

Université de Montréal

**Profil temporel de l'efficacité du traitement visuel en
reconnaissance d'objets et de visages**

par Roxanne Ferrandez

Département de psychologie, Université de Montréal
Faculté des arts et des sciences

Mémoire présenté
en vue de l'obtention du grade de Maîtrise
en psychologie

août 2019

© Roxanne Ferrandez, 2019

Université de Montréal
Département de psychologie, Faculté des arts et des sciences

Ce mémoire intitulé

Profil temporel de l'efficacité du traitement visuel en reconnaissance d'objets et de visages

Présenté par

Roxanne Ferrandez

A été évalué par un jury composé des personnes suivantes

Gregory West
Président-rapporteur

Martin Arguin
Directeur de recherche

Franco Lepore
Membre du jury

Résumé

Les variations d'efficacité du traitement visuel dans le temps ont été étudiées par échantillonnage temporel aléatoire. Vingt-quatre adultes ont identifié des stimuli composés de bruit blanc visuel et d'images d'objets familiers (expérience 1) ou de visages célèbres (expérience 2). Le ratio signal-bruit variait à travers le temps selon des fonctions d'échantillonnage générées par l'intégration d'ondes sinusoïdales de différentes fréquences (5 à 55 Hz) et de phases et amplitudes aléatoires. Des vecteurs de classification (VC) temporels ont été calculés en soustrayant la somme pondérée des ratios signal-bruit associés aux mauvaises réponses de celle associée aux bonnes réponses. Des images de classification (IC) temps-fréquence ont été obtenues en appliquant la même procédure aux résultats d'analyses temps-fréquence réalisées sur la fonction d'échantillonnage de chaque essai. Les VC temporels des deux expériences sont très variables entre les participants. Par contre, les IC temps-fréquence sont remarquablement similaires à travers les participants (cohérence inter-sujets de .93 et .57 pour l'expérience 1 et 2 respectivement). Des comparaisons par test t nous indiquent de nombreuses différences entre les IC temps-fréquence des objets et visages familiers, mais aussi des objets non familiers et des mots analysés dans des études précédentes. Ainsi, ces IC sont sensibles à la classe de stimuli présentés, mais aussi à la familiarité de ces derniers. Les résultats témoignent d'une variation rapide dans l'efficacité de l'encodage visuel durant les 200 premières millisecondes d'exposition au stimulus et suggèrent que les IC du domaine temps-fréquence reflètent un aspect hautement fondamental du traitement visuel, hypothétiquement rattaché aux oscillations cérébrales.

Mots-clés : traitement visuel, échantillonnage temporel, objets familiers, visages familiers, perception, oscillations cérébrales

Abstract

Variations in visual processing effectiveness through time were investigated using random temporal stimulus sampling. Twenty-four adults named photographs of either familiar objects (experiment 1) or famous faces (experiment 2). Stimuli were made by a linear combination of the target image and high density white visual noise. Signal-to-noise ratio varied throughout the 200 ms stimulus duration. A new temporal sampling function was generated on each trial by the integration of random amplitude and phase sinusoidal waves of frequency between 5 and 55 Hz (in 5 Hz steps). Temporal classification vectors (CV) were calculated by subtracting the weighted sum of the signal-to-noise ratio associated to errors from that associated to correct responses. Time-frequency classification images (CI) were obtained by applying the same procedure on the outcome of time-frequency analyses applied to the sampling functions of each trial. In both experiments, the temporal CVs were highly variable across participants, but the time-frequency CIs were remarkably similar across participants (inter-subject coherence of .93 and .57 for experiments 1 and 2 respectively). T-tests revealed multiple differences between the time-frequency CIs obtained with familiar objects and faces, but also with non-familiar objects and words analyzed in previous studies. Therefore, these CIs are sensitive to stimulus type, but also to stimulus familiarity. The present results indicate rapid variations of visual encoding effectiveness in the initial 200 ms of stimulus exposure and suggests that the time-frequency CIs tap a highly fundamental aspect of visual processing, hypothetically linked to brain oscillations.

Keywords : visual processing, temporal sampling, familiar objects, familiar faces, perception, brain oscillations

Table des matières

RÉSUMÉ.....	II
ABSTRACT.....	III
TABLE DES MATIÈRES	IV
LISTE DES FIGURES.....	VI
LISTE DES SIGLES.....	IX
INTRODUCTION	1
MISE EN CONTEXTE	2
OSCILLATION ET PERCEPTION	2
TECHNIQUE D'ÉCHANTILLONNAGE TEMPOREL.....	4
OBJECTIFS.....	8
HYPOTHÈSES.....	10
EXPÉRIENCE 1	11
MÉTHODOLOGIE	11
<i>Participants.....</i>	11
<i>Stimuli.....</i>	11
<i>Matériel.....</i>	12
<i>Procédures.....</i>	12
<i>Analyses statistiques.....</i>	14
RÉSULTATS.....	16
<i>Information générale.....</i>	16
<i>Vecteur de classification temporel.....</i>	17
<i>Image de classification temps-fréquence.....</i>	18
<i>Image de classification phase-amplitude.....</i>	20
DISCUSSION.....	20
EXPÉRIENCE 2	26
MÉTHODOLOGIE	26
<i>Participants.....</i>	26
<i>Stimuli.....</i>	26
<i>Matériel.....</i>	27

<i>Procédures</i>	27
<i>Analyses statistiques</i>	28
RÉSULTATS.....	28
<i>Information générale</i>	28
<i>Vecteur de classification temporel</i>	28
<i>Image de classification temps-fréquence</i>	29
<i>Image de classification phase-amplitude</i>	31
DISCUSSION.....	31
COMPARAISON ENTRE TYPES DE STIMULI	36
MÉTHODOLOGIE	36
<i>Analyses statistiques</i>	36
RÉSULTATS.....	36
<i>Vecteurs de classification temporels</i>	36
<i>Images de classification temps-fréquence</i>	37
<i>Images de classification phase-amplitude</i>	46
DISCUSSION.....	46
DISCUSSION GÉNÉRALE	54
CONCLUSION	59
RÉFÉRENCES	I
ANNEXE A	XI
ANNEXE B	XII
ANNEXE C	XIII

Liste des figures

- Figure 1.* Vecteur de classification temporel moyen représentant les variations de l'efficacité du traitement visuel à travers le temps en reconnaissance d'objets familiers (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau). 18
- Figure 2.* Image de classification temps-fréquence moyenne représentant les variations de l'efficacité du traitement visuel à travers le temps et les fréquences temporelles en reconnaissance d'objets familiers (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau). 19
- Figure 3.* Image de classification moyenne dans le domaine phase-amplitude représentant les données utilisées pour créer les fonctions d'échantillonnage de chaque essai lors de l'expérience sur les objets (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau). 20
- Figure 4.* Vecteur de classification temporel moyen représentant les variations de l'efficacité du traitement visuel à travers le temps en reconnaissance de visages familiers (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau). 29
- Figure 5.* Image de classification temps-fréquence moyenne représentant les variations de l'efficacité du traitement visuel à travers le temps et les fréquences temporelles en reconnaissance de visages familiers (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau). 30
- Figure 6.* Image de classification moyenne dans le domaine phase-amplitude représentant les données utilisées pour créer les fonctions d'échantillonnage de chaque essai lors de l'expérience sur les visages (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau). 31
- Figure 7.* Profils temporels moyens bruts des mots (bleu), des objets familiers (jaune), des objets non familiers (vert) et des visages familiers (mauve). 37
- Figure 8.* Comparaison statistique entre les profils temps-fréquence des objets familiers et non familiers; image de classification temps-fréquence moyenne brute des objets familiers (à gauche), image de classification temps-fréquence moyenne brute des objets non familiers (au

milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les objets familiers qu’avec les objets non familiers, et un * blanc lorsque la différence significative est en direction opposée. 39

Figure 9. Comparaison statistique entre les profils temps-fréquence des objets et des visages familiers ; image de classification temps-fréquence moyenne brute des objets familiers (à gauche), image de classification temps-fréquence moyenne brute des visages familiers (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les objets familiers qu’avec les visages familiers, et un * blanc lorsque la différence significative est en direction opposée. 40

Figure 10. Comparaison statistique entre les profils temps-fréquence des visages familiers et des objets non familiers; image de classification temps-fréquence moyenne brute des visages familiers (à gauche), image de classification temps-fréquence moyenne brute des objets non familiers (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les visages familiers qu’avec les objets non familiers, et un * blanc lorsque la différence significative est en direction opposée..... 42

Figure 11. Comparaison statistique entre les profils temps-fréquence des objets familiers et des mots; image de classification temps-fréquence moyenne brute des objets familiers (à gauche), image de classification temps-fréquence moyenne brute des mots (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les objets familiers qu’avec les mots, et un * blanc lorsque la différence significative est en direction opposée. À noter ; l’étendue des scores Z pour les images de classification est de -50 à 80 afin de mieux représenter la différence visuellement. 43

Figure 12. Comparaison statistique entre les profils temps-fréquence des visages familiers et des mots; image de classification temps-fréquence moyenne brute des visages familiers (à gauche), image de classification temps-fréquence moyenne brute des mots (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure

avec les visages familiers qu'avec les mots, et un * blanc lorsque la différence significative est en direction opposée. À noter ; l'étendue des scores Z pour les images de classification est de -50 à 80 afin de mieux représenter la différence visuellement. 44

Figure 13. Comparaison statistique entre les profils temps-fréquence des mots et des objets non familiers; image de classification temps-fréquence moyenne brute des mots (à gauche), image de classification temps-fréquence moyenne brute des objets non familiers (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les mots qu'avec les objets non familiers, et un * blanc lorsque la différence significative est en direction opposée. À noter ; l'étendue des scores Z pour les images de classification est de -50 à 80 afin de mieux représenter la différence visuellement. 45

Liste des sigles

EEG : Électroencéphalographie

IC : Image de classification

MEG : Magnétoencéphalographie

STM : Stimulation magnétique transcrânienne

VC : Vecteur de classification

Introduction

Depuis maintenant plusieurs années, nous savons que la perception visuelle varie à travers le temps, c'est-à-dire qu'à l'intérieur d'un bref intervalle de temps, par exemple une seconde, il y a des moments où nous semblons voir mieux que d'autres (Busch, Dubois, VanRullen, 2009; Dehaene, 1993; Latour, 1967; VanRullen et Koch, 2003; VanRullen, Reddy et Koch, 2005, 2010). Mais alors, que fait-il que nous sommes plus efficaces à percevoir visuellement notre environnement à un instant plutôt qu'un autre. Qu'est-ce qui contrôle ces changements au sein de notre capacité de traitement visuel? Qu'est-ce qui détermine le patron de ces variations? Y a-t-il un profil type de ce patron de variations dans la perception visuelle, peu importe la tâche? Bref, nous trouvons qu'il y a un manque dans l'état des connaissances actuelles en ce qui a trait à l'évolution temporelle de la capacité de traitement visuel étant donné les indications à l'effet que celle-ci n'est pas uniforme.

Ainsi, dans la présente étude nous nous intéresserons au « comment », c'est-à-dire comment notre perception varie dans le temps. Plus précisément, nous examinerons le décours temporel des changements dans la capacité de traitement de l'information visuelle auprès d'une population saine. Nous chercherons à obtenir des profils caractérisant les fluctuations de l'efficacité de ce traitement visuel. Pour ce faire, nous nous concentrerons sur la reconnaissance d'objets et de visages familiers. À travers l'usage de ces classes de stimuli, nous investiguerons les profils temporel, temps-fréquence et phase-amplitude de l'efficacité de l'encodage visuel, utilisant une technique d'échantillonnage temporel.

Mise en contexte

Oscillation et perception

L'invention de l'électroencéphalographie au début du 20^e siècle (Pidoux, 2010), nous a permis de faire l'une des plus grandes découvertes en neuroscience, soit la réalisation que l'activité neuronale oscille à travers le temps (Cohen, 2015). Depuis, des centaines de chercheurs ont travaillé et travaillent encore aujourd'hui à comprendre la raison d'être de ces oscillations, ce qui les crée, ce qui les modifie, leurs propriétés, leurs localisations, ce qu'elles provoquent, etc.

Entre autres, il existerait un lien étroit entre le rythme des oscillations neuronales et la perception visuelle. Plusieurs études démontrent l'implication d'un rythme alpha dans la vision, c'est-à-dire que l'activité cérébrale oscillant à une fréquence entre 8 et 13 Hz serait impliquée dans la détection consciente de stimuli visuels (Busch, Dubois, VanRullen, 2009; Mathewson, Gratton, Fabiani, Beck et Ro, 2009), l'inhibition des régions cérébrales non nécessaires lors de la perception visuelle (Rana et Vaina, 2014; Händel, Haarmeier et Jensen, 2010; Klimesch, Sauseng, et Hanslmayr, 2007), la perception d'illusions visuelles (Dugué, Marque et VanRullen, 2011; Lange, Keil, Schnitzler, van Dijk et Weisz, 2014; VanRullen, Reddy et Koch, 2006) et même la préparation du cerveau à recevoir l'information visuelle (Stenner, Bauer, Haggard, Heinze et Dolan, 2014).

