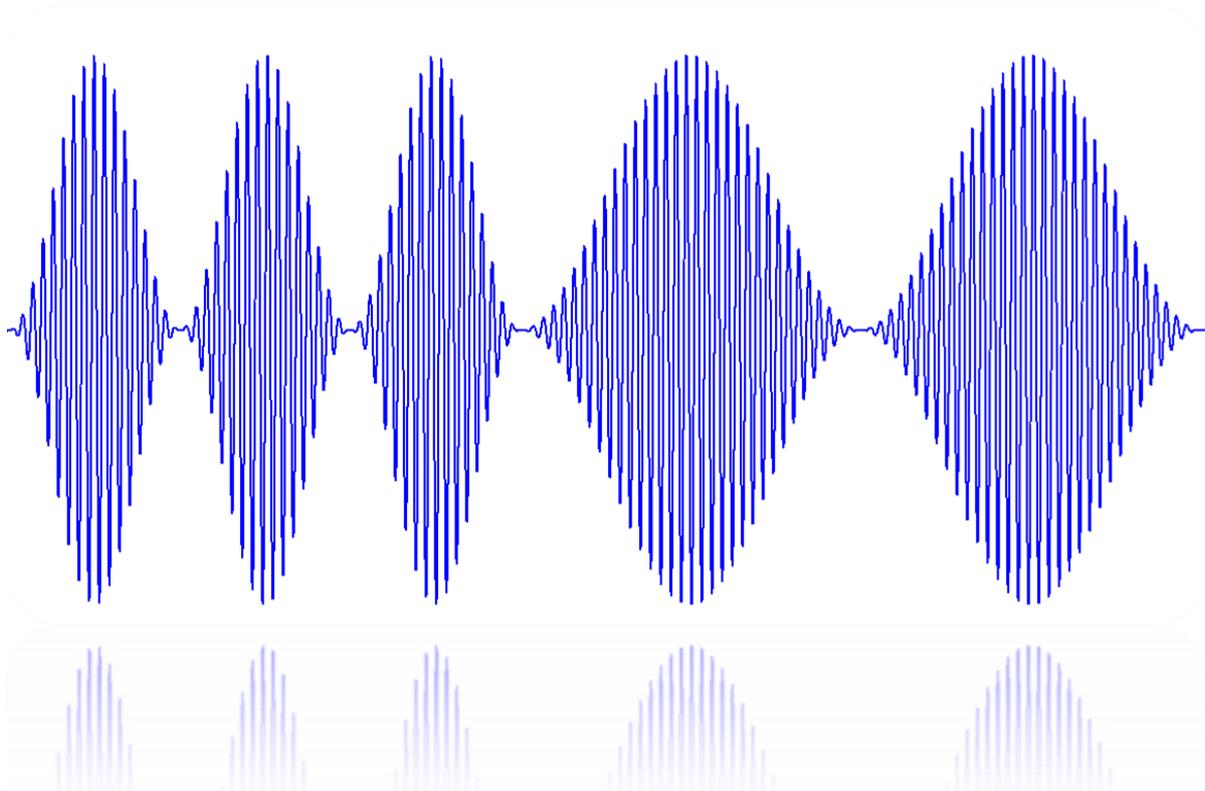


Rapport de stage

2^{ème} année de Master en Informatique Spécialité Image, Son, Vidéo

La sensibilité auditive à un changement
de fréquence de modulation



Stagiaire de l'INCI, CNRS :

Yann Bayle

Directeur de recherche à l'INCI, CNRS :

Laurent Demany

Responsable à l'Université de Bordeaux :

Paul Dorbec

Tuteur scolaire de l'ENSEIRB-MATMECA :

Myriam Desainte-Catherine

Période du stage :

2 Février au 3 Juillet 2015

Sommaire

1. Remerciements	3
2. Introduction.....	4
3. Présentation du laboratoire de recherche.....	5
4. Contexte	6
4.1. Le son	6
4.2. Le fonctionnement de l'audition humaine	7
4.3. Les Détecteurs de Changement de Fréquence (DCF)	8
4.4. La perception du rythme.....	8
4.5. L'étude	9
5. Première expérience : test de l'existence des DCR <i>via</i> la comparaison aux DCF	10
5.1. Matériel et Méthode.....	10
5.2. Programmation	13
5.3. Résultats.....	19
5.4. Discussion.....	20
6. Deuxième expérience : étude de la perception d'un changement de rythme.....	22
6.1. Matériel et méthode.....	22
6.2. Programmation	24
6.3. Résultats.....	24
6.4. Discussion.....	25
7. Conclusion	28
8. Perspectives.....	29
9. Sources	30
9.1. Bibliographie	30
9.2. Webographie.....	32

1. Remerciements

Mes premiers remerciements vont à mon directeur de recherche, le **Dr Laurent Demany**, qui m'a permis d'entrevoir le milieu rigoureux mais captivant de la Recherche. Un grand merci pour sa disponibilité, son soutien, son honnêteté et ses précieux conseils.

Je remercie ma tutrice de stage de l'ENSEIRB-MATMECA, **Madame Myriam Desainte-Catherine**, pour sa disponibilité ainsi que pour son aide dans ma recherche de ce stage.

Je remercie également les membres du laboratoire qui m'ont accompagné durant ce stage : **Marie-Caroline, Catherine et Max**, pour leur bonne humeur.

Je remercie les douze participants à mes deux expériences pour leur implication, leur compréhension ainsi que les nombreuses heures passées dans les cabines insonorisées.

A **Fidji**, pour sa présence, sa patience, son soutien sans faille et sa contribution à la relecture de l'ensemble de mon travail.

2. Introduction

« Nous pouvons définir le son comme un coup donné par l'air à travers les oreilles au cerveau et au sang et arrivant jusqu'à l'âme ». Si l'intérêt de Platon pour le son au IV^{ème} siècle av. J.-C. démontre une connaissance rudimentaire des mécanismes auditifs durant l'Antiquité, de nombreux phénomènes sonores impliquant le conscient et le subconscient demeurent encore aujourd'hui à découvrir et à comprendre.

La perception auditive permet aux Vertébrés d'appréhender leur environnement et surtout de communiquer entre eux. Le système auditif traite en effet automatiquement les informations provenant de multiples sources du milieu extérieur, qui sont indissociées lors de leur arrivée aux tympans. La psychoacoustique a permis depuis les années 1930 de découvrir de nombreux processus concernant la perception auditive, notamment en ce qui concerne le masquage fréquentiel qui est actuellement le principe fondamental de la compression audio. Un large champ des capacités de l'Homme à identifier les caractéristiques sonores d'une source parmi plusieurs autres demeure néanmoins très peu documenté.

Certaines capacités à identifier une source sonore ainsi qu'à la caractériser sont étudiées dans ce rapport. Il s'agit plus particulièrement de déterminer dans quelle mesure l'oreille humaine est sensible à un changement de fréquence de modulation. Il s'agira donc de déterminer si l'on peut prouver l'existence de Détecteurs de Changement de Rythme (DCR) dans l'oreille humaine.

La première partie est consacrée à la contextualisation de cette étude. La partie suivante décrit les deux expériences successives menées et les analyse afin de répondre à la problématique établie.

3. Présentation du laboratoire de recherche

L'Equipe Perception Auditive (EPA) est incluse au sein de l'Institut de Neurosciences Cognitives et Intégratives d'Aquitaine (INCIA), qui est l'Unité Mixte de Recherche (UMR) numéro 5287 du Centre National de la Recherche Scientifique (CNRS).

L'EPA est située sur le campus de Bordeaux-Segalen et dispose de deux cabines insonorisées par deux murs isolants ainsi que de l'équipement audio nécessaire au passage de tests psychoacoustiques par des participants humains.

L'EPA est composée de sept personnes, organisées comme présenté ci-dessous :

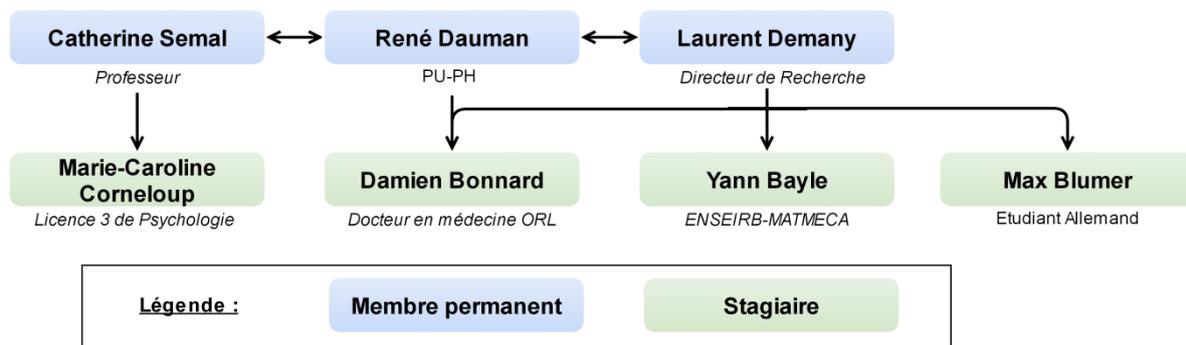


Figure 1 - Organigramme de l'EPA

Cette équipe se concentre sur l'étude du fonctionnement du système auditif dans le cadre de la psychophysique sensorielle, notamment en ce qui concerne les perception et mémoire auditives chez l'Homme. Les derniers travaux de cette équipe ont porté sur la présence de détecteurs automatiques de changement de fréquence chez l'Homme ainsi que sur les propriétés de la mémoire auditive.

