



HAL
open science

Mesure de l'attention auditive par électroencéphalographie

Johann Benerradi

► **To cite this version:**

Johann Benerradi. Mesure de l'attention auditive par électroencéphalographie. Informatique [cs]. 2019. hal-02285224

HAL Id: hal-02285224

<https://inria.hal.science/hal-02285224v1>

Submitted on 12 Sep 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Inria Sophia Antipolis - Méditerranée
Équipe Athena

Rapport de stage
Master 2 Sciences Cognitives

Mesure de l'attention auditive par électroencéphalographie

Stage réalisé du 15 mars au 15 septembre 2019

Auteur :
Johann BENERRADI

Encadrante :
Maureen CLERC
Référent :
Azim ROUSSANALY



Année universitaire 2018-2019

Table des matières

1	Contexte	1
1.1	Inria et l'équipe-projet Athena	1
1.2	Le projet MICADôme	1
1.3	Sujet du stage	2
2	Introduction à l'attention auditive	3
2.1	Fonctionnement du cerveau	3
2.2	Enregistrement de l'activité cérébrale par EEG	4
2.3	Étude des signaux EEG	5
2.4	Attention auditive en EEG	7
2.5	Hypothèses	9
3	Design expérimental	10
3.1	Objectif	10
3.2	Matériel	10
3.3	Calibration du volume des sons	12
3.4	Protocole expérimental	13
3.5	Implémentation de l'expérience	14
3.6	Réalisation des expériences	16
4	Méthodes d'analyse	17
4.1	Filtrage en fréquence	17
4.2	Découpage des essais	18
4.3	Visualisation des données	19
5	Résultats	21
5.1	Analyse des sons	21
5.2	<i>ASSR</i> pour les sons seuls en binaural	22
5.3	<i>ASSR</i> pour l'attention sur un son parmi deux	23
6	Discussion	25
6.1	Choix des stimuli audios	25
6.2	Périphérique audio et environnement	26
6.3	Interface de calibration	26
6.4	<i>ASSR</i> et détection de l'attention auditive	27
6.5	Pistes de recherche futures	27
7	Conclusion	29
	Annexes	32

Chapitre 1

Contexte

Ce stage de fin d'études d'une durée de 6 mois a été réalisé dans le cadre du Master 2 Sciences Cognitives de l'Institut des Sciences du Digital, Management et Cognition (IDMC) appartenant à l'Université de Lorraine. Il a été réalisé au sein de l'équipe Athena de Inria Sophia Antipolis - Méditerranée, dans le cadre du projet MICADôme. Ce chapitre va ainsi présenter Inria et l'équipe-projet Athena, le projet MICADôme, pour finir par le sujet du stage.

1.1 Inria et l'équipe-projet Athena

Inria, Institut national de recherche en informatique et en automatique, est né en 1967 à Rocquencourt, alors nommé simplement IRIA pour Institut de recherche en informatique et en Automatique [16]. Il a été créé dans le but de permettre aux sciences du numérique d'être au service de la société grâce au transfert technologique de l'innovation à l'industrie. Fort de ses 8 centres à Bordeaux, Grenoble, Lille, Nancy, Paris, Rennes, Saclay, et Sophia Antipolis dans lequel ce stage a été effectué, Inria emploie 2500 personnes et comporte 200 équipes-projets.

Inria a 5 domaines principaux de recherche :

- les mathématiques appliquées, le calcul, et la simulation ;
- l'algorithmique, la programmation, les logiciels, et les architectures ;
- les réseaux, les systèmes et services, et le calcul distribué ;
- la perception, la cognition, et l'interaction ;
- la santé, la biologie, et les sciences de la planète.

Le centre de recherche Inria Sophia Antipolis - Méditerranée est historiquement le troisième centre, construit après Rocquencourt et Rennes, et compte environ 500 employés, dont 360 scientifiques répartis entre Sophia Antipolis et son antenne à Montpellier. Ceci représente 33 équipes, dont Athena composée de 17 membres (hors stagiaires) dirigée par Rachid Deriche. Cette équipe se concentre sur l'étude du système nerveux central humain à partir de mesures non-invasives, plus spécifiquement l'Imagerie par Résonance Magnétique (IRM) de diffusion, l'électroencéphalographie (EEG) et la magnétoencéphalographie (MEG). En ce qui concerne l'EEG, Athena s'intéresse au temps réel avec les interfaces cerveau-ordinateur, et a notamment développé un clavier contrôlable par les ondes cérébrales : le P300-speller.

1.2 Le projet MICADôme

Le projet Musique Interactive Côte d'Azur (MICA) financé par l'Université Côte d'Azur a pour but de faire se rejoindre la recherche en sciences du numérique et en art. Il est porté par Michel Pascal du Conservatoire à Rayonnement Régional de Nice et réunit également le Centre Transdisciplinaire d'Épistémologie de la Littérature (CTEL), le Laboratoire d'Anthropologie et de Psychologie Cliniques, Cognitives et Sociales (LAPCOS), le Hublot (centre de création en arts numériques), le Centre International de Recherche Musicale (CIRM), et enfin l'équipe Athena de Inria. Ce projet a donné lieu à la création d'un dôme en 3D de 16 haut-parleurs avec pour but de l'étendre à 21 haut-parleurs (Figure 1.1).



FIGURE 1.1 – MICADôme au Conservatoire à Rayonnement Régional de Nice

Grâce à ce dispositif, il est alors possible d'étudier comment la spatialité du son peut être utilisée pour permettre une composition musicale immersive. Plus particulièrement au sein de l'équipe Athena, le but est d'étudier la réponse cérébrale à ce type de sons spatialisés mais également d'utiliser l'activité cérébrale comme modalité de contrôle de la composition musicale afin de produire une expérience interactive.

1.3 Sujet du stage

Le but de ce stage a été tout d'abord de développer une expérience de neurophysiologie comportementale afin d'étudier par EEG les caractéristiques de l'attention auditive. Ce travail allait de la création d'un protocole expérimental à l'implémentation de cette expérience sous forme d'un logiciel permettant de gérer différentes tâches tout en enregistrant les signaux EEG. Il s'en est alors suivi tout une phase d'acquisition des données en invitant des sujets à participer aux expérimentations. Dans un deuxième temps, ce stage a consisté à analyser les signaux enregistrés pour déterminer sur quel élément d'une scène auditive un sujet porte son attention.

Chapitre 2

Introduction à l'attention auditive

Dans ce chapitre, une introduction à l'attention auditive sera faite, en commençant de la base avec le fonctionnement du cerveau, en passant par l'EEG et l'étude des signaux, pour en arriver aux différents paradigmes de détection de l'attention auditive et leurs applications. Ceci nous mènera enfin à présenter les hypothèses de recherche qui ont été émises.

2.1 Fonctionnement du cerveau

Le cerveau est l'un des organes les plus importants du corps humain, cependant nous en savons très peu à son sujet tant il est complexe. Il est intéressant d'un point de vue philosophique, sociologique, scientifique, et médical. Chez l'humain son poids est en moyenne d'environ 1,3 kg [33]. Il est composé de 2 hémisphères (droit et gauche) reliés par un réseau de fibres appelé corps calleux, ainsi que du cervelet. Ce dernier est impliqué dans la motricité et notamment le maintien de la posture [12]. On peut diviser chaque hémisphère du cerveau en différentes régions appelées *lobes* (Figure 2.1).

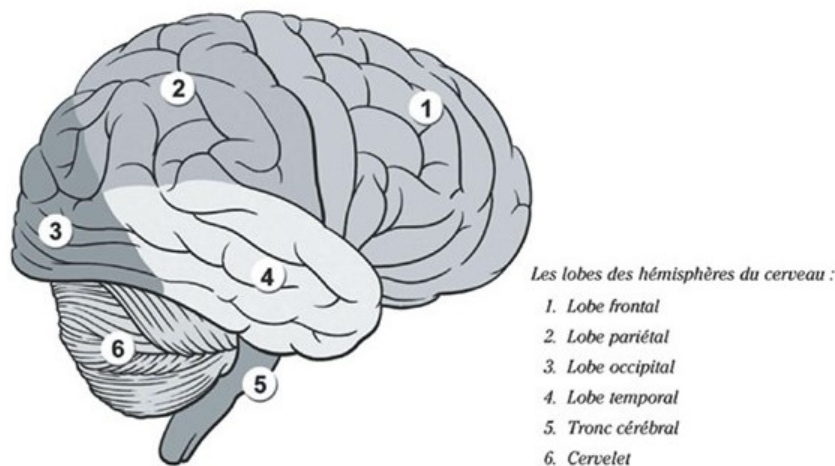


FIGURE 2.1 – Les différents lobes du cerveau (*Source : icm-institute.org*)

Chaque lobe a été décrit comme ayant une fonction particulière même si il est connu que la plus part des fonctions du cerveau impliquent plusieurs régions différentes [28]. Ces différentes fonctions ont pu être mises en évidence au cours des âges en étudiant comment certaines lésions du cerveau affectaient certaines fonctions. Ainsi on peut décrire le lobe frontal comme étant responsable du raisonnement, de la planification et de l'attention. Le lobe pariétal est responsable du traitement des informations sensorielles comme le toucher et la proprioception (sensibilité profonde). Le lobe occipital

est le lieu du traitement de l'information visuelle. Enfin le lobe temporal est responsable du traitement des informations auditives et visuelles (reconnaissance des visages), ainsi que des processus mnésiques avec l'hippocampe.

Lorsque l'on s'intéresse au niveau cellulaire, le cerveau est composé d'un nombre de cellules nerveuses, appelées neurones, autour de 100 milliards, connectées entre elles par des axones et dendrites, extensions du corps cellulaire (Figure 2.2).

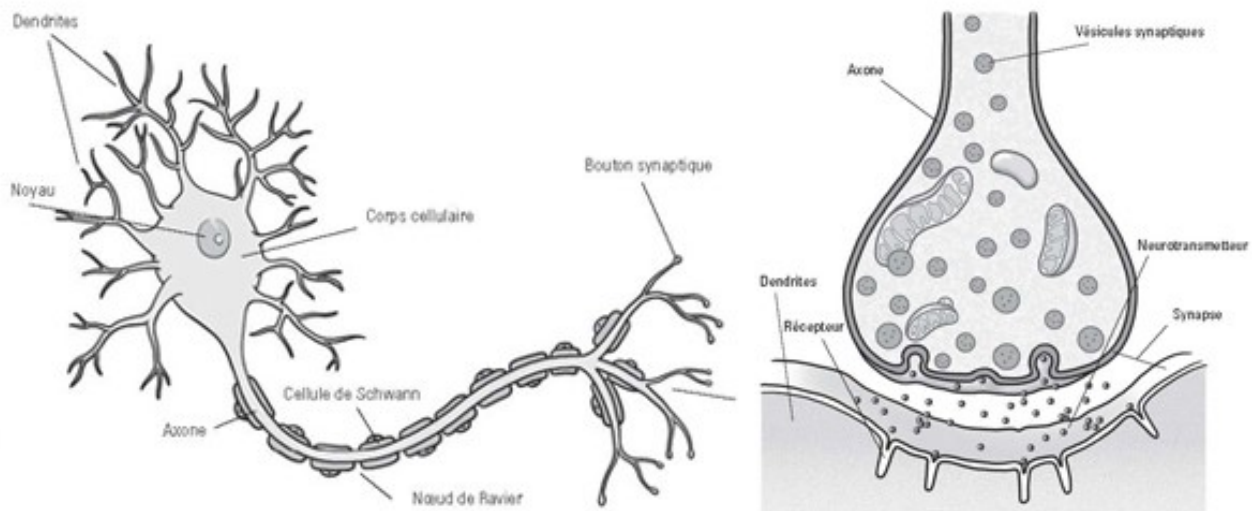


FIGURE 2.2 – Représentation schématique du neurone (*partie gauche*) et de la synapse (*partie droite*)
(Source : *icm-institute.org*)

La communication entre ces cellules nerveuses est réalisée grâce à des signaux électriques appelés *potentiels d'action*. Ces derniers parcourent les axones pour arriver jusqu'aux liaisons avec des dendrites appelées synapses. Ils vont donner lieu à une production de neurotransmetteurs (catécholamines, sérotonine, endorphines, ...) dans de la connexion synaptique. Ces neurotransmetteurs vont permettre la transmission de l'information nerveuse au niveau extracellulaire de l'extrémité de l'axone à la dendrite de la cellule suivante. Ces neurotransmetteurs vont pouvoir inhiber ou activer le neurone post-synaptique ce qui va permettre de continuer la transmission du signal électrique.

2.2 Enregistrement de l'activité cérébrale par EEG

La transmission des influx nerveux étant électrique, il est possible de la mesurer à l'échelle macroscopique grâce à une technique appelée l'électroencéphalographie (EEG). Le premier à avoir mesuré cette activité électrique est Richard Canton à la fin du XX^{ème} siècle, mais le premier électroencéphalogramme humain est réalisé par Hans Berger en 1924 grâce à une amplification des signaux [4]. L'EEG consiste à enregistrer l'activité électrique du cerveau grâce à des capteurs appelés *électrodes* qui peuvent être positionnés à la surface du cuir chevelu. Cette activité électrique étant cependant très faible, elle doit être amplifiée grâce à un amplificateur.



FIGURE 2.3 – Casque EEG non-invasif (*Source : www.maxisciences.com*)

Bien que sa résolution spatiale soit limitée, l'EEG est une technique de neuro-imagerie qui a pour avantage d'avoir une très bonne résolution temporelle comparée à d'autres techniques comme l'Imagerie par Résonance Magnétique (IRM) ou la Tomographie par Émission de Positons (TEP) [25].

L'EEG peut être intra-crânien ou non-invasif (Figure 2.3). En ce qui concerne les casques EEG non-invasifs, plusieurs types d'électrodes existent. Tout d'abord les électrodes avec gel, qui constituent la référence en terme d'EEG non-invasif, consistent à appliquer les électrodes avec un gel de solution saline qui permet d'améliorer la conduction électrique entre le crâne et les électrodes. Ce dispositif est très performant mais présente tout de même l'inconvénient d'être long à installer et de nécessiter de se laver les cheveux après utilisation. À l'inverse, il existe des électrodes dites sèches qui sont composées de petits pics métalliques qui permettent d'atteindre le cuir chevelu en passant au travers des cheveux. L'avantage de ce système est qu'il est très rapidement déployable et pratique à installer, mais il souffre cependant d'une moins bonne impédance électrique ce qui rends les signaux plus bruités et il peut parfois être assez inconfortable pour le sujet [14].

2.3 Étude des signaux EEG

L'étude de motifs des signaux EEG peut être très intéressante car elle peut donner des information sur l'état cognitif mais également l'état de santé d'un patient. L'EEG permet notamment de diagnostiquer des maladies comme l'épilepsie ou les troubles du sommeil, mais également la mort cérébrale [3].

Différents types d'ondes

Il est possible d'identifier et d'interpréter certains motifs également appelés *patterns* formés par les signaux EEG. Notamment, certaines oscillations à des fréquences particulières peuvent être reconnaissables.

On retrouve ainsi plusieurs types d'oscillations identifiables [24, 2] dont certains sont présentés Figure 2.4 :

- les ondes *gamma* entre 30 et 100 Hz sont caractéristiques d'une concentration intense ;
- les ondes *bêta* entre 12 et 30 Hz sont typiques d'un état de conscience alerte, et de réflexion active ;
- les ondes *alpha* au niveau occipital ou *mu* dans le cortex moteur (partie postérieure du lobe frontal), entre 8 et 12 Hz, sont caractéristiques d'un état de relaxation physique et mentale, elles sont typiquement visible en occipital lorsqu'un sujet ferme les yeux ;

- les ondes *thêta* entre 4 et 8 Hz sont visibles en EEG lors de l'endormissement et du sommeil léger ;
- enfin les ondes *delta* en dessous de 4 Hz sont typiquement présentes lors du sommeil profond.