Plus précisément, il semblerait que ce soit la phase de l'oscillation des ondes cérébrales qui module la performance en perception visuelle. La phase réfère à la position sur l'onde que forme l'oscillation, à un moment précis. Ce fait a été démontré par des études où les oscillations cérébrales étaient manipulées pour suivre un certain rythme et où la relation entre l'instant précis auquel la tâche était initiée, par la présentation du stimulus cible, et la phase de

ce rythme oscillatoire était modulée. De cette façon, ces études ont montré que la performance variait en fonction de la phase des oscillations cérébrales (Jaegle et Ro, 2014; Lakatos, Karmos, Mehta, Ulbert et Schroeder, 2008; Mathewson, Fabiani, Gratton, Beck, et Lleras, 2010; Mathewson, Prudhomme, Fabiani, Beck, Lleras et Gratton, 2012; Spaak, Lange et Jensen, 2014). Ce fait a aussi été démontré par l'utilisation d'une tâche d'attention visuelle où les performances de détection de stimuli fluctuaient à travers le temps en suivant la phase des oscillations cérébrales spontanées, d'une fréquence d'environ 7 Hz, juste avant l'apparition du stimulus (Busch et VanRullen, 2010). Ce résultat, soit que la phase des oscillations autour de 7 Hz prédirait les performances perceptuelles, a aussi été retrouvé dans une étude utilisant une tâche d'identification de contours. Dans cette étude, les mesures combinées d'électroencéphalographie (EEG) et d'imagerie par résonance magnétique fonctionnelle suggéreraient que les oscillations cérébrales créent des fenêtres de temps, de manière périodique à un rythme d'environ 7 Hz, où le transfert d'information à travers des circuits neuronaux distants serait possible (Hanslmayr, Volberg, Wimber, Dalal et Greenlee, 2013). Cela limiterait la perception visuelle à un fonctionnement rythmique. De plus, une étude s'intéressant à la mise à jour de l'information visuelle consciente, mesurée par l'effet de décalage de flash (« flash-lag effect »), a retrouvé de fortes corrélations entre les fluctuations de cette mise à jour à travers les essais et la phase des oscillations cérébrales spontanées mesurée conjointement (Chakravarthi et VanRullen, 2012). Ces oscillations variaient autour de 5-10 Hz au niveau occipital et de 12 à 20 Hz dans la région frontocentrale, suggérant ainsi un mécanisme périodique en deux étapes (Chakravarthi et VanRullen, 2012).

D'autres observations, quant à elles, suggèrent que la résolution spatiale de notre système visuel serait d'autant plus élevée que le rythme alpha des oscillations est rapide

(Samaha et Postle, 2015). De plus, le rythme des oscillations cérébrales pourrait être modifié, non seulement par certaines manipulations, mais aussi par des indices sur le moment d'apparition d'un stimulus qui modifieraient la phase des oscillations juste avant l'apparition du stimulus par des processus descendant, permettant ainsi une performance de détection visuelle optimale (Samaha, Bauer, Cimaroli et Postle, 2015).

L'importance du rôle des oscillations cérébrales au niveau du traitement perceptuel devient ainsi évidente, mais penchons-nous sur ce que cette affirmation signifie à la base. Une oscillation étant caractérisée par un mouvement de va-et-vient (Larousse, s. d.), cela veut donc dire que la perception visuelle fluctue à travers le temps en étant efficace à certains moments et inefficace à d'autres. Bien sûr, cela n'est pas nouveau, tenant compte des nombreuses études ayant démontré que la capacité de traitement visuel ne serait pas uniforme à travers le temps (Busch, Dubois, VanRullen, 2009; Dehaene, 1993; Latour, 1967; VanRullen et Koch, 2003; VanRullen, Reddy et Koch, 2005, 2010). Par contre, il nous reste encore à déterminer comment la perception visuelle évolue temporellement. En effet, la notion d'irrégularité dans la capacité à traiter l'information visuelle est une découverte importante qui demande à être approfondie, notamment pour décrire cette irrégularité. Puis, considérant la relation entre les oscillations cérébrales et la perception, ressortir le patron temporel d'efficacité du traitement visuel pourrait nous informer de manière indirecte sur les mécanismes oscillatoires impliqués dans celui-ci.

Technique d'échantillonnage temporel

Nous avons accès à une technique psychophysique, basée en partie sur la technique *Bubbles* temporelle développée en 2005 par McCabe, Blais et Gosselin (voir également Vinette, Gosselin et Schyns, 2004; pour une étude antérieure ayant utilisée un échantillonnage

dans les domaines spatial et temporel), nous permettant d'étudier l'évolution temporelle de l'encodage visuel ainsi que les fréquences temporelles donnant lieu à un meilleur encodage. Cette technique implique la présentation d'images superposées de bruit blanc, où le rapport signal sur bruit varie à travers le temps selon un profil temporel construit par l'intégration d'ondes sinusoïdales de différentes fréquences avec des amplitudes et des phases aléatoires. Ensuite, il est possible de créer des vecteurs ou des images de classification (selon le nombre de dimensions échantillonnées) pour chaque participant en soustrayant les caractéristiques temporelles des fluctuations du rapport signal-bruit des essais associés aux erreurs de celles des essais associés aux bonnes réponses. Puis, le profil général est obtenu en moyennant les vecteurs ou les images de classification à travers tous les participants.

Cette technique a d'abord été utilisée par Blais, Arguin et Gosselin (2013) dans leur article « Human visual processing oscillates: Evidence from a classification image technique ». Ces chercheurs ont utilisé une tâche de reconnaissance de visages non familiers où ils présentaient les stimuli pendant 200 ms et faisaient varier le ratio signal-bruit à travers le temps selon des fréquences de 5 à 20 Hz (par pas de 5 Hz). Ils ont trouvé, entre autres, que les performances étaient meilleures lorsque le ratio signal-bruit variait à un taux de 10 à 15 Hz. De plus, leurs données démontrent que certaines phases seulement sont corrélées avec les réponses correctes, suggérant ainsi que l'échantillonnage d'information fait par le cerveau se synchronise avec l'apparition du stimulus. Finalement, au niveau temporel, trois moments ressortent de leurs analyses comme étant bénéfiques pour la reconnaissance, soit entre l'apparition du stimulus et 25 ms, à environ 75 ms et entre 133 ms et 150 ms. Ces résultats vont dans le sens d'un traitement visuel qui se ferait de manière oscillatoire.

Ensuite, la technique a été reprise par Arguin (2018) pour évaluer l'encodage visuel de mots. Ainsi, dans cette étude, les participants devaient lire des mots superposés de bruit blanc et présentés pendant 200 ms. Par contre, le ratio signal-bruit variait à travers le temps selon l'intégration d'ondes sinusoïdales de fréquence entre 5 et 55 Hz (par pas de 5 Hz) avec des phases et des amplitudes aléatoires. Les analyses démontrent au moins deux mécanismes oscillatoires impliqués, un premier se manifestant aux environs de 43 ms et un autre vers 157 ms, où l'efficacité du traitement visuel semblait être plus grande. Pour le mécanisme plus précoce, les meilleures performances étaient obtenues avec des variations du ratio signal-bruit selon des fréquences entre 15 et 30 Hz, tandis que pour le plus tardif, ce sont les fréquences aux environs de 55 Hz qui ressortaient comme significativement associées à une bonne performance.

Puis, la même technique que pour la reconnaissance de mots a été appliquée à des objets non familiers (Arguin et Massé, 2019). Dans cette étude, les participants devaient d'abord apprendre à reconnaître six objets synthétiques 3D de formes non familières (voir Annexe C pour des exemples de stimuli utilisés), chacun étant présenté sous quatre points de vue possibles pour un total de 24 stimuli. Durant la phase d'expérimentation, les participants devaient indiquer la lettre associée à chaque objet (lui servant ici de nom) à l'aide du clavier. Les résultats dans le domaine temporel démontrent qu'en moyenne, la performance était meilleure lorsque les objets étaient visibles du tout début de la période de stimulation jusqu'à 36 ms et entre 168 ms et 200 ms. La reconnaissance était aussi significativement moins bonne lorsque les objets n'étaient visibles qu'entre 44 ms et 150 ms suite au début de la période de stimulation. Pour ce qui est des analyses dans le domaine temps-fréquence, les données montrent que dans une même période de temps, il y a des fréquences qui sont moins efficaces

pour la reconnaissance et des fréquences qui favorisent la reconnaissance. Il y aurait trois mécanismes oscillatoires impliqués. Le premier est associé à une haute efficacité du traitement visuel alors que les fréquences oscillatoires du rapport signal-bruit ont une haute amplitude dans l'étendue 30-35 Hz pendant les périodes 75-125 ms et 150-200 ms après le début de la présentation de la stimulation. Cependant, cette même étendue de fréquences oscillatoires est associée à une faible efficacité du traitement visuel dans l'intervalle 100-150 ms. Un deuxième mécanisme se manifeste à travers les oscillations signal-bruit de 10-15 Hz, pour lesquelles on constate une grande efficacité du traitement perceptif au moment de l'apparition du stimulus ainsi qu'entre 150 et 200 ms, et inversement, une faible efficacité pour les périodes 50-100 ms et 125-175 ms. Finalement, un troisième mécanisme est révélé par les fréquences de 50-55 Hz. Celui-ci est associé avec une grande efficacité dans les intervalles 50-100 ms et 125-175 ms, mais avec une faible efficacité de 75 à 125 ms, de 100 à 150 ms et 150 à 200 ms.

Ainsi, les résultats des trois études décrites précédemment démontrent que l'efficacité du traitement visuel réalisé par les participants varie de façon marquée à travers la durée de présentation des stimuli (200 ms). De plus, on remarque que les profils temporels sont très différents selon la classe de stimulus à reconnaître (visage non familier, mot et objet non familier). Dans tous les cas, les images de classification reposant sur une caractérisation des profils d'échantillonnage temporel en termes de l'amplitude et de la phase des ondes sinusoïdales (de fréquence entre 5 et 55 hertz inclusivement, par pas de 5 Hz) dont ils sont constitués sont toutes parfaitement uniformes (i.e. aucune valeur significativement différente de 0). Par contre, une caractérisation de ces profils d'échantillonnage temporel en termes des résultats de leur analyse temps-fréquence permet de constater : 1- des images de classification moyennes qui sont très distinctes selon la classe de stimulus à reconnaître et, 2- des images de

classification individuelles qui sont très semblables d'un participant à l'autre (cohérence inter-sujets, calculée par la moyenne de toutes les paires possibles de corrélations entre les images de classification individuelles: 0,87 pour mots et 0,67 pour objets non familiers). Ces observations nous permettent de valider les études énonçant que la perception visuelle est variable sur le plan temporel. Les résultats indiquent également qu'un code phase-amplitude ne parvient pas à capturer cette évolution temporelle du traitement visuel alors qu'un code temps-fréquence semble particulièrement apte à cette fin. Pour ce qui est des vecteurs de classification basés directement sur les profils d'échantillonnage temporel, il s'agit plutôt d'un cas intermédiaire. En effet, il semble qu'un code temporel puisse apporter davantage d'information quant aux moments privilégiés pour l'encodage visuel, mais qu'il y ait une plus grande variabilité entre les participants (cohérence inter-sujets de 0,18 pour mots et de 0,03 pour les objets non familiers).

Objectifs

Somme toute, la capacité de traitement visuel varie à travers le temps à la manière d'oscillations cérébrales et les mécanismes oscillatoires retrouvés dans la perception visuelle sont dépendants du type de stimulus. Cela étant dit, que se passe-t-il lorsque nous regardons des objets ou des visages qui sont familiers? Ces types de stimuli font partie de notre vie quotidienne et sont donc une bonne référence lorsqu'il est question d'étudier la perception visuelle. Cependant, les études précédentes ont utilisé des objets et des visages non familiers, c'est-à-dire des stimuli qui devaient être appris par les participants avant de faire la tâche de reconnaissance. Cela pourrait avoir un impact sur la façon dont on perçoit les choses. Nous pensons donc qu'il vaut la peine d'investiguer davantage la question en reprenant la technique

d'échantillonnage temporel et en l'appliquant à l'identification d'objets et de visages, cette fois-ci, familiers.

L'objectif principal de cette étude sera donc de déterminer le décours temporel (en termes de temps, de fréquences temporelles, de phase et d'amplitude) que suit la capacité à traiter l'information visuelle sur une courte période de temps (200 ms) suivant l'apparition d'une cible. Comme objectif secondaire, nous souhaitons comparer les différents décours temporels obtenus avec différents types de stimuli (différentes classes de stimuli et différents niveaux de familiarité). Ce deuxième objectif nous permettra d'approfondir notre compréhension des facteurs influençant les variations de la capacité de traitement de l'information. Ainsi, dans la présente étude, nous tenterons de répondre aux quatre questions suivantes :

1. Quels sont les profils temporel et temps-fréquence de l'efficacité du traitement visuel lors de la reconnaissance d'objets et de visages familiers?
2. Est-ce que le profil phase-amplitude est uniforme pour chaque stimulus, comme dans les études précédentes?
3. Est-ce que les profils sont, encore une fois, différents entre chaque type de stimulus?
4. Est-ce que la familiarité d'un stimulus modifie la façon dont celui-ci est perçu? (Ou plus précisément, est-ce que les profils temporel et temps-fréquence sont différents entre les objets familiers et non familiers, et entre les visages familiers et non familiers?).