4. Contexte

4.1. Le son

Un son est une vibration mécanique qui se propage de manière longitudinale. Il peut être produit par un instrument de musique, une voix humaine ou tout autre élément présent dans l'environnement. La fréquence et le niveau de pression constituent les caractéristiques physiques d'un son. L'intensité perçue, la hauteur, le timbre et le rythme en constituent les principales caractéristiques perceptibles.

La fréquence exprimée en Hertz (Hz) correspond au nombre de répétitions par seconde de la vibration mécanique provoquée par le son. La hauteur correspond au caractère plus ou moins aigu d'une fréquence. L'oreille humaine est capable de distinguer des sons entre 20 Hz et 20 kHz, ce qui équivaut respectivement à des sons grave et aigu. La note la plus grave de la guitare basse est par exemple de 41 Hz et la note la plus aiguë du piano est de 4186 Hz.

Le niveau de pression est exprimé en décibel (dB) SPL, qui signifie *Sound Pressure Level* ou niveau de pression acoustique. Celui-ci exprime le rapport de la pression acoustique par rapport à la valeur de référence de 0 dB SPL. Ce dernier correspond au *minimum* de pression acoustique perçue par l'oreille qui est de 20 μ Pa. L'intensité perçue se mesure quant à elle en sone et dépend de la pression acoustique ainsi que de la fréquence. Pour exemples, la pression acoustique reçue lorsqu'une personne chuchote à 1 mètre est de 20 dB tandis que celle-ci est environ de 110 dB dans une discothèque, ce qui correspond respectivement à 0,2 et 125 sones.

Le timbre est à différencier de la hauteur dans la mesure où il représente la répartition des fréquences du spectre sonore ainsi que leur évolution temporelle. Le timbre permet d'identifier la nature d'une source sonore. Une même note de musique jouée par un piano ou une trompette n'aura ainsi pas le même timbre.

Le rythme représente enfin la perception de la périodicité d'un son. Le son produit par le mouvement des pales d'un hélicoptère constitue un exemple de rythme aisément perceptible.

Une légère variation de l'un de ces paramètres est perçue par l'oreille humaine et peut modifier de façon importante la perception du son. L'une des conséquences de cette dernière observation est qu'il est possible d'identifier et de suivre plusieurs sources sonores dans un même environnement.

Les sons présents dans la nature ne sont que rarement composés d'une seule fréquence. Il est toutefois nécessaire, afin d'étudier la perception auditive, d'utiliser des sons composés de peu de fréquences et de conserver une homogénéité concernant les autres paramètres des différents sons.

4.2. Le fonctionnement de l'audition humaine

Lorsqu'un son est produit et capté par l'oreille, plusieurs mécanismes sont mis en jeu au sein de celle-ci et permettent par la suite une représentation mentale de ce son.

La psychoacoustique étudie la perception des *stimuli* sonores par le cerveau. De nombreux processus demeurent encore aujourd'hui à analyser concernant les mécanismes de perception sous-jacents.

Le système auditif humain fonctionne par banc de filtre (Munkong, 2008), c'est-à-dire que chaque fibre du nerf auditif répond à la présentation d'une fréquence. Les fréquences aiguës sont analysées au niveau de la base de la cochlée et les fréquences graves au niveau de l'apex, comme on peut le constater sur la figure 2. Cette relation entre un point de l'espace et une fréquence est qualifiée de tonotopique.

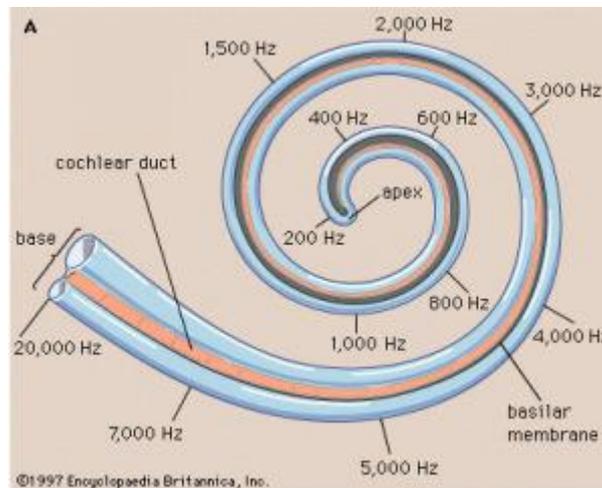


Figure 2 - Schéma représentant la cochlée (Source : Encyclopedia Britannica)

Les fibres du nerf auditif n'ont pas une précision parfaite et répondent également à des fréquences proches de leur fréquence caractéristique mais de manière moins importante. Leur sélectivité fréquentielle en Hertz augmente lorsque la fréquence caractéristique diminue, c'est-à-dire que deux fréquences graves proches sont plus précisément différenciables que deux fréquences aiguës proches. L'analyse des rythmes est également tonotopique (Giraud, 2000). Le colliculus inférieur est par exemple plus sensible à des rythmes compris entre 50 et 120 Hz tandis que le cortex auditif primaire répond mieux à des rythmes compris entre 10 et 20 Hz.

Un son constitue un spectrogramme pour le système auditif, c'est-à-dire qu'il constitue une évolution d'une ou de plusieurs fréquences dans le temps. Lorsqu'un son composé de plusieurs fréquences est présenté à l'oreille, la fusion spectrale complexifie la perception individuelle des fréquences (De Cheveigné, 1997). Cet effet n'est observé que lorsque toutes les fréquences sont synchrones, c'est-à-dire qu'elles commencent et se terminent au même instant. De plus, cet effet est constaté pour des fréquences proches ou non et l'effet est plus important lorsque les fréquences possèdent des relations harmoniques.

4.3. Les Détecteurs de Changement de Fréquence (DCF)

Dans un contexte de fusion spectrale, il est difficile d'identifier chaque fréquence indépendamment et il semble donc impossible de détecter la modification d'une fréquence qui n'a pas été perçue au préalable. La possibilité de ce dernier point a néanmoins été prouvée et est optimale pour un changement de fréquence d'environ un demi-ton. (Demany et Ramos, 2005). Un demi-ton correspond à un rapport de fréquence de $2^{1/12}$ soit environ 1,059. La fonction des DCF est l'établissement de liens entre des sons successifs afin de les regrouper. Ces DCF sont activés en permanence, ne nécessitent pas de concentration de la part de l'individu et permettent à ce dernier de détecter automatiquement les modifications de son environnement sonore. L'existence de ces DCF autorise le suivi simultané de plusieurs sources sonores. Les DCF sont mis en jeu dans la représentation de la fréquence des sons provenant des deux oreilles (Carcagno et al., 2011), c'est-à-dire qu'ils ne sont pas sensibles à la localisation du son. Afin d'étudier les DCF, on peut par conséquent présenter le même *stimulus* aux deux oreilles. Aucun équivalent des DCF n'a actuellement été découvert en ce qui concerne les changements de rythme. L'existence d'un équivalent des DCF dans le domaine temporel a fait l'objet d'une étude au cours de ce stage.

4.4. La perception du rythme

La perception du rythme ne semble pas être propre à l'audition (Levitan, 2014). Certains mécanismes pourtant ne s'expriment qu'à travers l'audition, comme par exemple le regroupement de sources sonores partageant un rythme identique (Moore, 1992). La présence de plusieurs rythmes simultanés crée toutefois des interférences empêchant une perception nette de chacun de rythmes distincts (Moore, 1992). Cet effet, dénommé MDI pour *Modulation Discrimination Interference* est d'autant plus important que les rythmes sont proches les uns des autres ou bien possèdent des relations harmoniques. Dans les deux expériences menées, il a été nécessaire de prendre en compte cet effet en établissant une distance minimale entre chaque rythme et en évitant les relations harmoniques.

Il est certain que dans un contexte de présentation rythmique des sons, des tâches de discrimination sont plus aisées que lorsque le contexte ne laisse transparaître aucun rythme (Riess Jones *et al.*, 1981). Bien que l'impact direct de l'attention rythmique n'a pas été étudié, cette attention permet à l'auditeur de séparer différentes sources (Lefkothea-Vasiliki *et al.*, 2011). Lorsqu'une source a clairement été identifiée, il est possible de suivre son évolution grâce à son rythme intrinsèque. Le système auditif est un modèle de prédiction temporelle aux suites isochrones d'évènements et toute rupture rythmique est détectable (Schröger *et al.*, 2014) et permet de repérer la fin de l'émission de sons par une source. Les effets d'une accélération ou d'un ralentissement sur la sensibilité à la rupture rythmique n'ont pas encore été caractérisés, c'est donc ce que proposent les deux expériences présentées dans le présent rapport.

4.5. L'étude

Le système de perception auditive humain est capable de détecter automatiquement des changements de fréquence. L'un des intérêts de cette étude est de déterminer dans quelle mesure le système de perception auditive humain est sensible à une modification de rythme, donc s'il existe des DCR dans l'oreille humaine. Afin d'étudier cette possibilité, le rythme peut être étudié sous plusieurs formes. L'étude de la sensibilité au rythme *via* une modulation d'amplitude sinusoïdale est pertinente (Formby, 1985) et a été choisie pour les deux expériences présentées dans ce rapport. La modulation d'amplitude consiste à créer une variation périodique du volume sonore d'un signal. Les deux expériences de ce rapport mesurent la sensibilité à un changement de rythme et utilisent des paradigmes similaires à ceux décrits pour l'observation des DCF (Demany et Ramos, 2005). La tâche proposée aux participants consistera à écouter un *stimulus* sonore. Un paramètre de celui-ci sera modifié lors de chaque présentation et le sujet déterminera s'il a perçu le changement en répondant à une question fermée.