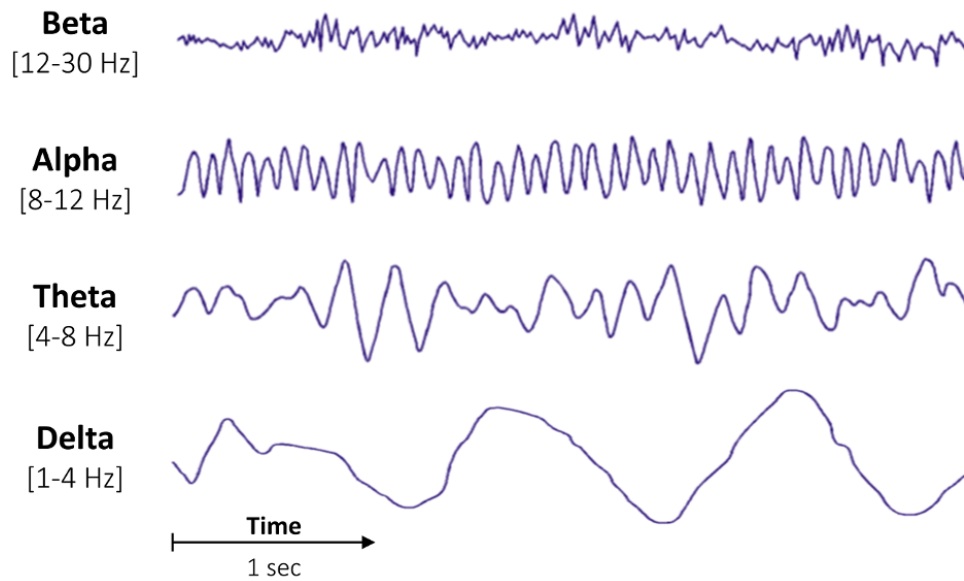


FIGURE 2.4 – Les différents types d'ondes en EEG (*Source : raphaelvallat.com*)

Potentiels évoqués

On peut également détecter en EEG des structures appelées *potentiels évoqués*, qui sont décrits par une modification du potentiel électrique en réponse à un stimulus externe [22]. La survenue d'un potentiel évoqué est synchronisée à la survenue du stimulus qui peut notamment être sensoriel (stimulus visuel donnant un potentiel évoqué visuel, auditif donnant un potentiel évoqué auditif, ...). Ces potentiels évoqués sont cependant généralement difficiles à détecter tels quels en raison d'un faible rapport signal sur bruit. Il est donc généralement nécessaire de moyenniser plusieurs signaux identiques et de les synchroniser sur le stimulus afin de pouvoir les interpréter [20].

En ce qui concerne les potentiels évoqués visuels et auditifs, ils sont composés d'ondes négatives et d'ondes positives comme on peut le voir sur la Figure 2.5. On peut typiquement reconnaître des structures particulières caractéristiques des différents potentiels évoqués [20] comme par exemple :

- l'onde P100 (ou P1) qui est positive survient 100 ms après l'apparition d'un stimulus visuel ;
- l'onde N100 (ou N1), est la première onde négative, 100 ms après l'apparition du stimulus et reflète un processus de discrimination des stimuli avec une focalisation de l'attention ;
- l'onde N170, onde négative apparaissant après 170 ms, est caractéristique de la reconnaissance des visages ;
- l'onde P300 (ou P3) est une onde positive apparaissant 300 ms après un stimulus rare attendu parmi d'autres stimuli non-cibles.

Bien que ces structures de potentiels évoqués aient été beaucoup décrites pour les stimuli visuels, on en retrouve également avec des stimuli auditifs.

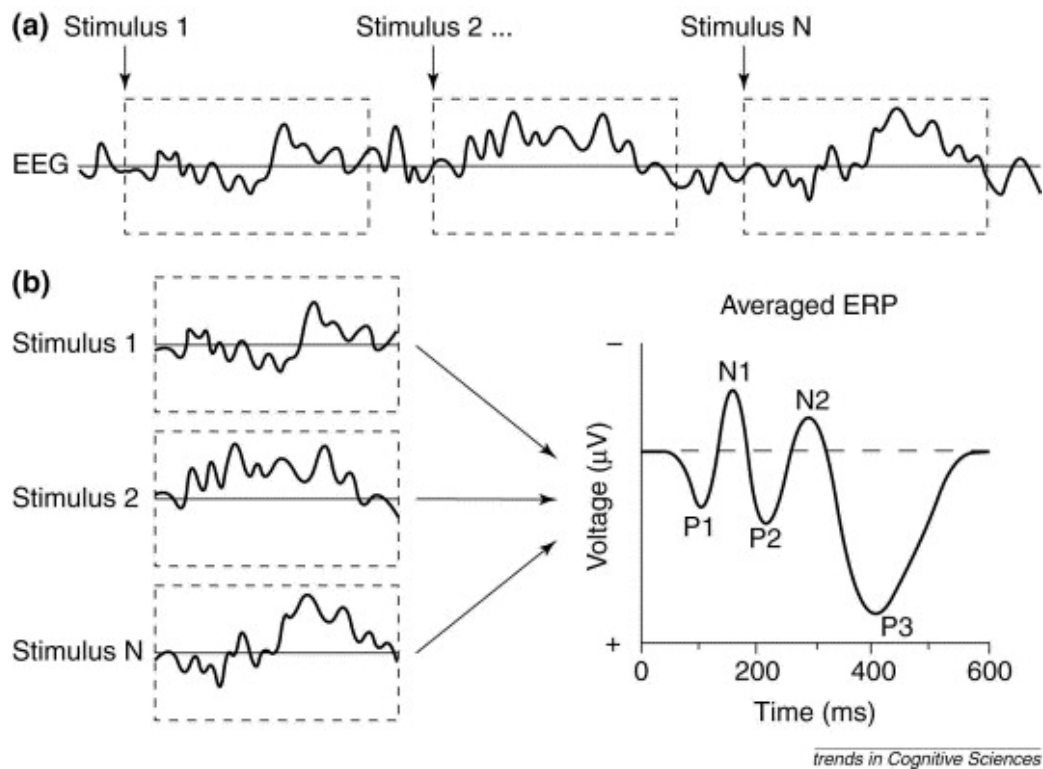


FIGURE 2.5 – Décomposition des potentiels évoqués (Source : Luck et al. 2000)

2.4 Attention auditive en EEG

Lorsque des stimuli sensoriels sont présentés à une fréquence constante, il se produit au niveau des signaux EEG ce que l'on appelle une réponse *steady-state*. Dans le cas de stimuli auditif, on l'appelle *réponse steady-state auditive* ou *ASSR* pour *Auditory Steady-State Response* [31, 7].

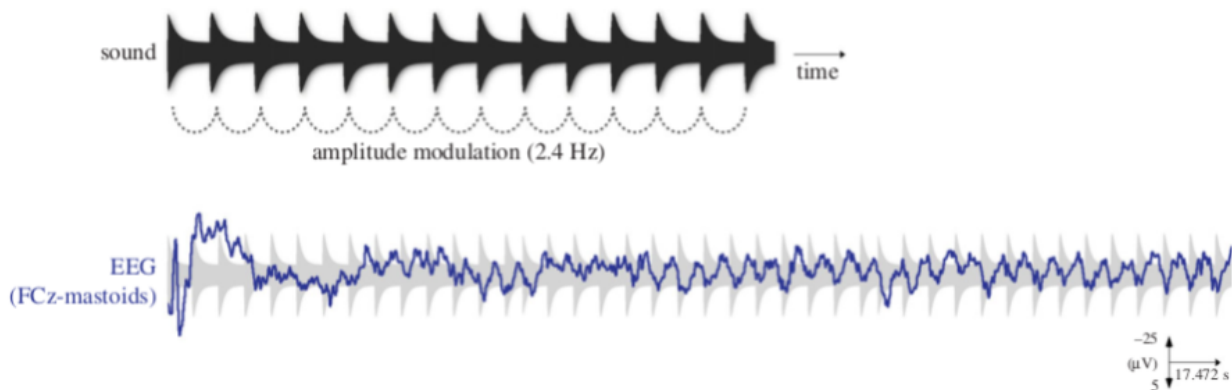


FIGURE 2.6 – Comparaison du signal audio à l'EEG (Source : Nozaradan et al. 2014)

Ceci est illustré par la Figure 2.6 où les stimuli présentés à une fréquence constante sont traduits en EEG sous la forme d'une modulation d'amplitude qui présente des accentuations à une fréquence spécifique. On remarque que le tracé EEG suit à peu près la même oscillation que l'amplitude du son. Ceci est d'autant plus visible Figure 2.7 avec une représentation de la densité des fréquences. On observe alors une dominance de la fréquence de modulation du son au niveau de l'EEG avec des pics aux mêmes fréquences.

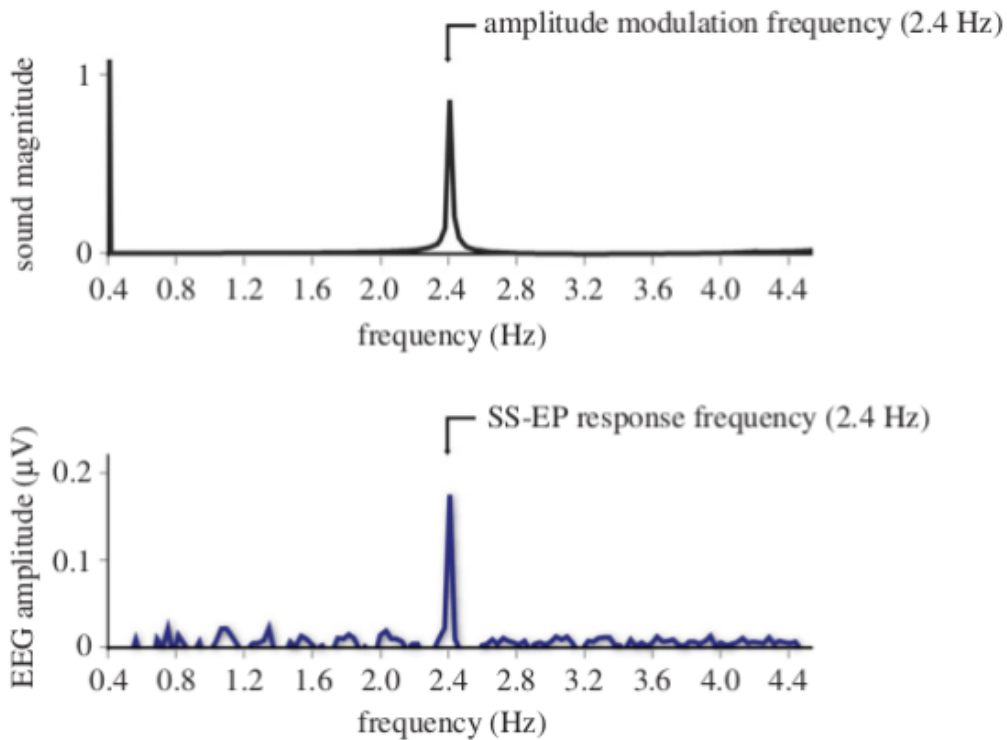


FIGURE 2.7 – Décomposition spectrale du signal audio comparée à celle de l’EEG (*Source : Nozaradan et al. 2014*)

Ces *ASSRs* peuvent être plus ou moins visibles. John et al. [18, 17] montrent que lors d’une modulation d’amplitude sinusoïdale ou exponentielle, au plus il y a de différence entre les amplitudes les plus hautes et les amplitudes les plus basses, au plus l’*ASSR* sera prononcée. Ceci est d’autant plus visible si les amplitudes les plus basses retombent à 0 entre chaque maximum donnant ainsi des sons séparés de silences.

Schafer et al. [32] montrent en 2011 que l’on retrouve une réponse *steady-state* à la même fréquence que le rythme qu’un sujet écoute mais que l’on retrouve également une réponse à la fréquence de la mesure. Ainsi on peut identifier à partir des EEG si le sujet écoutait un rythme binaire ou ternaire (une note accentuée sur 2 ou une note accentuée sur 3). De plus ils ont étendu cette expérience à l’imagination de mesure. Lors de l’écoute d’un rythme dont toutes les notes sont accentuées de la même manière, ils ont pu identifier si le sujet imaginait un rythme binaire ou ternaire à partir de la fréquence des *ASSR* en EEG. Des résultats similaires sont retrouvés par Nozaradan et al. [27].

Dans la même lignée Shuai et al. [34] montrent que l’*ASSR* peut être utilisée afin de détecter l’attention sur des événements sonores marquant dans un environnement bruyant.

Ce type de méthode de détection de l’attention peut être également appliqué à la détection de l’attention à un discours parmi plusieurs : c’est le phénomène *cocktail party* décrit par Colin Cherry [6], qui repose sur l’exemple que nous sommes capables de nous concentrer sur le discours d’une personne même s’il y a d’autres personnes qui parlent autour lors d’une réception. Srinivasan et al. [35] utilisent par exemple des analyses fréquentielles afin d’identifier sur quel discours parmi deux un sujet porte son attention.

Il existe également d’autres techniques afin de classifier l’attention auditive comme par exemple des techniques de séparation de sources comme le montrent Deng et al. [10] avec l’attention sur des répétitions de syllabes. On trouve également dans la littérature l’analyse de scènes auditives en s’intéressant à la prédictibilité des sons [1].

D’autres travaux ont plutôt mis l’accent sur l’étude du temps que met le sujet pour identifier que la scène auditive est composée de plusieurs sons différents et pour qu’il arrive à les différencier. Sussman-Fort et Sussman [36] s’intéressent notamment pour cela à la composition même des potentiels évoqués comme l’onde P100 et l’onde N100. Kondo et al. [19] étendent également ce principe lors de mouvement

de la tête qui induisent une réorganisation de la perception de la scène auditive.

De manière générale la littérature se recoupe pour dire que l'attention auditive peut être détectée de manière significative dans les zones du cerveau frontale et centrale (électrodes F, FC et C en EEG) [32, 35], mais également au niveau temporal en IRM fonctionnelle [11].

La détection de l'attention auditive peut avoir plusieurs applications. La principale application de l'étude de l'ASSR est la détection de la surdité chez l'enfant mais également l'évaluation de la qualité de l'audition chez les adultes.

Lorsque ces principes sont utilisés pour être traités en temps réel, on est dans le cadre des interfaces cerveau-ordinateur ou *BCI* pour *Brain-Computer Interface* [21]. Certaines *BCIs* utilisent ainsi les signaux EEG afin de notamment contrôler de la musique [23, 13], mais du travail reste à faire avant d'être capable de concevoir une *BCI* capable de classifier l'attention auditive en temps réel afin de réaliser des actions sur un ordinateur. C'est le challenge que s'est fixé l'équipe Athena dans le cadre du projet MICADôme.

2.5 Hypothèses

Sur la base de tous ces principes d'étude de l'attention auditive, nous avons décidé de nous concentrer principalement dans un premier temps sur l'*Auditory Steady-State Response*.

Les hypothèses de recherches sont les suivantes :

- l'écoute d'un son à une fréquence de modulation d'amplitude précise induit une réponse *steady-state* à la même fréquence ;
- lors de l'écoute de sons d'amplitudes modulées à des fréquences différentes, le fait de porter son attention sur un son précis induit une plus forte réponse *steady-state* à sa fréquence de modulation d'amplitude ;
- ce principe peut être utilisé pour classifier l'attention auditive en temps réel.

Chapitre 3

Design expérimental

Ce chapitre va aborder toutes les étapes de la conception de l'expérience que nous avons réalisée, de la définition des objectifs à la réalisation des expérimentations avec des participants.

3.1 Objectif

Afin de vérifier les hypothèses énoncées précédemment, il a fallu élaborer un protocole expérimental, dans lequel il est possible de maîtriser un maximum de paramètres afin d'avoir des résultats rigoureux. Afin de réussir à détecter l'attention auditive uniquement à partir des signaux EEG, il a été décidé de donner des instructions d'attention aux participants pendant qu'ils écoutent différents sons en simultané de telle sorte à recueillir des données labellisées qui seront utilisées pour mettre en évidence des spécificités de l'attention.

