures multipistes. L'avantage est de pouvoir gérer indépendamment chaque piste lors de la postproduction, mais avec en contrepartie une gestion plus complexe du projet.
- Toujours réaliser des back-ups sur des supports différents

## Les défauts de la captation

- Les saturations qu'on évite par l'observation et le réglage du volume d'enregistrement (peak-mètres) ; on teste d'abord avec le volume maximum de la source
- Les bruits divers (on les entend par un contrôle au casque fermé)
- Le vent dans le micro (éliminé par les bonnettes anti-vent)
- Les frottements des micros cravate sur les vêtements
- Les bruits du câble sur la perche qu'on évite en l'enroulant autour
- Les plosives dans les micros, à proximité (évitées avec un filtre anti-pop devant le micro)
- Les inductions des lignes électriques (50 Hz) et des téléphones
- Les prises de son trop distantes qui donnent un son imprécis et étriqué, ou en intérieur noyé dans la réverbération de la pièce
- Les prises de son trop proches qui donnent une sensation de manque de réalisme



*Figure 79 Placé près du micro, le filtre anti-pop absorbe les impulsions de la voix avant leur arrivée sur la membrane*

### L'effet de proximité

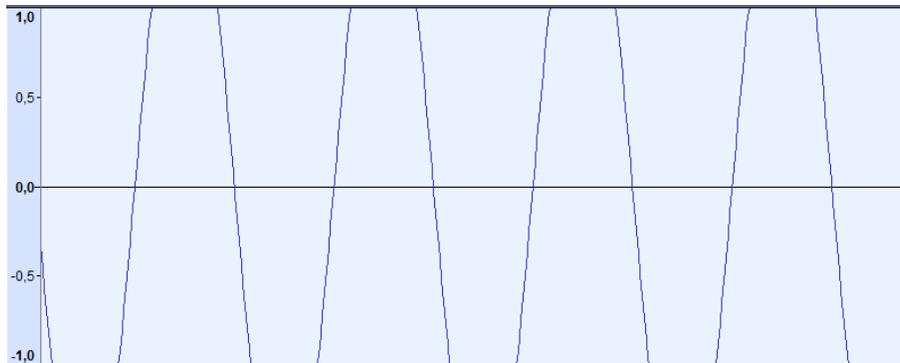
- C'est l'amplification artificielle des graves
- Particulièrement sur les enregistrements de voix
- Se produit à faible distance d'un micro directif (le micro omnidirectionnel n'y est pas sujet)
- Quand le micro comporte un filtre coupe-bas intégré, on doit l'enclencher pour les voix
- Utiliser un filtre coupe-bas lors des traitements ultérieurs pour obtenir un son plus naturel



*Figure 80 Le switch coupe-bas enclenché sur le micro*

## La saturation

- Elle se produit en analogique ou en numérique quand on dépasse le volume maximum :
- L'analogique est un peu plus tolérant pour les excès de volume, mais en numérique, les volumes maximum du signal n'ont plus de variations ; ils sont uniquement à 0 dB FS et produisent un effet très désagréable dû à la création d'harmoniques aigües
- La qualité du son est dégradée
- **Les enregistrements saturés sont inutilisables**
- **La solution : être toujours attentif au réglage de volume ; un volume trop faible est préférable à un volume trop fort.**



*Figure 81 Une sinusoïde saturée en numérique. Les maximums et minimums sont en plateau et n'ont plus d'évolution.*

## Le calage temporel : le Time Code

Après de multiples séquences de prises de vues et de sons enregistrées séparément, au montage on peut se trouver face à un problème : comment déterminer quelles prises de sons correspondent à quelles images ? Et surtout, comment les caler précisément ? Et après un tournage multicaméras, lors du montage, comment afficher en configuration multiécrans toutes les prises vidéo synchronisées ?

Un code temporel standardisé, appelé **Time Code** ou **Timecode** a été pensé et créé avec ces objectifs, en tant qu'élément commun aux images et au son. Ce code existe sous deux formes : **LTC** (Linear ou

Longitudinal Time Code) et l'ancien **VITC** (Vertical Interval Time Code), remplacé par le **DVITC** (Digital Vertical Interval Time Code), ou par l'**ATC** (Ancillary Time Code) depuis l'arrivée de la vidéo numérique.

Généralement, quand on parle de Timecode, il s'agit de sa forme LTC.

Le code **LTC** est un signal audio constitué d'une alternance de deux fréquences, l'une des deux correspondant à un 0 et l'autre par un 1. La position de chaque image en heure/minute/seconde/trame (ou « frame », une image) est définie par des blocs numériques qui représentent chacun des équivalents binaires des 8 chiffres décimaux du time code, donc quatre blocs de chiffres, de gauche à droite les heures, minutes, secondes et images, en incrémentation.

Le signal de ce time code LTC étant sous forme audio, il peut s'enregistrer sur une piste d'un enregistreur audio ou de caméra. En tant que signal audio, il peut aussi se transmettre, et de deux manières : soit pas câble, soit par liaison HF quand la distance émetteur/récepteur est importante. Certaines caméras professionnelles sont en effet équipées d'émetteurs/récepteurs HF. De petits boîtiers générateurs de time code à bas prix sont aussi commercialisés, utilisables comme référence pour une distribution vers l'ensemble des enregistreurs vidéo et audio.



*Figure 82 La fenêtre Timecode du Zoom F8*

Le Time Code fait office de compteur de la position des images dans un film ou dans une prise de vues par rapport à son début. Chaque image a donc son propre time code. On notera que notre système de télévision comporte 25 images par seconde, (les 2 derniers chiffres du time code de 00 à 24), alors que le cinéma en comporte 24 ou 48 par seconde.

Comme on le sait, une prise ou un film ont un début et une fin. Si un time code LTC est enregistré simultanément sur la piste son des caméras et sur l'une des pistes d'un magnétophone multipistes, on dispose d'un point commun entre toutes les captations, et d'un calage facile dans le logiciel de montage.

Même si le time code reste utilisé dans les productions multi-caméras à gros budgets, la facilité de calage visuel des éléments sonores dans les logiciels audio/vidéo le rend moins indispensable qu'auparavant, d'autant plus que l'enregistrement numérique du son et de l'image sont d'une grande

stabilité dans la durée. Un clap visuel et sonore suffit souvent pour trouver la position précise d'une prise de sons par rapport aux images.



*Figure 83 Un clap Timecode, le Denecke TS-TCB. La fenêtre inférieure programmable en Wi-Fi affiche les indications des scènes et des prises*

Le travail de la scripte aura aussi son importance dans le repérage des prises.

On peut aussi fournir à un intervenant extérieur, compositeur, sound designer, une version vidéo du projet avec une superposition du time code incrustée en haut ou en bas de l'image. De cette manière, un réalisateur peut leur donner des indications sur les moments de début et de fin d'une musique ou d'un effet sonore. On appelle « cues » ces indications qu'on fournit sous forme de liste.

Par ailleurs, la piste Timecode d'une vidéo est facilement convertible en **Midi Time Code (MTC)**. Ce MTC est utilisable par un logiciel de composition musicale doté de cette fonction et qui va se synchroniser dessus.

Le logiciel Plural Eyes a été conçu pour caler automatiquement les fichiers sons par rapport à une références. Son développement a été abandonné depuis que Premiere Pro, Final Cut Pro X, DaVinci Resolve et [Avid](#) Media Composer ont intégré nativement cette fonction.

Le LTC ne se lit que pendant le défilement du support où il est présent. Il en est autrement pour le Timecode VITC, qui lui est inscrit dans la partie invisible des images d'une vidéo. De ce fait, le Timecode se lit même à l'arrêt, ou au ralenti, ce qui n'est pas possible avec le LTC. En revanche, au contraire du LTC, le VITC ne peut pas être enregistré en même temps que la vidéo.

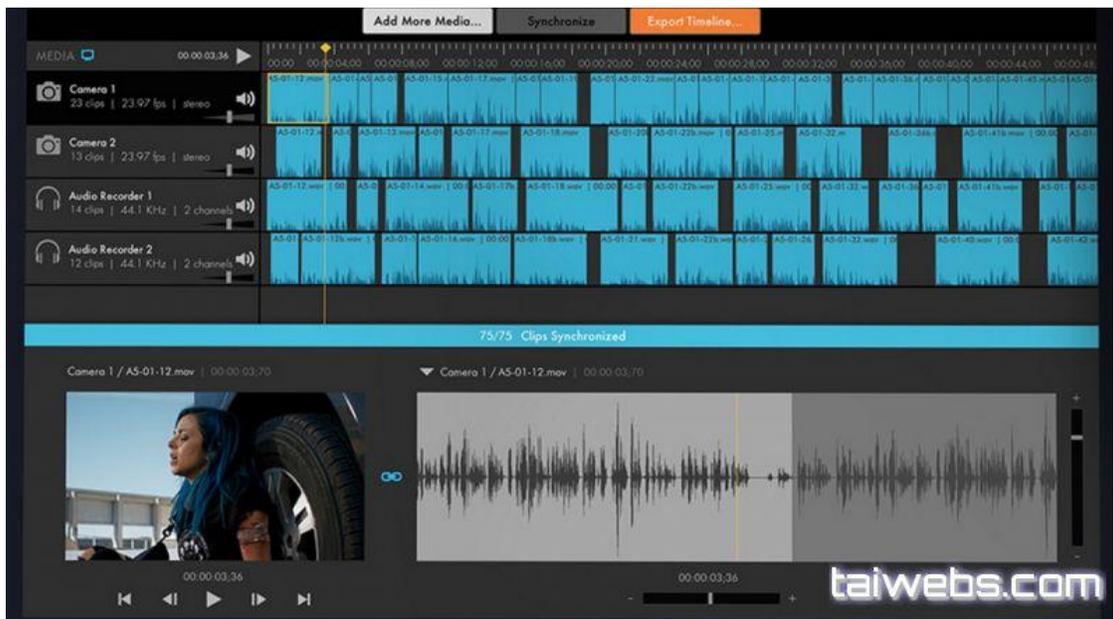


Figure 84 Le logiciel Plural Eyes de Red Giant Software permet de caler automatiquement plusieurs prises de son par rapport à une prise de référence (par exemple celle du micro caméra).

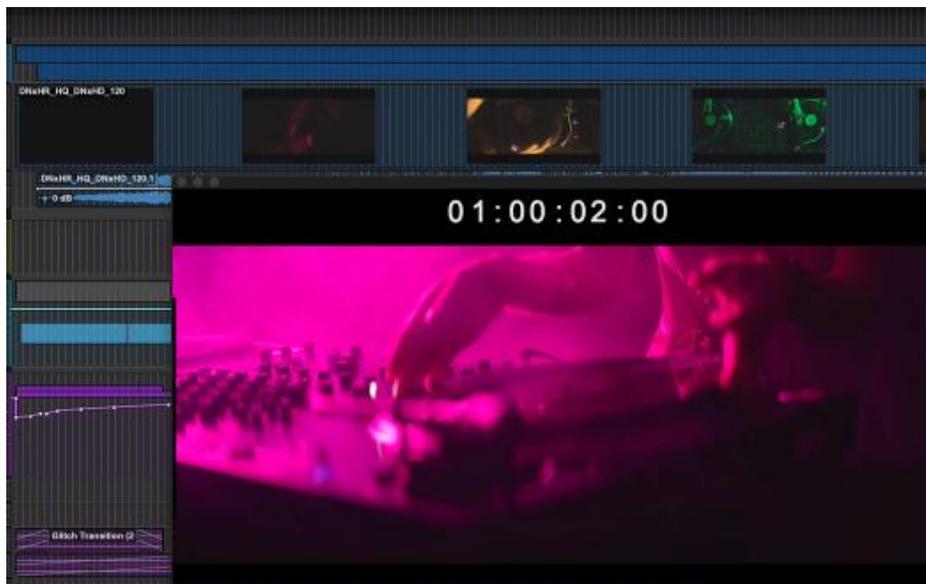


Figure 85 L'affichage du time code dans la version Ultimate de Pro Tools

Sur une machine le time code est soit configuré en mode « maître » (il pilote les autres), soit en mode « esclave » (il se cale sur les autres). On doit donc déterminer laquelle sera maître pour toutes les autres. Un générateur de time code séparé peut servir de maître.



*Figure 86 Le TC-1 de Deity, un générateur de time code miniature Bluetooth.*

## Le monitoring

On appelle monitoring l'ensemble des procédés d'écoute et de contrôle du son. « to monitor » en anglais veut dire « surveiller ». Le mot monitoring est associé à l'écoute de travail par rapport à l'écoute récréative. On utilise des enceintes ou des casques.

### Le monitoring par haut-parleurs

Pour prendre en compte l'incapacité pour un haut-parleur de reproduire correctement l'ensemble des fréquences, on en a combiné plusieurs sous forme de « moniteurs », qu'on appelle aussi des « enceintes », ce dernier mot étant plus utilisé en Hi-Fi. Il s'agit de caisses rigides dans lesquelles on fixe les haut-parleurs.

Ces moniteurs comportent deux haut-parleurs (un pour les graves et un pour les aigus) ou plus, par exemple un troisième pour les médiums. En spécialisant chaque haut-parleur pour ses capacités de reproduction, on peut obtenir un résultat optimum, l'objectif étant de restituer le plus fidèlement possible l'ensemble des fréquences contenues dans les sons qu'on écoute.

Pour générer le son, on conduit le signal complet au moniteur, dans lequel un filtrage le sépare en différentes bandes de fréquences, chacune pour un haut-parleur, lui-même mis en vibration grâce à son propre ampli. Ces moniteurs amplifiés, ou moniteurs « actifs », les plus courants dans le commerce, ont largement pris la place des moniteurs « passifs » qui nécessitent une amplification extérieure.

Il existe deux catégories de moniteurs de studio : les moniteurs principaux et les moniteurs de proximité.

Les moniteurs principaux sonorisent l'ensemble du studio. Ils sont capables de produire une puissance importante, par exemple semblable à celle d'une installation de cinéma. Le son qu'on perçoit est dépendant de leurs caractéristiques et de l'acoustique de l'endroit.



*Figure 87 Un exemple de moniteur principal de studio, le Genelec 1238 CF. De haut en bas, le haut-parleur des aigus, celui des médiums et deux haut-parleurs de graves.*

Les moniteurs de proximité sont destinés à une écoute à faible distance de l'opérateur, entre 1 et 2 mètres. Ils présentent l'avantage d'être peu dépendants de l'acoustique de la pièce puisqu'on entend principalement le son direct. Mais leurs faibles dimensions empêchent de les équiper de haut-parleurs de taille suffisante pour une bonne reproduction des sons graves. Pour compenser cet inconvénient, on leur ajoute éventuellement un caisson de sub-graves. Ce dernier peut être posé sur le sol, à l'endroit qu'on choisit, les sons graves n'étant pas directs.



*Figure 88 Des moniteurs de proximité posés sur le bord d'une console.*



*Figure 89 La Genelec 8038, un moniteur de proximité, comme ceux de l'ILOI. La connectique se trouve à l'arrière, ainsi que certains réglages d'ajustements de fréquences graves, en fonction de leur position.*

La connexion des sources audio sur les moniteurs est réalisée par prise XLR ou un jack 6,35 mm

Pour un travail en multicanaux, par exemple en 5.1, on doit disposer de 3 moniteurs identiques à l'avant et 2 moniteurs à l'arrière qui peuvent être de dimensions inférieures.

C'est par le réglage du volume d'écoute sur une console qu'on détermine la puissance acoustique émise par le moniteur.

## Le monitoring au casque

En dehors de son rôle de contrôle des sons lors de leur capture, le casque stéréo peut aussi être un principe de monitoring en studio. Son utilisation isole complètement de l'acoustique de la pièce et des bruits extérieurs, sauf si il est conçu pour être « semi-ouvert ». Il permet donc une très bonne concentration. Des détails qui passent inaperçus sur des moniteurs peuvent être révélés lors d'une écoute au casque. Par exemple en travail stéréo les erreurs de panoramiques gauche/droite sont plus évidentes que sur haut-parleurs.