La compréhension de la sensibilité humaine aux rythmes permettrait notamment d'approfondir les connaissances concernant la perception des sentiments d'autrui *via* les intonations vocales. Les sentiments se traduisent en effet, dans le domaine de la parole, par des variations rythmiques *via* la modulation d'amplitude (Mencatini *et al.*, 2014). La modulation d'amplitude présente dans la voix se situe généralement entre 0,6 à 12 Hz (Payton *et al.*, 1999). L'unité rythmique utilisée en musique se situe généralement entre 40 et 200 Battements par Minute (BPM), ce qui correspond respectivement à 0,6 et 3 Hz. Ces gammes de rythme seront donc prises en compte mais des rythmes plus rapides seront également étudiés afin de comparer leur impact sur la perception auditive humaine des rythmes.

Dans le domaine de l'audition, l'étude la plus semblable à celle effectuée lors de ce stage concerne la perception des événements sonores réguliers qui surviennent ou disparaissent de l'environnement sonore (Riess Jones *et al.*, 1981). L'étude d'un rythme qui évoluerait n'a pourtant pas encore été étudiée. Il s'agit donc de savoir s'il est possible pour l'Homme de détecter une accélération ou une décélération rythmiques de façon automatique, c'est-à-dire sans que l'attention soit impliquée.

Les deux expériences qui suivent tentent de fournir une réponse à cette question.

Des exemples de *stimuli* sonores sont présentés sur la page web yannbayle.free.fr afin d'aider le lecteur dans la compréhension des deux expériences.

5. Première expérience : test de l'existence des DCR via la comparaison aux DCF

La détection d'un changement de fréquence porteuse, c'est-à-dire de hauteur, a démontré un aspect pré-attentionnel et il s'agit dans cette étude de déterminer s'il en est de même en ce qui concerne un changement de fréquence de modulation, c'est-à-dire de rythme. Cette étude consisterait donc à savoir s'il existe des DCR comme il existe des DCF dans l'oreille humaine.

Cette première expérience tente de fournir des réponses à ces questions en comparant les performances obtenues pour une détection de changement de fréquence porteuse avec celles obtenues pour une fréquence de modulation. Il s'agit ici d'évaluer le rôle de l'attention dans la perception de la modulation d'amplitude.

5.1. Matériel et Méthode

5.1.1. Stimuli

Cette expérience utilise des sons purs définis par :

$$S_p(t) = A_p * \sin(2 * \pi * f_p * t + \varphi_p)$$

- où :
- A_p est l'amplitude du signal,
 - f_p est la fréquence du son,
 - φ_p représente la phase à l'origine du signal,
 - t est le temps.

Le paramètre t possède une précision de $\frac{1}{F_e}$, où F_e est la fréquence d'échantillonnage fixée à 44100 Hz. Chaque son pur est ensuite modulé en amplitude par une modulante définie par :

$$S_m(t) = \sin(2 * \pi * f_m * t + \varphi_m)$$

- où :
- f_m est la fréquence du son,
 - φ_m représente la phase à l'origine du signal,
 - t est le temps.

Le signal final est donc défini par :

$$S(t) = A_p * \sin(2 * \pi * f_p * t + \varphi_p) * (1 + \sin(2 * \pi * f_m * t + \varphi_m))$$

Le signal $S(t)$ sera par la suite nommé S et un accord A sera défini comme la somme vectorielle de quatre S . Les quatre S composant A posséderont chacun des f_p et f_m différentes. Afin d'étudier l'importance de l'attention dans la détection d'un changement de fréquence porteuse ou de modulation, l'expérience met en jeu quatre conditions qui sont résumées dans le tableau ci-dessous :

Modifications	Fréquence porteuse (P)	Fréquence de modulation (M)
S puis A (S-A)	Condition 1	Condition 3
A puis S (A-S)	Condition 2	Condition 4

Tableau 1 - Présentation des quatre conditions de l'expérience 1

La condition 1 et 2 mettent en jeu un S, respectivement précédé ou suivi d'un A. Les f_p de l'accord sont choisies aléatoirement entre 300 et 3000 Hz. Les f_m correspondantes sont sélectionnées aléatoirement entre 4 et 50 Hz. Les f_m doivent être séparées au *minimum* par 800 cents soit huit demi-tons et au *maximum* par 1100 cents soit onze demi-tons. Il en est de même en ce qui concerne les f_p . Lors de chaque nouvel essai, de nouvelles fréquences sont choisies aléatoirement. Dans les conditions P, la f_p de S est sélectionnée de façon équiprobable parmi l'une des quatre fréquences porteuses de l'accord et est augmentée d'un certain nombre de cents C_p . La f_m correspondante demeure inchangée. Les troisième et quatrième conditions impliquent de modifier la fréquence de modulation et de conserver une fréquence porteuse identique. Un exemple de spectrogramme obtenu pour chaque condition est présenté dans le tableau ci-dessous et représente l'évolution des fréquences du signal en Hertz en fonction du temps en secondes :

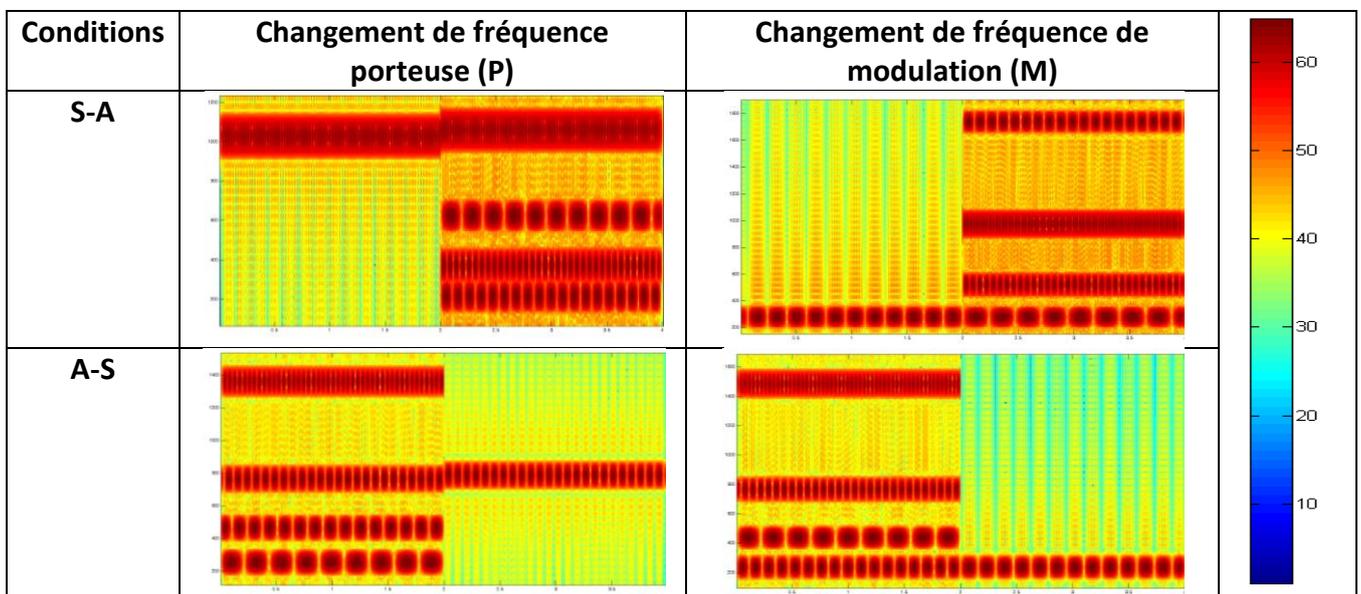


Tableau 2 - Exemples de spectrogrammes obtenus pour chaque condition de l'expérience 1

La couleur rouge représente la présence d'une fréquence et la couleur bleue représente l'absence de fréquence. On peut constater sur le graphique précédent que dans la condition S-A, la première moitié du *stimulus* est constituée d'un seul S et que la deuxième moitié du stimulus est composée de quatre S. Pour la condition A-S, cet ordre est inversé. Pour la condition P, S est supérieur ou inférieur au S correspondant dans l'accord, ce qui indique un décalage fréquentiel entre les deux composantes du *stimulus*. Dans la condition M, les deux S se situent au même niveau, ils partagent donc la même fréquence porteuse mais leur rythme est différent. Lorsque le rythme augmente, les raies apparaissent en effet plus rapprochées, et lorsque le rythme ralentit, la distance entre deux raies augmente.

La durée de S est de deux secondes, celle de A l'est également et il n'existe pas d'intervalle de silence entre S et A. La phase de S_p ainsi que celle de S_m se situent à zéro lors du changement entre S et A, afin d'éviter toute discontinuité dans le signal audio. Une discontinuité dans le signal audio produit un bruit pouvant potentiellement perturber

l'identification d'un changement de fréquence. Il est donc nécessaire de pallier l'apparition de toute discontinuité dans le signal.

La tâche du sujet dans la condition P est de juger si la fréquence porteuse perçue a subi une augmentation ou une diminution, ce qui se traduit par la perception d'une modification de hauteur. Dans la condition M, le sujet doit juger si la fréquence de modulation a également subi une augmentation ou une diminution, ce qui se traduit par une modification du rythme perçu.