Afin de commencer le plus simplement possible, nous avons décidé de débiter par l'écoute de 2 sons en simultané. Le principe général de l'expérience est de faire écouter au participant ces 2 sons en même temps et de lui dire de se concentrer sur un en particulier.

3.2 Matériel

L'équipe Athena a la capacité de réaliser des expériences grâce à une salle d'expérimentation. Elle dispose de tout le nécessaire pour l'acquisition EEG. Plus spécifiquement, pour les expériences réalisées au cours de ce stage, nous avons utilisé un casque EEG ANT Neuro WaveGuard de 32 électrodes fonctionnant avec du gel. Ce casque est relié à un amplificateur ANT Neuro Refa8 avec une référence moyenne. Cela signifie qu'au signal de chaque électrode est soustrait la moyenne de toutes les électrodes. Les signaux sont ensuite récupérés de l'amplificateur par connexion USB à un PC qui permet une acquisition en temps réel grâce au logiciel OpenViBE version 2.1.0. L'acquisition se fait à 2048 Hz.

OpenViBE¹ est un logiciel gratuit et open source développé par Inria pour faire des interfaces cerveau-ordinateur et des neurosciences en temps réel. Il permet de créer des *scénarios* qui permettent sous la forme d'une interface graphique de connecter un client d'acquisition à différents modules de traitement des données ou de gestion d'expériences. Il permet notamment d'afficher des instructions visuelles ainsi que de jouer des sons, tout en conservant les *triggers*, c'est à dire les instants des différents événements par rapport aux signaux EEG. Enfin, il permet d'inclure des script en différents langages de programmation tels que Lua, MATLAB, Python 2 ou C++.

Pour réaliser ces expériences auditives, un casque dynamique de monitoring supra-auriculaire Sennheiser HD 25 fermé est utilisé afin de bien isoler du bruit extérieur. Lors des expériences, celui-ci est positionné de telle sorte à ne pas toucher les électrodes, et aucun câble ne passe dans l'arceau afin de ne pas perturber les signaux avec du bruit électrique.

1. <http://openvibe.inria.fr/>

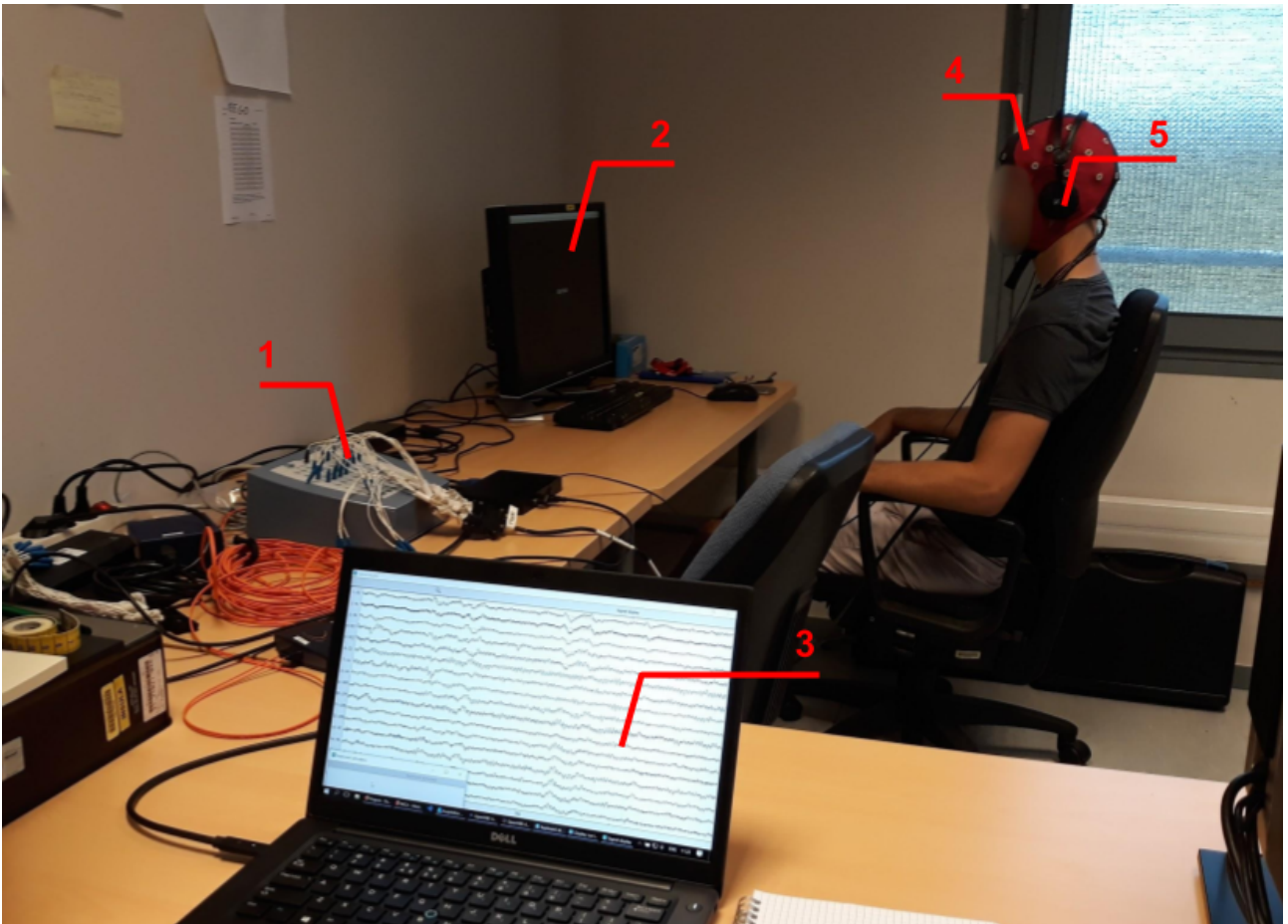


FIGURE 3.1 – Configuration du matériel pour l'expérience

1 : amplificateur ; 2 : écran pour les instructions ; 3 : ordinateur d'acquisition avec *OpenViBE* ;
4 : casque EEG 32 électrodes ; 5 : casque audio

Les sons utilisés pour les expériences sont deux mélodies que nous avons composées avec le logiciel MuseScore version 3.2.3. La première est un pattern de 12 notes de guitare à une fréquence de 2,075 Hz. La seconde est un pattern de 12 notes à 3,3 Hz. Les signaux audios sont présentés Figure 3.2. Ces fréquences ont été choisies de telle sorte que le début des patterns ne sont synchronisés que très rarement, rendant très difficile la tâche de reconnaître une mélodie de concert lorsque les deux sons sont joués en simultanément. De plus, les fréquences de modulation en amplitude sont assez éloignées de fréquences connues qui pourraient parasiter les signaux, comme les dérives lentes en dessous de 1 Hz et les ondes *mu* aux alentours de 10 Hz. Un extrait de la partition composée est disponible en annexe.

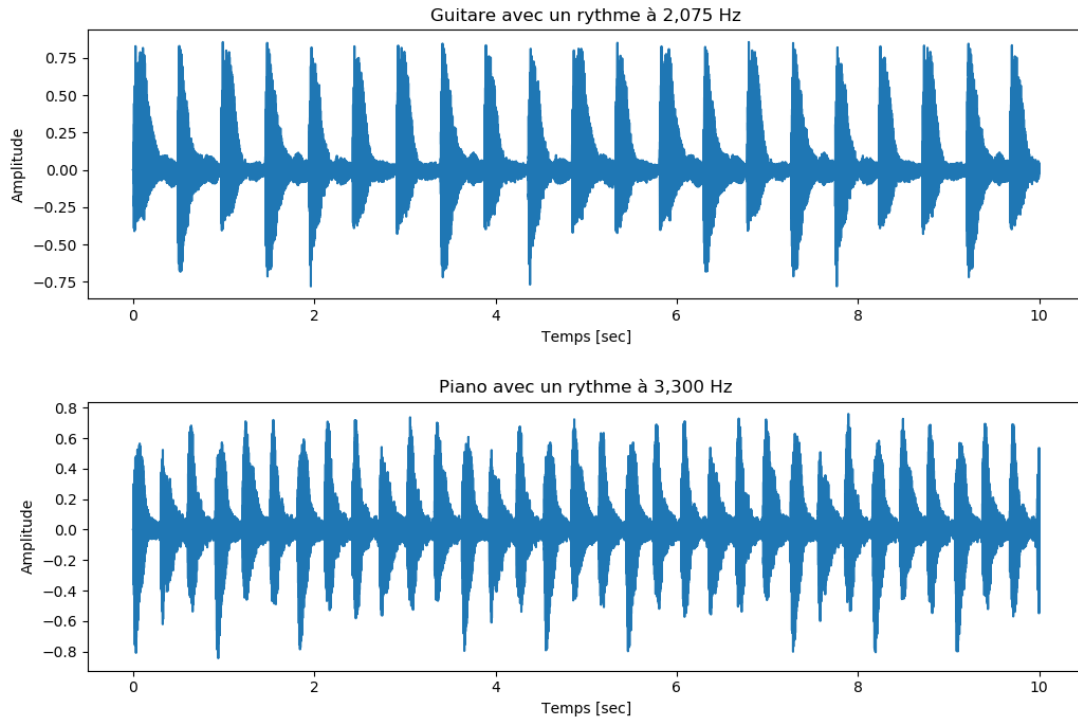


FIGURE 3.2 – Signaux des mélodies : guitare modulé en amplitude à 2,075 Hz et piano modulé en amplitude à 3,300 Hz

3.3 Calibration du volume des sons

Un phase de calibration précédant l'expérience est nécessaire afin d'équilibrer le volume des sons pour chaque participant. En effet, chaque personne a une sensibilité différente aux différentes fréquences et pour chaque oreille. Cette calibration permet au participant d'être capable de porter son attention aussi facilement sur un son que sur l'autre.

L'interface de calibration a été réalisée en utilisant les technologies web HTML, CSS et JavaScript avec l'aide du framework CSS Bootstrap pour être *responsive* (Figure 3.3). La gestion des flux audios a été faite grâce à la Web Audio API.

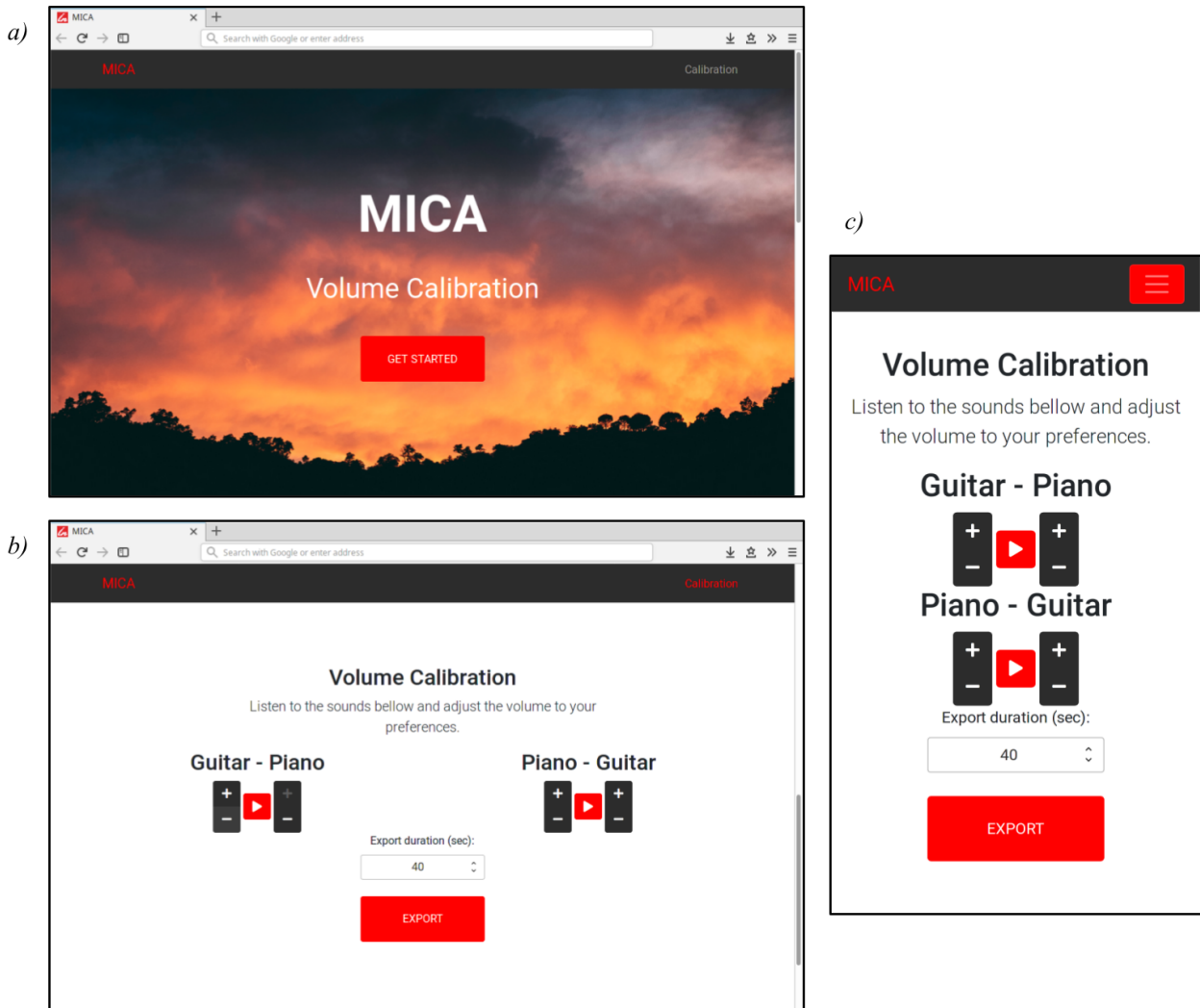


FIGURE 3.3 – Interface de calibration de l'expérience
 a) haut de la page, b) bas de la page, c) affichage en mode mobile

Cette interface permet, lors de l'utilisation d'un casque audio en stéréo, de jouer un son différent de chaque côté et de pouvoir modifier le volume de chacun indépendamment de l'autre. Elle permet ensuite d'exporter les sons mixés au format WAV dans un dossier pour qu'ils soient lus par OpenViBE et utilisés pour l'expérience.

3.4 Protocole expérimental

L'installation des électrodes est plutôt longue et fastidieuse et peut prendre aux alentours d'une demi heure pour 32 électrodes. Il faut en effet introduire grâce à une seringue un gel de composition saline au niveau du contact de chaque électrode avec le scalp afin d'améliorer la conductivité électrique. Afin de simplifier la procédure et ne pas imposer une expérimentation trop longue aux participants, nous avons décidé d'utiliser uniquement les 17 électrodes les plus utiles pour la mesure des ASSR notamment dans les régions frontales et centrales (F, FC et C) comme présenté Figure 3.4.

L'expérience est divisée en 2 blocs identiques, chacun étant composé de 10 tâches ou essais de 40 secondes chacun. Ceci permet de réaliser une même condition plusieurs fois afin d'améliorer la consistance grâce par exemple à un moyennage, mais également d'être dans la possibilité de se passer de certains essais si un problème est survenu pendant celui-ci ou si les signaux sont trop bruités. Au tout début de l'expérience, une *baseline* est mesurée pendant 15 secondes. C'est simplement une mesure de l'activité cérébrale au repos, sans aucune tâche. La durée totale de l'expérience est d'une vingtaine

de minutes, auxquelles il faut ajouter la calibration des volumes et l'installation du dispositif EEG au préalable, d'une durée d'environ 30 minutes.