L'expérience 1 du présent mémoire portera sur la reconnaissance d'objets familiers alors que l'expérience 2 portera plutôt sur la reconnaissance de visages déjà connus des participants. Puis, une section complète sera consacrée à la comparaison entre les types de

stimuli, incluant ceux des expériences 1 et 2 ainsi que ceux des études antérieures présentées précédemment (visages non familiers, mots et objets non familiers).

Hypothèses

Évidemment, nous nous attendons à retrouver des variations temporelles dans la capacité de traitement visuel, et ce sur les profils temporel et temps-fréquence des objets et des visages familiers. En nous fiant aux études précédentes, nous pensons retrouver des différences importantes à travers les types de stimuli, soit entre les profils des objets familiers et ceux des mots et des visages, et entre les profils des visages familiers et ceux des mots et des objets. Puis, nous nous attendons à retrouver des profils phase-amplitude uniformes pour les deux stimuli étant donné que ce code de valeurs ne semble pas apte à capturer l'évolution temporelle du traitement visuel. Cependant, nous n'avons pas d'hypothèses spécifiques pour ce qui a trait aux propriétés exactes des profils qui seront obtenus, ni pour notre quatrième question de recherche concernant la comparaison entre les stimuli familiers et non familiers. Il serait présomptueux de s'essayer à prédire les caractéristiques mêmes des profils et si ces derniers seront semblables ou non aux profils des stimuli non familiers, étant donné que cette étude figure dans les premières de ce genre et que, jusqu'à présent, chacune se démarque de l'autre.

Expérience 1

Méthodologie

Participants. Douze jeunes adultes (âgés entre 19 et 35 ans, âge moyen de 23 ans, 7 femmes) ont participé à cette étude. Ils avaient tous une vision normale ou corrigée à la normale. Ils étaient tous cognitivement et neurologiquement sains. Chacun a signé un consentement éclairé et reçu une compensation financière de 35\$ pour sa participation. Toutes les procédures ont été effectuées avec l'approbation du comité d'éthique de l'Université de Montréal.

Stimuli. Les stimuli se composent de 330 images d'objets familiers obtenus de la « Bank of Standardized Stimuli » (BOSS; voir Annexe A pour des exemples d'images utilisées) et choisis selon les normes de Brodeur, Kehayia, Dion-Lessard, Chauret, Montreuil, Dionne-Dostie et Lepage (2012), qui ont été recueillies auprès de participants francophones. Trois informations ont été utilisées afin d'effectuer la sélection des stimuli, soit la proportion de personnes disant ne pas connaître l'objet du tout sur le nombre de celles ayant donné une réponse quelconque, la proportion de personnes disant ne pas connaître le nom de l'objet sur le nombre de celles ayant donné une réponse, et la proportion de personnes qui avaient le nom sur le bout de la langue sans pouvoir le retrouver sur le nombre de celles ayant donné une réponse. Ces trois valeurs ont été additionnées afin de constituer un index du niveau de difficulté à reconnaître le stimulus en question et celui-ci a servi à la sélection des images, en commençant par les valeurs les plus faibles. 300 images ont été utilisées dans les blocs expérimentaux et 30 autres pour le bloc de pratique. Parmi les 300 images des blocs d'expérimentation, 208 avaient un index de difficulté de 0, 77 avaient un index de 3, 14 de 4

et une de 5. Parmi les 30 images du bloc de pratique, il y en avait une avec un index de 5 et 29 avec un index de 7. Dans la BOSS, les images d'objets sont placées sur un fond blanc dans un cadre de 2000 x 2000 pixels. Durant la tâche expérimentale, chaque image était présentée en niveaux de gris et exposée avec une résolution de 640 x 640 pixels. À l'intérieur de ce cadre, la taille maximale des objets était de 17,4° d'angle visuel horizontalement et de 17,1° d'angle visuel verticalement, et la taille minimale des objets était de 2,6° d'angle visuel horizontalement et de 3,3° d'angle visuel verticalement.

Matériel. Les stimuli sont présentés sur un écran à DEL ASUS VG248QR full HD à un taux de rafraîchissement de 120 Hz. Les participants sont maintenus à une distance de 57 cm de l'écran à l'aide d'une mentonnière. Le programme expérimental est écrit en MatLab, avec l'utilisation des extensions de « Psychophysics Toolbox » (Kleiner et al, 2007). Les images sont présentées pendant 200 ms et du bruit blanc leur est superposé. La plage de bruit aléatoire utilisée est constante pour un essai et varie aléatoirement à travers les essais. Le rapport signal-bruit varie à travers le temps selon une fonction d'échantillonnage créée par l'intégration d'ondes sinusoïdales de fréquence entre 5 et 55 Hz (par pas de 5 Hz) avec des amplitudes et des phases aléatoires. Par contre, l'énergie de la stimulation est maintenue constante, c'est-à-dire que toutes les images constituant la séquence présentée au cours d'un essai ont un contraste RMS normalisé à 1. Une nouvelle fonction d'échantillonnage est créée à chaque essai. Le rapport signal-bruit est normalisé dans l'étendue 0 (seul le bruit visuel est visible) à 0,5 (le signal est obstrué partiellement par le bruit visuel).

Procédures. La tâche utilisée consiste à reconnaître visuellement des objets familiers présentés sur un écran d'ordinateur. Le déroulement de l'étude se fait sur deux séances

d'environ 1h10 chacune. La première séance inclut un bloc de pratique suivi de deux blocs expérimentaux et la deuxième séance inclut deux autres blocs expérimentaux. Le bloc de pratique comporte 30 essais, soit 30 images d'objets et chaque bloc expérimental contient 300 essais, soit une répétition des mêmes 300 images. 10 ordres différents des images du bloc expérimental avaient été préparés et préalablement assignés aux participants de manière aléatoire. En complétant quatre blocs expérimentaux, chaque participant produit un total de 1200 essais.

Un essai commence par une plage carrée de bruit visuel vide de 17,7 cm par 17,7 cm pendant 1250 ms. Ensuite, une croix de fixation blanche apparaît au centre de ce carré pendant 250 ms, puis se retire, laissant le carré de bruit visuel vide comme au début. 150 ms après l'apparition de la croix de fixation, un son pur de 14 ms à 900 Hz et 70 dB est émis pour indiquer au participant que la cible est sur le point d'apparaître, ce qui est le cas 100 ms après. L'image de la cible est superposée de bruit blanc visuel et le rapport signal-bruit varie selon la fonction d'échantillonnage établie pour cet essai sur une durée de 200 ms. Ensuite, seule la plage de bruit visuel est visible de manière à effacer la trace du signal cible. Les participants n'ont pas de contrainte de temps pour donner leur réponse. Les instructions données aux participants sont simplement de nommer, à voix haute, l'objet présenté. Ils ont la possibilité de dire le nom de l'objet, un synonyme, le contexte dans lequel l'objet est utilisé, la façon dont il est employé ou de répondre qu'ils n'ont aucune idée de ce que peut être l'objet en question. Les réponses sont jugées bonnes ou mauvaises à un niveau de catégorisation de base (Marques, 2011), puis enregistrées par le chercheur qui tient avec lui une liste des objets. Une bonne réponse est suivie d'un feedback sonore de 1000 Hz (son aigu) alors qu'une tonalité de

300 Hz (son grave) suit une réponse erronée. De cette façon, les participants peuvent essayer de corriger leurs réponses lors des blocs expérimentaux suivants.

La performance de reconnaissance est maintenue à 50% en ajustant l'amplitude de la dégradation de l'image de la cible. Pour ce faire, une plage de bruit blanc est superposée à l'image de la cible et est intégrée dans la portion « signal » de la stimulation. Celle-ci est indépendante de la plage de bruit blanc constituant la portion « bruit » du stimulus. L'amplitude du bruit superposé à l'image de la cible peut varier de 0 à 127. Le maximum possible de 127 correspond à 50% du maximum de luminance que peut produire l'écran. Au début de l'expérience, l'amplitude de ce bruit est fixée à 64. Puis, à partir du 11^e essai, si le participant a eu une performance inférieure à 50% dans les 10 essais précédents, l'amplitude du bruit est diminuée d'un pas, et si le participant a eu une performance supérieure à 50%, l'amplitude du bruit est augmentée d'un pas. La valeur initiale du pas est de 16 et celle-ci diminue de moitié à chaque fois que la modification du niveau de bruit change de direction, et ce jusqu'à l'atteinte d'une valeur de 1.

Analyses statistiques. Les données sont analysées de manière comparable à celles des études d'échantillonnage temporel menées précédemment par Arguin (2018) et Arguin et Massé (2019). Ainsi, des images ou vecteurs de classification reposant sur la performance des participants (réponses correctes vs erreurs) sont faits à partir des différentes caractérisations des fonctions d'échantillonnage utilisées à chaque essai. Ainsi, les vecteurs/images de classification ont été construits soit à partir des fonctions d'échantillonnage brutes, soit selon les valeurs de phase et d'amplitude des ondes sinusoïdales à partir desquelles elles ont été constituées, ou encore à partir des résultats d'une analyse temps-fréquence de ces fonctions d'échantillonnage. Pour cette analyse temps-fréquence, les fonctions d'échantillonnage ont été

segmentées en fenêtres de temps relativement à la fréquence temporelle analysée. Plus spécifiquement, les fréquences de 5 à 55 Hz (par pas de 5 Hz) sont analysées sur chaque fonction d'échantillonnage. La durée de présentation du stimulus est divisée en fenêtres de Hamming contenant jusqu'à un maximum de quatre cycles de la fréquence analysée et se chevauchant à 50% (i.e. une fenêtre de 5 à 20 Hz, trois fenêtres de 25 à 40 Hz et cinq fenêtres de 45 à 55 Hz). Un filtre anti-repliement passe-bas (59 Hz) est ensuite appliqué sur ces fenêtres temporelles et finalement, une analyse de Fourier est effectuée sur ces dernières.

Pour toutes ces caractérisations des profils d'échantillonnage temporel, les valeurs associées aux erreurs ont été additionnées et pondérées (selon leur nombre) pour ensuite être soustraites de la somme pondérée des données caractérisant les fonctions d'échantillonnage associées à une bonne réponse. Des images ou vecteurs de classification reflétant l'efficacité du traitement visuel réalisé par les participants dans les domaines temporel, phase-amplitude et temps-fréquence ont ainsi été calculés pour chaque participant.

Sur la base d'un processus de « bootstrapping », les vecteurs/images de classification ont été convertis en scores Z . Cette technique consiste à générer aléatoirement 1000 images de classification qui utilisent les mêmes masques que ceux présents dans l'expérience, en les faisant correspondre aléatoirement aux réponses correctes ou aux erreurs. Pour chacun de ces vecteurs ou images de classification aléatoires, on retient la valeur moyenne des points qu'on y retrouve, puis la moyenne des moyennes qu'elles ont produites et leur écart-type, procurant une référence pour la transformation des données brutes en scores Z . Une fois les données converties en scores Z , nous avons procédé à des corrélations inter-sujets deux par deux au sein de chacun des domaines analysés afin de déterminer la cohérence inter-sujets. Puis, nous avons effectué un test t à échantillon unique pour chaque domaine sur l'ensemble des

corrélations obtenues afin de savoir si la valeur de l'indice de cohérence est significativement différente de 0.

Ensuite, les vecteurs/images de classification moyens (à travers tous les participants) ont été lissés en fonction d'un opérateur de forme gaussienne dont la variance est ajustée afin de révéler au mieux possible les phénomènes apparents dans les résultats. Dans le cas présent, la variance est de 1 pour le vecteur de classification temporel et de 1,25 pour les images de classification phase-amplitude et temps-fréquence. Enfin, les données à l'intérieur du même vecteur de classification moyen ou de chaque image de classification moyenne sont reprises et comparées les unes avec les autres pour déterminer si elles sont mutuellement significativement différentes, et ce en commençant par les données aux extrêmes. Le seuil du test statistique Pixel (Chauvin, Worsley, Schyns, Arguin et Gosselin, 2005), dont le seuil de signification était fixé à $\alpha = 0,05$, a été utilisé pour cette opération. Cette procédure offre une représentation graphique des vecteurs/images de classification dans laquelle toutes les valeurs distinctes qui y sont représentées sont significativement différentes les unes des autres. Dans le cas où les valeurs en question ne sont pas significativement différentes les unes des autres, celles-ci ont été moyennées et c'est cette valeur moyenne qui est illustrée sur les figures représentant les vecteurs/images de classification.

Résultats

Information générale. La moyenne de bruit visuel inclut dans la portion « signal » du stimulus pour obtenir une performance de 50% correcte, incluant tous les participants et tous les essais, est de 115,8 (taux moyen de réponses correctes : 49,2%).