On doit prendre en compte que chaque modèle de casque a son propre rendu sonore. Par exemple le très connu Sony MDR-7506 a un rendu très supérieur à la moyenne dans les aigus.



*Figure 90 Le MDR-7506 Sony.*

L'écoute au casque n'est pas sans inconvénients. On entend le son dans la tête plutôt que dans l'air, une sensation artificielle. C'est en partie la raison pour laquelle le volume des graves perçus de cette façon peut être trompeur, inférieur à la réalité sur haut-parleurs.

Bien entendu, par son écoute de deux sources uniquement, le casque stéréo est exclu du travail en multicanaux. Il est bien sûr indispensable pour le travail et le contrôle en binaural.

Pour bénéficier d'une écoute objective, il est important de se créer une empreinte sonore individuelle sur les moniteurs de travail, en écoutant d'autres réalisations, avec un volume le plus constant possible, y compris au casque. C'est ce qui permet d'éviter des erreurs telles que celles qui concernent les graves ou les aigus, ou même les médiums, là où notre oreille est la plus sensible.

## **La postproduction**

C'est la phase du projet où le son définitif se construit, en parallèle avec la finition du film. C'est donc l'étape cruciale pour l'obtention d'un son de qualité et en synergie avec la réalisation.

Entre autres, c'est l'étape du montage image, donc l'établissement de la structure définitive du film. De multiples sources de sons convergent vers cette étape, et vont nécessiter une sélection attentive pour leur intégration dans le projet final.

C'est aussi l'étape où sont réenregistrés les dialogues si nécessaire, processus assez long qui implique une grande précision pour le calage sur les originaux et des comédiens talentueux. Ce processus est facilité par l'utilisation de stations de montage numériques pour le calage sur les voix d'origine. Une fois enregistrés, ces dialogues doivent aussi passer par la phase du montage.

C'est aussi la phase du projet où le bruiteur peut intervenir.

Il existe plusieurs niveaux de structures de postproduction, en fonction du type de projet et du budget. Dans la structure de base, une seule personne va effectuer l'ensemble des tâches : montage image et son, mixage, etc.

Dans les projets de plus grande envergure, le travail du son peut même être délégué à plusieurs équipes. Musique, dialogues, ambiances sonores, effets spéciaux, peuvent confiés à plusieurs équipes encadrées par un superviseur.

À noter que pour les films en images de synthèse, tous les sons doivent être créés en postproduction.

# Les opérations de postproduction

## La préparation de la matière sonore avant le mixage

- Sélection des prises de la production
- « Nettoyage » des prises (suppression des éléments indésirables)
- Editing (raccordement d'éléments, séparation de plages...), fade-in, fade-out (fondus d'entrée ou de sortie) crossfades (transitions progressives entre 2 parties) ; les fades-in ou out sont fréquemment utilisés pour éviter les coupures brusques du son qui produisent un effet audible dérangeant. Quand on a le choix de plusieurs courbes de fades, le type « equal power » a un effet plus naturel que le type « linear ».

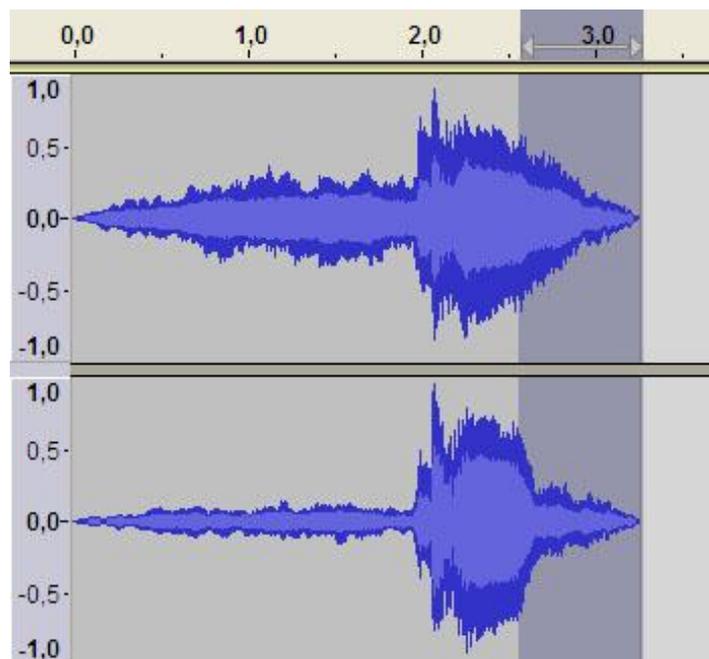


Figure 91 Un fade-in (arrivée progressive au début) et fade-out (disparition progressive à la fin).

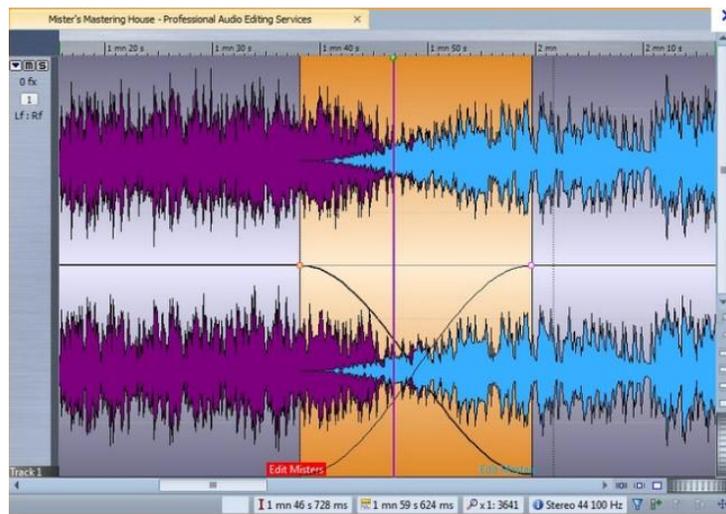


Figure 92 Un crossfade (transition entre 2 éléments). La fin du 1<sup>er</sup> est mélangée avec le début du suivant.

## Les traitements et effets audio

Nous savons que les vibrations de l'air qui produisent des sons peuvent être captées de différentes manières et transformées en variations de courant.

À partir de ce moment, on peut les manipuler à volonté, les transformer plus ou moins radicalement. On donne parfois le nom de « matière sonore ». à au mélange de ces vibrations électroniques. Cette dénomination imagée montre qu'à ce stade de l'audio, on peut penser le son comme un matériau de base, à partir duquel on va en créer un autre, modifié selon notre volonté et nos besoins, un peu comme on peut modifier la forme d'un morceau de pâte à modeler. Le résultat sera toujours un objet en pâte à modeler, mais différent du morceau d'origine.

### Pourquoi transformer le son ?

Transformer le son n'a rien d'obligatoire. Si on le fait, c'est un choix et une décision. Plusieurs raisons peuvent y conduire à partir de l'original :

- Son volume global n'est pas optimal
- Ses variations de volume sont trop importantes
- Son contenu en fréquences de convient pas : trop ou pas assez de graves ou d'aigus (ou toutes autres fréquences intermédiaires)
- Le bruit de fond est excessif
- La captation présente des défauts : bruits divers (effets du vent, volume trop faible...).

Nous ne reviendrons pas sur l'attention particulière qu'on doit apporter à toute prise de son. Certains défauts à cette étape peuvent être impossibles à corriger. D'autres peuvent l'être partiellement, mais en y passant un temps excessif dans une réalisation dont la durée est toujours limitée.

Sur un plan général, traiter un son a pour objectif de l'optimiser en vue de son utilisation dans le contexte du projet. Et c'est cette optimisation que nous allons détailler.

Pour illustrer cette section, nous utiliserons comme exemples les plug-ins de traitements et effets disponibles dans le logiciel Pro Tools de la salle son. On retrouvera les mêmes principes avec des traitements logiciels issus d'autres marques.

## **Hardware ou software ?**

Les traitements et effets sont monophoniques, stéréophoniques ou en multicanaux, selon leur conception de base et les fichiers à traiter. En fonction de leur format ou pour obtenir un résultat donné, on choisit les traitements appropriés. Certains effets peuvent utiliser un fichier monophonique et produisent un effet stéréophonique ou multiphonique.

Depuis les origines des traitements audio, ils ont été construits en hardware avec des composants électroniques, d'abord avec des lampes, puis avec des transistors, et finalement avec des circuits intégrés analogiques, puis numériques. Le principe de base : un appareil par traitement.

Les plus anciens modèles, même avec leurs imperfections, conservent toujours leurs adeptes. Mais leur utilisation est soumise à des contraintes : raccordements complexes par câbles à divers points des circuits de consoles par des « patches » (des racks d'interconnexions), difficultés de répétition ultérieure des réglages, risques de pannes, et réparations par des spécialistes.

Tout a changé avec l'apparition des équivalents numériques sous forme de logiciels. Certains d'entre eux ont été conçus en tant qu'éléments autonomes (stand-alone en anglais), mais la plupart sont destinés à être insérés dans un environnement logiciel complet, en principe dans une console virtuelle.

Pratiquement tous les équivalents des traitements et effets hardware, anciens ou récents, ont été créés sous forme logicielle, jusqu'à même simuler précisément leurs imperfections.

Différence majeure entre les versions hardware et celles logicielles, dans le premier cas, on a besoin d'un appareil par traitement, alors que dans le second, quand on dispose d'un plug-in, on peut en utiliser en même temps autant d'instances qu'on le souhaite, avec pour chacune ses propres réglages.

Deux catégories de plug-ins peuvent se rencontrer : les traitements dits « destructeurs » en plug-ins et ceux en temps réel, en hardware ou en plug-ins.

### **Les traitements destructeurs**

Ils modifient les fichiers ou des sélections partielles de façon durable. Mais en principe on peut toujours récupérer les originaux plus tard si nécessaire. Le traitement destructif est très utile dans un travail de préparation des multiples éléments à intégrer dans un mixage ou un jeu vidéo avant cette étape. En particulier on peut de cette manière les homogénéiser autant que possible (volumes, graves, médiums, aigus, etc.). Un mode de pré-écoute permet d'entendre préalablement le résultat des traitements. Une fois le réglage adopté, on procède au rendu, et l'ancien fichier est transformé en un autre.

Il existe des traitements par lots (batch) qui transforment automatiquement un groupe de fichiers audio avec des réglages identiques.

### **Les traitements temps réel**

Ils diffèrent des précédents dans la mesure où ils exercent leurs fonctions en permanence, sur toute la durée d'un projet. Autre intérêt : généralement, en versions plug-ins leurs réglages sont automatisables. Ils peuvent donc varier progressivement pendant le déroulement du projet. Par exemple, à différents moments d'un mixage, on peut avoir besoin d'augmenter ou diminuer le volume d'un dialogue.

Autre possibilité avec l'automatisation : la mise en service ou hors service d'un traitement à des moments précis.

Une fois ces variations captées, lors de la relecture elles sont répétées toujours identiquement. Et elles peuvent encore être affinées. On a du mal à imaginer qu'avant cette possibilité, toute variation de réglages pendant un mixage devait être manuelle sans rectification possible après coup !

Et surtout, l'ensemble des réglages et de leurs variations sont enregistrés avec la session de travail, y compris toutes les automatisations. C'est le « total recall », qui était très laborieux quand tous les réglages des consoles et périphériques devaient être retrouvés lors de retours en studio pour des modifications de mixages, opération appelée « recall ».

Les traitements audio peuvent s'appliquer soit à des fichiers mono, soit simultanément à tous les éléments des fichiers stéréo ou multicanaux. Dans certains logiciels, on peut choisir la version du plug-in de traitement, soit celle en traitement destructeur (Audiosuite dans Pro Tools), soit la version temps réel insérable dans une piste.

### **Les pré-réglages ou « presets »**

La plupart des traitements audio en plug-ins proposent des réglages prédéfinis appelés presets. Il s'agit d'une base, sachant que c'est toujours l'utilisateur qui décide si ces propositions lui conviennent. En effet, à partir de ces presets, on peut toujours les ajuster pour les adapter au contexte. Les nouveaux réglages sont enregistrés avec la session.

### **Le by-pass**

Il s'agit d'un switch qui permet de couper l'action d'un traitement, pour pouvoir comparer le résultat avec ou sans celui-ci, ou simplement pour le désactiver.

### **Le gain**

Ce réglage permet d'ajuster le volume qui entre dans le traitement et celui qui en sort. C'est entre autres ce qui permet une meilleure comparaison de son action, celle-ci pouvant modifier la perception de volume dans certains cas

### **Insertion ou Départ auxiliaire (« Insert » ou « Send »)**

La différence entre ces deux notions est très importante. Lors de l'utilisation d'un effet ou d'un traitement, on devra décider de sa position dans la chaîne audio. Certains traitements ne peuvent être utilisés qu'en les insérant dans le chemin d'un signal sonore, alors que d'autres offriront le choix entre une insertion ou un départ auxiliaire. Les deux utilisations, Insertion (Insert) ou Départ auxiliaire (Send) concernent le travail sur une console de mixage, quand elle comporte ces Inserts ou ces Sends.

Alors que l'insertion se comprend facilement (on insère le traitement), la notion de départ auxiliaire nécessite une explication. Pour cette dernière, on envoie par un réglage de départ pour chacune des pistes choisies une quantité plus ou moins grande du volume d'une piste vers un effet global, souvent un effet de réverbération. Cet effet est inséré sur une piste séparée des autres, nommée « Piste auxiliaire » ou « Auxiliary Track ». Celle-ci reçoit donc un mélange de tous les départs de pistes et leur applique cet effet. Le résultat est réinjecté dans le mixage. C'est ce principe qui permet d'obtenir un effet de réverbération cohérent, puisque cette réverbération virtuelle correspond à une seule salle.

Au contraire, l'insertion de plusieurs réverbérations différentes en insertion sur les pistes peut produire une confusion sonore.

## **Le travail sur le spectre**

C'est avec le **filtrage** et l'**égalisation** qu'on peut modifier globalement le spectre sonore, c'est-à-dire modeler la répartition des graves/médiums/aigus dans un son, grâce aux **égaliseurs** et aux **filtres**.

On trouve souvent les termes anglais **Equalizer** ou l'abrégié **EQ**.et **Filter**.  
Le filtrage et surtout l'égalisation font partie des opérations les plus courantes en audio.

## **Pourquoi modifier le spectre sonore ?**

Comme nous l'avons évoqué précédemment, l'audio en tant que version électrique du son peut être considéré comme de la matière sonore qu'on peut modeler à volonté pour la rendre utilisable au mieux dans un certain contexte. Différents outils sont utilisables pour y parvenir, en particulier pour modifier la répartition des fréquences dans les sons quand c'est nécessaire. En effet, certains sons bruts ne sont pas utilisables de façon optimale dans le contexte sans modification de leur spectre.

## **Filtrage et égalisation**

Le filtrage et l'égalisation sont les bases du travail sur le spectre (voir ci-dessus), grâce aux filtres et aux égaliseurs.

L'égalisation est un travail d'augmentation ou d'atténuation d'une zone de fréquences, alors que le filtrage concerne une élimination d'une partie du son.

Il faut noter qu'une certaine confusion existe entre le filtrage, l'égalisation, les filtres et les égaliseurs. Le mot « filtre » est généralement utilisé en tant qu'élément d'un égaliseur qui en comporte plusieurs.

Égaliseurs ou filtres existent sous forme hardware ou software.

On les trouve soit par éléments séparés avec une seule fonction, soit sous forme de combinaisons de fonctions dans un seul élément.

Leurs actions sont représentées sous forme de courbes.

Des éléments des paramètres de réglages se retrouvent en commun dans différents outils :

### **La fréquence :**

C'est la valeur en Hertz où l'action sur le spectre se produit.