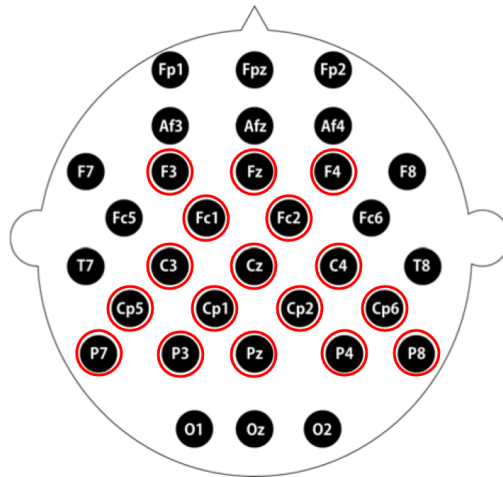


FIGURE 3.4 – Topologie des électrodes utilisées

Chaque bloc est composé des tâches suivantes :

- son de guitare à gauche et piano à droite sans attention pendant 40 sec
- son de piano à gauche et guitare à droite sans attention pendant 40 sec
- son de guitare seul binaural (des deux cotés) pendant 40 sec
- son de guitare à gauche et piano à droite avec attention à gauche pendant 40 sec
- son de piano à gauche et guitare à droite avec attention à droite pendant 40 sec
- son de piano seul binaural (des deux cotés) pendant 40 sec
- son de piano à gauche et guitare à droite avec attention à gauche pendant 40 sec
- son de guitare à gauche et piano à droite avec attention à droite pendant 40 sec
- son de guitare à gauche et piano à droite changement attention pendant 40 sec
 - attention à gauche pendant 10 sec
 - attention à droite pendant 10 sec
 - attention à gauche pendant 10 sec
 - attention à droite pendant 10 sec
- son de piano à gauche et guitare à droite changement attention pendant 40 sec
 - attention à gauche pendant 10 sec
 - attention à droite pendant 10 sec
 - attention à gauche pendant 10 sec
 - attention à droite pendant 10 sec

Les seuls critères d'inclusion des participants étaient de ne pas avoir de problèmes d'audition.

3.5 Implémentation de l'expérience

Le scénario de l'expérience est réalisé en partie avec le *designer* de OpenViBE. Il permet d'utiliser différentes boîtes qui permettent d'accélérer le processus de développement en évitant d'implémenter toute la partie interfaçage avec les périphériques d'entrées et sortie (clavier, amplificateur EEG, carte son, ...). Il se comporte un peu comme une API avec laquelle on peut interagir depuis un script en Lua (langage de programmation impératif, utilisé notamment pour la programmation embarquée). Comme présenté Figure 3.5, le scénario de l'expérience est divisé en deux parties. La partie de gauche gère la connexion au client d'acquisition qui permet de récupérer les signaux et de les enregistrer. Elle permet également de les afficher en temps réel avec des filtres simples pour améliorer la lisibilité. La partie droite correspond à la gestion des tâches et est centrée autour du script Lua qui commande toute l'expérience. Il reçoit les entrées du clavier qui permettent de contrôler l'expérience par exemple pour faire des pauses. Il contrôle en sortie l'affichage des instructions visuelles (Figure 3.6) avec la boîte

nommée *Display cue image*, et l'écoute des différents fichiers sons avec des *Sound Players*.

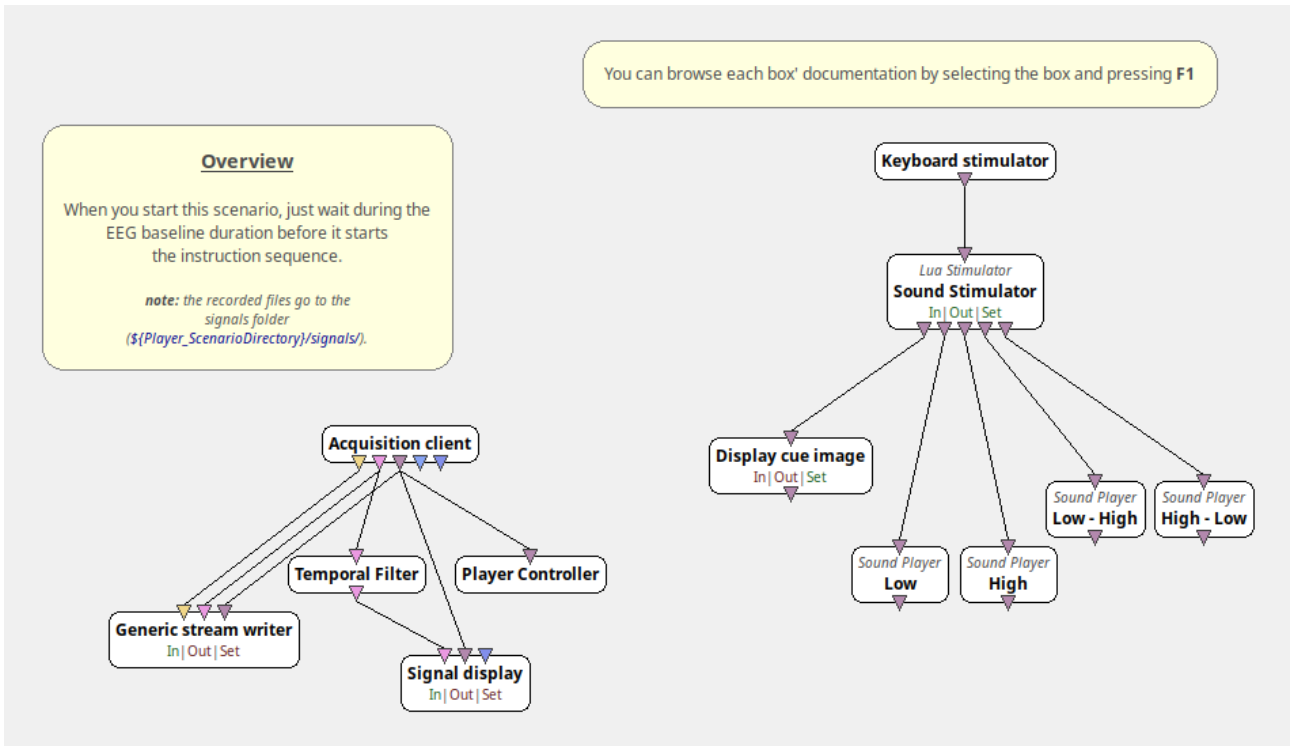


FIGURE 3.5 – Scénario OpenViBE de l'expérience

Les *triggers*, permettant d'annoter les différents événements temporellement par rapport aux signaux EEG, sont envoyés par le script Lua par TCP/IP au client d'acquisition qui s'occupe de les intégrer aux signaux EEG. Les boîtes d'affichages des instructions ainsi que celles permettant de jouer les sons envoient également des *triggers* par TCP/IP lorsqu'elles sont déclenchées.

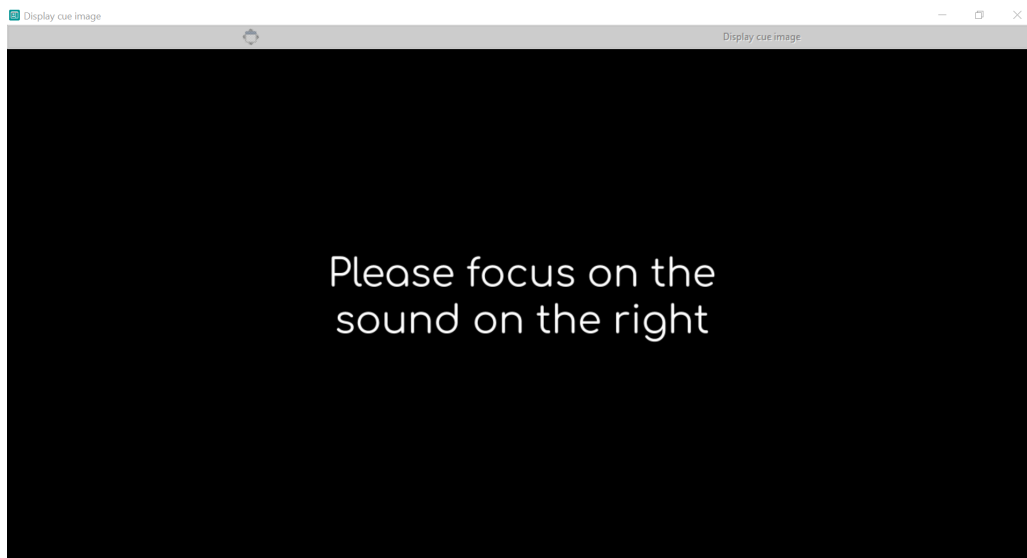


FIGURE 3.6 – Instruction visuelle affichée grâce à OpenViBE (boîte *Display cue image*)

Ce scénario enregistre au final les données EEG avec les *triggers* au format binaire OpenViBE (fichier en *.ov*), que l'on peut ensuite convertir au format GDF afin de procéder aux analyses en langage Python.

3.6 Réalisation des expériences

Les expériences ont donc été réalisées dans une salle d'expérimentation à Inria Sophia Antipolis - Méditerranée. 9 sujets (4 femmes et 5 hommes) ont pris part aux expérimentations et ont réalisé le protocole précédemment décrit. Chaque expérience a duré un peu moins d'une heure.

Chaque participant a au préalable lu et signé un formulaire de consentement éclairé (disponible en annexe) après que le scénario lui ait été expliqué et qu'il ait pu poser des questions à propos de l'expérience.

Chapitre 4

Méthodes d'analyse

À partir des données collectées, le but va être de trouver dans les EEG des caractéristiques qui permettent d'identifier sur quel son le sujet portait son attention. L'analyse des données collectées lors des acquisitions, se décompose en deux phases principales. La première est la préparation des données qui consiste à filtrer les signaux afin d'enlever les bruits parasites qui peuvent être présents en EEG, à synchroniser les différents essais de mêmes conditions, et à supprimer les essais trop bruités et non utilisables. La seconde phase consiste à appliquer différents traitements du signal afin de visualiser les signaux correspondant à chaque condition sous différentes formes.

4.1 Filtrage en fréquence

Le premier traitement à réaliser est le filtrage fréquentiel. Il est très important car il permet d'atténuer des perturbations courantes que l'on peut retrouver en EEG. Ces perturbations incluent notamment, le bruit électrique lié au secteur et les oscillations lentes des signaux également appelées *drifts*.

Lorsque l'on veut filtrer les signaux à des fréquences dont on connaît au préalable les valeurs, il est possible d'utiliser des filtres de Butterworth [5]. Ils permettent d'atténuer les fréquences à partir d'un certain seuil (fréquences de la bande de coupure), tout en conservant les autres fréquences (fréquences la bande passante).

Un filtre passe-bas est ainsi défini comme laissant passer les fréquences inférieures à la fréquence de coupure et atténuant les fréquences supérieures. Un filtre passe-haut laisse quant à lui passer les fréquences au dessus de la fréquence de coupure et atténue les fréquences inférieures. En combinant un filtre passe-haut et passe-bas on peut obtenir un filtre passe-bande qui laisse passer les fréquences dans un intervalles et atténue le reste, ou un filtre coupe-bande qui atténue uniquement un intervalle de fréquences. Le degré d'atténuation ou gain peu varier grâce à l'ordre du filtre : plus l'ordre est élevé, plus l'atténuation est forte dans la bande de coupure.

Un filtre de Butterworth est défini par sa fonction de transfert H qui dépend de l'ordre et de la fréquence de coupure. Elle est définie comme suit :

$$|H_n(j\omega)| = \frac{1}{\sqrt{1 + (\omega/\omega_c)^{2n}}} \quad (4.1)$$

Avec :

H_n la fonction de transfert du filtre d'ordre n ,

ω la fréquence angulaire en radian par seconde, $\omega = 2\pi f$ avec f fréquence en Hz

ω_c la fréquence angulaire de coupure en radian par seconde, $\omega_c = 2\pi f_c$ avec f_c fréquence en Hz.

On peut visualiser le gain en fonction de la fréquence angulaire grâce à un diagramme de bode (Figure 4.1).

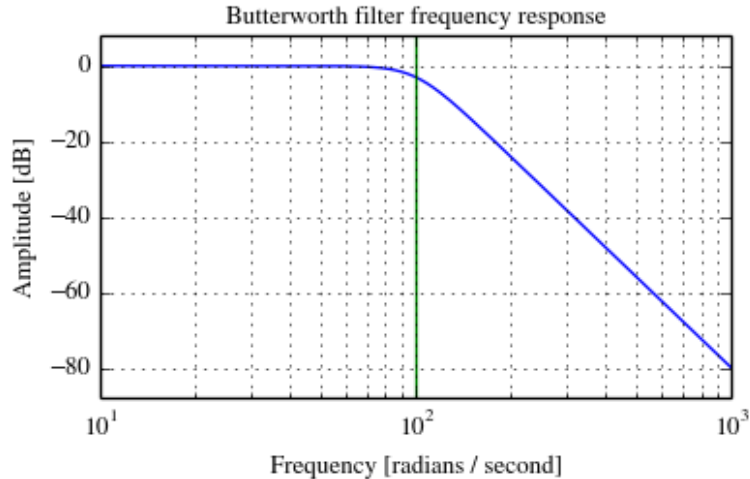


FIGURE 4.1 – Diagramme de bode d’un filtre passe-bas d’ordre 4 avec une fréquence angulaire de coupure de 100 radians par seconde (*Source : docs.scipy.org*)

Dans le cadre de notre expérience, nous avons tout d’abord le bruit électrique lié au secteur qui est un élément perturbateur classique en enregistrement EEG. Il est caractérisé par des oscillations rapides à 50 Hz. Il a ainsi été filtré par un Butterworth coupe-bande d’ordre 4 de 49 à 51 Hz.

Ensuite, il est nécessaire de filtrer les basses fréquences qui sont une simple dérive des signaux appelée *drift*, et dans notre cas nous avons décidé de filtrer également les fréquences au dessus de 9 Hz. En effet, nous nous intéressons ici uniquement aux fréquences 2,075 Hz et 3,3 Hz, et filtrer au dessus de 9 Hz permet d’atténuer les ondes *mu* aux alentours de 10 Hz qui peuvent être présentes car le sujet est assis et immobile. Ceci laisse également une marge pour pouvoir étudier les harmoniques à 4,150 et 6,6 Hz . Pour ces raisons, il a donc été appliqué un filtre passe-bande de 1 à 9 Hz.

4.2 Découpage des essais

Les signaux ont ensuite été découpés en segments appelées *epochs* pour extraire uniquement les moments correspondant aux tâches d’intérêt grâce aux *triggers* laissés par OpenViBE : c’est ce qu’on appelle en faisant un anglicisme l’*epoching*. Un aperçu des *triggers* sur les signaux est présenté Figure 4.2. Ce découpage est fait grâce à la librairie Python MNE¹ (*MEG + EEG Analysis & Visualisation*) qui permet de lire le fichier au format GDF enregistré et de le scanner pour retrouver les triggers spécifiés.

1. <https://mne.tools/stable/index.html>

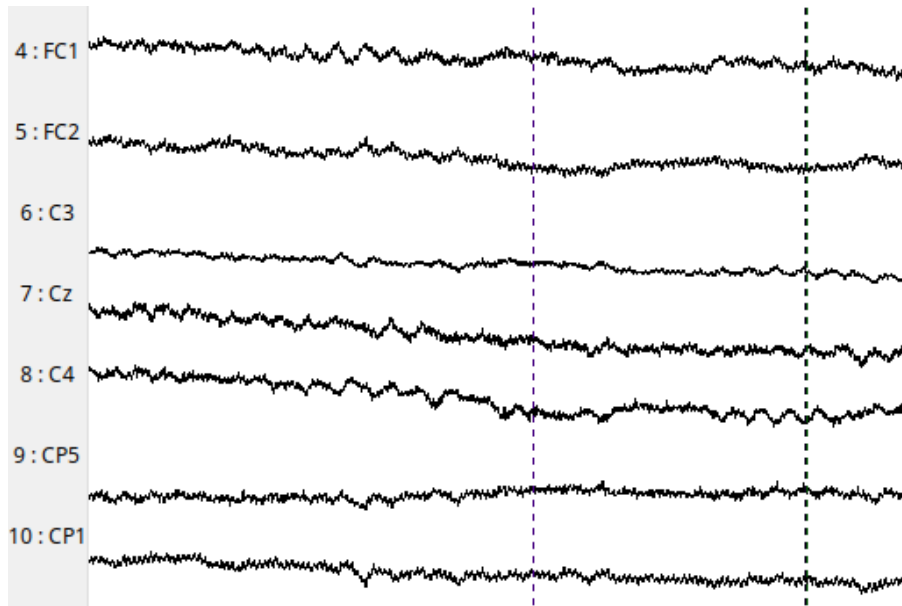


FIGURE 4.2 – Signaux EEG avec les *triggers*

Une fois extraits et synchronisés grâce aux *triggers*, les *epochs* sont analysés pour rejeter ceux qui sont trop bruités. On se base ici sur l'écart entre le minimum et le maximum pour chaque électrode dans un *epoch*. Si cet écart n'est pas compris entre la moyenne des écarts de tous les *epochs* plus ou moins trois fois l'écart type, l'*epoch* est rejeté.