Vecteur de classification temporel. Sur la base du test Pixel (Chauvin, Worsley, Schyns, Arguin et Gosselin, 2005), une différence de 1,65 entre les scores Z obtenus aux différents moments de la présentation du stimulus est requise pour les considérer comme significativement différents l'un de l'autre. Le vecteur de classification moyen du domaine temporel (Figure 1) affiche ainsi neuf paliers (ou plateaux) significativement ($p < 0,05$) différents. Ce dernier montre qu'à l'apparition du stimulus, l'efficacité du traitement visuel est associée à des scores Z négatifs, qui diminuent encore davantage pour atteindre le point le plus faible à 29 ms au score Z de -2,75. Ensuite, ces scores grimpent de plus en plus jusqu'aux alentours de 80 ms, où l'efficacité atteint un plateau maximal (79-88 ms). Puis, les scores Z redescendent jusqu'à 0 autour de 100 ms pour ensuite remonter graduellement vers un second pic s'étalant de 146 à 163 ms. Finalement, l'efficacité du traitement visuel descend rapidement et retourne à des scores Z négatifs, et ce jusqu'à la fin de la période d'exposition de la cible. La cohérence inter-sujets (moyenne des corrélations inter-sujets) est de 0,009 et n'est donc pas significativement différente de 0 ($t(65) = 0.32, ns$).

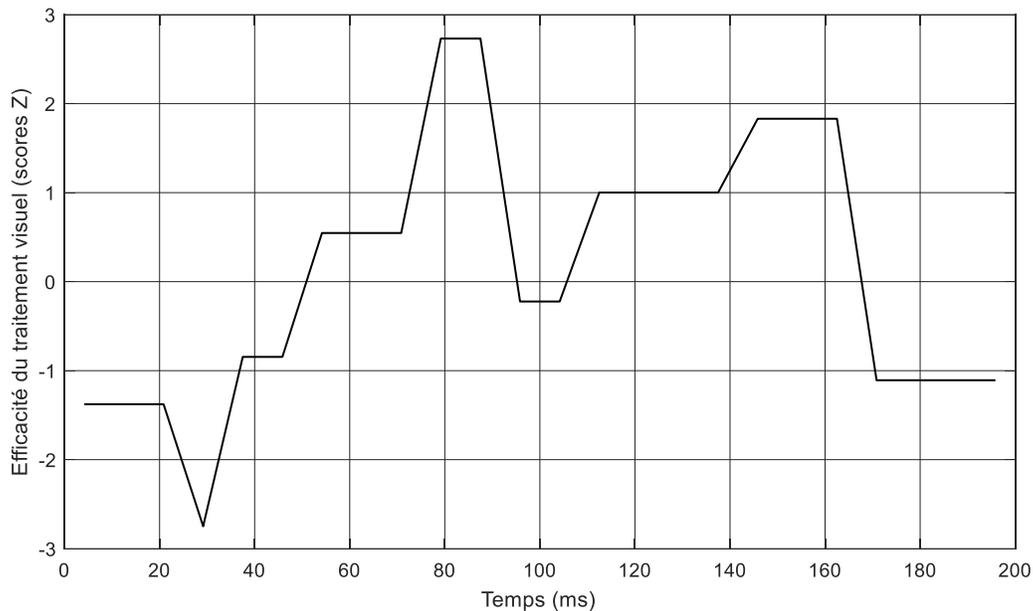


Figure 1. Vecteur de classification temporel moyen représentant les variations de l'efficacité du traitement visuel à travers le temps en reconnaissance d'objets familiers (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau).

Image de classification temps-fréquence. Dans le domaine temps-fréquence, l'axe temporel de l'image de classification moyenne (Figure 2) est un peu différent par rapport au vecteur de classification. L'étendue de 200 ms, pendant laquelle l'image cible est présentée, est divisée en cinq sections se chevauchant mutuellement sur 50% de leur étendue. Pour cette image de classification, le test Pixel prescrit une différence minimale de 4,01 entre les scores Z pour les considérer comme significativement ($p < 0,05$) différents. De 0 à 67 ms, toutes les fréquences de 5 Hz à 55 Hz favorisent l'encodage visuel, mais principalement celles de 25 à 35 Hz qui atteignent des scores Z plus élevés que 40. Entre 33 et 100 ms, les fréquences de 5 à 35 Hz sont associées avec une grande efficacité du traitement visuel, tandis que les fréquences de 45 et 50 Hz sont associées avec une moins bonne efficacité. De 67 ms à 133 ms, les

fréquences de 5 à 25 Hz sont favorables à l'encodage visuel, tandis que les fréquences de 30 à 55 Hz sont moins propices à l'encodage visuel. De 100 ms à 167 ms, seulement les fréquences de 5 à 20 Hz sont associées avec un traitement visuel efficace, alors que les fréquences de 30 à 55 Hz sont moins avantageuses pour l'encodage visuel. Finalement, de 133 ms à 200 ms, les fréquences de 5 à 20 Hz sont encore une fois propices à l'encodage, mais aussi les fréquences de 35 à 55 Hz. Les cellules temps-fréquence n'ayant pas été mentionnées sont considérées comme neutres, ou bien représentant une efficacité moyenne, étant donné le trop grand rapprochement entre les scores Z y étant obtenus et un score Z de 0 (identifié selon le seuil Pixel). L'indice de cohérence inter-sujets est de 0,93, ce qui est significativement différent de 0 ($t(65) = 151.76, p = .001$).

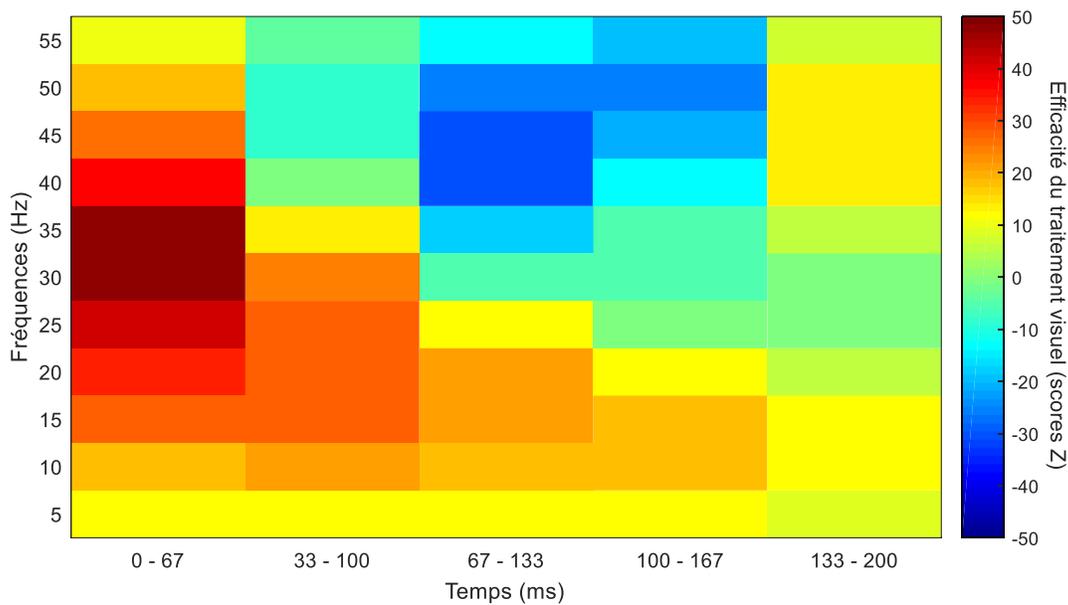


Figure 2. Image de classification temps-fréquence moyenne représentant les variations de l'efficacité du traitement visuel à travers le temps et les fréquences temporelles en reconnaissance d'objets familiers (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau).

Image de classification phase-amplitude. L'image de classification moyenne dans le domaine phase-amplitude (Figure 3) ne contient aucun contraste significatif selon le test Pixel ($p < 0,05$; Chauvin, Worsley, Schyns, Arguin, Gosselin, 2005). C'est pourquoi cette image de classification est complètement uniforme, c'est-à-dire qu'elle n'affiche qu'une seule grande zone de valeur unique (ou plateau).

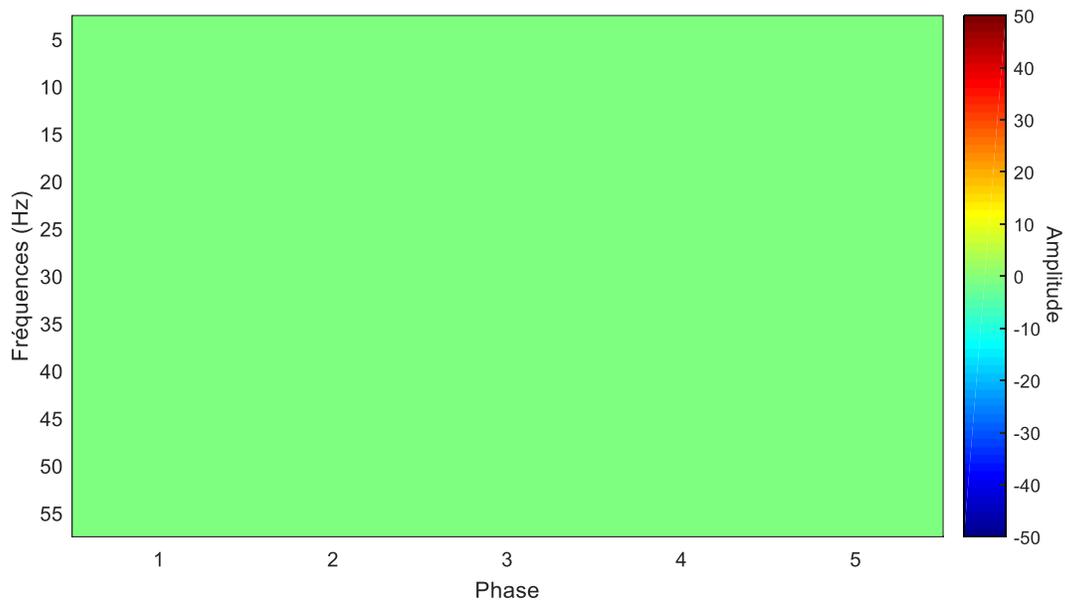


Figure 3. Image de classification moyenne dans le domaine phase-amplitude représentant les données utilisées pour créer les fonctions d'échantillonnage de chaque essai lors de l'expérience sur les objets (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau).

Discussion

En reconnaissance d'objets familiers, le traitement visuel serait, donc, globalement moins efficace au début de l'apparition du stimulus. Ensuite, cette efficacité augmenterait pour atteindre un pic autour de 80 ms, pour ensuite redescendre et enfin remonter vers un second

pic près de 145 ms. Puis, le traitement visuel diminuerait en efficacité pour se terminer à un niveau d'efficacité moyen à la fin de la présentation du stimulus. De ce profil, nous pouvons comprendre que le cerveau prend un certain temps avant de pouvoir encoder l'information visuelle efficacement sur l'objet qui vient d'être perçu. Cela concorderait avec une étude d'Ebaid et Crewther (2019) dans laquelle le seuil temporel de présentation minimal moyen pour identifier une image, dans ce cas-ci entre un poisson, un camion et un papillon, chez de jeunes adultes était de 49,30 ms. Cela voudrait dire que nous aurions besoin d'un minimum d'environ 50 ms pour arriver à reconnaître un objet et, dans nos données, ce point temporel correspond approximativement au moment où les scores *Z* atteignent 0, c'est-à-dire où l'efficacité du traitement visuel augmente jusqu'à atteindre un niveau moyen. Sinon, le fait que notre cerveau soit idéalement prêt à recevoir l'information visuelle aux alentours de 80 et 145 ms reste encore à élucider. Néanmoins, il faut interpréter ces résultats avec précaution considérant que la cohérence inter-participants dans le domaine temporel est nulle.

Au niveau du domaine temps-fréquence, nous avons vu que la plus haute efficacité de traitement visuel se retrouve à 0-67 ms pour des stimuli oscillant entre 25 et 35 Hz. L'amplitude de cette efficacité décline ensuite de 67 à 167 ms tout en se décalant vers de plus basses fréquences (10-15 Hz). Les fréquences autour de 35-55 Hz entre 67 et 167 ms sont associées à la plus faible efficacité de traitement visuel. Par contre, certaines de ces fréquences (40-50 Hz) amènent une grande efficacité à la fin de la présentation du stimulus (133-200 ms). Ainsi, au début et un peu à la fin de la présentation du stimulus, les hautes fréquences sont propices au traitement visuel, signifiant que lorsque le ratio signal-bruit variait rapidement à ces moments, la performance de reconnaissance des participants augmentait. En lien avec ceci, des études EEG ont démontré l'implication d'activité dans la bande de fréquences gamma

induite lors de reconnaissance d'objets; de 30 à 90 Hz entre 230 et 300 ms (Gruber, Malinowski et Müller, 2004), de 30 à 80 Hz entre 200 et 350 (Martinovic, Gruber et Müller, 2007) et de 35 à 90 Hz entre 180 et 350 ms (Martinovic, Gruber, Hantsch et Müller, 2008). Considérant que les performances perceptives sont modulées par le rythme des oscillations cérébrales, cela pourrait peut-être expliquer pourquoi les fréquences de 40 à 50 Hz sont associées avec une grande efficacité de traitement de 133 à 200 ms. Pour deux de ces études, l'activité gamma survient après 200 ms, mais rappelons qu'il existe un délai entre l'occurrence d'un événement dans le champ visuel et l'activité cérébrale qui en résulte. Il est probable aussi que les hautes fréquences propices à l'encodage soient en lien avec de l'activité gamma évoquée, plutôt qu'induite, par la présentation d'un objet; non spécifié (Gruber, Malinowski et Müller, 2004), de 20 à 30 Hz entre 0 et 200 (Martinovic, Gruber et Müller, 2007) et de 30 à 40 Hz entre 50 et 150 ms (Martinovic, Gruber, Hantsch et Müller, 2008), ce qui pourrait expliquer pourquoi les 20 à 40 Hz entre 0 et 67 ms sont les fréquences associées avec la plus grande efficacité de traitement visuel. Par contre, cela va en contradiction avec le fait que les hautes fréquences (> 35 Hz) soient associées avec une faible efficacité de traitement visuel entre 67 et 167 ms.