### **La sélectivité (Q)**

C'est l'action plus ou moins concentrée autour d'une fréquence. Plus la valeur de ce paramètre est faible, plus son action est concentrée ou « serrée », ou « pointue ». À l'inverse, plus elle est élevée, plus l'action est large, plus elle est globale

### **La pente (dB/octave).**

C'est l'efficacité d'un filtre au-dessus ou au-dessous de la fréquence choisie

### **Le gain**

On ajuste le volume d'entrée et celui de sortie.

## Les différents types de filtres :

- **Le filtre passe-bas** ou « **low-pass** », ou **high-cut** : Il laisse passer les basses fréquences, donc coupe les hautes fréquences (c'est pourquoi on l'appelle aussi coupe haut ou « high-cut ». C'est sa fréquence d'action qui détermine sa zone de filtrage, et sa pente qui définit son efficacité, en dB par octave (en général 6, 12 ou 24 dB/octave, parfois plus).



Figure 93 Un filtre passe-bas. Ici, le type de filtre est positionné sur « Low-Pass », la pente est de 12 dB/octave et la fréquence réglée sur 4,75 kHz

- **Le filtre passe-haut** ou **filtre high-pass** ou **low-cut** : Il laisse passer les hautes fréquences, donc coupe les basses fréquences (c'est pourquoi on l'appelle aussi coupe bas ou « low-cut ». C'est sa fréquence d'action qui détermine sa zone de filtrage, et sa pente qui définit son efficacité, en dB par octave (en général 6, 12 ou 24 dB/octave, parfois plus).



Figure 94 Un filtre passe-haut. Ici, le type de filtre est positionné sur « High-Pass », la pente est de 12 dB/octave, et la fréquence est réglée sur 1 000 Hz

- **Le filtre passe-bande**, ou **filtre sélectif** ou encore « **peak filter** » accentue ou diminue le contenu sonore centré sur fréquence déterminée. Un réglage « Q » élargit ou rétrécit sa zone

d'action, le rendant plus ou moins sélectif. Plus sa valeur est élevée, plus la courbe d'action est étroite, et plus l'effet du filtre est concentré sur la fréquence choisie. Le réglage de gain détermine l'amplification ou l'atténuation de la zone de filtrage.



Figure 95 Un filtre passe-bande. Ici, le type de filtre est positionné sur « Peak », le facteur Q est de 1.00, la fréquence est réglée sur 1.08 kHz, et l'action du filtre est de +3.5 dB

- **Le filtre éliminateur** ou « notch filter » sert à réduire très fortement une zone de fréquences très précise. La zone d'action plus ou moins sélective se règle par la valeur « Q ». La valeur de Gain n'est pas réglable, l'atténuation du filtre étant maximum pour ce type de filtre.



Figure 96 Un filtre éliminateur (Notch). Ici, le type de filtre est positionné sur « Notch », le facteur Q est de 1.00, et la fréquence est réglée sur 1000.0 Hz.

Ce type de filtre est utilisé en combinant plusieurs pour éliminer le bruit capté par le rayonnement des lignes électriques à 50 Hz et ses harmoniques.

- **Le filtre en plateau**, ou « shelving filter » atténue ou augmente l'intensité de toutes les fréquences situées au-dessus de la valeur de réglage pour un **high-shelf** ou au-dessous pour un **low-shelf**. Le réglage de gain détermine l'amplification ou l'atténuation de la zone de fréquences concernée



Figure 97 Un filtre en plateau (Shelf). Ici, le type de filtre est positionné sur « High-Shelf », le facteur Q est de 1.00, la fréquence est réglée sur 1.56 kHz et l'atténuation des fréquences supérieures est de -6.4 dB.



Figure 98 Ici, on a un filtre Low-Shelf, avec une valeur de Q de 1.00, la fréquence est réglée sur 101.0 Hz, et l'atténuation des fréquences inférieures est de -6.4 dB.

Différents types de filtres combinés sont proposés et permettent un travail très complet.

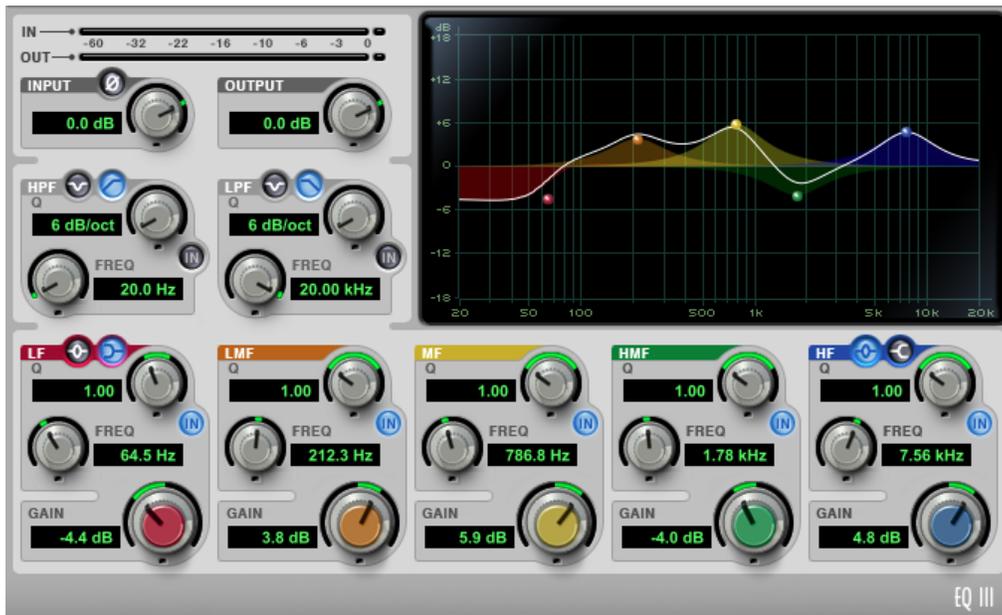


Figure 99 Un égaliseur très complet, comprenant différents types de filtres activés. Les zones colorées correspondent aux modifications du spectre : plus de volume au-dessus de la ligne médiane et moins en-dessous.

Les consoles d'enregistrement et de mixage comportent une section d'égalisation plus ou moins complexe par voie. Au plus simple, on trouve une coupure progressive des graves et des aigus, de même qu'un réglage des médiums. Pour les modèles plus performants, comme ci-dessous, d'autres fonctions sont présentes.

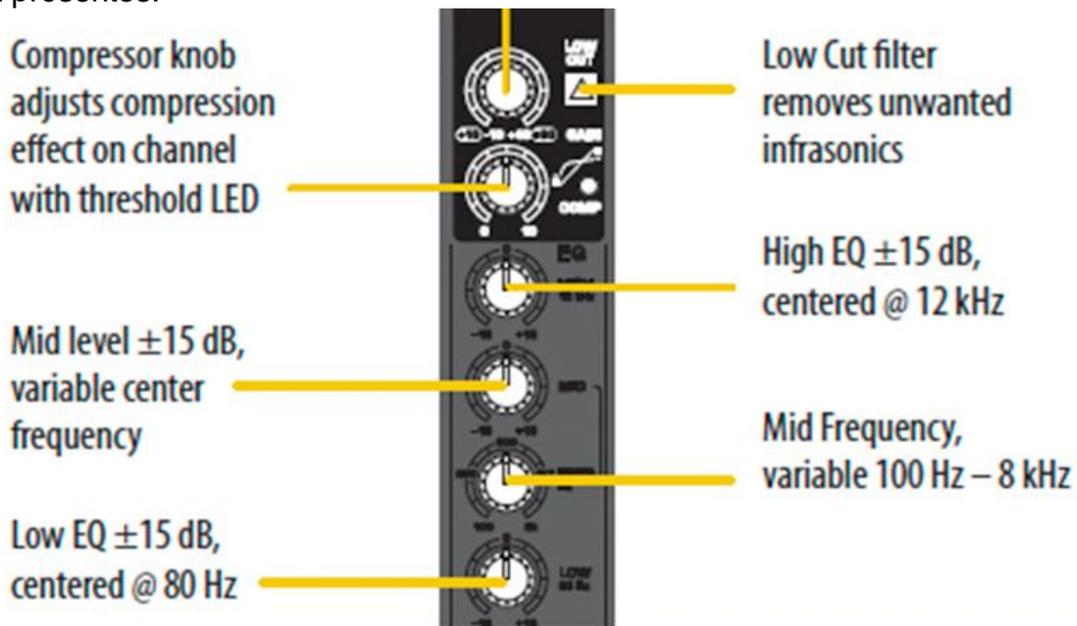


Figure 100 Un exemple d'égaliseur sur une voie de console hardware. De haut en bas, on voit le switch du filtre coupe-bas (ou passe-haut), le filtre d'aigus (ici centré sur 12 kHz et réglable entre + et - 15 dB, le filtre hauts-médiums à fréquence réglable et action comprise entre + et - 15 dB, le filtre bas médiums, réglable entre 100 Hz et 8 kHz et action comprise entre + et - 15 dB

## L'égaliseur graphique

Il doit son nom au fait que son action est directement en relation avec le spectre sonore. Il existe sous forme hardware ou logicielle. Ses fréquences d'action sont généralement réparties en tiers d'octave. Il est souvent utilisé en sonorisation pour corriger les irrégularités du système de diffusion et l'accrochage Larsen.

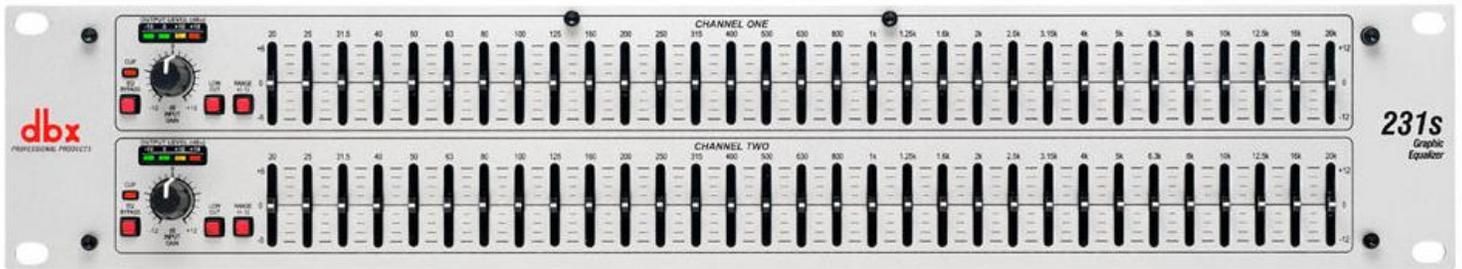


Figure 101 Un égaliseur graphique hardware en 1/3 octave à 2 canaux.

## L'égalisation logicielle sophistiquée

Les traitements audionumériques ont permis l'apparition d'égaliseurs complexes qui ne fonctionnent pas par bandes de fréquences individuelles, mais par un travail global sur l'ensemble du spectre. Ils utilisent le principe de l'analyse de Fourier. Celle-ci, une analyse de spectre en temps réel, est capable d'établir une moyenne dans la durée et de la faire matcher avec celle d'un autre son. On prend comme référence un son ou un ensemble de sons et on applique son spectre global à un autre, ce qui produit un spectre similaire.

Ce procédé est très utile quand une voix de comédien a été enregistrée dans des conditions différentes (distance, type de micro) de lui donner une sensation de similitude.



Figure 102 L'égaliseur CurveEQ de Voxengo dispose d'une fonction de mise en correspondance de spectres audio pour obtenir un profil audio similaire

## Les utilisations des filtres et égaliseurs

Diverses circonstances peuvent amener au besoin de traitement des fréquences d'une source sonore, pour en particulier :

- Faire matcher les différents éléments quand des dialogues ont été enregistrés à des distances différentes. Dans ce cas les graves et les aigus sont différents d'une prise à l'autre.
- Augmenter la clarté des dialogues en augmentant les hauts-médiums (entre 4 kHz et 8 kHz).
- Compenser une perte d'aigus lors de la captation par micro cravate quand il est masqué par le tissu
- Supprimer ou atténuer les défauts de captation, tels l'effet de proximité d'un micro, l'effet du vent, le rayonnement du courant à 50 Hz...
- Créer des effets sonores, comme le son téléphone
- Simuler une ambiance sonore perçue à l'extérieur d'une salle
- Améliorer la qualité sonore de sources anciennes ou de dialogues d'archives
- Augmenter les graves pour certains effets sonores
- Plus généralement lors des mixages, gérer les contenus en fréquences des différentes sources

**Quand on maîtrise mal les réglages d'un égaliseur, surtout s'il est complexe, pour éviter les erreurs il est recommandé de s'en tenir à un coupe-bas, un coupe-haut et éventuellement un filtre intermédiaire.** La captation d'un micro est une bonne base de travail et un excès de réglages peut produire des résultats inappropriés.

## Le travail sur la dynamique

### Pourquoi intervenir sur la dynamique ?

Comme nous l'avons vu précédemment, la dynamique représente le relief sonore entre les volumes forts et les volumes faibles.

La dynamique de sons d'origine captés peut rendre leur écoute difficile si elle nécessite une augmentation lors des faibles volumes et une diminution lors des forts volumes.

De plus, les fortes variations de volume dans un projet sont préjudiciables à leur gestion, puisque la plage dynamique utile doit être contenue dans certaines limites, aussi bien dans un mixage que lors de l'écoute. On peut noter que la norme EBU LUFS est imposée pour le respect d'un écart de dynamique bien défini dans un projet destiné à la télévision, lors de la fourniture aux chaînes du master PAD (Prêt à diffuser).

De plus, la gestion instantanée des volumes par la compression peut permettre une meilleure intelligibilité des dialogues. En effet, la parole consiste en variations de volume très rapides pendant la prononciation d'un mot ou d'une phrase. En régulant ces variations, les dialogues sont plus compréhensibles, surtout quand d'autres éléments sonores s'y superposent.

Par ailleurs, pour les mêmes raisons, d'autres types de sons peuvent mieux s'intégrer parmi d'autres si on gère correctement leur dynamique dans la continuité.

L'une des options dont dispose l'ingénieur du son pour gérer la dynamique sur la durée d'un enregistrement ou d'un programme est l'action manuelle sur le réglage de volume. Il l'augmente pendant les faibles volumes et le diminue pendant les forts volumes. Les traitements de la dynamique peuvent le faire de façon automatique.

Les outils pour le traitement de la dynamique sont le compresseur, le limiteur, le noise-gate et le dé-esseur.

Le **compresseur** ou **compressor** réduit la dynamique, le **limiteur** ou **limiter** en fait de même, mais de façon radicale. Le limiteur est une version extrême du compresseur. Compresseur et limiteur peuvent co-exister dans le même plug-in. L'**expandeur** ou **expand** augmente la dynamique.

Le **noise-gate**, ou simplement **gate** (ou porte de bruit) a pour rôle d'atténuer les parties des sons au-dessous d'un certain volume pour masquer le bruit de fond dans les silences.

Le **dé-esseur** compresse les aigus de la voix quand ils sont excessifs.

Comme avec les paramètres d'égalisation, ceux illustrés ci-dessous pour les traitements de la dynamique se retrouvent sur d'autres modèles.

## Le compresseur mono-bande

Avec un compresseur mono-bande, la réduction de volume se produit globalement sur l'ensemble des fréquences depuis les graves jusqu'aux aigus. Ses différents réglages déterminent son mode d'action.

Le rôle du compresseur est de régler les volumes automatiquement en permanence pour une source sonore. On peut comparer son action à une personne qui utilise un bouton de volume pour l'augmenter ou le diminuer au fur et à mesure, mais avec un temps de réaction bien plus rapide.

Le compresseur mesure en permanence le volume qu'il reçoit pour agir en fonction de celui-ci.



Figure 103 Le compresseur/limiteur de la série Digirack de Pro Tools, ici en version stéréo

La fenêtre centrale affiche la courbe de transfert du compresseur, indiquant la réduction de la dynamique au-dessus du réglage de seuil. L'axe horizontal indique les volumes d'entrée et l'axe vertical les volumes de sortie en relation. Au-dessous du seuil, la pente à 45 degrés indique que la sortie du compresseur est proportionnelle à l'entrée, ce qui signifie une absence de compression.