Afin de vérifier si chaque participant était en mesure de réaliser l'expérience, ceux-ci ont au préalable subi deux entraînements éliminatoires. Ces-derniers ont permis de vérifier la capacité des participants à détecter un changement de fréquence porteuse. Les *stimuli* présentés dans le paragraphe précédent, sans modulation d'amplitude cependant, ont été présentés aux sujets. Au cours de l'entraînement 1, A était composé uniquement de trois composantes et le changement de fréquence était de 100 cents. En ce qui concerne l'entraînement 2, A comprenait quatre composantes et la modification de fréquence était de 50 cents. L'objectif consistait à déterminer si la fréquence augmentait ou diminuait, c'est-à-dire respectivement si elle devenait plus aiguë ou plus grave. La perception des dix participants a été testée sur cent présentations du *stimulus* au cours de chaque entraînement. Le score de réussite requis afin de participer à l'expérience était de 90%. Dix participants ont atteint ce score, un onzième cependant n'a pas réussi les entraînements et n'a donc pas été retenu pour l'expérience.

L'étape suivante a ensuite consisté à réaliser une condition contrôle afin de vérifier l'aptitude des participants à détecter correctement une modification de fréquence de modulation. Cette phase n'était pas éliminatoire. Durant cette condition, le même *stimulus* a été utilisé mais A a été remplacé par un S. La tâche pour les participants consistait à déterminer si la fréquence de modulation de S augmentait ou diminuait. Les augmentations ou diminutions étaient de 500 cents. L'ampleur de ce changement sera par la suite appelé delta. Cette condition sera nommée S-S par la suite.

Il est possible de mesurer l'impact de l'attention dans la détermination d'un changement de rythme, en comparant les performances de la condition S-A avec celles de la condition A-S. Dans la condition A-S, les performances devraient être égales ou inférieures à celles obtenues pour la condition S-A. Dans cette dernière condition, l'auditeur peut en effet se concentrer sur la fréquence qui sera modifiée. Dans la condition A-S, l'auditeur ignore au préalable la composante qui sera modifiée. Afin de comparer efficacement les conditions P et M, il a donc été nécessaire d'ajuster le delta de P et de M afin que l'ensemble des participants aient les mêmes performances dans S-A. Il a alors été possible de mesurer la diminution de performances dans les conditions P et M pour A-S. Le pourcentage de réussite de cette tâche d'ajustement a été fixé à 75 %.

Une heure a été consacrée, par chaque participant, aux entraînements 1 et 2, deux heures l'ont été pour la condition S-S, quatre heures pour la recherche du delta et cinq heures pour les conditions S-A et A-S, ce qui a donc demandé un total de douze heures de participation pour chacun des sujets.

5.1.2. Procédure

La perception des participants a été testée dans une cabine insonorisée à doubles parois.

L'objectif a consisté, pour chaque participant, à répondre à une question fermée relative aux *stimuli* entendus. Les deux réponses possibles étaient écrites sur deux boutons virtuels. Dans la condition P, le premier bouton s'intitulait « Baisse » et le second « Monte », ce qui correspondait respectivement à une diminution et une augmentation de la fréquence porteuse. Dans la condition M, le premier bouton s'intitulait « Ralentit » et le second « Accélère », ce qui correspondait respectivement à une diminution et une augmentation de la fréquence de modulation. Le temps de réponse était illimité.

Au lancement du programme, le participant devait sélectionner l'une des quatre conditions définies dans la partie précédente.

Un essai est défini comme la présentation au participant d'un *stimulus* relatif à la condition choisie puis de l'acquisition de la réponse du participant. Une réponse correcte génère une coloration verte du bouton sélectionné, une réponse fautive en génère la coloration rouge, les participants étaient ainsi informés de la justesse de leurs réponses. Chaque essai est présenté 400 ms après que le participant a répondu à l'essai précédent et un bloc regroupe 50 essais. Chaque participant a réalisé huit blocs dans chacune des quatre conditions.

5.1.3. Participants

Dix participants ayant reçu une éducation musicale et ne présentant pas d'anomalies auditives ont pris part à cette expérience. La moyenne d'âge était de 22,3 ans, l'écart-type de 1,7 an et la répartition des genres était de sept femmes pour trois hommes. Les participants ont signé un accord de consentement de participation et ont été rémunérés.

5.2. Programmation

Le développement des *stimuli*, de l'interface graphique ainsi que l'analyse des résultats ont intégralement été effectués par le stagiaire à l'aide du logiciel Matlab R2012a, exceptée l'analyse de variance qui est décrite dans la partie [5.2.3. Analyse des données](#).

5.2.1. Génération des *stimuli*

Cinq fonctions, présentées sur la figure 3, sont utilisées afin de générer les *stimuli*. Le générateur de fréquences fournit un S. Le générateur d'accords renvoie un A et gère les relations *minimum* qu'il doit exister entre les différentes fréquences utilisées pour les quatre

S de A, comme décrit dans la section [5.1.1.Stimuli](#). Le générateur d'essais fournit le signal final composé du A précédemment généré et d'un S positionné avant ou après celui-ci. Les paramètres de ce dernier sont modifiés en fonction de la condition choisie par l'utilisateur. Le générateur de blocs gère la succession des 50 essais ainsi que l'enregistrement de l'intégralité des paramètres dans un fichier pour l'analyse *a posteriori*. La fonction *main* récupère enfin les informations concernant les choix de chaque utilisateur *via* l'interface graphique. Ces informations comprennent l'ordre S-A ou A-S, la modification de fréquence porteuse ou de modulation, le delta ainsi que les initiales du participant.

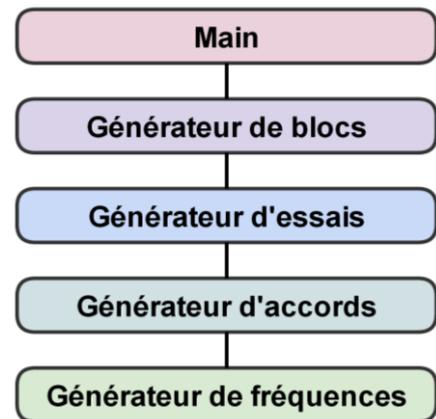


Figure 3 - Fonctions utilisées

Lors de l'élaboration du *stimulus*, il n'est pas possible de juxtaposer les *buffers* S et A puisqu'une discontinuité surviendrait lors de la transition de S vers A et de A vers S. Une telle discontinuité proviendrait des phases des sinus qui, lors du passage de l'un vers l'autre, seraient toutes différentes. Il est par conséquent nécessaire d'imposer à toutes les fréquences porteuses ainsi qu'à toutes les fréquences de modulation de se situer à la phase zéro lors de la transition entre S et A ou entre A et S. Cette contrainte implique donc qu'au début et qu'à la fin du *stimulus* les phases seront également toutes différentes de zéro. Afin de pallier toute discontinuité due à ce phénomène, il est nécessaire d'appliquer une rampe au début ainsi qu'à la fin du *stimulus* et qui permet de conduire progressivement le volume total à zéro. La durée de cette rampe a été fixée à 50ms. Cette durée est suffisamment importante afin de ne pas générer de discontinuité audio.

La présentation des sons aux participants est réalisée *via* un casque Sennheiser HD650, sélectionné pour sa courbe de réponses en fréquence. Cette courbe implique le fait que le casque possède la particularité de délivrer la même intensité sonore quel que soit le courant fourni à ses bornes. Cette particularité est uniquement vraie concernant la gamme de fréquences utilisée dans les deux expériences décrites dans ce rapport. Le niveau *maximum* de chaque S était fixé à 53 dB SPL.

Un fichier *.mat* est créé pour chaque bloc et contient l'ensemble des paramètres utilisés pour le bloc, soit :

- Les initiales du participant
- Les positions relatives de S et de A (S-A ou A-S)
- La fréquence à modifier
- La valeur du delta en cents
- L'horodatage, c'est-à-dire la date et l'heure de début et de fin du bloc
- Les réponses de l'utilisateur
- Les fréquences porteuses et de modulation utilisées pour A et S

Les réponses de l'utilisateur sont enregistrées dans un vecteur avec la convention présentée dans le tableau 3 qui suit :

Modification		Réponse de l'utilisateur	
		« Monte »	« Baisse »
Réalisée par le programme	Augmentation	1	2
	Diminution	4	3

Tableau 3 - Convention d'enregistrement des réponses de l'utilisateur

Ce codage permet un calcul rapide du nombre de réponses correctes puisque les choix corrects sont représentés par des nombres impairs. Afin d'obtenir le nombre de réponses correctes, il suffit d'effectuer une division *modulo 2* de chaque élément du vecteur et de réaliser la somme de l'ensemble des éléments du vecteur. Afin d'expliquer cette méthode de calcul, considérons le vecteur de réponses suivant : [1, 2, 3, 4]. La division modulo 2 donne le résultat suivant : [1, 0, 1, 0]. La somme des éléments du vecteur est de 2 et 4 éléments sont présents au total dans le vecteur, le pourcentage de réponses correctes est donc de 50 %.

5.2.2. Interface graphique

L'interface doit présenter une taille de police suffisamment importante pour que le texte puisse être lu par le sujet depuis un écran situé à l'extérieur de la cabine insonorisée et à travers une vitre de plusieurs centimètres d'épaisseur.

L'interface met en jeu deux fenêtres qui permettent respectivement de sélectionner les paramètres du bloc et de répondre à l'essai en cours.

La première fenêtre présentée ci-dessous permet à l'utilisateur de choisir s'il désire effectuer l'entraînement 1, l'entraînement 2 ou bien le Test Principal.