Pour ce qui est des basses fréquences (< 20 Hz), qui sont toujours associées à un traitement visuel efficace peu importe le moment de présentation du stimulus, il est plus difficile de trouver un lien avec la reconnaissance d'objet comme telle. Cependant, plusieurs études rapportent des mécanismes oscillatoires de basses fréquences dans la perception. Entre autres, nous pouvons revenir aux études démontrant que la perception serait régulée selon un fonctionnement rythmique à 7 Hz dans la détection de stimuli (Busch et VanRullen, 2010) et dans l'identification de contours (Hanslmayr, Volberg, Wimber, Dalal et Greenlee, 2013). À

cela, nous pourrions ajouter la démonstration d'un déploiement attentionnel modulé périodiquement à environ 7 Hz lors d'une tâche de recherche visuelle difficile (Dugué, McLelland, Lajous et VanRullen, 2015). Tout en restant dans de basses fréquences, les tâches en recherche visuelle peuvent, néanmoins, entraîner des résultats quelque peu différents. Par exemple, en faisant varier le ratio signal-bruit à un rythme de 2 à 20 Hz (par pas de 2 Hz) sur les stimuli d'une tâche de recherche visuelle de conjonction couleur-orientation, Dugué et VanRullen (2014) ont trouvé que l'efficacité attentionnelle était maximale à 9-10 Hz. Puis, une étude utilisant l'EEG et la stimulation magnétique transcrânienne (STM), toujours lors d'une tâche de recherche visuelle, a trouvé que la phase des oscillations cérébrales à 6 Hz serait différente entre les essais réussis et non réussis, mais surtout, serait opposée entre l'activité pré-stimulus et post-stimulus. De plus, la STM appliquée sur le cortex occipital ferait ressortir un patron périodique d'interférence aussi de 6 Hz (Dugué, Marque et VanRullen, 2015).

Des fréquences dans la bande thêta ont aussi été rapportées dans des études utilisant une variante de la tâche Egly-Driver (Egly, Driver et Rafal, 1994). Ainsi, dans une étude comparant la performance comportementale avec l'activité cérébrale mesurée par EEG intracrânien, des chercheurs ont trouvé que l'échantillonnage attentionnel à la position indiquée ce fait à un rythme de 4 Hz (Helfrich et al., 2018). Ce rythme était aussi retrouvé dans leur deuxième expérience mesurant l'attention implicite et dans les deux cas, l'activité corticale de haute fréquence aux électrodes répondant à l'indice fluctuait à environ 4 Hz aussi (Helfrich et al., 2018). Avec le même type de tâche, Fiebelkorn, Saalman et Kastner (2013), ont trouvé que les performances de détection variaient à environ 4 Hz lorsque la cible se trouvait sur l'objet qui n'était pas indicé, alors qu'elles variaient à environ 8 Hz lorsque la cible se trouvait

sur le même objet que l'indice. Fiebelkorn et Kastner (2019) expliquent dans leur revue de littérature que ce rythme thêta représenterait les variations dans les performances sensorielles et les mouvements d'exploration des yeux. Un rythme thêta (3,5-4 Hz) a aussi été retrouvé dans une étude sur une tâche d'identification lorsque les deux stimuli étaient de chaque côté de l'écran, mais lorsque les deux stimuli étaient du même côté le rythme de l'échantillonnage sensoriel était dans la bande alpha (11 Hz; Senoussi, Moreland, Busch et Dugué, 2019). Puis, nous avons aussi présenté une étude mentionnant que la mise à jour de l'information visuelle, mesurée par le « flash-lag effect », varierait en association avec les oscillations cérébrales occipitales autour de 5-10 Hz et frontocentrales autour de 12-20 Hz (Chakravarthi et VanRullen, 2012).

Donc, il est évident que les basses fréquences jouent un rôle important dans la rythmicité observée sur les performances perceptives et cela s'arrime bien avec nos résultats, ayant aussi trouvé qu'en reconnaissance d'objets, les participants semblent bénéficier d'un retour de l'information pertinente (le signal) de manière périodique par cycle de 20 Hz et moins, et ce potentiellement en tout temps. Cependant, selon la façon dont nous avons analysé nos données, effectuant une transformation de Fourier sur toute l'étendue de 200 ms pour les fréquences de 5 à 20 Hz, il est normal de retrouver une constance dans les scores Z à travers l'étendue de présentation du stimulus. Il serait intéressant de voir si l'efficacité du traitement visuel reste toujours élevée au niveau des basses fréquences avec une présentation des stimuli plus longue.

En résumé, il est possible que l'utilisation d'objets familiers en tant que stimulus évoque et/ou induise de l'activité gamma qui serait captée à certains moments sur le profil temps-fréquence d'efficacité du traitement visuel. Puis, il est aussi possible qu'en parallèle, la

perception visuelle, indépendamment du stimulus présenté, soit modulée par des oscillations cérébrales de basses fréquences, par exemple dans les aires visuelles de bas niveau.

Finalement, pour ce qui est du profil phase-amplitude, nous avons confirmé notre hypothèse, c'est-à-dire que l'image de classification complètement uniforme démontre que ce code de valeurs n'est pas apte à capturer l'évolution temporelle du traitement visuel.

Expérience 2

Méthodologie

Participants. Dix-huit jeunes adultes (âgés entre 20 et 35 ans, âge moyen de 22,7 ans, 11 femmes) ont participé à cette étude. Six ont été exclus de l'étude : trois avaient une connaissance insuffisante des célébrités dont les visages étaient utilisés comme stimuli (1% et moins de bonnes réponses dans le premier bloc expérimental comparativement aux autres participants qui avaient, en moyenne, 19,5% de bonnes réponses dans ce bloc), deux ont complété l'expérience alors que le programme expérimental comportait une erreur (l'ajustement de la difficulté à travers les essais ne fonctionnait pas), et un participant n'a effectué que la première séance et n'était plus disponible pour terminer l'expérience. Les données de 12 participants (âgés entre 20 et 35 ans, âge moyen de 23,5 ans, 8 femmes) ont été conservées et analysées.

Comme pour la première expérience, tous les participants avaient une vision normale ou corrigée à la normale. Ils étaient tous cognitivement et neurologiquement sains. Chacun a signé un consentement éclairé et reçu une compensation financière de 40\$ pour sa participation. Toutes les procédures ont été effectuées avec l'approbation du comité d'éthique de l'Université de Montréal.

Stimuli. Pour avoir des visages qui soient familiers de tous, nous avons utilisé des images de célébrités provenant de la banque de visages célèbres construite par Butler, Blais, Gosselin, Bub et Fiset (2010). De cette banque, nous avons choisi les 130 visages que nous considérons être les plus connus. Chaque visage est en niveau de gris et présenté sous une perspective frontale. L'expression des visages est soit joyeuse ou neutre. Cent visages (63

hommes, 37 femmes) ont été utilisés pour les blocs expérimentaux et 30 autres visages (20 hommes, 10 femmes) pour le bloc de pratique. Chaque image était exposée avec une résolution de 256 x 256 pixels. À l'intérieur de ce cadre, la taille maximale des visages était de 7° d'angle visuel horizontalement et de 7,1° d'angle visuel verticalement, et la taille minimale était de 4,3° d'angle visuel horizontalement et de 5,6° d'angle visuel verticalement.

Matériel. Le matériel utilisé pour l'expérience 2 était exactement le même que celui de l'expérience 1.

Procédures. La tâche utilisée consistait à reconnaître visuellement des visages familiers présentés sur un écran d'ordinateur. L'étude s'est déroulée sur trois séances. La première séance comprenait un bloc de pratique suivi de deux blocs expérimentaux et durait environ une heure. Les deuxième et troisième séances incluaient chacune deux blocs expérimentaux et étaient d'une durée d'environ 45 minutes. Le bloc de pratique comportait 30 essais, soit 30 images de visages, et chaque bloc expérimental comprenait 200 essais, soit une répétition des mêmes 100 images présentées dans un ordre aléatoire. Comme dans la première expérience, 10 ordres différents des images du bloc expérimental avaient été préparés et préalablement assignés aux participants de manière aléatoire. En complétant six blocs expérimentaux, chaque participant a produit un total de 1200 essais.

Un essai débutait par une plage carrée de bruit visuel vide de 7,1 cm par 7,1 cm. Le reste de la procédure était identique à celle de l'expérience 1. Cependant, dans l'expérience 2, les participants devaient nommer, à voix haute, le nom de la célébrité présentée. S'ils ne connaissaient pas son nom ou n'arrivaient pas à le retrouver, ils avaient la possibilité de dire le nom d'un personnage interprété par l'acteur ou l'actrice cible (ex. Harry Potter pour Daniel

Radcliff), le nom d'un autre acteur/personnage avec lequel la célébrité a joué (ex. la patronne de James Bond pour Judi Dench), un ou des films dans lesquels la célébrité a joué (ex. l'acteur principal dans « Top Gun » pour Tom Cruise), ou des chansons écrites par la célébrité (ex. le chanteur de « Lose Yourself » pour Eminem). Ces critères d'identification ont été repris de la procédure utilisée par Butler, Blais, Gosselin, Bub et Fiset (2010).

Analyses statistiques. Les données de l'expérience 2 ont été analysées exactement de la même façon que celles de l'expérience 1.

Résultats

Information générale. La moyenne de bruit visuel inclut dans la portion « signal » du stimulus pour obtenir une performance de 50% correcte, incluant tous les participants et tous les essais, est de 12,0 (taux moyen de réponses correctes : 30,1%).

Vecteur de classification temporel. Sur la base du test Pixel (Chauvin, Worsley, Schyns, Arguin et Gosselin, 2005), une différence de 1,65 entre les scores Z est requise pour les considérer comme significativement différents l'un de l'autre. Le vecteur de classification moyen du domaine temporel (Figure 4) affiche ainsi neuf paliers (ou plateaux) significativement ($p < 0,05$) différents. Ce dernier montre que dans les 54 premières ms de présentation du stimulus, le niveau d'efficacité du traitement visuel reste stable à un score Z légèrement sous 0, puis augmente, formant un petit pic positif à 63 ms, pour immédiatement redescendre à un score Z de 0. Ensuite, à partir de 88 ms les scores Z grimpent jusqu'à atteindre un plateau d'efficacité maximale de 104 à 121 ms, pour ensuite redescendre abruptement vers des scores Z assez négatifs où s'affiche la plus faible efficacité de traitement

visuel qui s'étend de 146 à 163 ms. Enfin, l'efficacité remonte, se terminant à un score Z légèrement sous 0 comme au début de l'apparition du stimulus. L'indice de cohérence inter-sujets est de -0,02 et n'est pas significativement différent de 0 ($t(65) = -0.92, ns$).

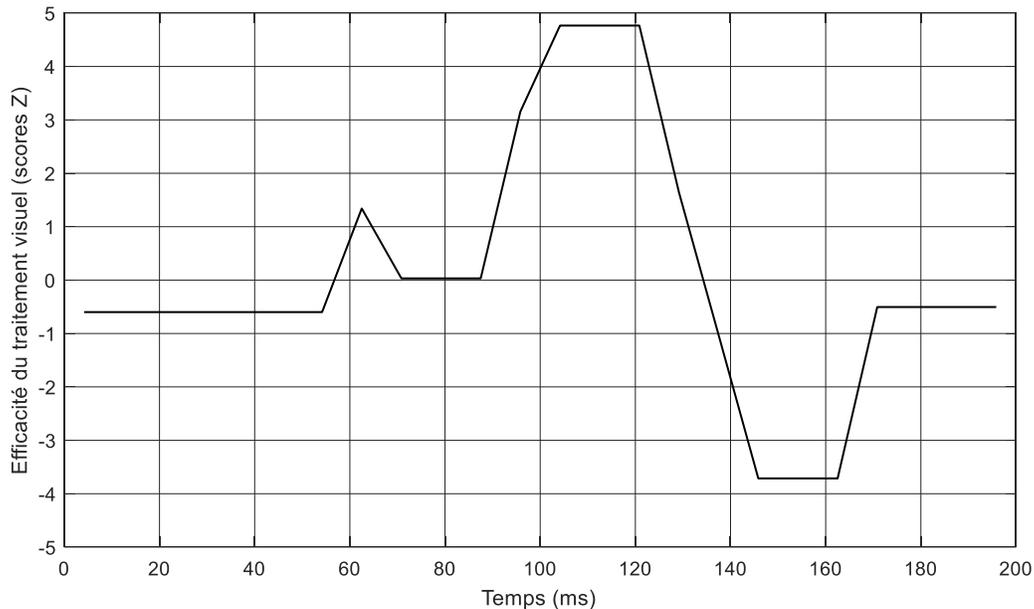


Figure 4. Vecteur de classification temporel moyen représentant les variations de l'efficacité du traitement visuel à travers le temps en reconnaissance de visages familiers (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau).