## Les affichages d'un compresseur

Avec l'exemple ci-dessus d'un compresseur stéréo (les voies d'entrée et de sortie sont doubles), on trouve ::

- **Levels** : est le niveau d'entrée gauche/droite. Il s'affiche sous forme instantanée. Le triangle indique le seuil d'action (THRESH pour **Threshold**).
- **Out** : est le niveau de sortie gauche/droite. Au-dessous des lettres « OUT » deux voyants s'allument quand une saturation se produit
- **GR (Gain Reduction)** : est l'indication de l'action du compresseur, plus ou moins intense en fonction de l'amplitude de l'affichage
- **Knee** détermine un arrondi plus ou moins anguleux de la courbe de compression qui a pour effet de rendre plus ou moins progressive l'action du compresseur autour du threshold, son niveau d'action
- **Attack** est le réglage du délai d'action du compresseur qui commence à agir après ce délai. Plus le temps d'attaque est faible plus vite le compresseur réagit aux variations d'entrée. Plus il est long, plus il réagit avec retard et plus les débuts des fichiers sont favorisés puisque leurs volumes n'ont pas encore été réduits par le compresseur
- **Gain** est le moyen utilisé pour compenser de manière continue la réduction de gain du compresseur, par exemple pour comparer son action avec l'original
- **Ratio** : est le taux de compression, autrement dit son intensité d'action. Il se règle en fraction avec le chiffre 1 en dénominateur. Le chiffre du numérateur est l'augmentation de volume d'entrée en dB qui correspond à une variation de 1 dB en sortie. Par exemple, un réglage de 3 :1 ou 3/1 donne une augmentation de volume de 1 dB quand il augmente de 3 dB en entrée. Poussé à l'extrême, un réglage de 100/1 correspond à la fonction Limiteur, avec lequel le volume de sortie ne peut dépasser le seuil. Une valeur inférieure à 1 au numérateur indique un mode d'expansion avec lequel tout volume au-dessus de la valeur du seuil est augmenté. Ainsi, par exemple avec un réglage de 0.5/1, une augmentation de volume de 0,5 dB dans le compresseur se traduit par une augmentation de 1 dB en sortie.
- **Release** est le temps que met le compresseur pour arrêter son action quand le niveau a baissé. Certains compresseurs ont un mode « Auto » avec lequel le temps de Release s'adapte automatiquement au contenu pour produire une compression optimisée.
- **Thresh (Threshold)** est le seuil de volume à partir duquel le compresseur entre en action et réduit le volume. Son réglage peut-être actionné par le bouton rotatif, ou pas la flèche triangulaire sur le réglage de volume d'entrée
- **Side-Chain** est le moyen (ici commun aux différents types de traitements de dynamique) dont on dispose pour rendre le traitement de dynamique sensible au contenu en fréquences, soit directement celles du signal, soit celles d'une source extérieure. On dispose de plusieurs types de filtrages pour sélectionner la gamme de fréquences choisies pour activer le traitement de dynamique.

## Le noise-gate ou gate.

Traduit par « porte de bruit », c'est un traitement destiné à atténuer plus ou moins fortement le bruit de fond dans les parties silencieuses d'un enregistrement. Cette « porte » s'ouvre ou se ferme en fonction du volume du signal. Le principe de base est de considérer que le bruit de fond est peu audible pendant les forts volumes, alors qu'il se remarque pendant les faibles volumes. Le noise-gate détecte les volumes en permanente. Au-dessus d'un certain seuil (le Threshold), le signal passe au-travers, alors que lorsque le volume passe au-dessous, le résultat en sortie est atténué en fonction de la valeur « RANGE ». Dans cette fonction, la valeur « ATTACK » règle le temps de réaction du noise-gate et celle

de « RELEASE » détermine le temps que met le traitement pour couper le son ou l'atténuer. Les réglages sont à déterminer pour que le résultat soit le plus naturel possible.



Figure 104 La version Expander/Gate de la série des traitements de dynamique de Pro Tools

## Le compresseur multi-bandes

Contrairement au compresseur mono-bande qui gère le volume global, le compresseur multi-bandes divise l'ensemble du spectre sonore en plusieurs bandes (généralement 4 ou 5) pour appliquer à chacune une compression différente et indépendante des autres. Exemple : lors de l'arrivée d'un son grave à volume élevé dans un compresseur mono-bande, celui-ci produit une baisse de niveau en même temps sur toutes les fréquences : graves, médiums, aigus. Avec un compresseur multi-bandes, on peut par exemple décider de compresser la partie « graves » mais pas les autres fréquences, ce qui évite l'effet de pompage décrit ci-dessus. En tout cas on peut compresser différemment, plus ou moins, les différentes bandes de fréquences.

C'est ce genre de compresseur que les radios FM utilisent principalement, pour modeler et personnaliser le son de leur station.

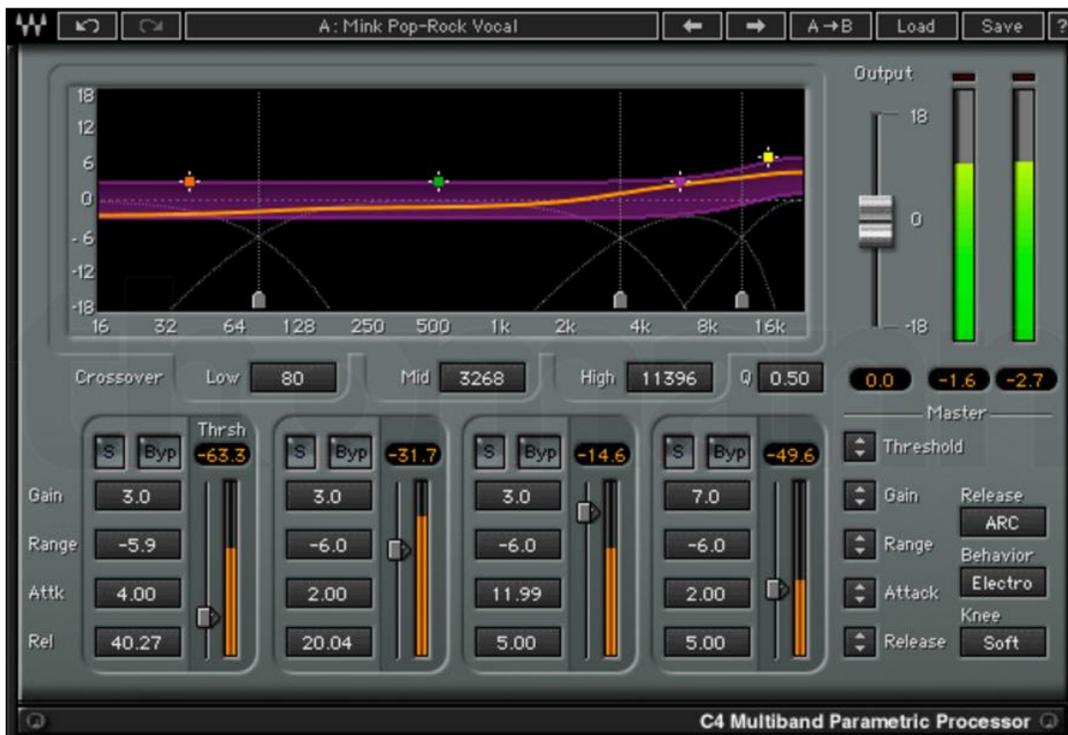


Figure 105 Le compresseur à 4 bandes C4 de Waves avec les réglages de chaque bande

## Quelques remarques sur l'utilisation des compresseurs

- Pour éviter l'exagération des attaques des mots sur les voix, on privilégiera les temps d'attaque très courts. Pour améliorer la compréhension d'un texte, on ajustera le temps de release autour de 100 ms, avec un taux de compression de 3 :1 à 4 :1.
- Avec un temps d'attaque long, les sons graves ayant généralement plus de volume électrique que les sons aigus, ils déclencheront le compresseur plus souvent, et auront donc tendance à favoriser les aigus qui le déclencheront plus rarement.
- Pour un nivelage de volume global et un effet efficace mais discret, on donnera la préférence à un taux de compression élevé et un long temps d'attaque et de Release.
- Des réglages inadaptés peuvent se remarquer et être inélégants
- Avec certains réglages incorrects, on peut obtenir un effet désagréable de « pompage », une variation de volume pendant la présence d'un son, notamment lors de l'activation du compresseur par un son grave intense.
- Il faut savoir se passer d'une compression quand elle n'est pas nécessaire, ou simplement quand on souhaite conserver le « naturel » d'un son, en particulier dans une séquence émotionnelle d'un dialogue.

## Le limiteur

Comme nous l'avons vu, il s'agit d'un compresseur extrême, qui empêche le volume de dépasser une valeur maximum. En dehors des versions logicielles, on le trouve sur les matériels d'enregistrement. On l'active quand on craint les dépassements de volume d'enregistrement afin d'éviter les saturations et distorsions. Mais son action excessive dénature le son et elle est néfaste.

La fonction limiteur peut être obtenue avec un compresseur logiciel en réglant au maximum le taux de compression (Ratio), le seuil (Threshold) à la valeur du maximum à ne pas dépasser, et le Release à la

valeur correspondant au signal. Pour respecter les nuances dans une voix d'acteur, on privilégie un temps de release d'une seconde ou plus. Cette durée est fixe dans les enregistreurs ou les caméras.

## Le dé-esseur

Il s'agit d'un compresseur qui effectue une compression dans la zone des aigus. Son utilisation se trouve dans la réduction des « sifflantes » de la parole, notamment dans les mots comprenant des « s ».



Figure 106 Le dé-esseur de la série Digirack

Avec le dé-esseur ci-dessus, la réduction de gain dans les aigus est effective quand on enclenche « HF ONLY ». Le bouton « LISTEN » permet d'écouter l'effet du traitement seul. On règle la fréquence d'action avec « FREQ » et l'intensité de son action avec « RANGE ». « GR » indique la réduction de gain qui se produit pour ces sifflantes.

## La réverbération

La réverbération est le nom donné au phénomène des réflexions multiples du son sur des surfaces environnantes réfléchissantes. Le même nom est utilisé pour le procédé qui recrée ce phénomène de façon artificielle.

Très utilisé en production musicale et en audiovisuel, l'effet de réverbération a plusieurs utilités : recréer l'acoustique d'un lieu de tournage quand une captation sonore a été réalisée de façon neutre à proximité d'un micro (doublage de voix, micro cravate) ou créer des ambiances artificielles (effets sonores en science fiction ou en jeu vidéo).

Auparavant uniquement disponibles sous forme physique (modèles à plaque, à ressort), de nombreux modèles existent aujourd'hui sous forme logicielle, chacun avec ses spécificités. Deux types de

réverbérations existent : la réverbération algorithmique et la réverbération à convolution. Les deux types existent en versions stéréo ou multicanaux.

La réverbération algorithmique simule les espaces acoustiques par synthèse, grâce à des calculs numériques par des modules de programmes spécifiques. Le résultat peut être convaincant. Mais elle n'a pas pour objet d'être réaliste.

La réverbération à convolution recrée des espaces acoustiques existants. Pour cela, elle utilise des échantillons sonores, appelés « impulsions », captés dans des endroits particuliers, cathédrale, salle de concerts, intérieur de voiture ou autres. Par un processus de calculs complexes qu'on appelle la convolution, on reproduit sur n'importe quel son l'ambiance acoustique qui correspond à un endroit existant. Par son réalisme, ce type de réverbération est largement adopté dans les milieux du film.

Un exemple : Audio Ease, célèbre pour sa réverb Altiverb à convolution propose une version spécialisée dans la simulation d'acoustiques de pièces, *Indoor*, avec le même principe et un réalisme excellent. Les positions des sources sonores et d'écoute sont choisies librement, soit de façon fixe, soit en mouvement, dans l'espace stéréo jusqu'au Dolby Atmos. On comprend donc l'intérêt d'une vérité sonore reconstituée pour le doublage de dialogues ou les bruitages de studio.



Figure 107 Le plug-in Indoor de Audio Ease

La même marque propose aussi sa légendaire Altiverb multicanaux à convolution incluant de nombreux espaces de multiples dimensions.



Figure 108 L'Altiverb 7 XL d'Audio Ease

La plupart des logiciels audio incluent au moins un modèle de réverbération, avec des qualités différentes. Un exemple : le modèle D-Verb ci-dessous qui accompagne le logiciel Pro Tools.



Figure 109 Le plug-in D-Verb fourni avec Pro Tools

La réverbération peut produire son effet en stéréo ou en surround à partir d'un fichier mono, en fonction de la version choisie (mono/mono, ou mono/stereo). C'est très utile pour placer dans l'espace par exemple les effets produits par un bruiteur dans un studio où le son enregistré est neutre et monophonique.

## Les réglages d'une réverbération

Ceux qui suivent concernent précisément le modèle ci-dessus, mais on les retrouve de façon similaire sur la plupart des autres plug-ins de réverbération.

### Gain

C'est le volume d'entrée dans l'effet

### Types de réverb

Un choix de différents types d'espaces acoustiques, avec pour chacun 3 choix : SMALL (Petit), MEDIUM (Moyen) ou LARGE (Grand)

### PRE-DELAY

C'est le délai d'action de la reverb. Plus on augmente le délai, plus on se rapproche virtuellement de la source sonore, puisque dans la réalité le son met plus longtemps à revenir. À l'inverse, avec un délai court on a la sensation d'être proche des réflexions, donc loin de la source.

### DECAY

Le temps d'amortissement de la reverb en secondes. Il concerne la simulation d'espaces plus ou moins réverbérants.

### HF CUT

Ce réglage permet de couper plus ou moins les aigus, pour obtenir une réverbération plus moins « claire ». Une coupure des aigus correspond à une salle avec des surfaces absorbantes.

### DIFFUSION

Définit le degré de densification, dans le temps, de l'écho initial. Les valeurs modérées ou faibles donnent un effet plus naturel que les valeurs élevées.

### LP FILTER

Un filtrage qui coupe plus ou moins la partie grave des sons réverbérés en fonction du réglage. Il est utile quand le son de base est chargé en graves qui alourdissent le son réverbéré.

### MIX

Dosage du rapport entre le son direct, donc non réverbéré, appelé « DRY » (complètement à gauche), et la réverbération totale, appelée « WET », (complètement à droite).

En mode de départ auxiliaire (voir ci-dessus), le réglage doit être en position WET maximum. En mode insert la position se règle en fonction de la quantité de réverb par rapport au son brut.

## Le denoising

Parmi les défauts fréquents qu'on cherche à éliminer, les bruits indésirables sont en bonne place. L'un d'eux est le bruit de fond. Il peut être de différentes natures du type souffle, du type rayonnement constant du réseau électrique. Ou bien des bruits intermittents se produisent par la proximité d'un smart phone. Il peut aussi s'agir de divers bruits parasites pendant une prise de voix par exemple : une sonnerie de téléphone, un objet qui tombe, un étternement, le passage d'une voiture ou autre.

Selon le type de bruit, différents traitements peuvent être utilisés pour les éliminer, à condition que leur volume ne soit pas nettement supérieur à celui du son principal.

En commun, à la base de ces traitements, on utilise l'analyse spectrale. Pour traiter le bruit de souffle, on prend une « empreinte » du bruit sur une partie où il est seul. On analyse cette empreinte et on applique à l'ensemble du signal le filtre sous forme du « profil » qui en résulte. En principe, on peut écouter séparément le bruit qu'on supprime pour vérifier que le son principal n'est pas entamé par le traitement. On doit être attentif à ne pas exagérer la réduction pour éviter les effets artificiels de type gazouillis qui se produisent dans ce cas. Waves a été l'une des premières à utiliser ce principe avec sa série X-Noise. La même marque a développé d'autres techniques pour parvenir à des résultats optimisés en utilisant des procédés d'intelligence artificielle, comme avec Clarity Vx Pro, spécialisé dans la réduction des bruits pour les voix.