Enfin, les *epochs* restants sont moyennés parmi tous les sujets par conditions ce qui permet d'améliorer le rapport signal sur bruit. Cela donne au final 40 secondes de signal moyenné pour chaque électrode et pour chaque condition.

4.3 Visualisation des données

Pour visualiser les résultats de cette expérience d'*Auditory Steady-State Response*, il est nécessaire de faire des analyses fréquentielles.

Les analyses en fréquences se basent sur le principe que toute onde peut être décomposée en une somme de sinusoides : c'est le principe des séries de Fourier.

La transformation de Fourier est une méthode qui permet d'obtenir le spectre de densité d'une fonction intégrable. Elle est donnée par la fonction suivante :

$$\mathcal{F}(f) : \omega \mapsto \hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt \quad (4.2)$$

Avec :

f fonction intégrable,

t en secondes,

ω la fréquence angulaire (ou pulsation) en radian par secondes.

On peut ainsi grâce à cette méthode déterminer quelles fréquences sont prédominantes dans une onde plus complexe. Ce principe est donc très intéressant dans le cadre de l'analyse de l'*ASSR* pour détecter la densité de fréquence dans les tracés EEG à la fréquence de stimulation.

Bien que la transformation de Fourier soit une méthode très utile, l'intégration d'une fonction de manière continue est impossible avec un ordinateur, qui ne travaille qu'avec des données discrètes. Une solution est donc d'utiliser une transformation de Fourier discrète. Elle est ainsi calculée avec une certaine résolution. Un des algorithmes qui permet de calculer cette transformation de Fourier discrète sur ordinateur de manière rapide est l'algorithme de transformation de Fourier rapide ou *FFT* pour *Fast Fourier Transform* [8]. Il présente l'avantage d'avoir une complexité en $O(n \log(n))$ alors qu'une

transformation de Fourier discrète simple a une complexité en $O(n^2)$. Cet algorithme est très facile d'utilisation car il est implémenté en Python 3 dans la librairie SciPy².

Une autre méthode qui permet d'obtenir la densité des fréquences d'une onde est l'estimation spectrale avec la méthode de Welch [37]. Elle consiste à découper le signal en segments se chevauchant partiellement et à appliquer une transformation de Fourier rapide sur chacun. Ces transformations de Fourier vont ensuite être moyennées pour obtenir l'estimation spectrale. L'avantage de cette technique est qu'elle permet d'atténuer le bruit lié à des fréquences aléatoires qui ne sont pas présentes tout au long du signal. Cependant, cette réduction du bruit se fait aux dépens de la résolution en fréquence. En effet, plus le segment est long, plus la résolution sera importante. Cette méthode est également implémentée dans la librairie SciPy.

Les analyses sont donc faites avec la méthode de Welch sur les *epochs* de 40 secondes moyennés par classe. Une fenêtre de 20 secondes est utilisée avec un recouvrement de 10 secondes.

Les sons sont également analysés afin de s'assurer que les modulations sont aux bonnes fréquences. Pour cela on extrait l'enveloppe du signal grâce à une transformée de Hilbert. On réalise ensuite une transformation de Fourier rapide (*FFT*) sur cette enveloppe pour étudier la densité spectrale.

2. <https://www.scipy.org/>

Chapitre 5

Résultats

Dans ce chapitre vont être présentés les résultats des analyses des sons ainsi que des signaux EEG.

5.1 Analyse des sons

Dans un premier temps, l'analyse des signaux audios a été faite. Les résultats de la transformation de Fourier de l'enveloppe sont présentés Figure 5.1. On y remarque que les sons présentent bien les fréquences qui ont été établies en terme de rythme. En effet, la mélodie de guitare présente sa plus grande densité de fréquence aux alentours de 2,1 Hz, alors que la mélodie au piano présente sa plus grande densité aux alentours de 3,3 Hz.

Ce qui est également intéressant est que ces sons présentent également des harmoniques avec des densités élevées à des multiples de la fréquence primaire. Au plus ce multiple est grand, au plus la densité est faible. Lorsque l'on s'intéresse à la bande de fréquence de 1 à 9 Hz, on retrouve pour la guitare modulée en amplitude à 2,075 Hz des harmoniques de densité décroissante à 4,150 Hz, 6,225 Hz, et 8,3 Hz. Pour le piano modulé en amplitude à 3,3 Hz, on retrouve des harmoniques à 6,6 Hz.

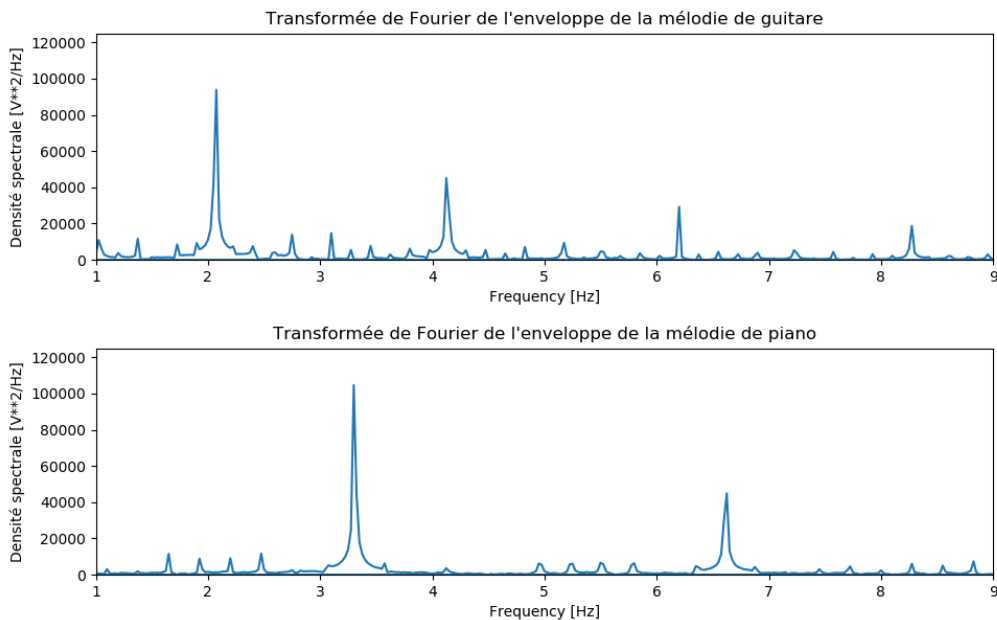


FIGURE 5.1 – Transformées de Fourier des enveloppes des mélodies de guitare et de piano

5.2 ASSR pour les sons seuls en binaural

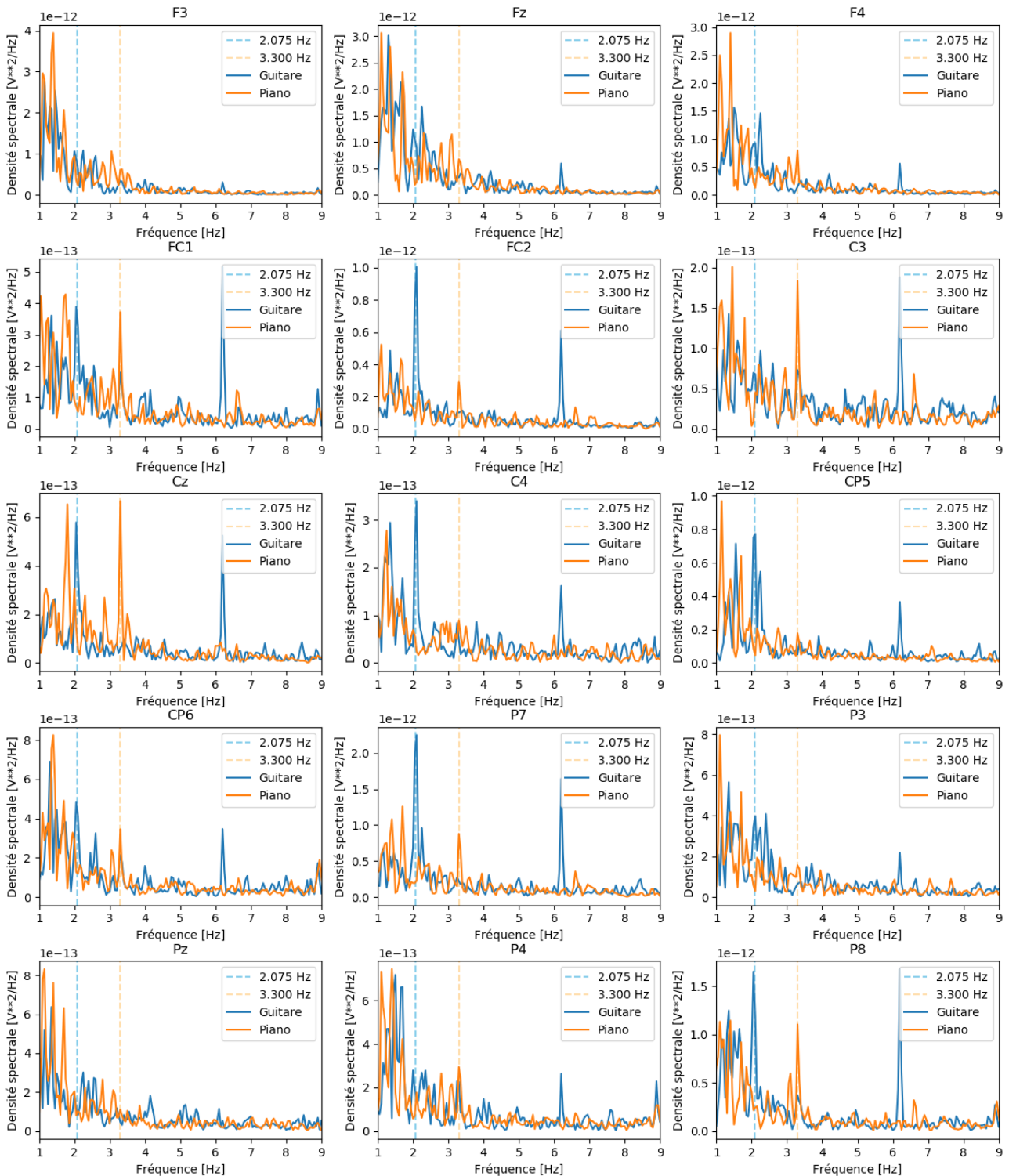


FIGURE 5.2 – Estimation spectrale avec la méthode de Welch pour chaque électrode (électrodes moyennées parmi tous les sujets)

Guitare : écoute de la mélodie de guitare au rythme de 2,075 Hz

Piano : écoute de la mélodie de piano au rythme de 3,300 Hz

L'analyse des ASSRs en réponse à l'écoute d'un son seul donne des résultats intéressants présentés Figure 5.2. Grâce à une estimation spectrale avec la méthode de Welch, on retrouve en effet sur plusieurs électrodes des pics de densité aux fréquences correspondant à celle du rythme des mélodies.

La mélodie de guitare à un rythme de 2,075 Hz donne de très bonnes réponses *steady-state* notam-

ment sur les électrodes fronto-centrales (FC1 et FC2), sur les électrodes centrales (Cz et C4), sur les électrodes centro-pariétales (CP5 et CP6), ainsi que les électrodes pariétales P7 et P8. Les harmoniques à 6,225 Hz semblent également donner une réponse *ASSR* importante sur plusieurs électrodes dont plus particulièrement FC1 et FC2, les électrodes centrales Cz, C3 et C4, ainsi que P7 et P8.

La mélodie au piano à un rythme de 3,3 Hz donne elle aussi des *ASSRs* très marquées sur certaines électrodes. Ceci est visible sur les électrodes fronto-centrales (FC1 et FC2), les électrodes centrales avec Cz et C3, en centro-pariétal sur CP6, ainsi que sur P4, P7 et P8 en pariétal. Les harmoniques à 6,6 Hz sont cependant moins visibles que pour la guitare même si on remarque des densités fréquentielles plus élevées à cette fréquence sur FC1 et C3.

5.3 *ASSR* pour l'attention sur un son parmi deux

Les *ASSRs* obtenues lors des tâches attentionnelles sont cependant plus mitigées. Les estimations spectrales avec la méthode de Welch sont présentées Figure 5.3. Il est en effet très difficile de tirer des conclusions quant à l'attention auditive à partir de ces estimations spectrales. En effet, des pics de densité correspondant aux fréquences des rythmes des mélodies sont présent sur certaines électrodes, au niveau frontal et central. Ceci concerne en particulier les électrodes FC1, FC2, Cz, C3 ou C4, et ce sont des régions qui correspondent à ce que l'on peut trouver dans la littérature en terme d'*ASSR*. Il ne semble cependant pas y avoir de discrimination claire entre les différentes conditions. En effet, la condition sans tâche d'attention a été choisie comme témoin ici et celle-ci ne semble pas présenter des densités de fréquences particulièrement plus faibles que lors des tâches d'attention.

Les harmoniques sont aussi présentes aux fréquences 6,225 Hz et 6,6 Hz mais là encore, il ne semble pas possible de discriminer clairement les différentes conditions.

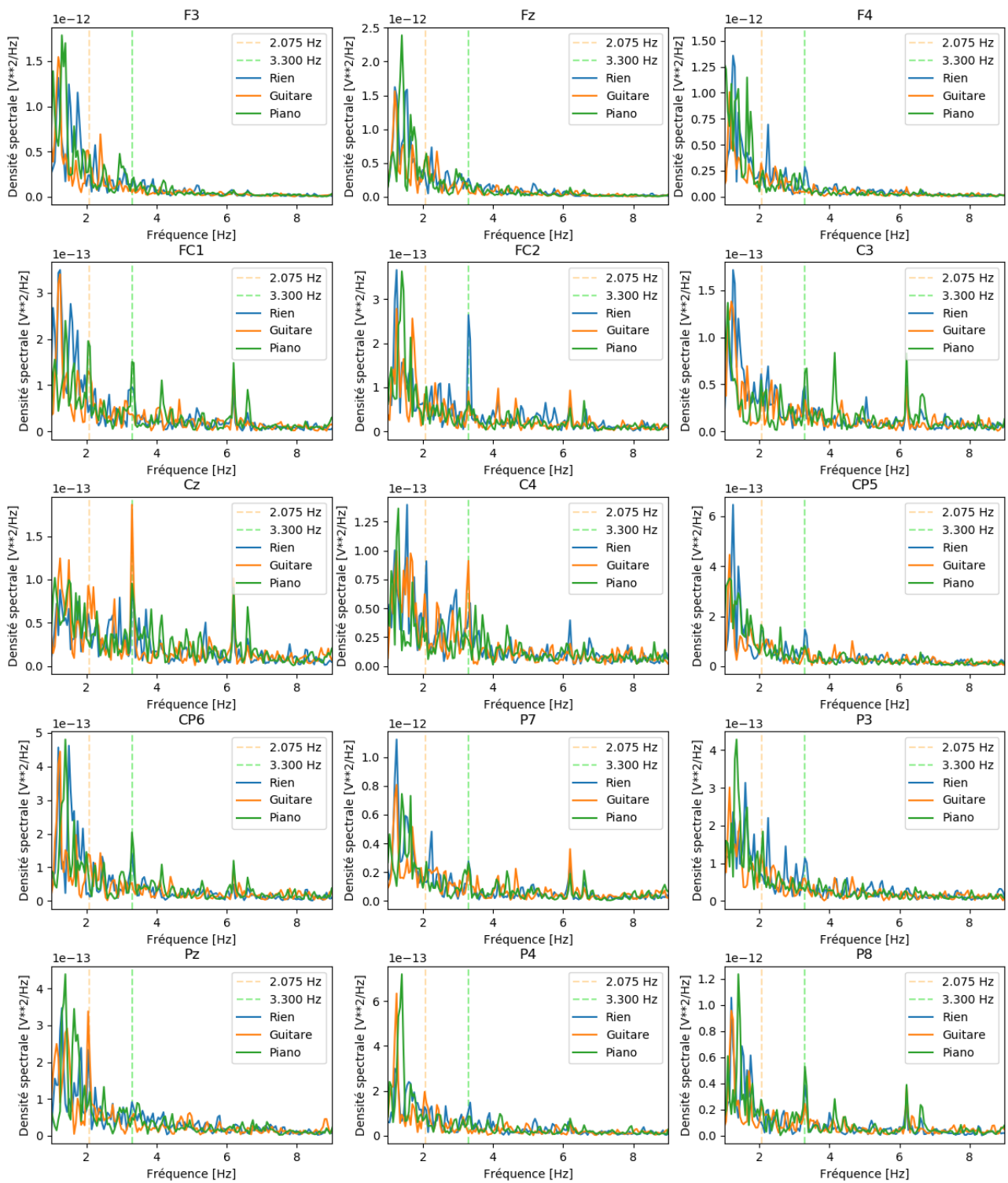


FIGURE 5.3 – Estimation spectrale avec la méthode de Welch pour chaque électrode (électrodes moyennées parmi tous les sujets)

Guitare : écoute des deux sons avec attention sur la guitare au rythme de 2,075 Hz

Piano : écoute des deux sons avec attention sur le piano au rythme de 3,300 Hz

Rien : écoute des deux sons sans tâche d'attention

Chapitre 6

Discussion

La discussion va permettre d'aborder différents points concernant les choix au niveau du protocole mais également de commenter les résultats obtenus.

