

Synthèse
documentaire

Le format mp3 et les livres audio



Sven Dierick
Institut Lallemand
Synthèse documentaire

Introduction

L'audio livre, cet objet changeant le premier sens, ou le sens premier diront ces détracteurs, du livre. La vue pour l'ouïe. L'audio livre cassette, cet objet contenant pour beaucoup leurs premières comptines. L'audio livre, un livre mais autrement. L'audio livre des supports, mais divers. L'audio livre héritiers de la longue tradition orale des récits. Jusqu'à l'alphabétisation de masse, en commençant par la diffusion de l'imprimerie au 15^{ième} siècle pour se prolonger dans la loi obligeant la scolarité en Belgique depuis 1983, les récits et les histoires se transmettaient de bouches à oreilles. La lecture dans l'intimité a alors eu l'histoire que l'on connaît et que l'on ne va pas développer dans ce dossier.

Le livre que l'on va parcourir ici vient de l'autre côté de la branche littéraire, celle de l'orale. Afin de répondre aux besoins d'une frange de la population en incapacité de lire, de par une mauvaise vue, l'incapacité de tourner des pages ou simplement par manque de temps, le livre audio s'est développé. D'abord à la radio, pensons à la guerre des mondes d'Orson Wells en 1938, ensuite sur cassettes audio, puis sur les cd à partir des années 1980 pour se retrouver sur nos iPods, iPhones et autres lecteurs digitaux d'aujourd'hui. Le livre audio n'est donc pas à placer en face du texte imprimé mais bien à côté.

Qu'est-ce qu'un livre audio ?

Un livre audio ou audio-livre est le support sur lequel une lecture à voix haute d'un texte imprimé a été enregistrée. Par extension ce terme désigne également le sujet qui est présent sur le support, « LE livre audio », le texte en lui-même. Le support peut prendre diverses formes. Physique tout d'abord avec les cassettes audio et les CDs, puis sous la forme d'un fichier numérique « dématérialisé » tel le format mp3 ou le DAISY, qui peuvent être exploités par un lecteur dédié, comme un smartphone, un ordinateur ou un lecteur mp3 portable. Le compact-disc et les formats dématérialisés sont les seuls encore produits de nos jours. Ces deux types de supports sont comparables en ce qu'ils supportent l'audio de manière digitale en opposition aux cassettes qui ne supportent que l'audio analogique.

Faut-il la même qualité sonore pour un cd musical que pour un cd de livre audio ? Quelle qualité sonore pour le livre audio ? Quelles sont les avantages et inconvénients du son compressé. Le mp3 est-il la bonne réponse ? C'est à ces questions que tentera de répondre cette synthèse documentaire.

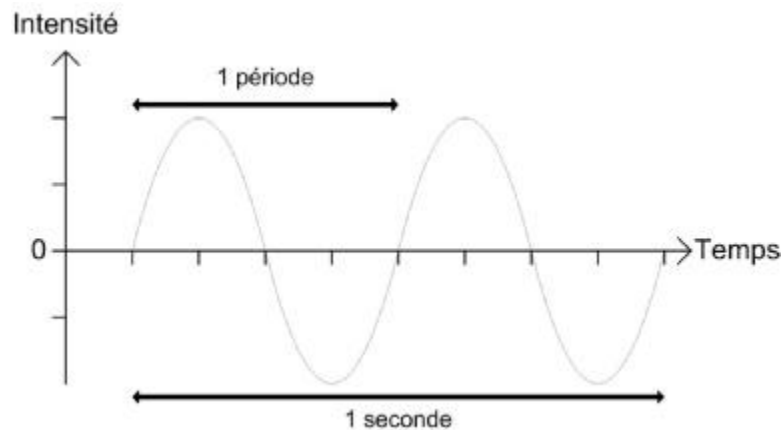
Table des matières

Introduction	1
L'audio.....	3
Qu'est-ce qu'un son ?.....	3
Qu'est-ce que l'audio analogique ?	3
Qu'est-ce que l'audio digital ?	4
La compression mp3	7
Bibliographie	8
Ressources en ligne	8
Ressources imprimés.....	8
Ressources audiovisuelles	8

L'audio

Qu'est-ce qu'un son ?