Figure 4 – Partie supérieure de la première fenêtre de l'interface

Si l'utilisateur choisit l'entraînement 1 ou l'entraînement 2, il est directement redirigé vers la seconde fenêtre. S'il choisit d'effectuer le Test Principal, il doit au préalable renseigner différents paramètres, indiqués dans la capture d'écran ci-après :

Yann 1

Informations

Veillez insérer vos initiales :

Test Principal

Fréquences

Choisissez si vous voulez modifier la : Hauteur
 Rapidité

Ordre

Choisissez le moment où vous voulez entendre l'accord : Son seul / Accord
 Accord / Son seul

Réglages

Delta de la fréquence porteuse en cents :

Nombre d'essais à réaliser :

Lancer le test

Figure 5 - Première fenêtre de l'interface

Afin que tout utilisateur puisse comprendre les paramètres de cette fenêtre, les termes « fréquence porteuse » et « fréquence de modulation » ont respectivement été modifiés en « Hauteur » et « Rapidité ». Cette vulgarisation permet à l'utilisateur d'identifier correctement et sans aucun doute la tâche à accomplir. De plus et comme indiqué dans les parties précédentes, le participant doit également choisir l'ordre dans lequel il entendra S et A. Le panneau inférieur de l'interface permet au participant d'indiquer son propre delta en fonction de la condition choisie.

Après avoir cliqué sur le bouton « Lancer le test », l'utilisateur peut écouter les sons et indiquer la variation qu'il perçoit sur la fenêtre présentée ci-après :



Figure 6 - Troisième fenêtre de l'interface

Cette fenêtre permet au participant d'indiquer ce qu'il a entendu et de constater sa progression. La ligne de texte disposée en haut de la fenêtre constitue une vérification des options sélectionnées afin de ne pas commettre d'erreurs lors de l'écoute.

5.2.3. Analyse des données

Comme décrit dans la partie [5.2.1. Génération des stimuli](#), l'intégralité des données utilisées pour chaque essai est enregistrée dans un fichier .mat, il est donc possible d'effectuer une analyse *a posteriori* détaillée de ces données. Cette partie décrit les analyses effectuées concernant cette expérience mais les graphiques représentant les résultats correspondant à chaque analyse sont présentés dans la partie suivante intitulée [5.3. Résultats](#).

La première analyse concerne le taux de réussite global, calculé comme indiqué dans la partie [5.2.1. Génération des stimuli](#). Ce taux de réussite est représenté en fonction de l'ordre de S et de A ainsi qu'en fonction du changement de P ou de M. Le programme d'analyse des données parcourt l'ensemble des fichiers .mat et calcule le pourcentage de réussite moyen pour chacune des quatre conditions.

L'index de sensibilité d' constitue également une mesure permettant d'évaluer les performances des participants. Cette dernière mesure la sensibilité perceptive d'une personne à détecter un événement particulier qui se produit dans un signal sonore. L'événement à détecter ici est un changement de fréquence. Il est primordial de noter que l'index de sensibilité est indépendant d'éventuel biais de réponses du participant, à la différence du pourcentage de réussite. Un d' proche de zéro indique que le participant répond aléatoirement alors qu'une augmentation d'un d' indique une augmentation de la capacité

du participant à percevoir correctement les changements de fréquence. Par exemple, en considérant que la modification à détecter est une augmentation de la fréquence, le tableau ci-après présente les différents cas possible en fonction de la modification réalisée par le programme et la modification choisie par l'utilisateur.

Modification		Choisie par l'utilisateur	
		« Monte »	« Baisse »
Réalisée par le programme	Augmentation	Succès	Echec
	Diminution	Fausse Alarme	Rejet Correct

Tableau 4 - Définition du vocabulaire utilisé en fonction de la modification réalisée par le programme et du choix de l'utilisateur

A partir de ce tableau, le pourcentage de réussite est calculé en additionnant les pourcentages de Succès et de Rejet Correct. L'index de sensibilité se calcule quant à lui selon la formule qui suit :

$$d' = Z(\text{Taux de Succès}) - Z(\text{Taux de Fausse Alarme})$$

où $Z(x)$ est l'inverse de la fonction de répartition de la loi Normale (Bonnet, 1986), avec $x \in [0,1]$.

La deuxième analyse proposée fournit un graphique représentant d' en fonction de l'ordre de S et de A ainsi qu'en fonction des fréquences de modulation utilisées.

Deux analyses supplémentaires ont été réalisées mais ne sont pas présentées dans le présent rapport puisqu'elles n'ont révélé aucune différence significative. Pour l'ensemble des participants, les performances n'ont pas été modifiées au cours des différents blocs, il n'existe par conséquent dans cette expérience pas d'effet d'apprentissage. L'influence du rapport entre la fréquence porteuse et la fréquence de modulation n'a également pas eu d'impact significatif sur les performances des participants.

Il est ensuite nécessaire de déterminer si les moyennes sont significativement différentes ou si les différences observées sont dues au hasard. Cette analyse s'effectue après le calcul des performances des participants en fonction des différents paramètres spécifiés précédemment. Une ANOVA, de l'anglais ANalysis Of VAriance, a été le modèle statistique appliqué afin de vérifier la significativité de chaque différence de moyennes.

Trois facteurs possédant chacun deux modalités ont été analysés et sont détaillés ci-dessous :

Facteur	Modalité 1	Modalité 2
Ordre relatif	S-A	A-S
Type de changement	P	M
Registre de fréquence	Bas	Haut

Tableau 5 - Détail des trois facteurs de l'expérience et de leurs modalités

Une condition correspond à la combinaison de modalités. Pour chaque condition, deux variables dépendantes ont été analysées par l'ANOVA et correspondent, pour chaque participant, au pourcentage de bonnes réponses ainsi qu'au d' . L'analyse a été effectuée à l'aide du logiciel édité par IBM SPSS, de l'anglais *Software Package for Statistical Analysis*.

Les résultats fournis indiquent les différences significatives pour chacune des variables indépendantes ainsi que pour les interactions entre ces dernières.

5.3. Résultats

Les analyses détaillées ci-dessous sont celles présentant des résultats significatifs. Les résultats obtenus suite aux entraînements ne sont en revanche pas détaillés puisque l'ensemble des participants a atteint 90 % de réussite lors de la réalisation de deux blocs consécutifs, ces performances ne peuvent donc pas être imputées au hasard.

5.3.1. Pourcentage de réussite et index de sensibilité d'

La figure 7 et 8 représentent respectivement le pourcentage de bonnes réponses et le d' moyennés sur l'ensemble des participants en fonction des conditions d'ordre entre A et S. Les résultats obtenus suite à la modification de P sont présentés en rouge, ceux obtenus suite à la modification de M sont présentés en bleu.

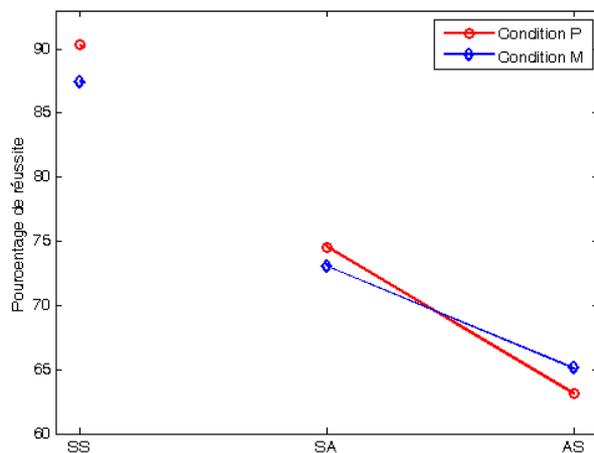


Figure 7 - Pourcentage de réussite des participants en fonction de la condition contrôlée S-S, de l'ordre S-A et de l'ordre A-S

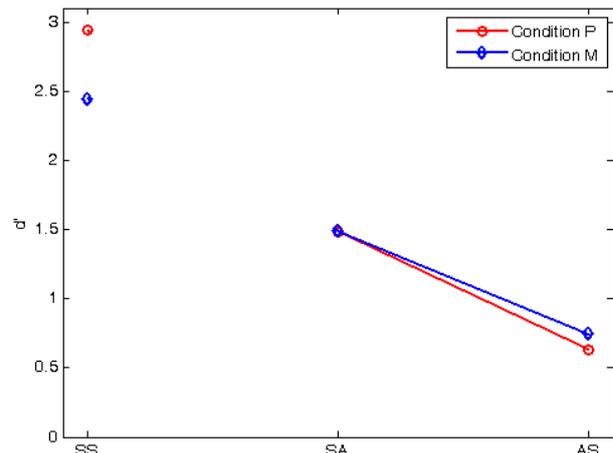


Figure 8 – Index de sensibilité d' des participants en fonction de la condition contrôlée S-S, de l'ordre S-A et de l'ordre A-S

La figure 7 montre un pourcentage de réussite moyen des participants dans la condition S-S de 89 %. Ce pourcentage est supérieur de 20 % à la condition S-A et est supérieur de 50 % à la condition A-S. Les différences de performances dans les conditions P et M ne sont pas significatives et ce quel que soit l'ordre S-S, S-A ou A-S.

L'index de sensibilité d' est en moyenne de 2,7 pour tous les participants pour la condition S-S. Cette sensibilité diminue de 80 % dans la condition S-A et chute de 380 % dans la condition A-S. En ce qui concerne l'ordre S-A, il n'existe pas de différence significative entre les performances de détection de changements de P ou de M. Ceci résulte de l'ajustement des performances décrit dans la partie [5.1.1 Stimuli](#). Il n'existe pas non plus de différence significative de d' dans la condition d'ordre A-S pour la détection d'une modification de P ou de M. Dans la condition S-S, le d' est 20 % supérieur dans la condition P par rapport à la condition M.