*Figure 110 Le plug-in X-Noise de Waves en action. En blanc, la courbe du profil du bruit, en rouge le signal d'entrée et en vert le résultat en sortie. La ligne verticale à gauche de la fenêtre principale montre la réduction du bruit.*

Autre principe qui s'est largement développé à partir de 2020 : l'édition spectrale. C'est une sorte de Photoshop de l'audio. On analyse le spectre de la section qui comporte un bruit qui apparaît à un moment donné. Puis dans l'image produite on cherche la représentation du bruit et on l'efface, comme sur une partie d'une photo.

On trouve ce type de suppression de bruits parmi les outils de la gamme RX10 de la société Izotope ou chez Steinberg.

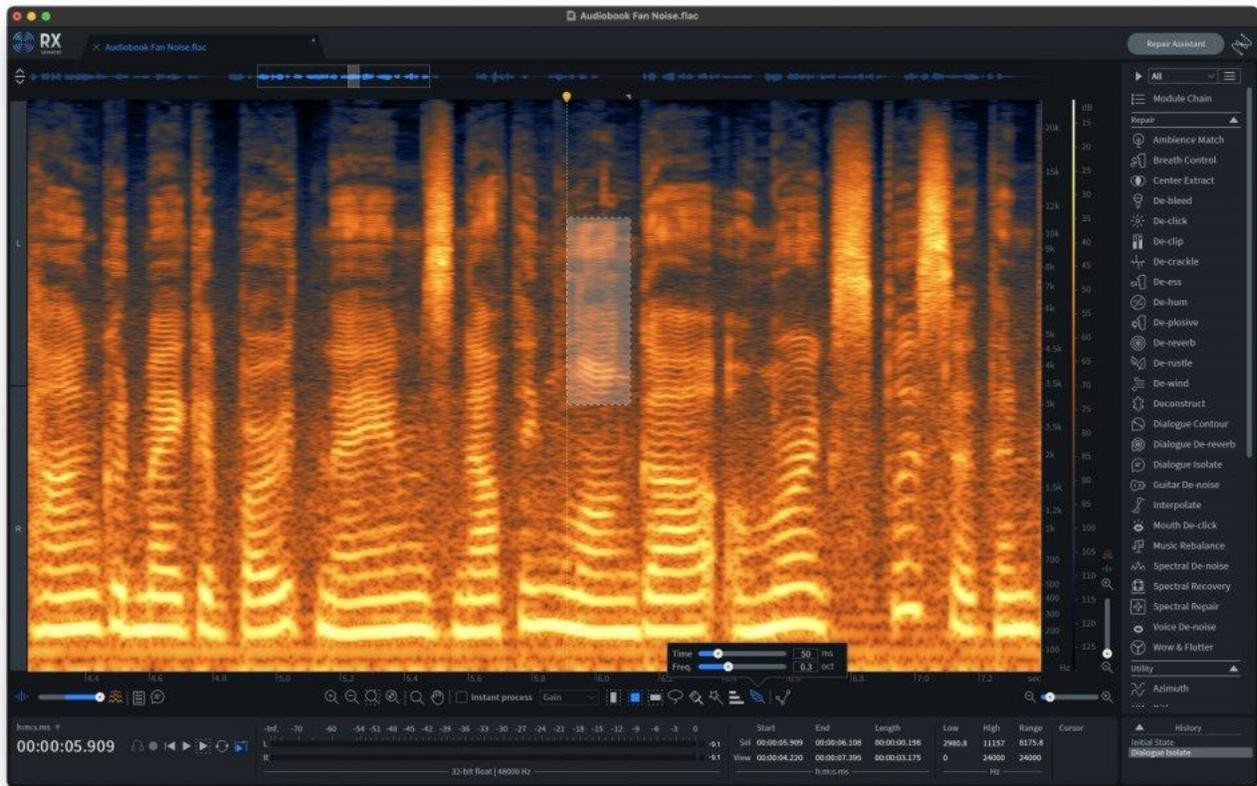


Figure 111 La page d'édition spectrale de RX10 d'Izotope. On voit sur l'image la sélection d'une zone où un bruit est présent. On va l'éliminer en l'atténuant avec les outils de traitements.

# Les effets temporels

## Le time-stretching

Ce traitement permet d'augmenter ou de réduire la durée d'un fichier audio, donc de l' « étirer » ou de le compresser (« to stretch » veut dire « étirer »). L'utilisation principale est l'ajustement de la durée pour l'adapter à un événement à l'image. Par ailleurs, en augmentant fortement la durée on obtient des sons complètement différents des originaux. Ce procédé est fréquemment utilisé pour la création d'effets sonores.

Anciennement ce traitement changeait la hauteur du son mais maintenant certains traitements de time-stretching proposent de la conserver.

PaulStretch, un traitement inclus dans le logiciel gratuit aussi Audacity en version simplifiée, permet des effets d'extension de durée spectaculaires (Audacity comprend aussi d'autres traitements temporels de très bonne qualité). PaulStretch existe aussi sous forme de logiciel séparé.

Bien sûr, un traitement de durée ne peut être réalisé en temps réel. C'est le résultat d'un calcul qui produit un fichier plus court ou plus long que l'original.

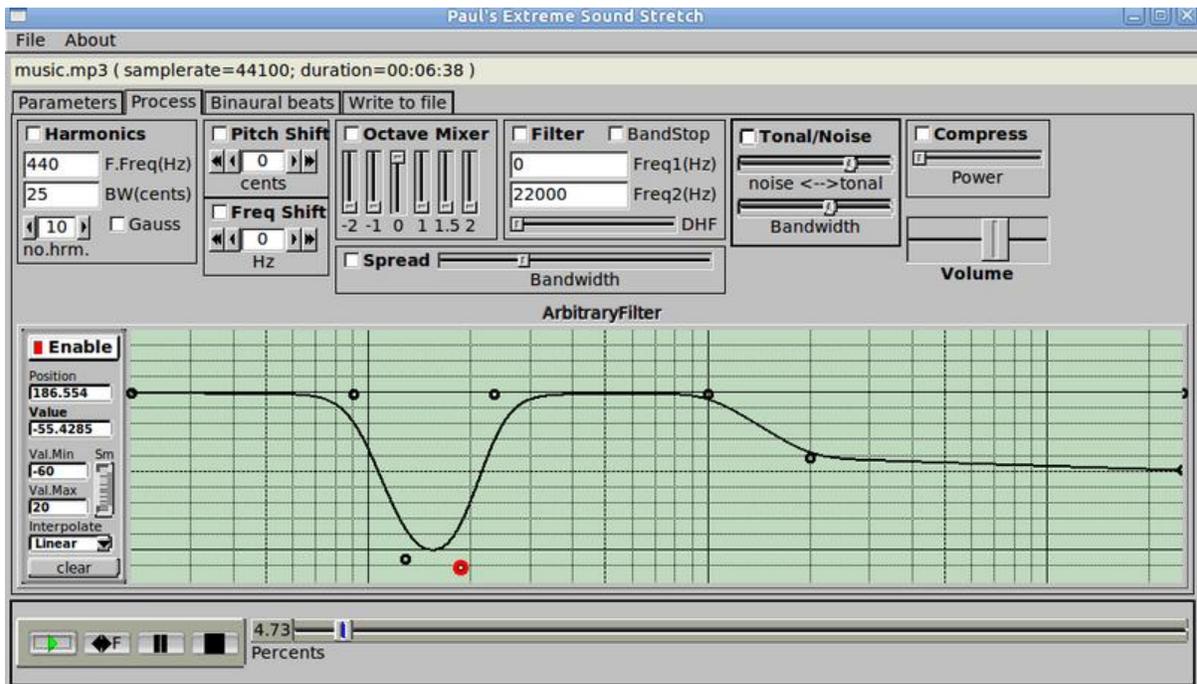


Figure 112 Paul Stretch en logiciel séparé, pour des extensions temporelles extrêmes

## Le pitch-shifting

Il s'agit d'un changement de hauteur. On décale l'ensemble des fréquences vers le haut ou vers le bas. La qualité du résultat est très variable. Elle dépend largement de celle de sa conception.



Figure 113 Le Graphic SoundShifter de Waves. Deux courbes sont proposées pour une variation évolutive : l'une pour la hauteur, l'autre pour la durée.

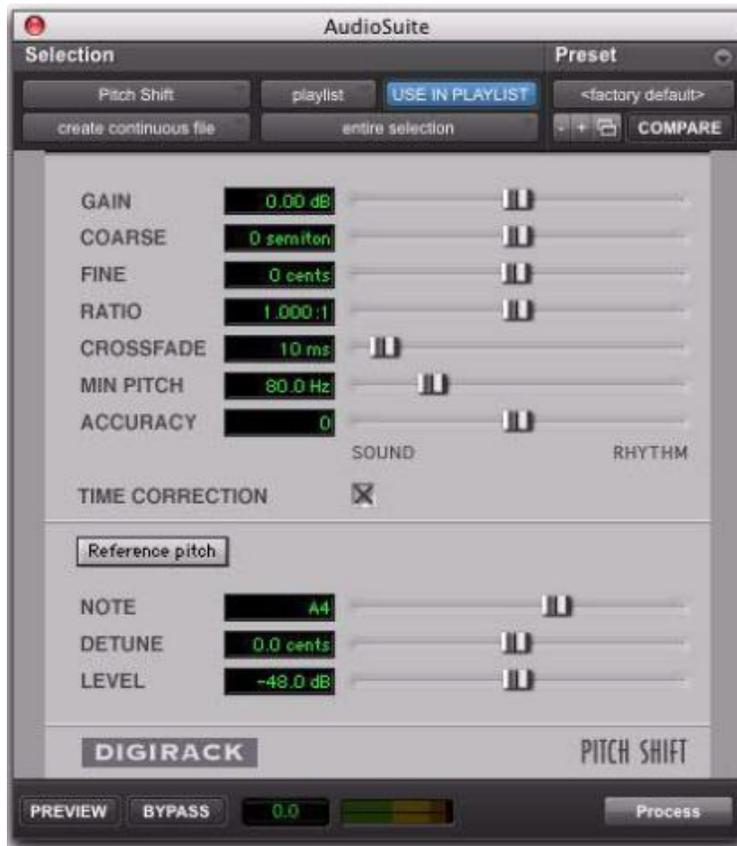


Figure 114 Le PITCH SHIFT de Pro Tools

En musique, le changement de hauteur est lié à la transposition. C'est pourquoi on trouve souvent des réglages de hauteur par demi-tons sur les traitements de pitch shifting. Il suffit de savoir qu'un demi-ton correspond à un changement de hauteur d'environ 6%, et une octave au double de la fréquence de base (ou à la moitié au-dessous).

Les réglages de traitement se retrouveront sur d'autres versions d'autres marques.

Ici le mode Preview permet l'écoute de l'effet avant traitement, mais d'autres traitements de hauteur fonctionnent en temps réel.

## Les réglages de pitch-shifting

Nous prenons comme exemple ceux de la figure ci-dessus pour le PITCH SHIFT de Pro Tools.

### Gain

Le volume d'entrée du traitement

### Coarse et Fine

Définissent en demi-tons et centièmes de tons la transformation de hauteur souhaitée, en + ou en moins

### Ratio

Le pourcentage de changement de hauteur par rapport à l'original. Au centre (1.000 :1) la hauteur est inchangée, plus à droite elle augmente et plus à gauche elle diminue.

**Crossfade, Min Pitch** et **Accuracy** sont des ajustements d'optimisation de la qualité sonore en fonction du type de son à traiter

**Time Correction** est un paramètre important. En principe, un changement de hauteur est associé à un changement de la durée du fichier. Plus on augmente la hauteur et plus le fichier raccourcit. Au contraire plus on la diminue et plus il s'allonge. En cochant la case Time Correction, on conserve la durée d'origine. Suivant le degré de transposition, la qualité sonore du résultat est plus ou moins dégradée. D'autres types de traitements de hauteur peuvent produire une meilleure qualité, notamment sur les voix, avec un rendu le plus naturel possible.

La partie **Reference Pich** donne la possibilité de prendre comme référence un fichier, de le faire correspondre à la hauteur d'un signal sinusoïdal pour procéder avec précision au changement de hauteur d'autres fichiers. Cette utilisation est surtout destinée aux traitements musicaux de réaccordage.

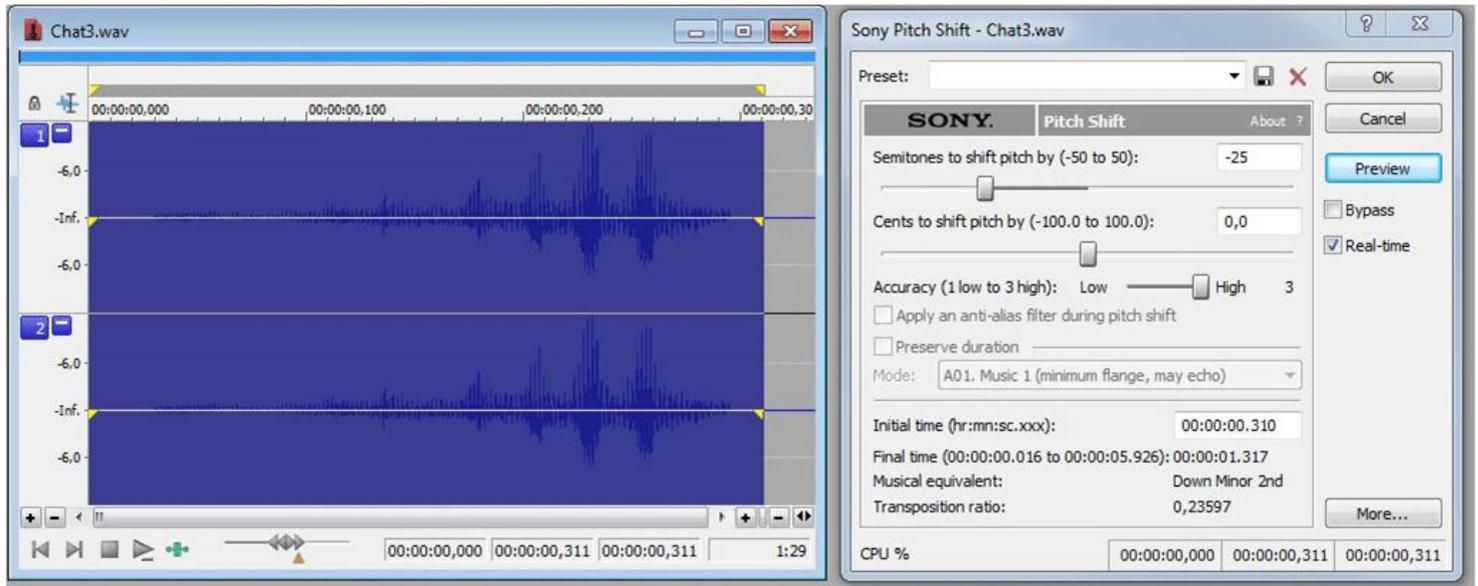


Figure 115 Le pitch shifting de Sound Forge

## Le délai (Delay) et l'Echo

À la base, le délai est une seule répétition du son, alors que l'écho est une série de répétitions du son qui s'affaiblissent jusqu'à disparaître. L'écho est utilisé pour des effets spéciaux.

Aussi bien le délai que l'écho sont utilisables en mode insert ou en départ auxiliaire. Dans ce dernier cas, selon le principe de base, plusieurs pistes pourront subir le traitement de délai avec un seul plug-in. Tout comme l'effet de réverbération, l'effet d'écho peut être produit en stéréo ou en surround à partir d'un son mono, en fonction de la version choisie (mono/mono, ou mono/stereo).



Figure 116 Le Mod Delay III de Pro Tools, un délai modulable, ici avec une entrée mono ainsi que la sortie

Ce Mod Delay III, un plug-in aussi tourné vers les utilisations musicales, peut néanmoins être intéressant pour créer des effets spéciaux. Pour cette raison, nous laisserons aux musiciens l'utilisation des fonctions Sync, Meter, Tempo, durées de notes et Groove.