Image de classification temps-fréquence. Pour l'image de classification temps-fréquence moyenne (Figure 5), le test Pixel prescrit une différence minimale de 3,96 entre les scores Z pour les considérer comme significativement ($p < 0,05$) différents l'un de l'autre. Cette image de classification affiche qu'au début de l'apparition du stimulus (0 à 67 ms), les fréquences de 30 à 55 Hz sont associées à une haute efficacité de traitement visuel. Au contraire, de 33 à 100 ms, les fréquences de 20 à 45 Hz sont associées à une faible efficacité

de traitement visuel, tout comme les fréquences de 20 à 40 Hz entre 67 et 133 ms. De 100 à 167 ms, les fréquences de 25 à 55 Hz sont associées à un traitement visuel efficace. Finalement, de 133 à 200 ms, les fréquences de 20 à 55 Hz sont aussi favorables au traitement visuel. Les cellules temps-fréquence n'ayant pas été mentionnées sont considérées comme neutres, ou bien représentant une efficacité moyenne, étant donné le trop grand rapprochement entre les scores Z y étant obtenus et un score Z de 0 (identifié selon le seuil Pixel). L'indice de cohérence inter-sujets est de 0,57, ce qui est significativement différent de 0 ($t(65) = 16.04$, $p = .001$).

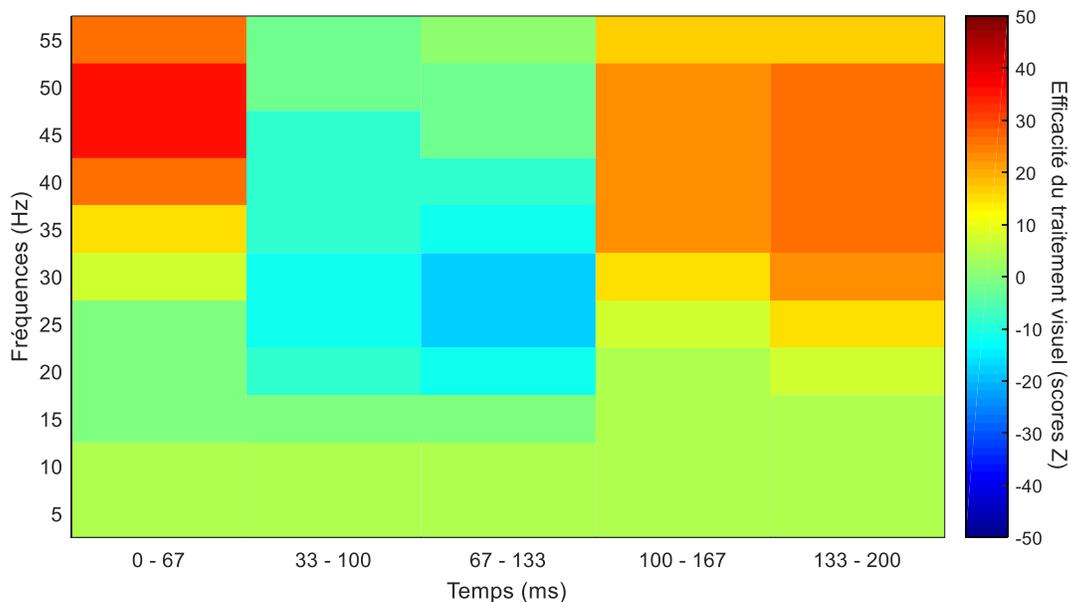


Figure 5. Image de classification temps-fréquence moyenne représentant les variations de l'efficacité du traitement visuel à travers le temps et les fréquences temporelles en reconnaissance de visages familiers (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau).

Image de classification phase-amplitude. L'image de classification moyenne dans le domaine phase-amplitude (Figure 6) ne contient aucun contraste significatif selon le test Pixel ($p < 0,05$; Chauvin, Worsley, Schyns, Arguin, Gosselin, 2005). C'est pourquoi cette image de classification est complètement uniforme, c'est-à-dire qu'elle n'affiche qu'une seule grande zone de valeur unique (ou plateau).

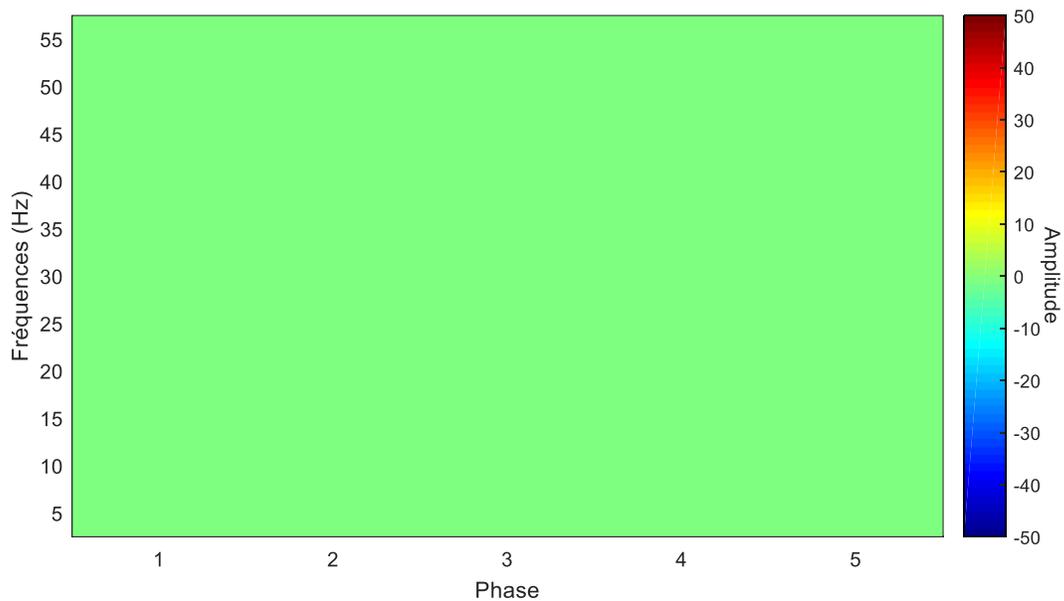


Figure 6. Image de classification moyenne dans le domaine phase-amplitude représentant les données utilisées pour créer les fonctions d'échantillonnage de chaque essai lors de l'expérience sur les visages (les scores Z voisins qui ne sont pas significativement ($p < 0,05$) différents l'un de l'autre ont été moyennés et reposent au même plateau).

Discussion

En ce qui a trait à la reconnaissance de visages familiers, nous avons vu que l'efficacité du traitement visuel semble être relativement moyenne jusqu'à environ 90 ms. Ensuite, cette efficacité augmente rapidement pour atteindre un pic autour de 115 ms, puis redescend tout aussi rapidement jusqu'à des niveaux d'efficacité significativement faible autour de 155 ms, et

finalement remonte légèrement vers une efficacité moyenne à la fin de la présentation du stimulus. Comme dans l'expérience 1, on peut penser que notre système visuel requiert un certain temps avant de pouvoir encoder l'information visuelle efficacement sur un visage qui vient d'apparaître dans notre champ visuel. Cependant, une étude sur la reconnaissance de visages célèbres a trouvé que le seuil minimal moyen de présentation est de 29,01 ms afin de bien reconnaître les visages (Kwon, Liu et Chien, 2016). D'un autre côté, les chercheurs ont choisi un taux de 80% de bonne réponse en tant que seuil. Il est possible qu'avant d'arriver à une performance parfaite (100%), le temps minimal de présentation se rapproche davantage de 90 ms. Ensuite, à partir de 90 ms nous pouvons remarquer que notre vecteur de classification moyen ressemble étrangement au tracé EEG obtenu lors de tâche de perception de visage. Évidemment, sachant qu'il y a un délai entre l'occurrence d'un événement dans le champ visuel et l'activité cérébrale qui en résulte, il ne s'agirait que d'une coïncidence. Il est tout de même intéressant de constater la concordance entre l'augmentation de l'efficacité du traitement visuel de 90 à 120 ms et la composante M100 (latence moyenne du pic à 105 ms, intervalle de 85 à 131 ms) mesurée avec une magnétoencéphalographie (MEG) et correspondant à la catégorisation correcte de visages (Liu, Harris et Kanwisher, 2002). Puis, la baisse observée dans l'efficacité aux alentours de 155 ms coïncide bien avec la M170 correspondant à l'identification correcte de visages selon les études utilisant la MEG (Liu, Harris et Kanwisher, 2002; Ewbank, Smith, Hancock et Andrews, 2008; Lueschow et al., 2015) ou la bien connue N170 liée à l'encodage structurel des visages selon les études utilisant l'EEG (Bentin, Allison, Puce, Perez et McCarthy, 1996; Eimer, 2000; Caharel, Courtay, Bernard, Lalonde et Rebaï, 2005). D'un autre côté, si la diminution de l'efficacité du traitement visuel observée sur notre vecteur temporel était liée aux composantes M170 et

N170, nous aurions théoriquement dû retrouver un pic positif dans l'efficacité du traitement aux alentours de 170 ms plutôt qu'un creux associé avec une faible efficacité. Dans tous les cas, de par l'état actuel des connaissances, il n'est pas possible de statuer quant à l'existence d'un réel lien entre ces changements dans l'efficacité du traitement visuel et l'activité corticale.

Par rapport au profil temps-fréquence, il semble que lorsqu'un visage est visible de manière périodique selon un rythme de haute fréquence (> 30 Hz), cela favorise la reconnaissance au tout début (0 à 67 ms) et à la fin de la présentation du stimulus, où les fréquences de 20-25 Hz et plus sont favorables au traitement visuel entre 100 et 200 ms. Puis, au contraire, les variations du ratio signal-bruit entre 20 et 40-45 Hz sont associées avec une faible efficacité à reconnaître des visages entre 33 et 133 ms. Pour ce qui est des basses fréquences, nous avons vu qu'elles ne sont ni associées avec une haute efficacité, ni avec une faible efficacité de traitement visuel, restant neutre ou, si vous préférez, à des niveaux d'efficacité moyens. Pourtant, dans l'étude de Vinette, Gosselin et Schyns (2004), où une technique d'échantillonnage temporel a aussi été utilisée, la performance de reconnaissance des visages variait selon un rythme de 10,6 Hz. De plus, dans une étude où la présentation d'une séquence de visages était contrôlée en suivant un rythme de 1 à 16 Hz, les chercheurs ont trouvé que la présentation de différents visages élicitait un plus grand potentiel évoqué visuel que la présentation d'un même visage répété, mais seulement lorsque les visages étaient présentés à un rythme entre 4 et 8,33 Hz avec une différence maximale à 5,88 Hz (Alonso-Prieto, Belle, Liu-Shuang, Norcia et Rossion, 2013). Ces résultats pourraient possiblement être expliqués par le fait que, dans ces études, les stimuli étaient des visages inconnus plutôt que familiers. Nous reviendrons sur la question dans la section « Comparaison entre types de

stimuli ». Puis, les autres études rapportant des mécanismes oscillatoires de basses fréquences dans la perception visuelle en général ne permettent pas non plus d'expliquer les résultats retrouvés ici pour les visages familiers. Nous tenterons de voir pourquoi, encore une fois, dans la section « Comparaison entre types de stimuli ».

Néanmoins, nous pouvons faire un parallèle avec des études en EEG/MEG. Entre autres, la présentation de visages « Mooney » activerait une forte réponse gamma avant 200 ms dans le gyrus fusiforme et le cortex occipital latéral adjacent (Lachaux et al., 2005), ce qui pourrait correspondre avec la haute efficacité de traitement visuel retrouvée au niveau des hautes fréquences à la fin de la présentation du stimulus dans notre étude. Une activation dans la bande gamma (25 à 100 Hz) en réponse à des visages a aussi été retrouvée dans une étude plus récente, entre 150 et 350 ms après l'apparition du stimulus, chez des participants contrôles (Liu et al., 2018), ce qui vient appuyer le point précédent. Par contre, cette dernière étude rapporte aussi une plus grande activation dans la bande alpha (8 à 14 Hz) entre 50 et 300 ms et dans la bande bêta (15 à 20 Hz) entre 100 et 200 ms lors de la présentation de visages comparativement à des objets (Liu et al., 2018), ce qui ne correspond pas nécessairement à nos résultats. Une implication des basses fréquences a aussi été démontrée dans une étude où une synchronisation reliée aux événements dans les bandes de fréquences thêta (4 à 7 Hz), alpha (8 à 12 Hz) et bêta (13 à 29 Hz) a été retrouvée dans les aires occipito-temporales entre 0 et 200 ms en réponse à des visages par rapport à des images brouillées (Sakihara, Gunji, Furushima et Inagaki, 2012). Ensuite, une autre étude a trouvé des modulations dans les fréquences en deçà de 30 Hz où la présentation d'images de visages provoquait une augmentation de l'amplitude entre 20 et 30 Hz autour de 80 ms suivie d'une diminution de l'amplitude entre 10 et 25 Hz dans le gyrus occipital inférieur gauche (avec un effet similaire dans le gyrus

occipital inférieur droit) (Foley, Rippon et Senior, 2017). Ces trois derniers résultats viennent à l'encontre de ce qui a été énoncé dans la présente étude en mettant l'emphase sur les basses fréquences. Encore une fois, deux de ces études n'ont utilisé que des visages inconnus et la troisième s'est concentrée spécifiquement sur la synchronisation des oscillations cérébrales. Notre étude étant assez unique en son genre, il est difficile de relier nos résultats avec ceux des autres études sur la perception de visages.