5.3.2. Impact du registre sur d'

La figure 9 ci-après détaille les performances obtenues dans la condition M en fonction de la gamme de fréquences de modulation utilisée. L'analyse effectuée ici est une scission en deux de la gamme de fréquences au niveau de la moyenne géométrique, ce qui correspond environ à 14 Hz pour les fréquences de modulation.

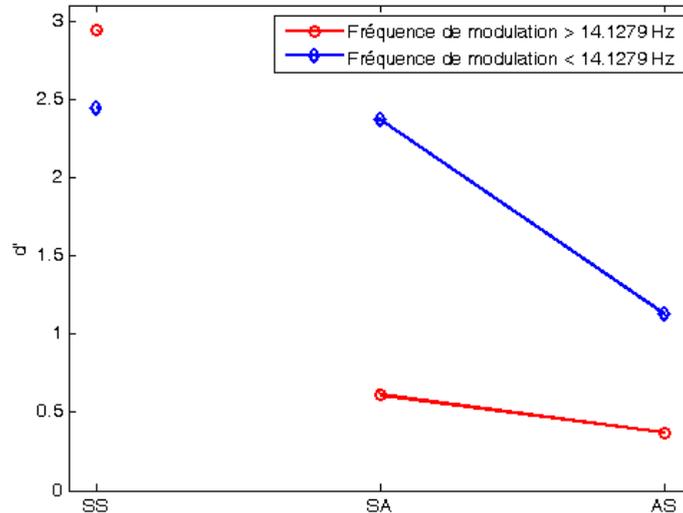


Figure 9 - Index de sensibilité d' en fonction de la condition d'ordre dans la condition M

Il n'existe pas de différence significative entre les performances obtenues pour les conditions S-S et S-A pour les fréquences basses de modulation. Les performances des fréquences de modulation élevées sont en revanche divisées par six entre les conditions S-S et S-A. Il n'existe enfin pas de différence significative dans les performances des conditions S-A et A-S pour les fréquences de modulation élevées alors que les performances pour les fréquences basses sont divisées par deux.

5.4. Discussion

Cette expérience a tenté de prouver l'existence des DCR en les comparant aux DCF. L'existence des DCF est avérée (Demany et Ramos, 2005) et leur manifestation se traduit par des performances similaires dans les conditions S-A et A-S. Les résultats de l'expérience 1 montrent une différence significative entre les ordres S-A et A-S dans la condition P, ce qui indique que les DCF ne se sont pas manifestés dans cette expérience. Cette absence de manifestation des DCF peut provenir du fait que l'existence de celles-ci a été prouvée en utilisant des *stimuli* composés de sons stationnaires, ce qui n'est pas le cas ici avec la modulation d'amplitude. Il est par conséquent impossible de comparer les DCF aux DCR et donc de confirmer ou non l'existence des DCR. Cette expérience prouve néanmoins que l'attention favorise grandement la détection d'un changement de fréquence de modulation.

Malgré les problèmes rencontrés, un effet remarquable transparait suite à l'analyse des fréquences utilisées. Lors de l'écoute d'une fréquence de modulation basse, il est en

effet aussi aisé de suivre son évolution lorsqu'elle est seule que lorsqu'elle est présentée dans A. Un effet délétère est en revanche observé en ce qui concerne les performances relatives à la détection d'un changement de fréquence de modulation élevée. Les MDI, qui sont les interférences dans la discrimination de fréquences de modulation, ne s'appliquent donc pas aux fréquences basses alors qu'elles s'appliquent aux fréquences élevées.

L'expérience qui suit propose une nouvelle approche afin d'étudier la sensibilité des auditeurs à un changement de fréquence de modulation d'amplitude.

6. Deuxième expérience : étude de la perception d'un changement de rythme

L'étude de l'effet de l'attention sur la perception d'un changement de rythme est poursuivie dans cette deuxième expérience. Des *stimuli* différents sont utilisés afin de rechercher un indice de l'existence des DCR.

Cette deuxième expérience étudie deux aspects différents du rythme afin de caractériser plus précisément la sensibilité auditive à un changement de fréquence de modulation.

Le premier aspect concerne l'asymétrie pouvant exister entre la détection d'une augmentation et d'une diminution de la fréquence de modulation. Si le participant est attentif à une modification de rythme, il est peu probable qu'une telle asymétrie puisse exister mais si des mécanismes de détection automatique tels que les DCR existent, alors la présence de cette asymétrie pourrait prouver leur existence.

L'impact d'une rupture temporelle au sein de la modulation constitue le deuxième aspect étudié. L'ajout d'une rupture temporelle *via* un silence à l'instant du changement informe l'auditeur de l'instant exact d'une telle modification. L'auditeur devrait donc pouvoir détecter plus aisément les changements. Toutefois, un silence provoque une rupture de rythme, la mémoire à court terme de l'auditeur sera donc mise en jeu et cette rupture temporelle entraînera donc très probablement une diminution des performances de détection d'un changement de rythme. S'il existe des DCR, alors les performances devraient être meilleures lorsqu'aucune rupture temporelle n'est présentée dans le *stimulus*. De plus, lorsqu'un silence est présenté au cours du *stimulus*, l'ensemble des composantes subit une rupture identique, ce silence devrait donc perturber l'activation des DCR. Lorsqu'aucun silence n'est présenté à l'instant du changement, l'intégralité des composantes de A poursuivent en revanche leur progression tandis que la composante modifiée subit une rupture temporelle. Une telle rupture devrait être aisément détectable par des DCR.

6.1. Matériel et méthode

L'objectif consiste ici à évaluer la capacité des participants à identifier la fréquence de modulation qui a été modifiée parmi les cinq fréquences présentes dans A. La différence de performances de la part des participants, obtenues lors d'accélération ainsi que de ralentissements rythmiques, sera évaluée. Lors de l'initiation de chaque bloc, l'auditeur sélectionne la caractéristique de ce bloc, c'est-à-dire si ce dernier contient des accélérations (a) ou des ralentissements (r). Ces modifications correspondent respectivement à des augmentations ou à des diminutions de fréquence de modulation.

Cette seconde expérience nécessite des matériel et méthode identiques à ceux décrits dans le cas de l'expérience 1, à l'exception des points présentés ci-après.

Le *stimulus* est ici composé de deux A successifs puis d'un son pur SP. A est composé de cinq P et de cinq M. SP se compose d'une fréquence porteuse identique à celle de l'un des cinq P de l'accord.

Une condition a été ajoutée à l'expérience afin d'identifier l'évolution des performances lors d'une rupture rythmique du *stimulus*. L'utilisateur sélectionne en effet lors de chaque initiation de bloc s'il souhaite entendre deux A séparés par un silence (s) de 100 ms ou si la transition entre ces A doit être continue (c).

La durée de A est ici de 2,5 s, un silence de 100 ms sépare le dernier A de SP et SP a une durée de 500 ms.

Le delta de M est ici fixé à 600 cents pour l'ensemble des participants.

La figure ci-dessous représente le spectrogramme d'un exemple de *stimulus*, c'est-à-dire l'évolution des fréquences au cours du temps.

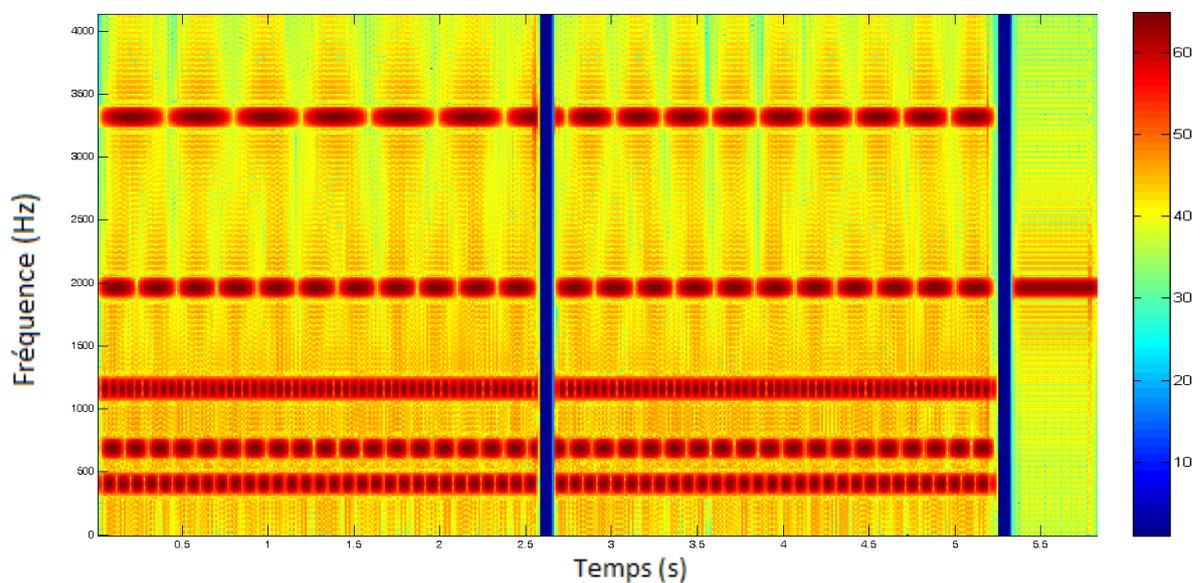


Figure 10 - Spectrogramme d'un exemple de *stimulus* de l'expérience 2

Les paramètres utilisés afin d'obtenir ce spectrogramme sont les suivants :

- SP après les deux A,
- Un silence entre les deux A,
- Une accélération durant le second A

Dans cet exemple, la composante la plus aiguë, c'est-à-dire celle située dans la partie supérieure du spectrogramme, accélère après le silence. Cette accélération est visible sur le spectrogramme *via* la longueur des raies rouges qui est inférieure dans le deuxième accord par rapport au premier. Il existe en effet six raies pour le premier accord et neuf raies pour le second accord. Le SP est au niveau de la quatrième composante et ne correspond pas à la composante ayant subi une modification.