Un son est provoqué par une vibration dans l'air. Cette vibration provient d'un déplacement d'air. L'air ainsi déplacé va créer successivement des pressions (crêtes) et des dépressions (creux), comme nous pouvons l'observer lorsque l'on laisse tomber un caillou dans l'eau. Selon la vitesse à laquelle les crêtes et les creux se succèdent le son va avoir un ton différent. Un cycle, ou période, est composé d'une succession d'une crête et d'un creux. La fréquence est le nombre de cycles en une seconde. Elle est exprimée en « Hertz » ou « Hz ». 1 Hz équivaut à un cycle en une seconde. Plus la fréquence est élevée plus le son va sonner aigue et plus la fréquence est basse plus le son va sonner grave. Un jeune humain dont l'oreille n'a pas encore été endommagée peut potentiellement entendre les fréquences comprises entre 20 et 20 000 Hz (ou 20 kHz pour KiloHertz). Au-delà nous allons parler d'ultrasons, en deçà d'infrasons.



Sur le graphique ci-dessus nous pouvons voir qu'une période est composée d'une crête et d'un creux et qu'il s'agit ici d'un son de 2 Hz car il y a 2 périodes en 1 seconde. L'intensité d'un son, autrement dit le volume, est exprimé en « décibels » ou « dB » et est indiqué par l'axe y.

Qu'est-ce que l'audio analogique ?



Lors des premiers enregistrements sonores une aiguille était mise en action par le déplacement d'air provoqué par un instrument ou une voix. Cette aiguille vibrait à la même vitesse que l'onde reçue et gravait des microsillons dans des cylindres en cire. Ces sillons représentaient donc à leurs tour la fréquence de vibration de l'aiguille et donc du son original. Si le son était fort, l'aiguille se déplaçait avec une plus grande amplitude mais toujours avec la même fréquence.

Dans l'autre sens pour restituer le son depuis ces cylindres, une aiguille était posée au début du sillon. En suivant ce sillon elle vibrait à une certaine fréquence mais sans avoir une assez grande amplitude pour permettre à toute une assemblée de l'entendre. Pour cela une sorte d'entonnoir nommé pavillon servait par réverbération à amplifier le son. Il s'agit là du fonctionnement du phonographe, représenté en photo ci-dessus. Ses successeurs, jusqu'au vinyle, ont améliorées les méthodes d'enregistrement et de retransmissions, mais n'ont pas changées fondamentalement le fonctionnement de l'audio. Pour cela il faudra attendre le compact-disc et l'audio dit « digital ».

Tout son produit l'est toujours sous forme analogique. C'est-à-dire qu'il résulte d'une vibration d'air pouvant être représentée sous la forme d'un graphique, comme nous l'avons vue page 3, avec une intensité, un période et une fréquence. De nos jours le son est capté par un microphone composé d'une membrane qui vibre à la fréquence du son et qui convertit cette vibration en signal électrique par un système complexe d'électro-aimants. Afin de diffuser le son, un signal électrique est transmis à un haut-parleur qui est lui aussi composé d'une membrane qui vibre et transmet cette vibration à l'air, la rendant audible. Le défaut majeur de l'audio analogique est la reproductibilité. A titre d'exemple essayer de copier à main levée le graphique précédant. Le moindre écart va être audible et à chaque copie les écarts vont s'accumuler jusqu'à obtenir un son complètement difformé. En ajoutant à cela la fragilité des supports on peut considérer que l'audio analogique n'est pas pérenne.

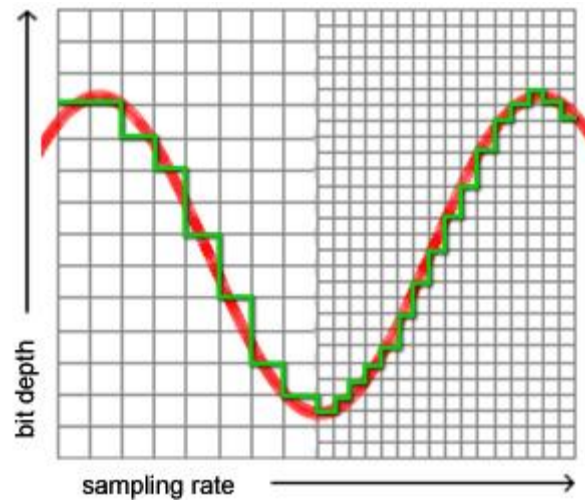
Qu'est-ce que l'audio digital ?