Enfin, au niveau du profil phase-amplitude, nous avons confirmé notre hypothèse. Tout comme pour les objets familiers, l'image de classification est complètement uniforme démontrant que ce code de valeurs n'est pas apte à capturer l'évolution temporelle du traitement visuel.

Comparaison entre types de stimuli

Méthodologie

Analyses statistiques. Des analyses additionnelles ont été effectuées afin de déterminer si les vecteurs/images de classification obtenus avec les objets familiers, les objets non familiers (Arguin et Massé, 2019), les visages familiers et les mots (Arguin, 2018) sont différents ou non. À cette fin, nous avons effectué des tests t avec correction de Bonferroni afin de maintenir le risque total d'erreur α à 0,05 pour chaque comparaison de vecteurs/images de classification entre deux expériences. Cette opération a été répétée avec toutes les paires possibles de stimuli (la comparaison entre les mots et les objets non familiers a été réalisée à titre de vérification). Puisque la procédure de l'étude sur les visages non familiers (Blais, Arguin et Gosselin, 2013) est un peu différente de celle utilisée avec les autres types de stimuli (les fréquences utilisées pour créer les fonctions d'échantillonnage déterminant la variation du ratio signal-bruit vont de 5 à 20 Hz au lieu de 5 à 55 Hz), il ne serait pas avisé de comparer statistiquement les profils obtenus dans cette étude avec ceux obtenus dans la présente étude. La comparaison entre visages familiers et non familiers n'est alors que qualitative et plutôt exploratoire.

Résultats

Vecteurs de classification temporels. Avec une correction de Bonferroni (24 comparaisons) pour un seuil $\alpha = 0,05$ pour l'ensemble des comparaisons, la valeur de p requise pour considérer une différence significative est de 0,002. Les contrastes entre les profils temporels moyens bruts obtenus avec les objets familiers, les objets non familiers et les

visages familiers (Figure 7; lignes jaune, verte et mauve respectivement) ne montrent aucune différence significative. Par contre, le profil temporel moyen brut obtenu avec les mots (Figure 7; ligne bleue) se distingue de ceux observés avec les trois autres types de stimuli, et ce à trois moments. Les scores Z pour les mots sont significativement plus élevés que ceux pour les objets familiers à 29 ms, de 46 à 54 ms et à 88 ms. Ils sont également plus élevés que ceux pour les objets non familiers à 54 ms ainsi que de 79 à 88 ms, et que ceux pour les visages familiers de 46 à 54 ms et de 79 à 88 ms.

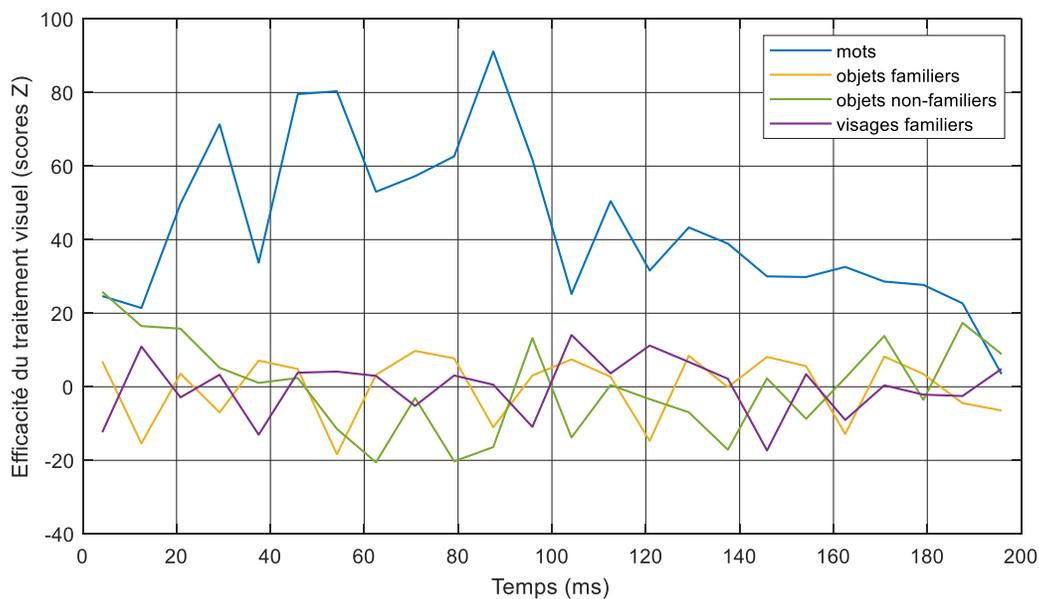


Figure 7. Profils temporels moyens bruts des mots (bleu), des objets familiers (jaune), des objets non familiers (vert) et des visages familiers (mauve).

Images de classification temps-fréquence. La comparaison entre les objets familiers et non familiers démontre 37 cellules temps-fréquence significativement différentes ($dl = 22$, $p < 0,0009$, avec correction de Bonferroni) à travers les 55 tests t réalisés (Figure 8). Parmi ces différences, quatre grandes étendues temps-fréquence peuvent être distinguées. Premièrement,

un ensemble de cellules correspondant à de basses fréquences (5-10 Hz) démontre une efficacité plus élevée en reconnaissance d'objets familiers que non familiers, et ce sur toute la durée de présentation du stimulus. Deuxièmement, un second ensemble de cellules correspondant à des fréquences intermédiaires à élevées (20 à 55 Hz) affiche un avantage significatif pour les objets familiers par rapport aux objets non familiers pour des périodes de temps limitées, commençant dès le début de la présentation du stimulus, pour les fréquences les plus élevées et pour des périodes de temps qui graduellement s'étalent ou se décalent dans le temps jusqu'à 167 ms pour les plus basses fréquences. Troisièmement, un ensemble de cellules représentant de hautes fréquences (40 à 55 Hz) démontre, au contraire, un avantage significatif pour le traitement visuel des objets non familiers relativement aux objets familiers de 33 à 167 ms, concernant initialement que le 45 Hz et progressant vers une plus grande étendue de fréquence (40 à 50 Hz), puis vers de plus hautes fréquences (45 à 55 Hz). Quatrièmement, un ensemble de cellules correspondant toujours aux fréquences élevées (45 à 55 Hz) indique ensuite un traitement significativement plus efficace avec les objets familiers en comparaison aux objets non familiers à la fin de la présentation du stimulus (133 à 200 ms).

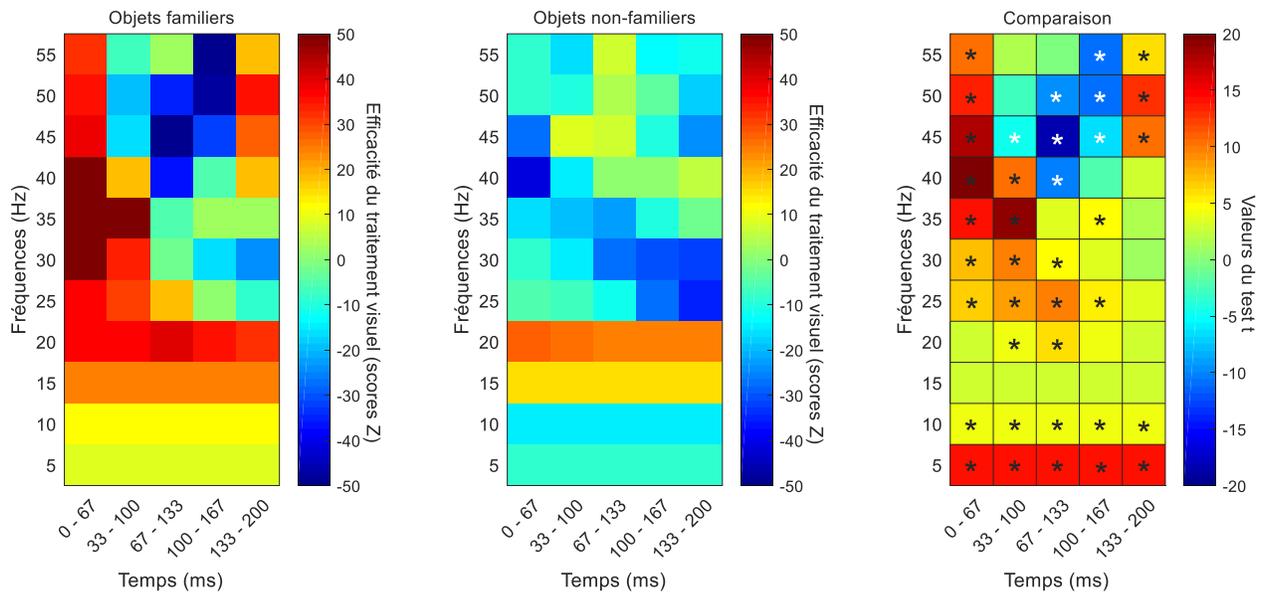


Figure 8. Comparaison statistique entre les profils temps-fréquence des objets familiers et non familiers; image de classification temps-fréquence moyenne brute des objets familiers (à gauche), image de classification temps-fréquence moyenne brute des objets non familiers (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les objets familiers qu’avec les objets non familiers, et un * blanc lorsque la différence significative est en direction opposée.

La comparaison entre les objets familiers et les visages familiers démontre 31 cellules temps-fréquence significativement différentes ($dl = 22$, $p < 0,0009$, avec correction de Bonferroni) à travers les 55 tests t réalisés (Figure 9). Parmi ces différences, trois grandes étendues temps-fréquence peuvent être distinguées. D’abord, un ensemble de cellules correspondant à des fréquences intermédiaires (15 à 40 Hz) démontrent une efficacité de traitement visuel plus élevée en reconnaissance d’objets familiers qu’en reconnaissance de visages familiers pour des périodes de temps limitées en ce qui concerne les fréquences les plus élevées et pour des périodes de temps qui graduellement s’étalent jusqu’à couvrir la durée complète de présentation du stimulus pour les plus basses fréquences. Ensuite, un second

ensemble de cellules correspondant à des fréquences plus élevées (30 à 55 Hz) et à des périodes de temps plus tardives (100 à 167 ms et aux alentours pour certaines fréquences) démontrent, au contraire, une efficacité de traitement visuel plus élevée avec les visages familiers qu’avec les objets familiers. Puis, deux cellules temps-fréquence isolées (50-55 Hz de 0 à 67 ms) indiquent également un avantage significatif pour les visages familiers relativement aux objets familiers.

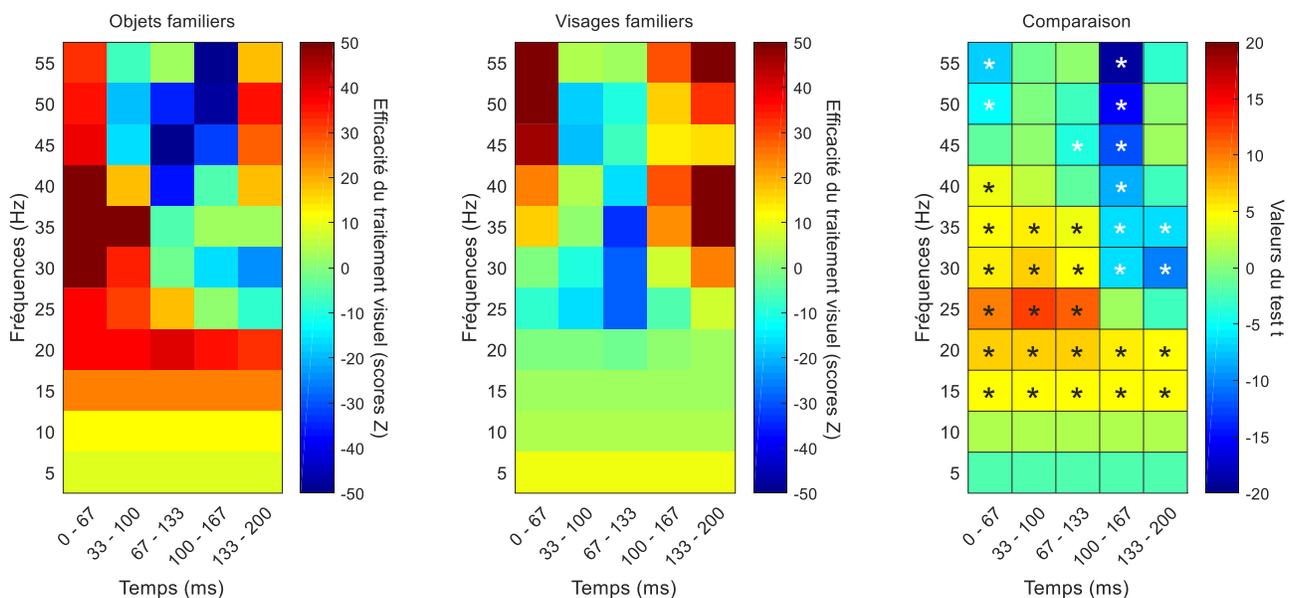


Figure 9. Comparaison statistique entre les profils temps-fréquence des objets et des visages familiers ; image de classification temps-fréquence moyenne brute des objets familiers (à gauche), image de classification temps-fréquence moyenne brute des visages familiers (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les objets familiers qu’avec les visages familiers, et un * blanc lorsque la différence significative est en direction opposée.