6.2. Programmation

Le programme décrit pour la réalisation de l'expérience 1 est également utilisé pour cette expérience. Certaines modifications lui ont néanmoins été apportées et sont explicitées ci-après.

L'interface graphique propose trois choix au participant, à savoir :

- L'ordre relatif entre SP et les deux A (SP-AA ou AA-SP),
- La durée de transition entre les deux A (0 ou 100 ms),
- La nature de la modification de rythme (Accélération ou Ralentissement)

La structure présentée en figure 3 a été réutilisée afin de créer le nouveau *stimulus*. Contrairement à la production de S et de A lors de l'expérience 1, l'algorithme utilisé pour l'expérience 2 permet de produire deux A ainsi qu'un SP non-modulé.

6.3. Résultats

L'analyse du d' est présentée ci-dessous, où AV et AP désignent respectivement le fait que SP est positionné avant et après les deux A :

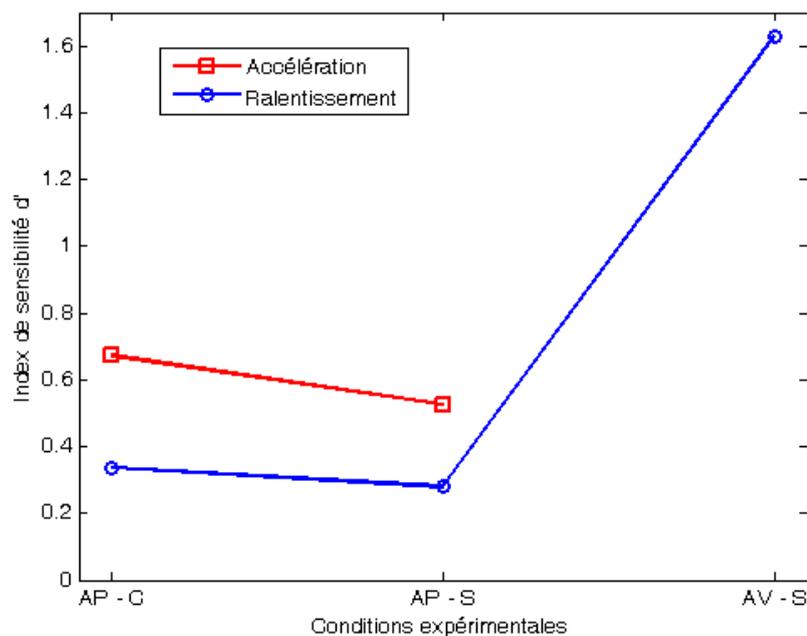


Figure 11 - Analyse de l'index de sensibilité d' en fonction de la position du SP (AV /AP) et de la condition Continuité (c) ou Silence (s)

Lorsque SP est positionné avant les deux A, les performances sont environ trois fois supérieures à celles obtenues lorsque SP est positionné après A. Ce résultat est obtenu en comparant les performances du participant pour la condition SP-AA avec celles pour la condition AA-SP.

Les performances obtenues par les participants lorsque la fréquence de modulation augmente sont deux fois supérieures à celles obtenues lorsque cette dernière diminue.

La différence de performances entre les conditions silence et continuité n'est pas significative.

Le détail des performances obtenues en fonction des fréquences de modulation est présenté dans la figure 12 ci-après :

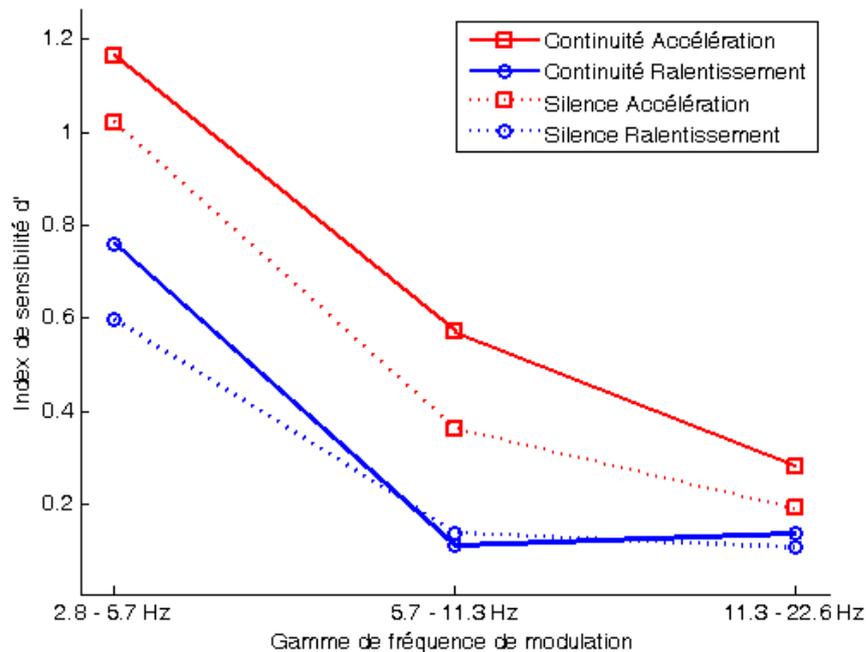


Figure 12 - Analyse de l'index de sensibilité d' en fonction du registre de la fréquence de modulation pour la condition AA-SP

A l'instar des résultats présentés dans la figure 11, il n'existe aucune différence significative entre les performances de la condition c et celles de la condition s. Quelle que soit la gamme de fréquences de modulation observée, les performances sont en moyenne deux fois supérieures pour la détection d'une accélération.

Le fait le plus remarquable concerne les performances qui sont, pour toute condition, environ deux fois meilleures pour des fréquences de modulation inférieures à 5,7 Hz que pour des fréquences de modulation comprises entre 5,7 et 11,3 Hz. Les performances observées dans la condition r pour les gammes de fréquences comprises entre 5,7 et 22,6 Hz ne sont pas significativement différentes. Dans la condition a, les performances pour la gamme de fréquences comprises entre 5,7 et 11,3 Hz sont en moyenne deux fois supérieures que pour la gamme de fréquences comprises entre 11,3 et 22,6 Hz.

6.4. Discussion

Tout comme décrit pour [la première expérience](#), les différences de résultats pour SP-AA ou AA-SP montrent que l'attention favorise grandement la détection d'un changement de rythme. Cette dernière observation constitue un premier indice en faveur de l'inexistence des DCR.

Les faibles performances obtenues pour la détection d'un ralentissement de rythme et pour les fréquences de modulation élevées indiquent un effet plancher, ce qui signifie que le système auditif humain n'est pas adapté à détecter ce type de changement.

L'absence de différence significative des performances entre les conditions c et s fournit un argument contre l'existence des DCR. Un détecteur automatique de changement de rythme serait en effet beaucoup plus stimulé par un rythme brutalement modifié que par un rythme modifié après un silence.

L'asymétrie observée dans les performances entre la détection d'une accélération et d'un ralentissement montre un effet encore jamais décrit. Les causes de cette asymétrie peuvent être expliquées par trois hypothèses.

La première hypothèse propose un rapprochement entre l'asymétrie observée dans l'expérience 2 à celle rapportée par Tajadura-Jiménez (2011) montrant que l'Homme est plus sensible à un son en approche plutôt qu'à un son qui s'éloigne. L'explication fournie peut toutefois difficilement être réduite à un avantage évolutif qui impliquerait qu'un son en approche constitue un avertissement de l'arrivée d'un prédateur. C'est pourquoi de nouvelles études devraient être menées afin de confirmer ou de compléter cette hypothèse.

Une deuxième hypothèse implique le delta utilisé. Ce dernier, fixé à 600 cents, est identique pour l'ensemble des participants, ce qui implique que la différence entre les fréquences lors d'une accélération ou d'un ralentissement ne sera pas la même en Hertz. Pour une fréquence de modulation initiale de 10 Hz par exemple, une augmentation de 600 cents produit une fréquence d'arrivée de 14,14 Hz alors qu'une diminution fournit une fréquence d'arrivée de 7,07 Hz. Dans le cas d'une accélération, une différence de fréquence de 4,14 Hz est donc obtenue alors que pour un ralentissement cette différence est de 2,93 Hz. Cette observation tend à montrer qu'il est plus aisé pour l'oreille humaine de détecter une modification importante de rythme plutôt qu'une modification plus faible de ce dernier. L'asymétrie des performances entre accélération et ralentissement pourrait être considérée comme étant la conséquence de cette différence qui, de plus, augmente avec la fréquence initiale. Une telle considération est néanmoins incorrecte puisqu'elle impliquerait que les performances des participants soient meilleures lors de la détection de rythmes rapides que lorsque ceux-ci sont plus lents, ce qui n'est expérimentalement pas le cas.