L'audio digital ou numérique est la transposition d'un son en langage binaire. Il est présent sous formes de fichiers numérique comme le mp3, le wmv ou le FLAC.

L'audio digital va être composé d'une suite de numéros, des 0 ou des 1, ce qui facilite et rend plus sûr la copie. Afin d'obtenir cette série numérique il va falloir convertir un signal analogique en signal digital à l'aide d'un « Analog-to-digital converter » ou « ADC ». Après que le son soit traité par le microphone l'ADC va s'occuper de découper en petites séquences, d'où le terme de séquençage ou d'échantillonnage, le signal reçu. Chaque séquence se verra attribuer une série de chiffres. Plus le taux d'échantillonnage sera haut, et donc plus le temps entre chaque échantillon est court, plus l'enregistrement obtenu sera de qualité et précis. Il est également possible d'attribuer un plus grand nombre de bits à chaque échantillon, c'est-à-dire une plus longue série de chiffres, afin d'en améliorer la qualité.

Lorsqu'on parle d'un certain nombre de bits on parle d'une série chiffrée composée de 1 et de 0. Quand on parle d'un encodage à 16 bits chaque valeur est composée de 16 0 ou 1. Cela permet donc 64 000 possibilités. Plus on augmente le nombre de bits, plus le nombre de possibilités est élevé et donc plus l'encodage peut être précis et de qualité, mais plus le fichier final sera lourd.

Chaque valeur ainsi obtenue sera enregistrée sous la forme d'un fichier. Pour la restitution du son, un « Digital-to-analog converter » ou « DAC » s'occupera de retranscrire ces valeurs en signal électrique.



Nous avons donc vu que plus le taux d'échantillonnage est haut et plus le nombre de bits est élevé plus un enregistrement augmente en qualité et en fidélité. Ceci peut être aperçu dans le graphique ci-dessus, notamment dans la partie de droite.

Mais jusqu'où faut-il échantillonner pour obtenir un son de bonne qualité ? Le théorème de Nyquist dit qu'il faut échantillonner à une vitesse deux fois supérieure à la fréquence la plus haute que l'on veut reproduire. L'oreille humaine entendant jusqu'à 20 000 Hz, il faut donc tabler sur un taux d'échantillonnage de 40 000 Hz, c'est à dire que l'on va vérifier 40 000 fois par seconde quelle valeur notre signal a. Ou autrement dit dans le cas du son le plus aigu que l'humain peut entendre la valeur sera notée 2 fois par période. Pour satisfaire aux normes d'un enregistrement sur cd, le son doit être échantillonné à 44,1 kHz. Afin de pouvoir retranscrire correctement les différences d'intensité d'un signal, il faut travailler sur 16 bits. Donc une suite de 16 numéros 0 ou 1 par valeur notée. Ce qui donne 64 000 possibilités par échantillon. Il s'agit là du traitement du son dont le but est d'être gravé sur un cd. Pour le cinéma l'on travaille sur 48 kHz et 24 bits.

Le défaut de combiner un grand nombre de bits et un haut taux d'échantillonnage est que les fichiers en résultants sont très lourds de par la quantité de données qu'ils doivent supporter. Le cd, premier support de stockage digital populaire, peut contenir à peu près 700 mégaoctets, 1 octet étant composé de 8 bits, il est donc possible de stocker un peu plus d'une heure de son stéréo échantillonné à 44,1 kHz composé 16 bits sur un cd.