6.1 Choix des stimuli audios

Le choix des stimuli audios a été l'une des principales difficultés lors de ce stage et a nécessité plusieurs ajustements.

Nous avons d'abord commencé par des sons purs à des fréquences de 111 Hz et 1000 Hz modulés en amplitudes avec des enveloppes sinusoïdales à respectivement 9,8 Hz et 38,2 Hz. Cependant, ces stimuli ne permettaient pas d'obtenir d'ASSR en EEG. Une explication possible est que les fréquences de modulations étaient trop élevées si bien qu'il était impossible de se concentrer sur le rythme. De plus les modulations sinusoïdales faisaient que le silence entre chaque note était infiniment court ce qui rendait la distinction du rythme très compliquée.

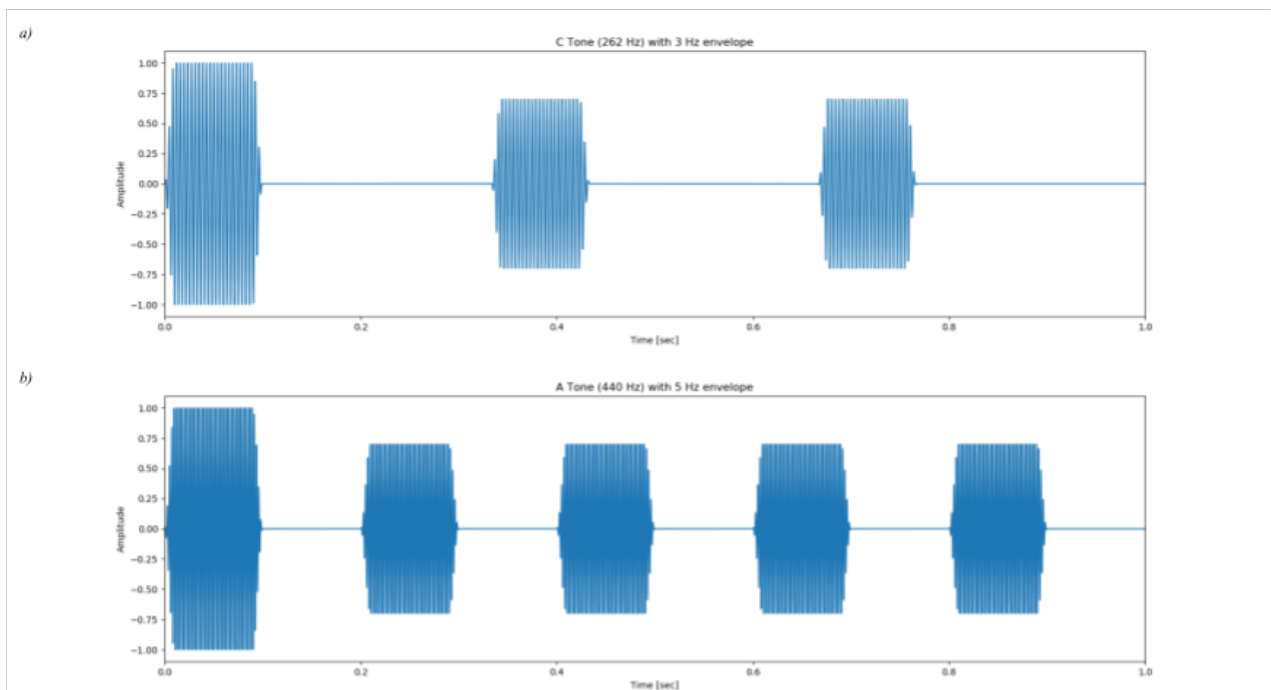


FIGURE 6.1 – Signaux audios de la première expérience
a) Do modulé à 3 Hz, b) La modulé à 5 Hz

De nouveaux sons ont donc été utilisés. Cette fois-ci étaient utilisés des sons à 262 Hz (note Do) et 440 Hz (note La) modulés en amplitudes avec une enveloppe en créneaux à respectivement 3 et 5

Hz (Figure 6.1). Ces sons ont permis une légère amélioration dans l'apparition des *ASSRs* mais les réponses ne ressortaient toujours pas clairement en EEG. Après un retour des sujets pilotes, nous avons réalisé que le fait d'entendre toujours la même note était rébarbatif et produisait un sentiment d'ennui qui allait à l'encontre de la tâche de concentration qui était demandée.

Nous avons donc décidé au final de composer nos propres mélodies avec des notes dont l'attaque est plus importante pour encore améliorer la sensibilité aux stimuli [18], et à des fréquences définies comme décrit dans le chapitre chapitre 3.

6.2 Périphérique audio et environnement

Un des doutes qui pourrait planer sur l'expérience est le fait qu'avoir un casque électro-acoustique à côté des électrodes pourrait perturber les signaux et même fausser les résultats.

Toutefois, des expériences ont été également réalisées avec des hauts-parleurs de qualité professionnelle au CIRM à Nice et des résultats similaires ont été retrouvés.

Le choix d'un casque audio été fait car il permet d'isoler correctement des bruits ambiants qu'il peut y avoir à l'extérieur de la salle d'expérience. L'idéal serait en effet d'avoir une pièce insonorisée. Une session d'expérience a donc été réalisée dans une chambre anechoïque à l'Hôpital de la Timone à Marseille dans l'équipe de Daniele Schön. Bien que cette configuration améliorerait le confort lors de l'expérience, le fait que les autres expériences aient été réalisées dans une salle un peu plus bruyante ne semble pas détériorer les résultats, l'essentiel étant que le sujet puisse se concentrer sur les sons de manière convenable.

6.3 Interface de calibration

L'interface de calibration a évolué au cours du temps afin d'implémenter une application qui soit la plus fonctionnelle et facile à utiliser.

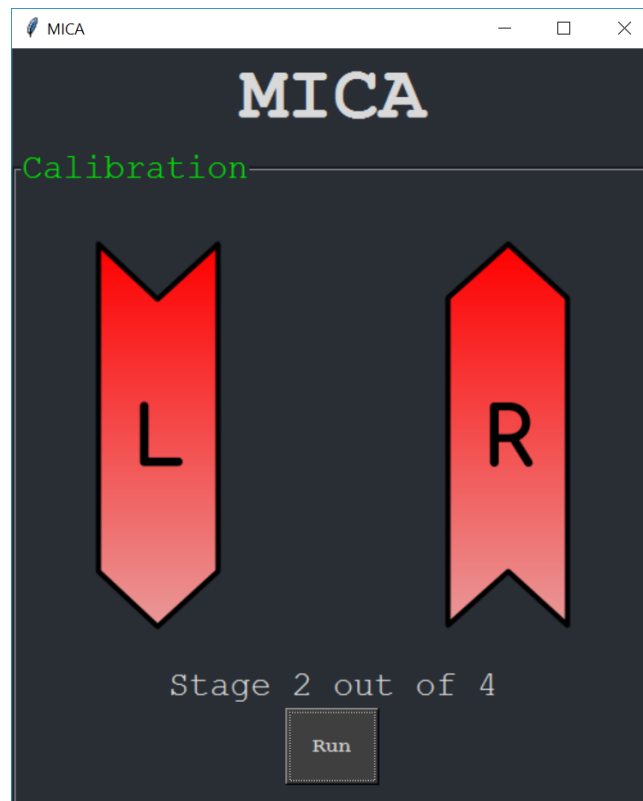


FIGURE 6.2 – Première version de l'interface de calibration en Python 3

Afin de rendre la tâche ludique, la première version jouait les deux différents sons en en baissant

un petit à petit tout en augmentant l'autre. Le sujet devait alors appuyer sur une touche lorsqu'il percevait les sons au même volume. Ce processus était répété plusieurs fois et les valeurs des volumes étaient moyennées pour au final exporter les sons.

Cette première version était à l'origine implémentée en Python 3 avec la librairie Tkinter (Figure 6.2). Ce choix avait été fait car ce langage est très adapté au traitement du signal et qu'il permettait de créer des sons sinusoïdaux modulés en amplitude facilement. Le volume était ensuite modifié grâce à la librairie Pydub¹.

Cependant au fil de l'avancement et de l'évolution du protocole, les sons sinusoïdaux s'étant transformés en mélodies, l'avantage qu'avait Python pour la production de sons n'entraînait plus en considération. De plus cette interface souffrait de plusieurs défauts. Premièrement, l'interface faite avec Tkinter était austère et ne fournissait que peu de souplesse. Ensuite, l'exécution de manière synchrone n'était pas évidente et optimisée avec Python. De plus, le système nécessitant un appui sur une touche à l'équilibre rendait la calibration longue et nécessitait de tout recommencer si l'on souhaitait modifier un son. Enfin, les librairies pour jouer les sons et gérer les frappes de clavier n'étant pas les mêmes sous Windows et sous Linux, le code devenait complexe si on souhaitait pouvoir l'exécuter sous ces deux systèmes d'exploitation.

Pour toutes ces raisons, le choix a été fait de réimplémenter l'interface en JavaScript. Cela a permis d'une part d'avoir une interface plus esthétique et adaptable à la taille d'écran (*responsive design*) avec l'utilisation d'un framework CSS. Ensuite, le JavaScript gère très bien l'exécution synchrone avec par exemple le système des *promises*. De plus l'application est exécutable sous n'importe quel système d'exploitation car elle fonctionne dans un navigateur web. Cela a donc au final donné l'interface qui est présentée en Figure 3.3.

6.4 ASSR et détection de l'attention auditive

Les résultats obtenus en ce qui concerne les *ASSRs* en réponse à des sons seuls binauraux (écoutés des deux oreilles) vont dans le sens de ce qui est retrouvé dans la littérature [27, 32]. Ils confortent dans l'idée que le protocole final et plus spécifiquement les stimuli auditifs sont adaptés pour produire de telles réponses.

Cependant, la partie avec attention donne des résultats qui semblent a priori peu satisfaisants pour réaliser notre objectif de classifier l'attention auditive en temps réel. Plusieurs facteurs peuvent expliquer ces résultats.

Tout d'abord, la durée de l'expérience fait qu'il est difficile de se concentrer du début à la fin, même si des moyens ont été mis en oeuvre du côté des stimuli pour les rendre moins ennuyeux.

Il se pose également le problème de l'étiquetage des données : si un sujet perd le fil et change son attention sur l'autre son, il n'est pas possible de le savoir autrement qu'avec des questionnaires après chaque essai. Cette procédure ne sera cependant que peu précise. Il serait peut être nécessaire de décomposer plus précisément les segments de 40 secondes afin de représenter la fréquence en fonction du temps afin de voir si les densités de fréquences évoluent au cours d'un essai.

Enfin les analyses réalisées sont plutôt simples et ne permettent peut être pas d'identifier des motifs dans l'EEG qui seraient caractéristiques de l'attention. Du travail doit encore être effectué afin de classifier ces données d'attention.

6.5 Pistes de recherche futures

Le rendu de ce rapport ne marquant pas la fin de ce stage, le temps restant va être consacré à implémenter des algorithmes d'apprentissage automatique afin de classifier les signaux pour savoir sur quel son le sujet se concentre.

Une piste actuellement explorée est l'analyse canonique des corrélations [15]. Elle permet à partir d'un signal cible, sorte de modèle, de classifier les signaux en fonction de leur corrélation à celui-ci.

1. <http://pydub.com/>

Les signaux cibles seront par exemple les signaux EEG dans les conditions d'écoute de chaque son seul qui permettront d'avoir des références. Ces références serviront pour classifier les signaux lors de l'écoute de plusieurs sons avec attention sur un de ceux référencé pour les signaux cibles. Une autre option pour les signaux cibles pourra être d'utiliser directement les signaux sonores de chaque son seul étant donné que leurs fréquences de modulations doivent normalement être celles des *ASSRs*.

Le but est à terme de pouvoir faire cette classification en temps réel, afin de concevoir une véritable interface cerveau-ordinateur qui donnera un *feedback* sur l'attention auditive en temps réel.

Une piste à explorer également dans le cadre cette fois-ci d'une autre expérience sera l'étude du paradigme P300-oddball auditif. Dans une scène auditive avec plusieurs bruits non-cible, l'attente d'un son cible entraîne, à sa survenue, une onde positive 300 ms après : c'est le P300 [7]. Ce type de paradigme est d'ores et déjà utilisé dans le cadre d'interfaces cerveau-ordinateur, notamment pour composer de la musique d'une manière innovante [30].

Chapitre 7

Conclusion

Au cours de ce stage, j'ai trouvé très formateur d'effectuer un travail de recherche sur une période étendue en partant de la naissance d'un projet. Il était ainsi très intéressant de faire prendre forme au projet, de la création du protocole expérimental jusqu'à l'analyse des résultats.

J'ai apprécié avoir une évolution par itération, proche d'une méthodologie *agile* dans laquelle l'implémentation d'une version de l'expérience était suivie de sa réalisation avec 2 ou 3 sujets pilotes qui donnaient un retour permettant d'améliorer le protocole lors de l'itération suivante. C'est au total 26 expériences qui ont été réalisées, dont 16 avec le protocole final.

J'ai également eu l'occasion de travailler en équipe dans un environnement stimulant intellectuellement. Ce travail en équipe s'est notamment traduit par l'utilisation d'outils de versionnage collaboratifs comme Git et GitLab. Il s'est également traduit par le partage et l'échange à travers des meetings et des workshops, comme l'UCA Deep Learning School 2019 (conférences et workshops sur l'apprentissage automatique), ou le meeting NeuroMod 2019 à Fréjus (rencontre autour des neurosciences et de la cognition) où j'ai pu faire une présentation (disponible en annexe). Il était également très intéressant de collaborer avec des spécialistes de la musique qui ont apporté une note de créativité au travail que j'ai réalisé.

J'ai pu au cours de ce stage bénéficier d'une grande autonomie, ce qui a été très formateur pour préparer à la réalisation d'une thèse.

Le travail réalisé au cours de ce stage donne des résultats qui semblent prometteurs même si beaucoup de challenges restent à surmonter afin de pouvoir classifier l'attention auditive en temps réel. Les applications du projet MICADôme s'annoncent cependant très intéressantes.