En ce qui nous concerne, les réglages utiles sont :

- La durée du délai par le bouton central, en millisecondes (ici 5 000 ms, soit 5 s)
- **FBK**, abréviation de **Feedback** est le taux de réinjection du délai sur lui-même. Il détermine le nombre de répétitions. Une seule répétition se produit avec un réglage de 0%. En rotation vers la droite on obtient de plus en plus de répétitions. En dessous de 0%, en rotation vers la gauche on obtient des effets bizarres, de tunnel ou autres.
- **LPF** coupe plus ou moins les aigus des répétitions qui perdent des aigus au fur et à mesure des répétitions
- **Rate** définit la vitesse de modulation de la durée du délai
- **Depth** règle l'amplitude de la modulation
- Le réglage **Dry/Wet** dose la proportion du signal direct, non traité, par rapport à l'effet total, sans le son de base

## Les transformations vocales

Différents moyens sont utilisables pour transformer une voix en une autre, complètement différente. On peut remplacer une voix d'homme ou de femme par une voix de femme ou d'homme ou d'enfant, supprimer plus ou moins l'expression en la rendant plus neutre, robotique.

Les applications vont de la création de voix de dessins animés à celles de science fiction.



Figure 117 Le plug-in TRAX de l'IRCAM est destiné à la transformation des voix.

## Les autres traitements

Des traitements de plus en plus sophistiqués et de plus en plus performants apparaissent progressivement. Certains d'entre eux se destinent particulièrement à l'univers de la post-production et facilitent le travail de réparation de fichiers audio problématiques.

Un exemple typique est celui de la société Izotope avec sa série « RX Postproduction Suite ». Dans celle-ci, une grosse partie appelée RX est dédiée au traitement des voix, et en particulier aux dialogues, notamment ceux des doublages (bruits de bouche, de respiration, etc.), suppression des bruits de fond divers, et même réduction de réverbération.

D'autres marques proposent aussi des outils d'amélioration sonore, comme par exemple CEDAR Audio pour la suppression des bruits.

L'un des exemples typiques est la suppression d'un bruit parasite superposé à un dialogue, par exemple une sonnerie de téléphone, en travaillant directement sur le spectre sonore de la partie incriminée, comme vu précédemment.

Cependant, lors des prises de son il est fortement recommandé de ne pas compter sur ces traitements ultérieurs pour des réparations de défauts. D'une part, les problèmes ne seront pas forcément résolus à 100% (par exemple l'effet du vent dans un micro), mais de plus ce travail supplémentaire pourra handicaper les opérations de post-production pour le temps qu'ils peuvent nécessiter.

On a donc largement intérêt à produire des enregistrements de la meilleure qualité possible, dès l'origine. Rien ne les remplacera parfaitement.

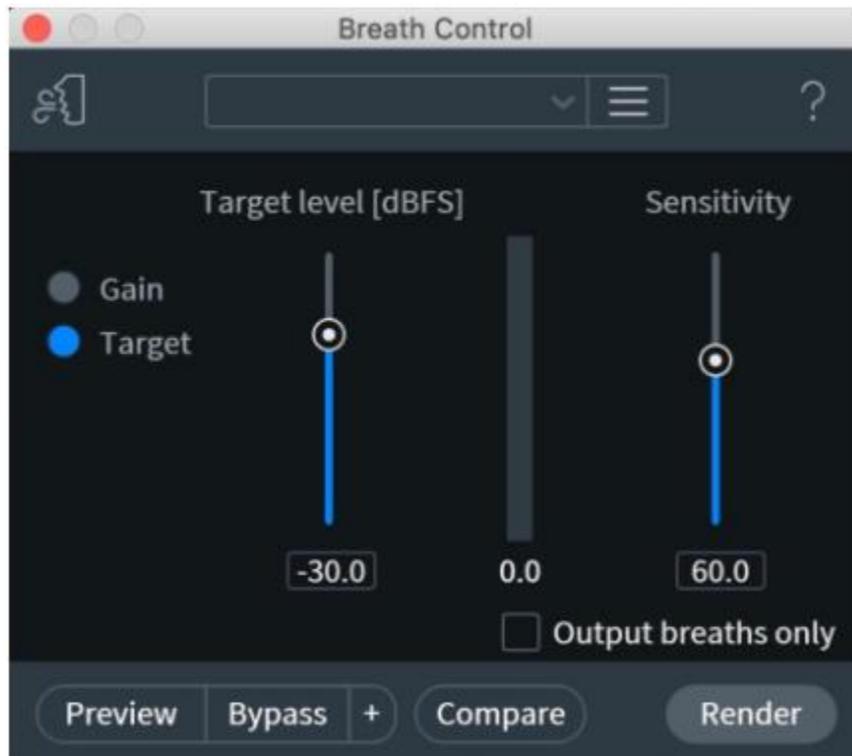


Figure 118 Le module *Breath Control* de traitement des respirations de RX Izotope

Autre traitement, l'inversion de la lecture d'un son, une fonction appelée « **Reverse** » est à inclure dans la panoplie des effets sonores.

Il faut aussi évoquer l'**effet Doppler**, cette variation de hauteur qui correspond au passage rapide à faible distance d'un véhicule. Il s'agit du phénomène physique décrit précédemment. C'est la variation de hauteur qu'on perçoit quand un véhicule passe rapidement devant soi.

Cet effet peut être créé artificiellement par des logiciels, par exemple pour produire un réalisme lors d'un passage de vaisseau spatial, ou en jeu vidéo pour des mouvements en images de synthèse. D'ailleurs Unity et Unreal Engine proposent tous deux des fonctions audio pour créer l'effet Doppler.

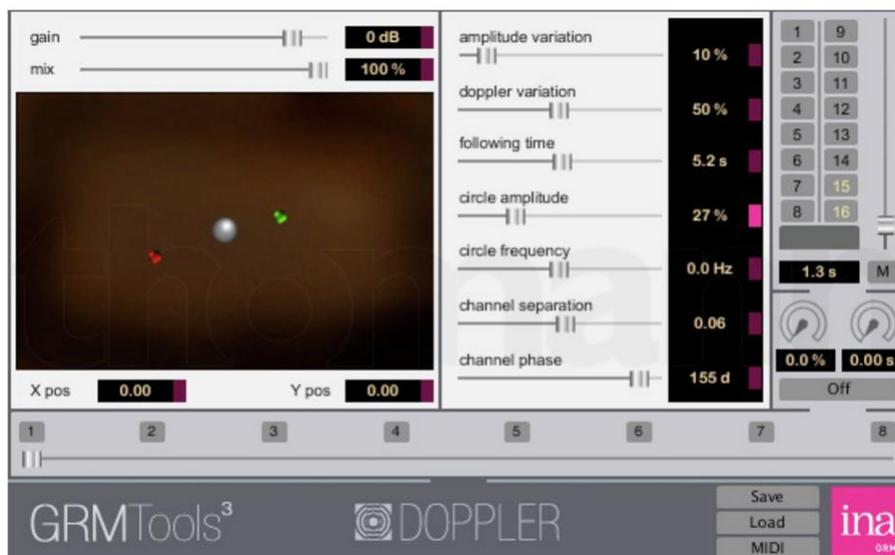


Figure 119 Dans la vaste collection des *GRM Tools*, *DOPPLER* simule l'effet du même nom.

# Le sound design et les effets spéciaux

En dehors des sons de la réalité, issus des prises sur le terrain ou recréés par un bruiteur, d'autres sons peuvent être nécessaires pour produire une sensation de réalisme sur des images irréelles. Quelques exemples de ce genre de situation :

- Le célèbre sabre laser de Star Wars
- Le passage d'un vaisseau spatial
- La voix d'un monstre
- Les mouvements motorisés d'un robot
- Les actions dans un dessin animé
- Les armes de science fiction
- Les sons d'interaction des jeux vidéo

Quand on se trouve dans une telle situation on a deux possibilités : soit puiser dans les banques d'effets sonores, soit les fabriquer si on en a le temps.

## Les collections d'effets spéciaux sonores

Appelées aussi des banques, ces collections comportent des effets sonores classés par catégories. Les meilleures sont payantes, et vendues soit par collections entières, soit à l'unité. Sur les sites qui les proposent, on cherche, généralement en anglais, ceux qu'on souhaite utiliser. La pré-écoute en ligne donne une idée du son.

Quelques un de ces fournisseurs d'effets spéciaux. Ils sont souvent fournisseurs de sons réels aussi :

Quelques exemples d'éditeurs :

- Blastwave
- Hollywood Edge
- Sound Ideas
- SFX Source
- Pro Sound Effects
- SoundDogs
- Freesound

## La creation des effets spéciaux

Appelée "sound design", la création des effets sonores est un vaste et passionnant univers de créativité et d'expérimentations. Il s'agit de sons imaginaires, donc en principe sans ressemblance avec des sons réels. Pour les réaliser, on dispose de deux directions possibles : la captation de sons suivie de leur transformation, ou bien la création directe par les synthétiseurs, quels qu'ils soient. Ceux-ci sont des

outils programmables qui produisent des sons par un ensemble complexe qui obéit à des règles bien définies.

Une version « hybride » est celle que proposent les « samplers », comme Kontakt ou UVI Workstation, qui utilisent des captures de sons réels qu'on peut déclencher avec un clavier ou par des événements MIDI inscrits dans la Timeline d'un logiciel. Ces samplers incluent souvent divers traitements en temps réel, dont une partie est similaire à ceux des synthétiseurs.

Une troisième voie intermédiaire consiste à utiliser toutes sortes de sons issus de bibliothèques existantes pour les travailler à volonté. Le sound designer va utiliser l'ensemble des moyens qui lui semblent appropriés pour obtenir l'effet souhaité (avec quelquefois d'heureuses surprises).

Tous les traitements décrits auparavant sont utilisables à volonté, avec ou sans modulations : réverbération, time stretching, pitch shifting écho/délai, flanger, Doppler, filtrages, transformation de voix...

Suivant le format sonore final et les besoins du contexte, les effets seront créés en monophonie, stéréo ou multicanaux. Par exemple, le passage d'un vaisseau spatial sera utilement réalisé en surround, et même en Dolby Atmos.

Le canal de sub-graves d'un cinéma peut être mis à profit pour renforcer la sensation de puissance d'un effet et le rendre plus impressionnant.

Souvent les effets sonores sont obtenus par la superposition de plusieurs sons mélangés, dont l'un pourra constituer le début, alors que les autres donneront le corps principal.

Les logiciels d'effets sonores sont aussi très utiles : ceux du Forum de l'IRCAM, ou les GRM Tools, ou d'autres, comme les traitements inclus dans les samplers.

Et comme nous l'avons vu précédemment, de nouveaux types de programmes informatiques apparaissent pour fabriquer une multitude d'effets sonores par des calculs en temps réel, sous l'appellation « audio procédural ».

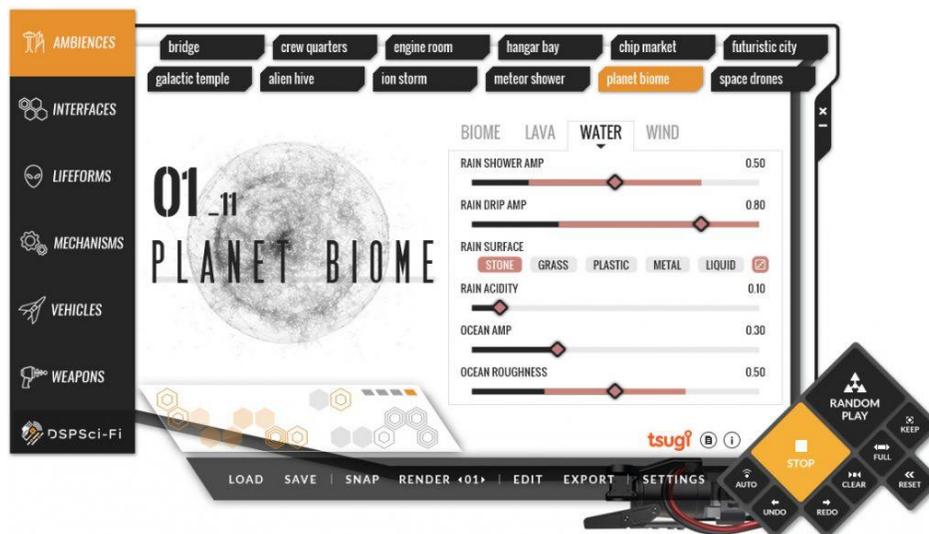


Figure 120 Une fenêtre de la section Science Fiction de la génération procédurale de Tsugi audio

# Le mixage

Le mixage est l'étape où on utilise tous les sons d'une production pour les mélanger. En anglais, « to mix » veut dire « mélanger ». C'est le moment où on va créer le son final du projet, celui qui sera encodé dans le format de diffusion définitif, Digital Cinema Package (DCP), MPEG4 ou autre. De très nombreuses décisions sont à prendre, aussi bien de nature technique qu'artistique.

En fin de travail, toutes les pistes séparées sont réunies sur une piste appelée « master », sous forme stéréo ou multicanaux.

Le mixage est le moment de la mise en commun de tous les éléments sonores d'un projet. Suivant son importance, le nombre de ces éléments, répartis sur autant de pistes d'une console de mixage, peut nécessiter jusqu'à plusieurs personnes ensemble.

Le mixage est un processus très créatif, tout en étant au service du scénario, sous contrôle du réalisateur. Tout doit être mis en œuvre pour intégrer le spectateur dans l'histoire, pour qu'il y soit immergé. On va capter son attention par tous les moyens possibles, lui créer des émotions, et maintenir son attention en éveil en créant des contrastes sonores variés. C'est bien sûr la structure du projet et son déroulement qui sont les bases de la philosophie du mixage. Par exemple, le montage est déterminant pour le résultat final.

À cette étape, la musique peut déjà avoir été composée et mixée, mais nécessiter des ajustements, en collaboration avec le réalisateur. Ou encore, une fois le montage terminé (ou des sections du montage), le compositeur peut à ce moment modifier sa composition en tenant compte des événements marquants du scénario, pour obtenir une correspondance entre certains éléments de la musique et les cuts dans le film.

Le rôle du réalisateur est essentiel, dans la mesure où il est le maître de son projet, et il dirige le mixage qui doit correspondre à sa vision d'ensemble.

## Les conditions de travail du mixage

Dans le cadre professionnel, le mixage s'effectue dans un « auditorium » (on dit souvent un « audi »). Équivalent d'un petit cinéma, c'est un local disposant du matériel de gestion du son (console de mixage et effets), des enceintes de bonne qualité produisant un volume sonore proche de celui des salles de cinéma, et un écran de dimensions suffisantes pour pouvoir juger correctement du rapport son/image. Les conditions sont réunies pour un travail de qualité, grâce au rendu sonore dans une salle dont l'acoustique est traitée pour éviter les résonances.



*Figure 121 L'auditorium de la Cité du Cinéma (Saint-Denis)*

## **Le volume sonore**

Au début de ce document, nous avons vu que notre perception des fréquences dépend du volume d'écoute. À faible volume, on entend surtout la zone de fréquences autour de 3 à 4 kHz.

Lors du mixage en écoutant à volume trop faible, on ne se rendra pas compte du résultat dans les graves et les aigus. À fort volume, la sensation graves/aigus est plus adaptée aux conditions d'écoute correspondantes, mais peut provoquer une fatigue auditive et même un risque de perte d'audition.

Des recommandations indiquent pour les enceintes principales un volume réglé à 85 dBA en mixage cinéma et 79 dBA en mixage télévision. La différence vient du fait que le volume d'écoute de la télévision est plus faible que celui au cinéma. Ce calibrage s'effectue avec un sonomètre.

Pour les mêmes raisons, on travaillera autant que possible à volume constant, afin d'éviter les changements de perception des fréquences en rapport avec le volume.

Notons qu'avec un casque, des volumes sonores très élevés peuvent être produits directement dans les oreilles, jusqu'à mettre aussi en danger l'audition. Et il est difficile de mesurer la pression acoustique produite par un casque !