En cas de lecture attentive d'un livre la vitesse de lecture atteint en moyenne 30 à 40 pages par heure. Cela veut dire que pour « Le nom de la rose » d'Umberto Eco¹ il faut compter au mieux 16 CDs. Donc soit le support n'a pas assez de capacité de stockage soit les fichiers audio sont trop lourds. C'est en essayant de réduire le poids des fichiers que nous en arrivons à nos bien connus fichiers et lecteurs mp3.

¹ Le nom de la rose [Texte imprimé] / Umberto Eco ; traduit de l'italien par Jean-Noël Schifano . – Grasset, 1982 . – 633p. . - EAN : 9782253033134

La compression mp3

Lorsqu'on parle de compression, on parle de compression par rapport à une référence. Pour le son la référence est le format CDA, donc un taux d'échantillonnage de 44.1 kHz et un encodage à 16 bits. Deux types de compression sont possible, la destructive, là une partie des données sont irrémédiablement irrécupérable, et la non-destructive, ici aucune donnée est théoriquement perdue. La compression mp3 utilise la méthode destructive.

Le brevet pour la technique de compression du mp3 est détenu en majorité par l'institut de recherche Allemand Fraunhofer depuis 1996. La compression va se baser sur toutes les connaissances de l'ouïe humaine. Il va partir du principe qu'il ne sert à rien d'enregistrer ce qu'il ne sait pas être entendu par une oreille humaine lambda. Ainsi il va supprimer les fréquences en dessous de 20 Hz et au-dessus de 20 kHz, ce qui peut entraîner un appauvrissement du son. Il va aussi à l'aide de logarithmes simplifier le codage au niveau de redondances des valeurs par exemple. Diminuer les bits ou le taux d'échantillonnage est également possible. Du fait de nombreuses possibilités de compression le mp3 peut exister en de nombreuses tailles. Pour les distinguer on va comparer leurs poids en kbits par seconde enregistré. Plus le nombre de kbits par seconde est lourd, plus le fichier sera lourd au final. Pour l'enregistrement musical un poids de 192 kbits/s pour la compression mp3 est recommandé au minimum.

Comme nous l'avons vu, la compression mp3 va supprimer les parties du spectre sonore inaudible, mais dans le cadre d'un audio livre où la seule plage de fréquence utilisée est celle de la voix, une suppression plus grande encore permettrait de gagner en espace de stockage sans se faire au détriment de la qualité sonore. Dans le cas d'un mp3 à 192 kbits/s il est possible de stocker 8 heures de son stéréo sur un cd classique de 700 Mo. En diminuant la qualité de compression du mp3 à 96 kbits/s il est possible d'encore faire doubler la capacité de stockage.

Pour reprendre l'exemple du livre cité précédemment, Le nom de la rose, en prenant la qualité minimale conseillée (192 kbits/s) il ne faut plus que 2 CDs au lieu de 16. En augmentant encore le taux de compression jusqu'à obtenir 96 kbits/s le livre entier ne tient plus que sur un cd.

Bibliographie

Ressources en ligne

Loi concernant l'obligation scolaire [loi] . – 1983 . – Disponible sur : [http://www.galilex.cfwb.be/document/pdf/09547_004.pdf] . – Consulté le 02/05/2017.

Valérie Lévy-Soussan, directrice d'Audiolib : « Un livre, ça s'écoute aussi » [article en ligne] / Julie Malaure . – Le point , 2009 . – Disponible sur : [<http://www.lepoint.fr/actualites-litterature/2009-06-29/valerie-levy-soussan-directrice-d-audiolib-un-livre-ca-s-ecoute/1038/0/356795>] . – Consulté le 02/05/2017.

Les livres audio aux frontières de la lecture [article en ligne] / Jonathan Journiac . – Evéne , 2005 . – Disponible sur [<http://eve-ne.lefigaro.fr/livres/actualite/livres-audio-livraphone-livrior-audible-gallimard-lecture-246.php>] . – Consulté le 02/05/2017.

Ressources imprimés

Guide de la prise de son pour l'image [texte imprimé] : Reportage, documentaire, fiction en radio et télévision / Vincent Magnier . – 2^e édition . - Paris : Dunod , 2011 . – ISBN : 978-2-10-053814-0

Ressources audiovisuelles