La détection de l'attention auditive grâce à une interface cerveau-ordinateur pourrait être utilisée afin de concevoir des manières ludiques d'apprendre à identifier les instruments d'un morceau de musique. On peut également penser à la conception de jeux auditifs dans lesquels la concentration sur un son réaliserait une action. Enfin, une application qui pourrait voir le jour dans le cadre de ce projet est la conception d'une interface cerveau-ordinateur permettant de composer de la musique en 3D. Cette interface permettrait d'identifier un son parmi un scène auditive sur lequel le compositeur porte son attention et de le sélectionner. Une interface tactile permettrait alors de déplacer ce son en 3D dans le dôme de haut-parleurs.

Platon disait :

La musique donne une âme à nos coeurs et des ailes à la pensée.

Ce projet pourrait cependant contribuer à donner des ailes à la musique grâce à la pensée à travers la technologie.

Remerciements

Je tiens à remercier Maureen Clerc, mon encadrante, pour m'avoir donné l'opportunité de réaliser ce stage, et pour m'avoir permis d'être autonome dans mon travail de recherche tout en étant de très bon conseil et en me guidant lorsque j'en avais besoin.

Merci à Daniele Schön d'avoir été d'une aide précieuse pour surmonter les difficultés liées aux stimuli auditifs utilisés pour les expériences.

Merci à Michel Pascal et Camille Giuglaris pour leurs suggestions intéressantes sur le plan musical.

Merci tout spécifiquement à Federica Turi, pour m'avoir transmis des connaissances pratiques mais également théoriques sur les expérimentations en EEG et le traitement du signal.

Je suis reconnaissant d'avoir passé ces 6 mois de stage dans l'équipe Athena dirigée par Rachid Deriche dont tous les membres sont très aimables et sympathique, participant à une bonne ambiance de travail.

Merci également aux encadrants du Master Sciences Cognitives de l'IDMC, notamment Manuel Rebuschi et Armelle Brun pour avoir oeuvré afin de rendre cette formation très intéressante et diversifiée, ce qui m'a permis de réaliser deux stages exceptionnels en Master 1 et en Master 2.

Enfin, je remercie Anne Guérin-Dugué, mon encadrante de stage de Licence, pour m'avoir transmis sa passion pour les neurosciences computationnelles et le traitement du signal.

Annexes

Annexe A : Partition des mélodies

$\text{♩} = 180$

Piano

Acoustic Guitar

8

2

Pno.

Guit.

8

3

Pno.

Guit.

8

Annexe B :
Formulaire de consentement éclairé



FORMULAIRE D'INFORMATION ET DE RECUEIL DE CONSENTEMENT ECLAIRE

Dans le cadre de leurs recherches, les chercheurs Inria sont amenés à demander à des personnes de participer à des études, nécessaires pour l'avancement de leurs projets de recherches.

Toute participation de personnes à une étude doit être précédée, d'une part de la transmission d'une note d'information et, d'autre part de la signature d'un formulaire de consentement éclairé.

Ce document présente le formulaire d'information et de recueil de consentement éclairé que le COERLE recommande de faire signer aux personnes qui participent à des recherches menées par les chercheurs Inria.

PARTIE II/ INFORMATIONS COMMUNIQUEES AUX PARTICIPANTS

Remerciements :

Votre collaboration est essentielle à la réalisation de notre projet et l'équipe de recherche tient à vous en remercier.

Avant de donner votre consentement éclairé, il est important que vous compreniez pourquoi la recherche est faite, ce qu'elle implique et quels sont vos droits et obligations.

PRESENTATION DU PROJET

IDENTIFICATION

Responsables du projet : Maureen Clerc et Théodore Papadopoulo

Equipe-projet et CRI : Athena, Inria Sophia Antipolis

Nom du projet : Test d'une interface cerveau-ordinateur

Vous êtes invités à prendre part à ce projet dont la finalité est de vérifier le fonctionnement d'une interface cerveau-ordinateur.

Les modalités de réalisation du projet seront les suivantes :

L'électroencéphalographie (EEG) est un dispositif permettant de mesurer le potentiel électrique simultanément en plusieurs points à la surface de la tête. Pour ceci, un bonnet semi-rigide est placé sur la tête du sujet, et un gel conducteur est appliqué afin de faire le contact électrique entre la peau et les électrodes. Cet examen ne comporte aucun risque dans l'état actuel des connaissances.

Les variations de potentiel électrique de l'EEG proviennent essentiellement de l'activité électrique cérébrale.

Une interface cerveau-machine extrait de ces signaux des informations qui seront transmises en temps réel à un dispositif extérieur dans l'objectif de le contrôler (exemple: le contrôle du mouvement d'un curseur sur un écran).

Les bénéfices attendus du projet sont les suivants : recueillir des données d'utilisation de l'interface.

Ce projet a été autorisé par le Comité Opérationnel d'Evaluation des risques légaux et éthiques d'Inria.

Pour de plus amples informations :

<http://www.inria.fr/institut/organisation/instances/coerle/presentation-du-coerle>

VOTRE PARTICIPATION AU PROJET

Modalités de votre participation : Votre participation est prévue pour une unique séance d'au plus deux heures. Un bonnet comportant une vingtaine d'électrodes sera positionné sur votre tête. Il faudra rester immobile durant l'utilisation de l'interface, et des pauses seront prévues où vous pourrez vous détendre sans ôter le bonnet EEG.

L'EEG sera mesuré au repos, et pendant des tâches où vous recevrez des stimulations sensorielles (visuelles, auditives ou tactiles). Il vous sera demandé de répondre à certaines questions des expérimentateurs au sujet de vos intentions de communication ou de votre ressenti sur l'interface cerveau ordinateur.

Votre participation à cette étude ne comporte aucun risque particulier.

Votre consentement ne décharge pas les organisateurs de la recherche de leurs responsabilités. Vous conservez tous vos droits garantis par la loi.

Participation volontaire au projet :

Après avoir lu et compris les informations contenues dans la partie « Présentation du projet », votre signature de ce formulaire indique que vous acceptez de participer au projet, et ce sans aucune contrainte ou pression extérieure.

S'il vous manque des informations nécessaires à votre prise de décision, n'hésitez pas à demander un complément d'informations aux personnes contacts (voir ci-après).

Retrait du projet à tout moment :

Vous êtes totalement libre de mettre fin à votre participation au projet à n'importe quel moment. Il vous suffit d'en informer les personnes contacts (ci-après). Vous n'aurez aucune justification à apporter à votre retrait et vous n'encourez aucune responsabilité.

Suite de votre retrait du projet :

L'équipe de recherche pourra continuer à utiliser les renseignements vous concernant jusqu'au retrait de votre consentement, sauf si vous vous y opposez en l'indiquant aux personnes contacts. Dans ce cas, les renseignements vous concernant seront détruits.

Demande d'informations au cours du projet :

Si vous avez des questions additionnelles sur le projet ou sur vos droits en tant que participant à ces travaux de recherche, n'hésitez pas à en faire part aux personnes contacts mentionnées ci-après.

Information concernant les résultats principaux de la recherche

Si vous souhaitez obtenir un résumé écrit des principaux résultats de cette recherche ainsi que de leur utilisation, veuillez en faire la demande aux personnes contacts.

CONFIDENTIALITE ET LOI INFORMATIQUE & LIBERTES

Inria est très attentif et vigilant concernant les questions de protection des données à caractère personnel (DCP), il s'est d'ailleurs doté d'un Délégué à la Protection des Données (DPD) qui est chargé de s'assurer de la conformité d'Inria avec le Règlement européen sur la Protection des Données (RGPD).

Ces données sont les suivantes : aucune

Ces données seront traitées de la manière suivante : sans objet

PERSONNES CONTACTS

Nom : Clerc
Prénom : Maureen
Qualité : Directrice de Recherche
N° de téléphone : 04 92 38 77 35
e-mail : maureen.clerc@inria.fr

Nom : Papadopoulo
Prénom : Théodore
Qualité : Directeur de Recherche
N° de téléphone : 04 92 38 76 01
e-mail : theodore.papadopoulo@inria.fr

PARTIE II/ RECUEIL DU CONSENTEMENT ECLAIRE DU PARTICIPANT

Je soussigné(e)reconnais avoir lu et compris le présent formulaire d'informations concernant le projetet consens volontairement à participer à ce projet de recherche.

Je reconnais avoir été informé que je peux mettre fin à ma participation au projet à tout moment sans aucun motif à donner, en contactant simplement les personnes contacts référencées ci-dessus.

Je reconnais que j'ai été informé sur le but et les modalités du projet, ainsi que les avantages, risques et inconvénients du projet.

Je reconnais avoir obtenu les réponses mes questions de manière claire et précise.

Je reconnais enfin avoir la possibilité de poser des questions additionnelles sur le projet ou sur mes droits, à tout moment pendant le projet, en communiquant avec les personnes contacts.

Fait en deux exemplaires dont un sera remis au participant et un sera conservé par le responsable du projet.

Pour le participant :
Date et signature

Pour le responsable du projet :
Date et signature :

Annexe C :
Support de présentation pour le meeting
NeuroMod 2019 à Fréjus

Measuring auditory attention with EEG

Johann Benerradi (*Université de Lorraine, Inria*)

Maureen Clerc (*Inria*)

NeuroMod, 1-2 July 2019 Fréjus



UNIVERSITÉ
CÔTE D'AZUR

Measuring auditory attention with EEG

Outline

- Introduction
- Tests
- Experimental design
- Future work
- Conclusion

Introduction

3

Introduction

MICA project

*(Musique Interactive Côte d'Azur)
Led by Michel Pascal (Conservatoire à
Rayonnement Régional, Nice)*

Interface between art and digital science

Use neural activity as an input modality for
music creation

Immersive experience with a sound dome



4

Introduction

Objectives:

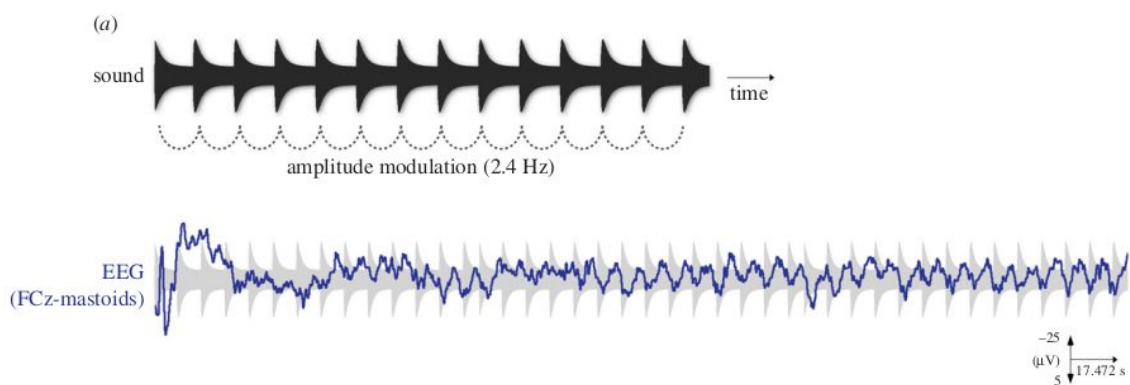
- Detect auditory attention based on EEG activity
- Design a Brain Computer Interface which could enable interaction with music

5

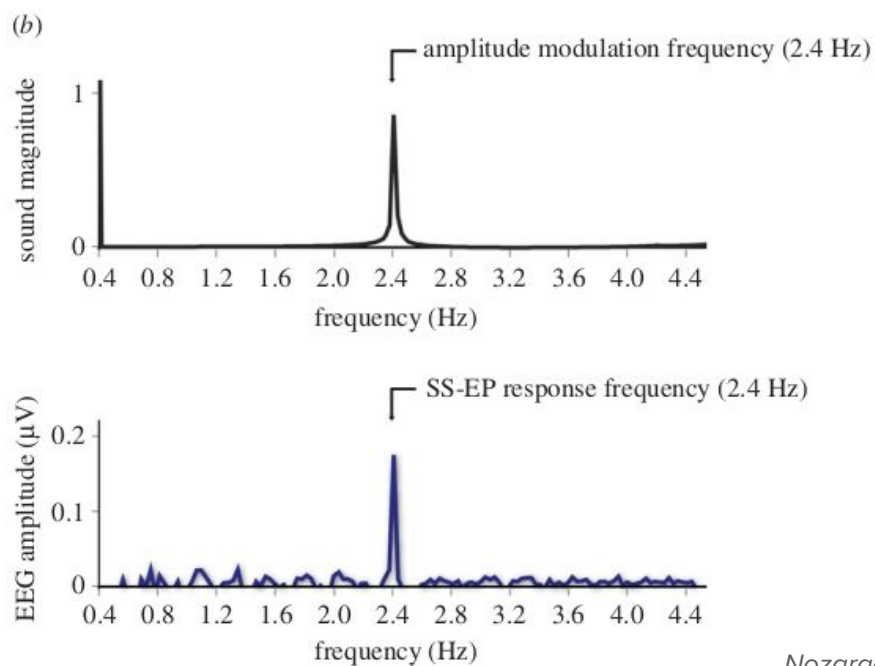
Introduction

Auditory Steady-State Response (ASSR) paradigm:

Synchronization of the EEG brain activity to the frequency of a periodically repeated auditory stimulus



6



Nozaradan 2014

7

Introduction

Auditory Steady-State Response (ASSR) paradigm:

John et al. 1998: ASSR with amplitude modulated sounds

Detection of rhythm:

Nozaradan et al. 2011: ASSR with beat and meter (binary and ternary)

Schaefer et al. 2011: ASSR with self-imposed rhythmic patterns

8

Introduction

Assumptions:

- The spectral density of EEG is higher at frequencies corresponding to the modulation frequencies of sounds
- The spectral density of EEG is higher at the frequency of the modulation of the sound of attention
- It is possible to classify the auditory attention based on those principles in real-time

A sound sample...