## **Enceintes ou casque stéréo ?**

Dans le cas d'un mixage stéréo, celui pratiqué à l'ILOI, on dispose d'un double choix : soit utiliser le studio son, soit travailler au casque. Dans le premier cas, en important les pistes son et image du projet dans une session Pro Tools, on se trouve dans les conditions d'un petit auditorium avec les enceintes Mackie ou Genelec pour l'écoute.

L'autre possibilité est le travail au casque. Elle comporte des avantages et des inconvénients. Les casques sont en quantité suffisante à l'ILOI pour être largement disponibles pour chacun (la salle son ne l'est pas toujours). Le casque isole du bruit extérieur et procure une bonne concentration et une indépendance de l'acoustique d'une salle. Plusieurs personnes peuvent travailler sur des projets différents sans se gêner mutuellement. Mais en contrepartie, le travail au casque sur une longue durée procure une fatigue qui dépend du réglage du volume et du fait que le son arrive directement dans les oreilles, sans la réalité d'une pièce. De plus, la plupart des casques ne permettent pas de percevoir les graves avec la même sensation qu'avec des haut-parleurs. On peut donc être amené à exagérer la place des sons graves dans le contenu du mixage. Et de plus, la sensation d'écoute stéréo est faussée par l'absence de centre virtuel (voir précédemment dans le document).

## **La destination du mixage**

Parmi les grands principes du mixage, on tente de fournir à l'auditeur un résultat sonore équilibré, avec un rendu correct entre graves et aigus. Un même mixage sera évidemment entendu très différemment entre une salle de cinéma et un poste de télévision, ou pire encore sur les haut-parleurs d'un ordinateur portable. En mettant à part ce dernier, en principe on devrait optimiser le mixage en fonction de sa destination, cinéma ou télévision. Par exemple, pour un téléfilm, il est recommandé de pratiquer des tests en écoutant le résultat sur un téléviseur, même si l'épreuve s'avère douloureuse !

## **Les grands principes du mixage**

Le mixage est autant une opération technique qu'artistique. On est à la fin des opérations de post production. Une grosse charge de responsabilité repose sur tous ceux impliqués de cette opération : faire en sorte essentiellement que le spectateur ait la sensation d'être « dans » le film.

Une perspective globale de la totalité de la partie son est aussi nécessaire à cette étape, pour permettre une dynamique suffisante, avec des parties faibles et d'autres très intenses. Ces décisions dépendent fortement du volume et du type de diffusion final.

L'un des critères principaux est de s'assurer que le son corresponde à l'esprit des images. En l'occurrence, une scène proche doit être entendue avec un son proche. Inversement, une scène éloignée doit être entendue avec un son éloigné, même si le son de base a été capté à proximité. On parle de « plans sonores ». La partie « perception des sons » de ce document donne des indications à cet effet, concernant les caractéristiques des sons en fonction de leur distance. Les traitements audio décrits ci-dessus sont autant d'outils pour parvenir à ce résultat.

Concernant la partie technique du mixage, l'un des points essentiels est de régler correctement le volume de travail permanent pour ne pas être tenté d'augmenter les volumes séparés des différents éléments. C'est ce qui permet d'éviter de réduire le volume général quand on s'aperçoit qu'il est excessif, rendant difficile la gestion des volumes relatifs.

## Les consoles de mixage

On utilise trois types de consoles de mixage :

- les consoles matérielles autonomes qui renferment toutes les circuits et traitements en interne
- les consoles matérielles sans traitements internes (ils sont effectués par l'ordinateur) ; il s'agit donc de télécommandes des fonctions du logiciel par des accès physiques. C'est le cas de la Digi003 du studio son de l'ILOI qui donne accès à toutes les fonctions du logiciel Pro Tools.



*Figure 122 La console Digi003 Digidesign du studio son*

- les consoles logicielles, aussi appelées DAW (pour Digital Audio Workstation, ou Station de travail audio numérique), où tous les traitements sont réalisés par l'ordinateur. Ces consoles logicielles sont des équivalents des modèles hardware à la différence que tout y est virtuel, y

compris l'accès à tous les réglages, accessibles par la souris et repose là aussi sur des calculs internes à un ordinateur



Figure 123 L'écran de travail sur Pro Tools : au premier plan la console de mixage, à l'arrière plan, la fenêtre d'édition



Figure 124 La console de mixage SSL Duality Fuse, à la fois munie de traitements internes et de fonctions de commandes de DAW



*Figure 125 La console Avid S4, une interface de commande pour Pro Tools ou autres stations DAW*



*Figure 126 La console Fairlight Da Vinci Resolve 18*

Les modèles ci-dessus sont des exemples de matériels utilisés en milieu professionnel en 2023 pour la gestion de projets de large envergure.

Les consoles de mixage en hardware sont maintenant majoritairement numériques avec leurs traitements en interne. Mais un certain nombre de modèles hardware sont des ensembles de commandes physiques des traitements numériques délégués à un ordinateur. Leur avantage est la manipulation directe des réglages sans obligation d'utiliser la souris. C'est le cas pour la console Digi003 de l'ILOI, dont le but principal est de piloter la plupart des fonctions du logiciel Pro Tools installé sur le Mac Pro du studio. Les 8 faders peuvent être affectés à un nombre de pistes très supérieur. Ils constituent une « Bank » de 8 et on peut basculer entre ces groupes de 8 voies par un switch. C'est ce qui permet d'accéder à un nombre élevé de pistes dans un espace réduit.

Cela dit, certains ingénieurs du son réalisent leurs mixages à la souris, « In the box » comme disent les anglo-saxons.

## Les stations de travail audio

Il s'agit des ensembles de matériels et de logiciels propres à l'audio numérique. En anglais, on les appelle « DAW », pour Digital Audio Workstations.

Vidéo et audio sont intégrés dans un seul logiciel. Ils comportent un mixeur et des traitements audio. On peut y ajouter des plug-ins audio supplémentaires (AAX, Audio Units ou VST)

- **Nuendo** Steinberg
- **Cubase** Steinberg
- **Pro Tools** Avid en position dominante dans les studios de postproduction
- Adobe **Audition**
- **Final Cut Pro X** Apple
- **Premiere** Adobe
- **Vegas Pro** Magix
- DaVinci **Resolve 17 – Studio** Blackmagic Design
- **Reaper** Cockos

Tout travail sur le son, dans toutes les étapes d'une production, peut pratiquement être réalisé avec d'excellents résultats sur n'importe lequel des logiciels cités ci-dessus. Les choix d'un logiciel par rapport à d'autres dépendent de plusieurs facteurs : échange de projets entre studios (standard de fait), coût, préférences personnelles. Un studio autonome n'a pas les problèmes de compatibilité que ceux qui fonctionnent en réseau sur le web ou en salles interconnectées.



Figure 127 Une fenêtre de travail de Resolve 17 Studio de DaVinci.

## Le travail de mixage

Le travail se réalise sur une console de mixage, réelle ou virtuelle, en correspondance avec la complexité du projet. Pour des projets peu complexes avec un nombre d'éléments réduits, on peut se contenter de modèles simples et obtenir un travail de haute qualité.

On est en synchro avec l'image. Pour éviter de charger inutilement le CPU de l'ordinateur de travail, on évite d'utiliser la vidéo finale. On peut se contenter d'une basse définition, puis associer le résultat sonore terminé à la vidéo définitive en résolution maximum.

On dispose de multiples pistes de sources diverses (dialogues, voix off, effets sonores, ambiances, musique...) qu'on va rassembler pour en faire un nombre de pistes correspondant à la destination finale (stéréo, 5.1, 7.1, 7.2, Dolby Atmos ou autre).

Dans une même session de travail, les pistes d'origine sont de natures diverses : mono, stéréo, ou multicanaux.

Le principe général du mixage est de doser chaque élément sonore pour lui donner sa place dans le résultat final. La base du travail est de régler chacun des volumes, tout en les gérant globalement pour qu'ils restent contenus dans une certaine limite électrique avant saturation, par exemple -18 dB FS ou 0 dB FS (selon la destination) et de relief dynamique, suivant les médias de destination. À ce propos, un film destiné à la télévision devra se conformer à la norme LUFS (aussi nommé LKFS) EBU avec un contrôle par affichage séparé pour en faire le contrôle permanent.

Cette contrainte de relief sonore n'existe pas pour la diffusion sur les réseaux sociaux ou sur YouTube/Vimeo. Cependant, il est utile de savoir que les plateformes de streaming vidéo, de même que celles de musique, appliquent des traitements de normalisation de volume sonore. Ainsi, un volume excessif est diminué lors de la lecture. Par conséquent, la recherche d'un fort volume par une utilisation

excessive de la compression de dynamique est rendue inopérante par cette tentative. Au contraire, le son des vidéos où la dynamique est mieux respectée peut même être entendu avec plus d'intensité. Des logiciels permettent d'évaluer l'impact du traitement du son sur les plateformes. Avec l'un d'eux, Loudness Penalty, par un glisser/déposer sur la page du site indique l'action sur le volume de YouTube, Spotify, TIDAL, Apple Music.



Figure 128 Le module de contrôle de la dynamique Clarity de TC Electronic

Ajuster les volumes des pistes est donc le rôle fondamental du mixage. C'est ce qui explique pourquoi parmi les différents réglages d'une console ceux qui concernent le volume des pistes, les **faders** ont la plus grande dimension.

Autre rôle du mixage, celui de définir la provenance des sons dans l'espace stéréo ou multicanaux du projet final. Quand les sources d'origine sont déjà dans l'un de ces formats, on n'aura pas à se soucier de leurs positions, puisqu'elles sont déterminées. En revanche, pour toutes les sources monophoniques, on devra soit les positionner, de façon statique ou en mouvement, au centre, ou entre les différentes sources de reproduction. Cette fonction de positionnement s'effectue avec le réglage de « **panoramique** », ou « **Pan** » ou encore « **Pan Pot** ». Le plus courant est le panoramique stéréo, sous forme de bouton rotatif, avec un repère sur le dessus. À fond à gauche, le son est dirigé vers le haut-parleur de gauche, alors qu'à fond à droite il est dirigé vers le haut-parleur de droite. Au centre, le son sera perçu devant l'auditeur.

Dans la grande majorité des mixages cinéma, les dialogues sont mixés au centre, avec éventuellement ajout de réverbération sous forme de simulation d'ambiance stéréo. Pour les autres éléments, on a toute liberté pour positionner les éléments dans l'espace gauche/droite/avant/arrière.

En mixage multicanaux, on peut utiliser un joystick qui positionne une source mono entre l'avant, le centre et l'arrière, soit de façon fixe, soit en mouvement.

L'un des principes de facilitation du mixage consiste à réaliser auparavant des pré-mixages, les stems. Il s'agit de regroupements de pistes de même nature et pré-mixées, dialogues, effets sonores, ambiances, musique. Le mixage de chaque stem peut être modifié si nécessaire lors du mixage final. L'intérêt est de disposer des groupes d'éléments prêts à être utilisés. Par exemple, réaliser un stem pour les dialogues est l'occasion de gérer les différentes prises qui peuvent provenir de captures à des moments et des circonstances différents. On aura donc besoin de les faire « **matcher** » autant que possible, avec les traitements nécessaires, EQ, volumes d'ambiances, etc. D'ailleurs, on a intérêt à séparer les groupes de pistes de voix live de ceux des voix off qui en principe devront être traités différemment.

Lors du mixage, même si des traitements en amont ont déjà été réalisés, ils restent encore possibles.

Rappelons qu'avant de passer à l'étape du mixage, il est indispensable de préparer soigneusement les fichiers et de tous les « nettoyer » (bruits, défauts divers) et les optimiser au maximum. Rien de plus démotivant et frustrant à cette étape que de perdre du temps avec des détails, avec un travail intense en perspective pour le mixage lui-même.

Parmi les traitements, l'un d'eux est très fréquent : la réverbération, surtout du type « convolution ». C'est elle qui permet de simuler les espaces acoustiques réels, comme nous l'avons vu précédemment. Elle est l'élément qui donne une crédibilité à des prises de proximité ou des doublages de voix pour les rendre conformes au contexte visuel.

Autres traitements fréquents à cette étape, l'égalisation et la compression de dynamique, chacun pour donner leur place aux différents éléments, chacun par rapport au son global et pour augmenter l'intelligibilité des voix.

Et en ce qui concerne la crédibilité générale des scènes, lors des opérations de mixage, on devra être attentif à intégrer des ambiances qu'on aura pris soin d'enregistrer précédemment, ou même recourir à des banques de sons.

Les effets spéciaux du sound design font partie des éléments à gérer au mixage. On a aussi intérêt à les grouper par catégories sur des pistes séparées pour faciliter leur gestion à cette étape.

En bref, tout ce qui sera de nature à rendre le mixage le plus simple possible (et le plus efficace) doit être entrepris en amont.

## Les dialogues

Les dialogues étant une part essentielle d'un projet, on y accorde une importance particulière lors du mixage. L'intelligibilité est l'objectif essentiel. La musique ou les effets sonores ne doivent pas distraire l'attention sur les voix. L'un des facteurs est aussi la « présence » qui est en relation avec les aigus. Un son de voix clair produit une meilleure compréhension des dialogues qu'un son étouffé. L'égaliseur est un outil précieux pour parvenir à ce résultat.

La compression de dynamique est un autre moyen pour améliorer l'intelligibilité des dialogues quand l'environnement sonore est chargé.

Les dialogues font partie des éléments émotionnels. Or, l'émotion exprimée par un dialogue est directement en rapport avec les variations de son volume, sa dynamique. Un compromis est à trouver pour conjuguer compression/émotion. Si on cherche à préserver la compréhension en même temps que l'émotion, on devra peut-être gérer les volumes des autres éléments pour laisser une place suffisante aux nuances de la voix. Pour autant, dans ce contexte la compression reste utilisable si elle est gérée raisonnablement.

## Les stems

À la fin des opérations précédentes, des groupes de mixages sont produits, chacun pour les catégories de sons principaux : dialogues, bruitages/ambiances, effets spéciaux et musique. Ces groupes sont souvent appelés « **stems** » (terme issu de **ST**EreO **M**aster, mais les stems sont souvent en multicanaux).

À partir de ces stems on produit le master son final dans le format audio nécessaire. Pour les versions internationales, on fournit le même mixage que la version d'origine, mais sans les dialogues qui font l'objet de stems séparés.

La fixation du mixage dans son format final est réalisée sur des pistes séparées de la session de travail, ou bien sur une machine séparée. À ce moment, on peut encore intervenir sur le son global en réglant le volume final ou en insérant des traitements appliqués sur la piste master. On appelle « mastering » ce traitement final.

## Le mixage en résumé

- C'est le mélange de tous les éléments sonores : dialogues, ambiances, effets spéciaux de sound design, musique
- C'est l'étape finale pour le son, la création des pistes définitives, stéréo ou multicanaux
- Le travail se fait en synchro avec l'image, généralement dans une station de travail informatique
- On utilise une console de mixage réelle ou virtuelle
- On crée le look sonore final du projet.
- De nombreuses décisions sont prises : volume des éléments, leur positionnement dans l'espace (stéréo, 5.1, 7.1, Atmos), gestion de la place des éléments les uns par rapport aux autres
- On applique tous les traitements requis : simulations d'ambiances acoustiques par réverbération à convolution, optimisation des sons (égalisations, compressions)
- Une perspective globale est nécessaire : l'évolution des niveaux sonores parties faibles/parties fortes
- On utilise souvent des pré-mixages (dialogues, effets sonores, ambiances, musique...)
- On veille particulièrement à l'intelligibilité des dialogues, en même temps qu'aux émotions qu'ils produisent
- À partir des pré-mixages on produit des stems qui vont permettre le mixage final.