Une dernière hypothèse permet d'expliquer partiellement l'asymétrie des performances obtenues lors des détections d'accélération et de ralentissement de rythmes. Le colliculus inférieur constitue un amas de neurones du cerveau permettant la représentation mentale des rythmes (Schreiner, 1988). Une accélération et une décélération de rythme provoquent respectivement une augmentation et une diminution de l'activité neuronale du colliculus inférieur. Or, une augmentation de l'activité neuronale est mieux détectable de façon automatique qu'une diminution (Cervantes Constantino, *et al.*, 2012), ce qui entraîne donc une asymétrie de perception des rythmes. Les neurones du colliculus inférieur sont par ailleurs activés uniquement lorsque le participant n'est pas attentif au *stimulus* qu'il entend. Lors d'une écoute attentive d'un *stimulus* en revanche, aucune

asymétrie de détection de rythme ne devrait subsister puisque d'autres zones du cerveau sont activées lorsque l'attention est mise en jeu. La présence d'une asymétrie de détection de rythmes pourrait par conséquent constituer un argument en faveur de l'existence des DCR. Toutefois, au vu des résultats obtenus, cet argument est uniquement valide pour des rythmes lents. En dépit de cette dernière restriction, cette hypothèse semble actuellement être la plus probable afin d'expliquer la présence de l'asymétrie entre la détection d'une accélération et celle d'un ralentissement de rythme.

7. Conclusion

Ce stage de fin d'études du Master Recherche à l'Université de Bordeaux ainsi que du diplôme d'ingénieur à l'ENSEIRB-MATMECA m'a permis, durant cinq mois, de mettre en pratique mes connaissances en traitement du signal acquises au cours de mon parcours universitaire. Ce stage m'a également permis de découvrir le milieu de la Recherche fondamentale en psychoacoustique et de me familiariser avec les méthodes utilisées dans ce domaine. Je souhaite actuellement orienter mon parcours vers un emploi dans la Recherche. Cependant, et bien que la Recherche fondamentale s'est avérée très instructive au cours de mon stage, la proportion de développement s'est révélée trop peu importante pour correspondre à mon projet professionnel. Je souhaite donc par la suite exercer une activité dans le domaine de la Recherche appliquée.

Mon intégration dans l'équipe du Dr Laurent Demany s'est avérée très rapide, de même que ma mise à niveau concernant les notions élémentaires de psychoacoustique, notamment *via* la consultation d'articles internationaux de référence dans ce domaine. J'ai pu réaliser deux expériences supervisées par mon directeur de recherche. Chaque expérience m'a ainsi permis d'approfondir ma connaissance du fonctionnement de la perception auditive du rythme par l'oreille humaine. Les résultats de la seconde expérience sont concluants et novateurs et peuvent de ce fait faire l'objet d'une publication. Ceci renforce ma motivation à poursuivre mon parcours dans le domaine de la Recherche. Les objectifs proposés au début de ce stage ont, de plus, été atteints puisque de nouvelles connaissances concernant la sensibilité auditive à une modification de fréquence de modulation ont été établies. Le cadre de ce stage, tant au niveau des objectifs à atteindre que du point de vue de la composition de l'équipe, a par ailleurs constitué pour moi l'opportunité de développer mes compétences techniques.

Cette expérience professionnelle au sein d'un laboratoire de recherche m'a permis d'appréhender certaines des difficultés majeures dans la réalisation d'une thèse puisqu'elle a mis mon autonomie, mes connaissances techniques ainsi que ma motivation à l'épreuve. Je tiens enfin à exprimer ma satisfaction d'avoir été le moteur de cette étude qui s'est déroulée dans d'excellentes conditions matérielles ainsi que ma gratitude envers l'équipe de recherche qui m'a fourni un environnement de travail agréable, propice aux discussions et aux échanges d'idées.

8. Perspectives

Les deux expériences proposées fournissent des informations relatives à la sensibilité auditive humaine et plus précisément à la capacité de l'oreille humaine à détecter des changements de fréquence de modulation, donc de rythme. Deux arguments en faveur de l'inexistence des DCR ont été proposés, un troisième cependant maintient un doute quant à leur existence. De nouvelles interrogations ont été soulevées suite aux données des expériences, notamment en ce qui concerne l'interprétation physiologique des résultats. Les réponses à ces interrogations permettraient par exemple de comprendre les raisons de l'asymétrie de la sensibilité humaine à des rythmes qui accélèrent ou décélèrent.

Les deux expériences menées offrent plusieurs pistes d'étude en psychoacoustique afin de caractériser plus en détail la perception des rythmes chez l'Homme. Afin d'explorer de telles pistes, une nouvelle expérience pourrait consister à soumettre la perception des participants à la succession de deux A dont les fréquences de modulation seraient échangées au sein d'un même A. Une deuxième expérience permettrait par ailleurs d'étudier l'impact de la détection d'une modification de fréquence porteuse lorsqu'une modification de fréquence de modulation s'effectue également.

L'expérience cependant sélectionnée et réalisée actuellement au sein du laboratoire consiste à reproduire l'expérience 2 en modifiant uniquement les fréquences porteuses. Il est probable que les résultats de cette nouvelle expérience permettent de comparer efficacement les DCF aux DCR.

9. Sources

9.1. Bibliographie

Alam, M., et al. (2013). Amplitude modulation features for emotion recognition from speech. *Interspeech conference*, 14th Annual Conference of the International Speech Communication Association.

Bonnet, C. (1986). Manuel pratique de psychophysique. *Armand Colin*, 59-86.

Carcagno, S., Semal, C., & Demany, L. (2011). Frequency-shift detectors bind binaural as well as monaural frequency representations, *Journal Of Experimental Psychology: Human Perception And Performance*, 37, 1976-1987.

Cervantes Constantino, F., et al. (2012). Detection of Appearing and Disappearing Objects in Complex Acoustic Scenes, *PLOS One*, 7, 1-13.

De Cheveigné, A. (1997). Harmonic fusion and pitch shifts of mistuned partials. *The Journal of the Acoustical Society of America*, 102, 1083-1087.

Demany, L., & Ramos, C., (2005). On the binding of successive sounds : Perceiving shifts in nonperceived pitches, *Journal of The Acoustical Society of America*, 117, 833–841.

Demany, L., et al. (2009). Tuning properties of the auditory frequency-shift detectors. *Journal of the Acoustical Society of America*, 126, 1342-1348.

Demany, L., et al. (2010). Fundamental differences in change detection between vision and audition, *Experimental Brain Research*, 203, 261–270.

Demany, L., et al. (2011). Implicit versus explicit frequency comparisons : two mechanisms of auditory change detection, *Journal of Experimental Psychology*, 37, 597–605.

Formby, C., et al. (1985). Differential sensitivity to tonal frequency and to the rate of amplitude modulation of broadband noise by normally hearing listeners, *Journal of the Acoustical Society of America*, 78, 70-77.

Giraud, A-L., et al. (2000). Representation of the Temporal Envelope of Sounds in the Human Brain, *Journal of Neurophysiology*, 84, 1588-1598.

Henry, M. J., et al. (2014). Entrained neural oscillations in multiple frequency bands comodulate behavior, *PNAS*, 111, 14935-14940.

Kollmeier, B., et al. (2007). Hearing – From Sensory Processing to Perception. *Springer*.

Lange, K., et al. (2010). Can a regular context induce temporal orienting to a target sound?, *International Journal of Psychophysiology*, 78, 231-238.

- Lefkothea-Vasiliki A., et al. (2011). The role of temporal regularity in auditory segregation, *Hearing Research*, 280, 228–235.
- Leviton, C. A., et al. (2014). Rate perception adapts across the senses: evidence for a unified timing mechanism, *Scientific Reports*.
- Mencattini, A., et al. (2014). Speech emotion recognition using amplitude modulation parameters and a combined feature selection procedure. *Knowledge-Based Systems*, 63, 68-81.
- Moore, B. C. J., et al. (1992). Modulation discrimination interference and auditory grouping, *Philosophical Transactions : Biological Sciences*, 336, 339-346.
- Munkong, R., et al. (2008). Auditory perception and cognition, *Signal Processing Magazine, IEEE*, 25, 98-117.
- Oetjen, A., et al. (2014). Spectro-temporal modulation masking patterns reveal frequency selectivity, *Journal of the Acoustical Society of America*, 137, 714-723.
- Pavani, P., et al. (2008). Change perception in complex auditory scenes, *Perception and Psychophysics*, 70, 619-629.
- Payton, K., et al. (1999). A method to determine the speech transmission index from speech waveforms, *Journal of the Acoustical Society of America*, 106, 3637.
- Riess Jones, M., et al. (1981). Evidence for Rhythmic Attention, *Journal of Experimental Psychology*, 7, 1059.
- Riess Jones, M., et al. (2012). Temporal aspects of stimulus-driven attending in dynamic arrays, *Psychological Science*, 13, 313-319.
- Schröger, E., Bendixen, A., Denham, S. L., Mill, R. W., Böhm, T. M., & Winkler, I. (2014). Predictive regularity representations in violation detection and auditory stream segregation: from conceptual to computational models, *Brain topography*, 27(4), 565-577.
- Strauss, M., et al. (2015). Disruption of hierarchical predictive coding during sleep, *PNAS*, 112, 1353-1362.
- Tajadura-Jiménez, A., et al. (2011). I-Space: the effects of emotional valence and source of music on interpersonal distance, *PLoS One*, 6, 26083.
- Viemeister, N. F., et al. (2002). Psychophysical and physiological aspects of auditory temporal processing. *Holmens Trykkeri*, 273-291.
- Yost, W. A., et al. (2008). Auditory perception of sound sources. *Springer*, 233-281.

9.2. Webographie

<http://journal.frontiersin.org/article/10.3389/fpsyg.2011.00110/full#>, consulté le 2 Mars 2015

<https://www.dynamicbrain.ca/syllable-stacks.html>, consulté le 3 Mars 2015

http://en.citizendium.org/wiki/Music_perception, consulté le 4 Mars 2015

<http://www-psych.stanford.edu/~lera/psych115s/notes/signal/>, consulté le 20 Avril 2015