Tests

11

First tests

Binaural sound with amplitude modulated at 3.33 Hz during 100 sec

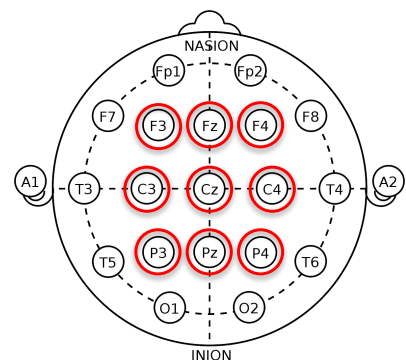
EEG acquisition with 9 gel-based electrodes thanks to OpenViBE

ANT Refa8 amplifcator with average reference

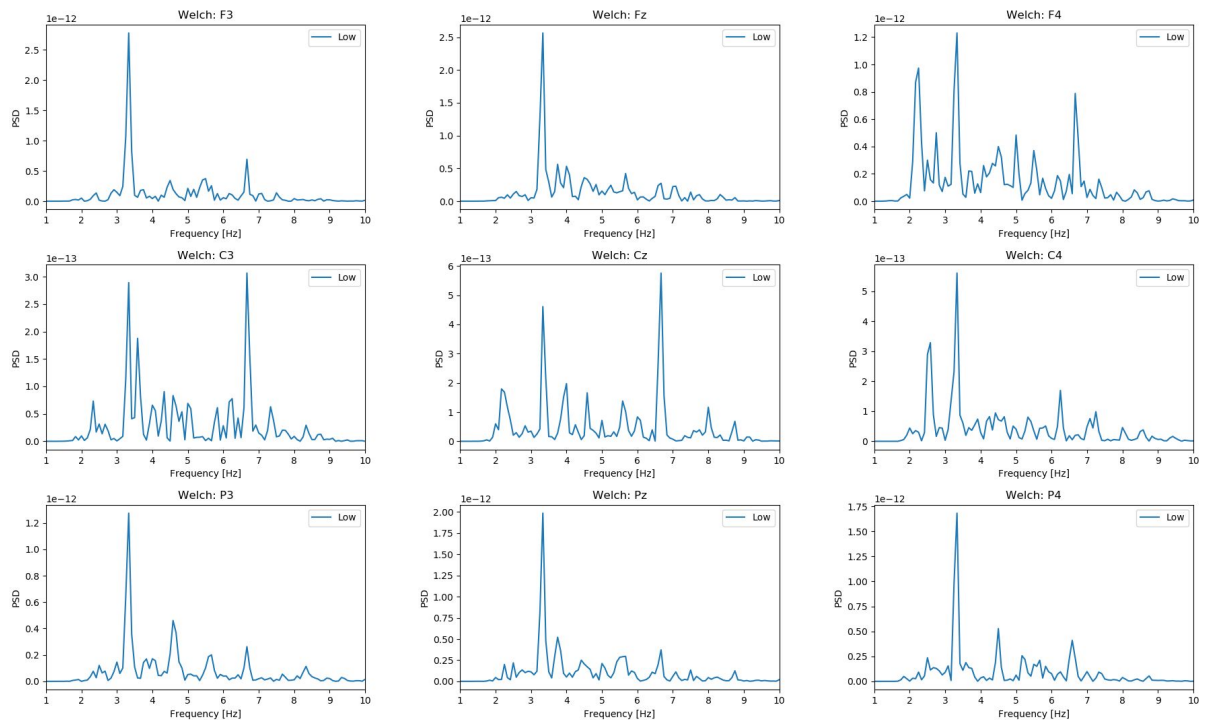
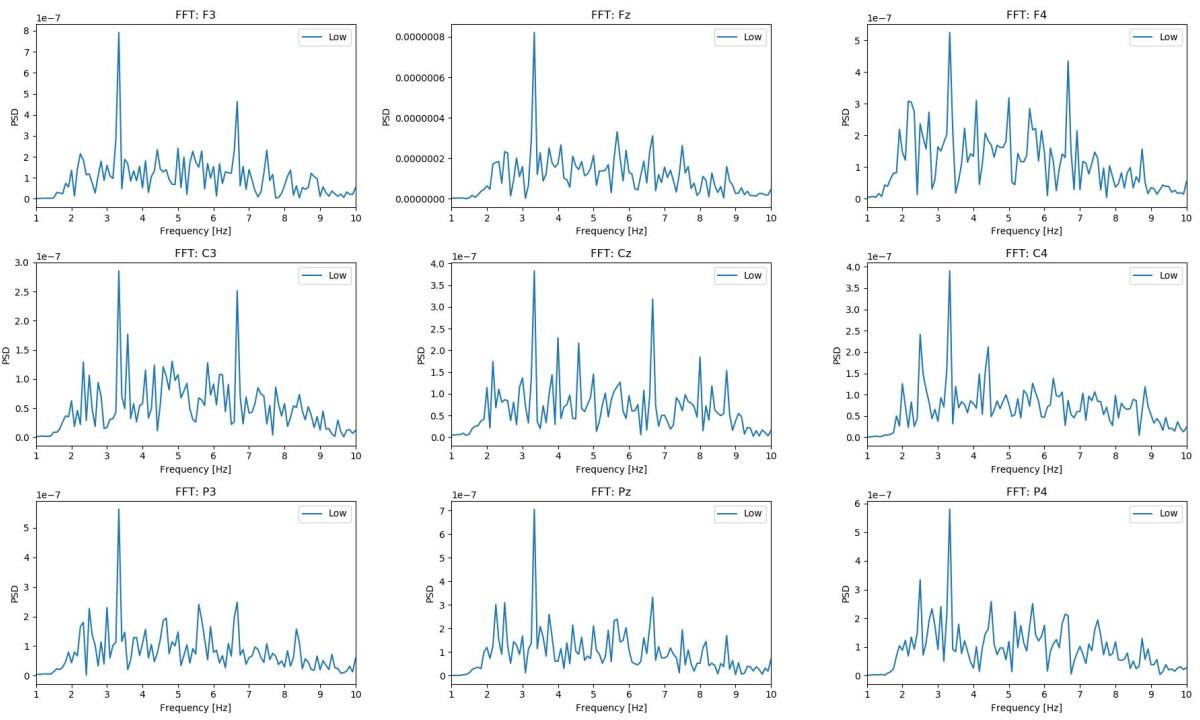
Analysis:

- band-pass Butterworth filter 2 to 9 Hz
- Fast Fourier transform
- Welch's method (12 sec sliding window)

*Special thanks to Daniele Schön
(Institut de Neurosciences des Systèmes, Marseille)*



12

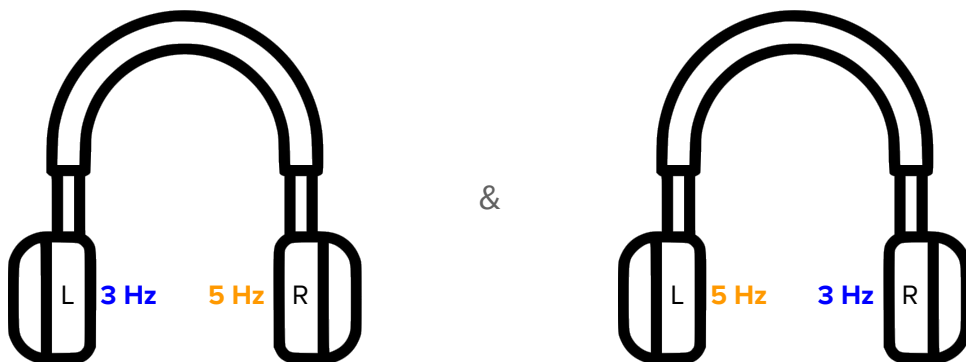


Experimental design

15

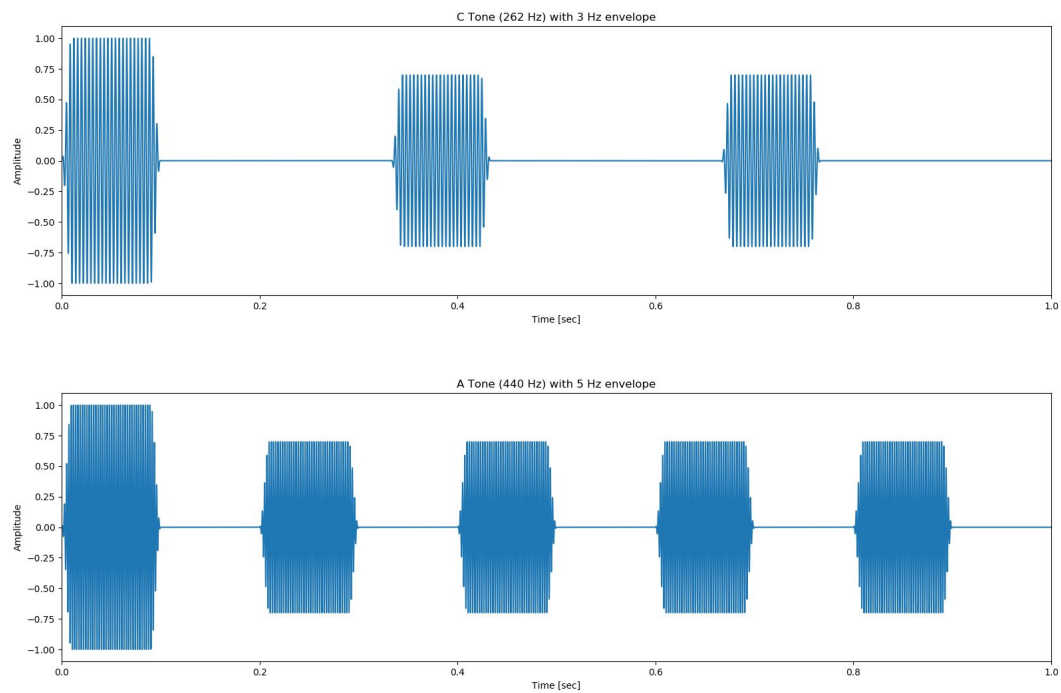
Experimental design

Generate 2 amplitude modulated sounds (Python): 3 Hz and 5 Hz



Volume calibration to have a balanced perception

16



17

Experimental design

EEG acquisition with 15 gel-based electrodes thanks to OpenViBE

ANT Refa8 amplifier with average reference

Conditions:

- single sound alone
- different sound on each side
- different sound on each side with attention task

18

Future work

19

Future work

Finish data collection

Do more work on the analysis to be able to differentiate attention on one sound among the two

From offline to real-time processing → Brain Computer Interface

P300 oddball paradigm:

→ a different and identifiable sound awaited by the listener leads to a specific EEG response 300 ms after it (*Shuai et al. 2014*)

20

Conclusion

21

Conclusion

Auditory Steady-State Response → attention to rhythmic patterns

P300 oddball → attention to salient musical effects

Promising perspectives of such Brain Computer Interfaces:

- Create unique music compositions controlled by the listener
- Create playful ways of teaching music
- Create sound based games
- ...

22

Thank you for your attention

johann.benerradi@inria.fr

Bibliographie

- [1] Alexandra Bendixen. Predictability effects in auditory scene analysis : a review. *Frontiers in neuroscience*, 8 :60, 2014.
- [2] Jonathan Benghozi. Classification des phases de sommeil. <https://support.dreem.com/hc/fr/articles/360019948151-Classification-des-phases-de-sommeil>.
- [3] Karla Blocka. Eeg (electroencephalogram). <https://www.healthline.com/health/eeg>.
- [4] Jeffrey W Britton, Lauren C Frey, JL Hopp, P Korb, MZ Koubeissi, WE Lievens, EM Pestana-Knight, and EK Louis St. *Electroencephalography (EEG) : An introductory text and atlas of normal and abnormal findings in adults, children, and infants*. American Epilepsy Society, Chicago, 2016.
- [5] Stephen Butterworth et al. On the theory of filter amplifiers. *Wireless Engineer*, 7(6) :536–541, 1930.
- [6] E Colin Cherry. Some experiments on the recognition of speech, with one and with two ears. *The Journal of the acoustical society of America*, 25(5) :975–979, 1953.
- [7] Maureen Clerc, Laurent Bougrain, and Fabien Lotte. *Brain-Computer Interfaces 1 : Methods and Perspectives*. John Wiley & Sons, 2016.
- [8] James W Cooley and John W Tukey. An algorithm for the machine calculation of complex fourier series. *Mathematics of computation*, 19(90) :297–301, 1965.
- [9] Susann Deike, Susan L Denham, and Elyse Sussman. Probing auditory scene analysis. *Frontiers in neuroscience*, 8 :293, 2014.
- [10] Siyi Deng, Ramesh Srinivasan, Tom Lappas, and Michael D’Zmura. Eeg classification of imagined syllable rhythm using hilbert spectrum methods. *Journal of neural engineering*, 7(4) :046006, 2010.
- [11] Lena-Vanessa Dolležal, André Brechmann, Georg M Klump, and Susann Deike. Evaluating auditory stream segregation of sam tone sequences by subjective and objective psychoacoustical tasks, and brain activity. *Frontiers in neuroscience*, 8 :119, 2014.
- [12] Institut du Cerveau et de la Moelle épinière. Cervelet. <https://icm-institute.org/fr/actualite/comprendre-le-cerveau-et-son-fonctionnement/>.
- [13] Mike Grierson, Chris Kiefer, and Matthew Yee-King. Progress report on the eavi bci toolkit for music : Musical applications of algorithms for use with consumer brain computer interfaces. 2011.
- [14] Christoph Guger, Gunther Krausz, Brendan Z Allison, and Guenter Edlinger. Comparison of dry and gel based electrodes for p300 brain–computer interfaces. *Frontiers in neuroscience*, 6 :60, 2012.
- [15] Harold Hotelling. Relations between two sets of variates. In *Breakthroughs in statistics*, pages 162–190. Springer, 1992.
- [16] Inria. Des balbutiements de l’informatique au règne du tout numérique. <https://www.inria.fr/institut/inria-en-bref/histoire-d-inria>.
- [17] M Sasha John, Andrew Dimitrijevic, and Terence W Picton. Auditory steady-state responses to exponential modulation envelopes. *Ear and hearing*, 23(2) :106–117, 2002.
- [18] Michael S John, Otavio G Lins, Brigitte L Boucher, and Terence W Picton. Multiple auditory steady-state responses (master) : stimulus and recording parameters. *Audiology*, 37(2) :59–82, 1998.

- [19] Hirohito M Kondo, Iwaki Toshima, Daniel Pressnitzer, and Makio Kashino. Probing the time course of head-motion cues integration during auditory scene analysis. *Frontiers in neuroscience*, 8 :170, 2014.
- [20] Emmanuelle Kristensen. *Méthodologie de traitement conjoint des signaux EEG et oculométriques : applications aux tâches d’exploration visuelle libre*. PhD thesis, Université Grenoble Alpes, 2017.
- [21] Fabien Lotte, Laurent Bougrain, Andrzej Cichocki, Maureen Clerc, Marco Congedo, Alain Rakotomamonjy, and Florian Yger. A review of classification algorithms for eeg-based brain–computer interfaces : a 10 year update. *Journal of neural engineering*, 15(3) :031005, 2018.
- [22] Steven J Luck. *An introduction to the event-related potential technique*. MIT press, 2014.
- [23] Tim R Mullen, Richard Warp, and Adam Jansch. Minding the (transatlantic) gap : An internet-enabled acoustic brain-computer music interface. In *NIME*, pages 469–472. Citeseer, 2011.
- [24] Muse. A deep dive into brainwaves : Brainwave frequencies explained. <https://choosemuse.com/fr/blog/a-deep-dive-into-brainwaves-brainwave-frequencies-explained-2/>.
- [25] César Noronha. A brief introduction to eeg and the types of electrodes. <https://www.brainlatam.com/blog/a-brief-introduction-to-eeg-and-the-types-of-electrodes-75>.
- [26] Sylvie Nozaradan. Exploring how musical rhythm entrains brain activity with electroencephalogram frequency-tagging. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 369(1658) :20130393, 2014.
- [27] Sylvie Nozaradan, Isabelle Peretz, Marcus Missal, and André Mouraux. Tagging the neuronal entrainment to beat and meter. *Journal of Neuroscience*, 31(28) :10234–10240, 2011.
- [28] The University of Queensland. Lobes of the brain. <https://qbi.uq.edu.au/brain/brain-anatomy/lobes-brain>.
- [29] OpenViBE. Openvibe, software for brain computer interfaces and real time neurosciences. <http://openvibe.inria.fr/>.
- [30] Andreas Pinegger, Hannah Hiebel, Selina C Wriessnegger, and Gernot R Müller-Putz. Composing only by thought : Novel application of the p300 brain-computer interface. *PloS one*, 12(9) :e0181584, 2017.
- [31] David Regan. Human brain electrophysiology. *Evoked potentials and evoked magnetic fields in science and medicine*, 1989.
- [32] Rebecca S Schaefer, Rutger J Vlek, and Peter Desain. Decomposing rhythm processing : Electroencephalography of perceived and self-imposed rhythmic patterns. *Psychological research*, 75(2) :95–106, 2011.
- [33] Futura Sciences. Comprendre le cerveau et son fonctionnement. <https://www.futura-sciences.com/sante/definitions/biologie-cervelet-13719/>.
- [34] Lan Shuai and Mounya Elhilali. Task-dependent neural representations of salient events in dynamic auditory scenes. *Frontiers in neuroscience*, 8 :203, 2014.
- [35] Ramesh Srinivasan, Samuel Thorpe, Siyi Deng, Tom Lappas, and Michael D’Zmura. Decoding attentional orientation from eeg spectra. In *International Conference on Human-Computer Interaction*, pages 176–183. Springer, 2009.
- [36] Jonathan Sussman-Fort and Elyse Sussman. The effect of stimulus context on the buildup to stream segregation. *Frontiers in neuroscience*, 8 :93, 2014.
- [37] Peter Welch. The use of fast fourier transform for the estimation of power spectra : a method based on time averaging over short, modified periodograms. *IEEE Transactions on audio and electroacoustics*, 15(2) :70–73, 1967.