## La finalisation

Lorsque le mixage est terminé, il ne reste plus qu'à lui donner le format de destination. Les opérations sont les suivantes :

- Mastering et mise à niveau selon la destination (Cinéma, télévision, Web)
- Encodage (compression mp4 pour le web, Dolby 5.1, 7.1, Atmos, DTS-X ou PCM pour le cinéma ; Digital Cinema Package (DCP))

En postproduction, le mastering est différent de celui de la musique enregistrée, où des traitements complexes d'optimisation, et notamment de compression globale, sont effectués pour donner le son final. Le son pour les images nécessite au contraire un respect des nuances entre les différents éléments, même si certains traitements séparés ont été effectués sur des sections séparées.

## La post-production en résumé

- Sélection des prises
- Nettoyage (suppression des éléments indésirables)
- Editing (raccordement d'éléments, séparation de pages...)
- Intégration de la matière sonore dans le projet (généralement sources multiples)
- Mixage

- Mastering et mise à niveau selon la destination (Cinéma, télévision, Web)
- Encodage (compression mp4 pour le web, Dolby 5.1, 7.1, Atmos, DTS-X ou PCM pour le cinéma ; Digital Cinema Package (DCP))

# Le son des jeux vidéo

Pourquoi avoir attendu la fin de ce document pour aborder le son des jeux vidéo ? Certainement pas pour le manque d'importance pour ce domaine. Dans le domaine des loisirs le jeu vidéo est arrivé plus tard que le cinéma et la télévision (l'un des tous premiers jeux a été "Computer Space", produit en 1971, un jeu d'arcade). Mais aujourd'hui, le jeu vidéo les a largement dépassés en termes d'investissements et de revenus, aussi bien qu'en nombre d'emplois. Que ce soit en ligne ou sous forme matérielle, en 2022 on a constaté que 200 milliards d'Euros de ventes de jeux vidéo ont été réalisées dans le monde pour 3,2 milliards de joueurs d'un âge moyen de 30 ans. Dans la même période, 120 000 personnes ont travaillé dans l'industrie du jeu vidéo.

En fait, la raison de l'emplacement de ce sujet dans le document vient du fait que malgré des méthodes de travail très différentes, le son des jeux vidéo possède de nombreuses similitudes avec les techniques décrites précédemment dans ce document : enregistrements divers, traitements, effets sonores, spatialisation, musique, mixage.

Pourtant, même avec ces points communs, le travail du son dans ce domaine du jeu reste une spécialisation et nécessite une formation spécifique qui peut cependant s'appuyer sur les bases des connaissances générales des travaux relatifs à l'audio. Ce document ne pourra qu'en aborder certaines généralités.

Aujourd'hui, nous sommes arrivés à un point où la conception d'un jeu commercial suffisamment attractif est souvent devenue impossible pour une personne seule, sauf pour des jeux très basiques. Trop de connaissances sont nécessaires en même temps. Les grosses entreprises de jeux vidéo ont constitué des équipes de spécialistes, dont les ingénieurs du son font partie.

Le son des jeux a atteint un niveau de qualité très élevé, semblable aux meilleurs standards audio actuels. Ce qu'on lui demande, c'est avant tout d'être au service du jeu et qu'il enrichisse le ressenti du joueur.

## Les bases de la conception sonore d'un jeu vidéo

### L'interactivité

Une action doit correspondre à un son qui la valide

### La diversité

Les sons répétitifs identiques sont ennuyeux. On doit pouvoir entendre des variantes. De nombreux échantillons sonores sont nécessaires, avec des variantes pour chacun. Des centaines de milliers de sons peuvent être inclus dans un grand jeu d'aventure.

### La gestion du mouvement

Le son doit être entendu en fonction des déplacements dans les différents espaces du jeu.

### La gestion de l'espace

Que ce soit sur haut-parleurs ou au casque, on cherche à donner au joueur la sensation qu'il se trouve immergé dans l'univers virtuel du jeu. Toutes les techniques d'immersion sonore sont mises à profit, en particulier au casque, avec les techniques binaurales (voir plus haut dans ce document).

## La musique

Son rôle sur l'esprit est similaire à son utilisation dans un film. Sa conception est différente, puisqu'elle doit prendre en compte l'interactivité des actions. Elle doit être suffisamment variée pour éviter l'ennui. Son tempo peut évoluer en fonction du suspense. On peut créer des musiques affectées à chaque personnage et des jingles en réponses à des actions. Grâce à des stems, une musique décomposée en parties séparées (thème, rythmique, accompagnement, voix) permet de nombreuses variations par toutes les combinaisons possibles.

Pour éviter les ruptures musicales trop brusques le compositeur doit prévoir des transitions musicales ou des fondus ou « fades » pour passer d'un thème à un autre. Le middleware peut être d'une aide précieuse pour gérer ces transitions de la manière la plus souple possible.

## Les effets sonores

Tout comme dans un film, les effets sonores ont une importance capitale, d'autant plus en jeu vidéo, où l'interactivité joue un rôle essentiel.

Un principe émergent, la synthèse procédurale, permet de créer de très nombreux effets sonores en mouvement, en particulier pour l'univers du jeu vidéo. Il ne s'agit plus d'utiliser des samples séparés, mais de créer les sons et leurs mouvements par programmation. On utilise des règles, des algorithmes et des interactions. De nombreuses variantes d'un même son peuvent être générées.



Figure 129 La création procédurale de sons par assemblages de modules dans le moteur génératif de Tsugi Audio.

## Le middleware

Les moteurs de jeu principaux, Unreal Engine, Unity, CryEngine comprennent des sections complètes dédiées à la gestion des sons et à leur interactivité, par la programmation. Mais le développement d'un jeu complexe peut nécessiter de séparer les activités de programmation de celles de la partie sonore, pour soulager le travail du programmeur. C'est ainsi que sont nés les Middlewares. Citons parmi les plus connus **Wwise** d'Audiokinetic **FMOD** de Firelight Technologies, **Tazman Audio/Fabric**. Ils comportent des fonctions similaires à celles des logiciels des DAW qui sont connus par tous ceux qui ont travaillé sur le son et sont bien plus complets et pratiques à utiliser que les sections audio des moteurs de jeu universels.

Ces middlewares sont des intermédiaires entre la programmation d'un jeu et sa conception sonore. Ils sont conçus avec pour objectif l'optimisation des échanges entre programmeur et sound designer par rapport au moteur de jeu. L'avantage est de libérer le programmeur de cette tâche. Le sound designer peut affecter ses sons aux événements du jeu sans être programmeur.

Les principales plateformes middleware :

- Firelight **FMOD**
- Audiokinetic **Wwise**
- Tazman Fabric
- **AstoundSound** (GenAudio)
- **Master Audio**
- **Miles Sound System**
- **Steam Audio**
- **Elias Software**



Figure 130 La section mixage de FMOD, très ressemblante avec celles des consoles de mixage générales.

# Les caractéristiques essentielles des matériels audio

Les matériels audio sont toujours plus ou moins sujets à des imperfections et ont leurs propres limites.

## **Le gain (ou amplification) en dB**

C'est la capacité pour un équipement d'augmenter le volume électrique d'un signal.

## **La bande passante en Hz**

C'est l'intervalle des fréquences qu'un matériel peut gérer correctement. On y ajoute souvent « +/- pour *plus ou moins* », afin d'exprimer les écarts d'exagération ou de diminution de certaines fréquences. En numérique, la bande passante dépend seulement de la fréquence d'échantillonnage en Hz.

## **La distorsion en %**

C'est une déformation du son par rapport au son d'origine.

## **Le rapport signal/bruit en dB**

C'est le rapport entre le volume utile et le bruit de fond généré par le matériel. Plus le bruit de fond est important, plus ce rapport est faible. Il s'agit donc de la dynamique utile. En numérique il dépend de la résolution, en bits.

## **La sensibilité**

C'est la capacité plus ou moins importante d'un matériel soit de produire un certain volume sonore à partir d'un volume électrique donné (haut-parleur, casque, écouteur), soit au contraire de produire un certain volume électrique à partir d'une certaine pression sonore (micros). Plus la sensibilité est grande plus le résultat utile est important.

## **L'impédance en Ohms ( $\Omega$ )**

Il s'agit de la caractéristique indiquant la capacité d'un matériel, par sa conception, à s'opposer au passage du courant, ou à produire un courant dont l'intensité dépend de cette impédance. Plus l'impédance d'une source est élevée, moins le courant qu'elle produit est intense.

On doit toujours connecter une sortie d'impédance faible à une entrée d'impédance forte. Sur une même sortie audio, avec un casque d'impédance 8 Ohms on produit plus de volume acoustique qu'avec un casque 16 Ohms ou 600 Ohms. Un micro à condensateur est à haute impédance et nécessite une adaptation interne pour son utilisation sur les matériels courants..

## La connectique audio

L'interconnexion audio entre les éléments physiques, si elle n'est pas hertzienne (Wi-Fi, Bluetooth) nécessite des câbles munis de connecteurs qui se raccordent aux matériels. Deux domaines sont concernés : l'audio analogique et l'audio numérique.

### Connectique de l'audio analogique



Figure 131 La prise XLR, ou prise Cannon



Figure 132 Le jack mono 6,35 mm



Figure 133 Le jack stereo 6,35 mm (ou "TRS")



Figure 134 La prise RCA ou « Cinch »



Figure 135 Les connecteurs 5 broches utilisés pour la connexion des micros stéréo

Deux types de liaisons co-existent : la liaison asymétrique et la liaison symétrique. La première utilise un câble unique entouré d'une gaine métallique, qu'on appelle le blindage, qui évite la captation des inductions diverses par les rayonnements parasites, surtout ceux du réseau électrique. La seconde nécessite deux câbles, entourés d'un blindage. Il s'agit d'une technique permettant d'annuler les captations parasites par opposition de phase entre elles (voir précédemment le principe).

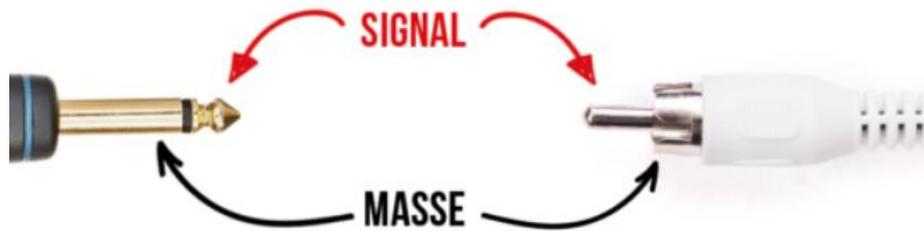
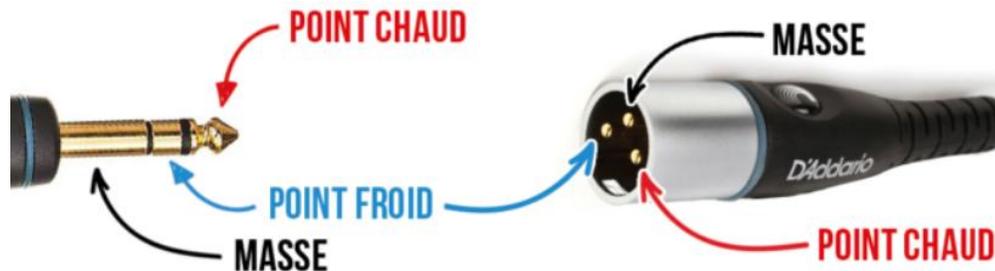


Figure 136 Les connexions d'une liaison asymétrique : le signal, à l'extrémité des connecteurs, et la masse, généralement raccordée au blindage du câble.

- La prise XLR est utilisée en audio professionnel, toujours pour les liaisons des micros. Elle est conçue pour les liaisons symétriques. Elle est aussi utilisée pour les raccordements aux moniteurs de studio.
- Le jack mono 6,35 mm (il équivaut à ¼ de pouce), transporte un seul signal. Il est surtout utilisé dans les domaines de la musique. Certains moniteurs offrent ce type de connexion.
- Le jack stéréo 6,35 mm transporte 2 signaux, généralement gauche/droite. C'est celui du branchement des casques stéréo. Pour cette utilisation, on trouve aussi les « mini-jacks » de 3,5 mm qui se raccordent aux prises 6,35 mm par un adaptateur. Par ailleurs, le jack stéréo est aussi très répandu avec les interfaces audio pour ordinateur, où il remplace les connexions XLR sur les

prises « Combo » qui combinent connectique XLR et Jack. Un jack mono peut aussi être utilisé sur ce type de prise Combo, par exemple pour une connexion d'une entrée ligne sur une interface audio.

- La connexion du type « TRS » (Tip Ring Sleeve) est le nom anglais du jack stéréo
- La prise RCA. Elle est monophonique. On l'utilise surtout en audio « grand public » (Home Cinéma, Hi-Fi)
- La prise XLR transporte un signal de niveau « ligne » au standard +4 dB, alors que la prise RCA est au standard -10dB. On en déduit qu'on a une différence de 14 dB entre un signal sur prise XLR et un autre sur prise RCA. On doit prendre en compte cette importante différence de volume en raccordant deux liaisons à standards différents.



*Figure 137 Les connexions d'une liaison symétrique : le signal circule sur une paire de câbles, connectés sur le « point chaud » (Tip) et le « point froid » (Ring). La masse (Sleeve) est aussi raccordée au blindage du câble.*

## Connectique de l'audio numérique

Le transfert des données numériques a longtemps été effectué en temps réel par la lecture de bandes ou de cassettes, grâce à des câbles spécifiques. Depuis la généralisation du stockage sur les disques durs ou statiques (USB, cartes mémoire), les transferts par câbles se sont raréfiés.

De plus, les traitements audionumériques hardware nécessitant des câbles d'interconnexion appropriés ont largement été remplacés par leurs équivalents logiciels. Les câbles entre modules de traitements plug-ins sont virtuels.

Certains modèles de caméras professionnelles disposent d'une entrée numérique sur prise XLR.

Il reste cependant encore des raccordements numériques de trois sortes :

- la prise XLR (identique à celle des connexions analogiques), qui transporte deux canaux (gauche/droite) au standard AES/EBU sur un câble blindé spécial
- la prise RCA (identique à celle des connexions analogiques), qui transporte deux canaux (gauche/droite) au standard S/PDIF sur un câble blindé spécial
- la prise TOSLINK de raccordement par fibre optique dont la connectique est spécifique, qui peut soit transmettre 2 canaux au standard S/PDIF, soit 8 canaux au standard ADAT

Précisons que le standard professionnel AES/EBU (prise XLR) est similaire au standard S/PDIF dans la mesure où chacun transporte 2 canaux. Le standard S/PDIF, très répandu en Hi-Fi pour les interconnexions numériques, utilise soit une liaison par câble (prise RCA) soit par fibre optique (prise TOSLINK).



*Figure 138 La connexion TOLINK filaire*



*Figure 139 Le câble Toslink optique*



*Figure 140 Le câble S/PDIF qui transporte l'audio par une liaison filaire blindée*



*Figure 141 La connexion AES/EBU*

## Les termes de base de l'audio. Récapitulatif.

On trouvera tous ces termes dans les différentes parties de ce document

- **Fréquence : Hz, kHz**
- **Volume : dB ; dB FS (peak-mètres), dBu (vu-mètres)**
- **Bande passante : Hz, kHz (à - 3 dB)**
- **Niveaux : micro, ligne (- 10 dB, + 4 dB)**
- **Symétrique, asymétrique**
- **Sensibilité (micros, casques, haut-parleurs)**
- **Directivité**
- **Dynamique : dB**
- **Gain : dB**
- **Impédance (Ohms, kilohms, mégohms)**

- **Phase**

## **L'audio numérique**

- **Principe de base : l'échantillonnage : Bytes, bits, bps**
- **Fréquences d'échantillonnage : 44,1 kHz, 48 kHz, 96 kHz, 192 kHz**
- **Définition : 16 bits, 24 bits, 32 bits : dynamique disponible**
- **Connectique : AES/EBU : XLR, S/PDIF : RCA/Optique**
- **Conversion de fréquences**
- **En vidéo : préférence uniforme pour le 48 kHz : éviter les erreurs 44,1 kHz/48 kHz**
- **Supports d'enregistrement : bandes numériques, disques durs, cartes mémoires flash**

© Michel Geiss 2023