La comparaison entre les visages familiers et les objets non familiers dévoile 26 cellules temps-fréquence significativement différentes ($dl = 22, p < 0,0009$, avec correction de Bonferroni) à travers les 55 tests t réalisés (Figure 10). Parmi ces différences, nous pouvons

distinguer cinq ensembles de cellules temps-fréquence. Premièrement, un ensemble de cellules correspondant au 5 Hz démontre une plus grande efficacité de traitement visuel en reconnaissance de visages familiers qu'en reconnaissance d'objets non familiers sur toute la durée de présentation du stimulus. Deuxièmement, un ensemble de cellules associées aux fréquences élevées (40 à 55 Hz) affiche aussi un avantage significatif avec les visages familiers relativement aux objets non familiers au tout début de la présentation du stimulus (0 à 67 ms). Troisièmement, un petit ensemble de cellules correspondant au 20 Hz démontre, au contraire, un avantage significatif avec les objets non familiers par rapport aux visages familiers de 0 à 133 ms, incluant aussi le 25 Hz à partir de 67 ms. Quatrièmement, une cellule isolée représentant le 45 Hz de 33 à 100 ms montre aussi une efficacité de traitement plus élevée pour les objets non familiers que pour les visages familiers. Cinquièmement, un ensemble de cellules correspondant à des fréquences intermédiaires à élevées (25 à 55 Hz) démontre un avantage significatif avec les visages familiers relativement aux objets non familiers vers la fin de la présentation du stimulus (100 à 200 ms).

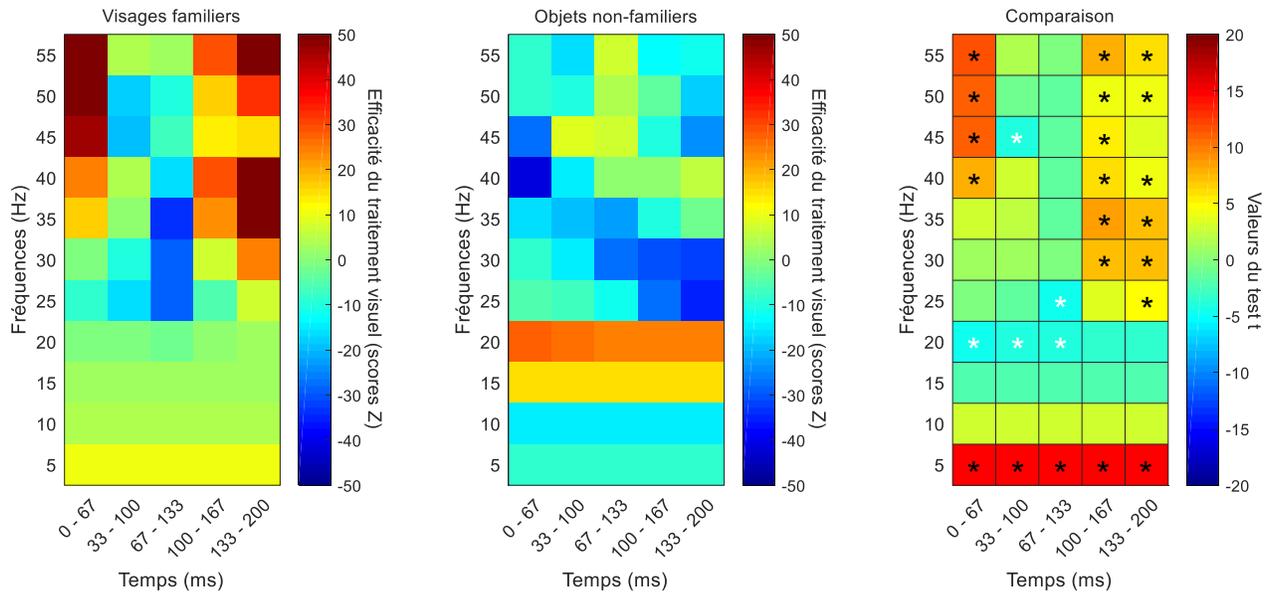


Figure 10. Comparaison statistique entre les profils temps-fréquence des visages familiers et des objets non familiers; image de classification temps-fréquence moyenne brute des visages familiers (à gauche), image de classification temps-fréquence moyenne brute des objets non familiers (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les visages familiers qu'avec les objets non familiers, et un * blanc lorsque la différence significative est en direction opposée.

La comparaison entre les objets familiers et les mots dévoile 25 cellules temps-fréquence significativement différentes ($dl = 18, p < 0,0009$, avec correction de Bonferroni) à travers les 55 tests t réalisés (Figure 11). Parmi ces différences, trois ensembles de cellules peuvent être dissociés, mais représentent tous une efficacité de traitement visuel plus élevée en reconnaissance de mots qu'en reconnaissance d'objets familiers. D'abord, nous avons un ensemble de cellules correspondant à de basses fréquences (5 à 15 Hz) qui s'étale sur toute la durée de présentation du stimulus, incluant également le 20 Hz au tout début (0 à 67 ms). Ensuite, nous avons un ensemble de cellules correspondant aux hautes fréquences (45 à 55 Hz) entre 33 et 133 ms. Puis, nous avons un troisième ensemble de cellules représentant des

fréquences intermédiaires (25 à 35 Hz) à des périodes de temps plus tardives (100 à 200 ms) qui progresse des plus basses fréquences vers les plus hautes.

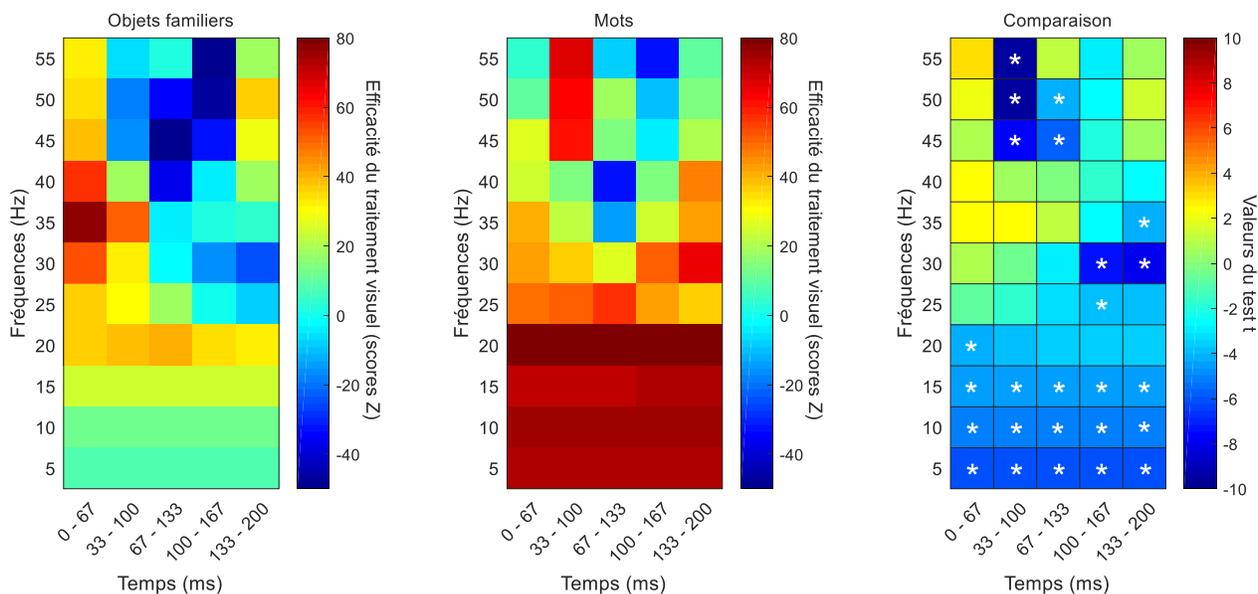


Figure 11. Comparaison statistique entre les profils temps-fréquence des objets familiers et des mots; image de classification temps-fréquence moyenne brute des objets familiers (à gauche), image de classification temps-fréquence moyenne brute des mots (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les objets familiers qu’avec les mots, et un * blanc lorsque la différence significative est en direction opposée. À noter ; l’étendue des scores Z pour les images de classification est de -50 à 80 afin de mieux représenter la différence visuellement.

La comparaison entre les visages familiers et les mots démontre 32 cellules temps-fréquence significativement différentes ($dl = 18, p < 0,0009$, avec correction de Bonferroni) à travers les 55 tests t réalisés (Figure 12). Parmi ces différences, nous pouvons dégager quatre ensembles de cellules. D’abord, un grand ensemble de cellules correspondant aux fréquences basses à intermédiaires (5 à 30 Hz) démontre une efficacité de traitement visuel plus élevée en reconnaissance de mots qu’en reconnaissance de visages familiers sur toute l’étendue de

présentation du stimulus pour les fréquences de 20 Hz et moins et pour des périodes de temps de plus en plus isolées en montant vers le 25 et 30 Hz. En ce qui concerne les hautes fréquences, nous avons un premier ensemble de cellules (50-55 Hz) au début de la présentation du stimulus (0 à 67 ms) affichant un avantage significatif avec les visages familiers relativement aux mots. Puis, nous avons un second ensemble de cellules (45 à 55 Hz) dans la période de temps suivante (33 à 100 ms) qui, au contraire, affiche un avantage significatif pour les mots par rapport aux visages familiers. Enfin, nous avons une cellule isolée (55 Hz de 100 à 167 ms) affichant un avantage significatif retournant aux visages familiers relativement aux mots.

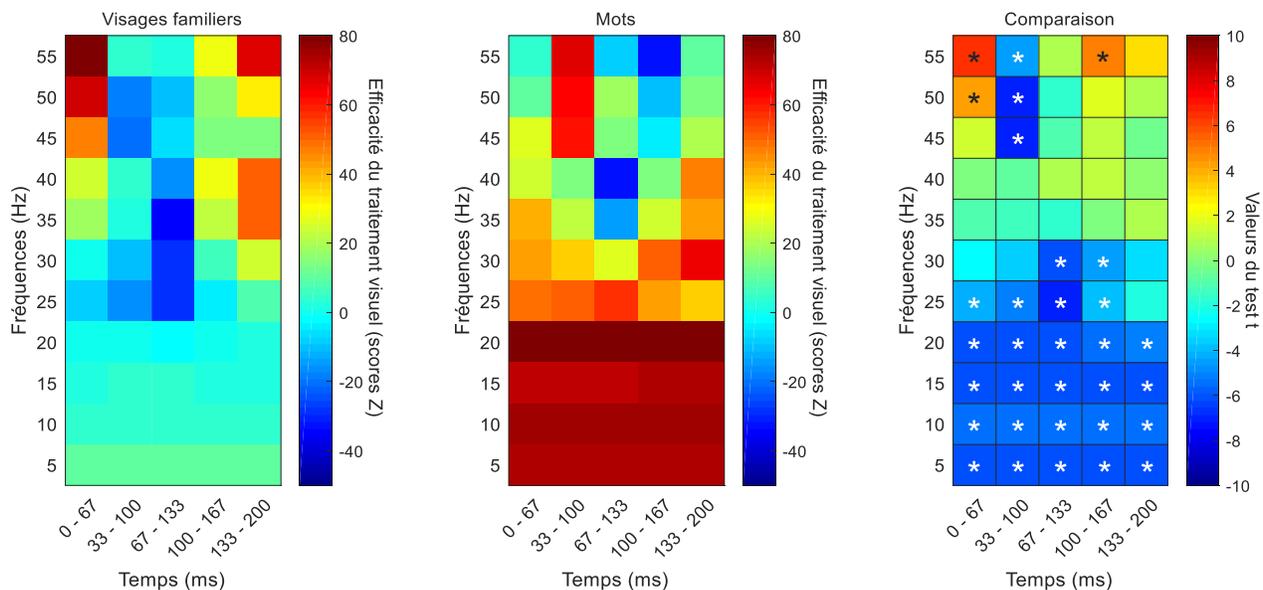


Figure 12. Comparaison statistique entre les profils temps-fréquence des visages familiers et des mots; image de classification temps-fréquence moyenne brute des visages familiers (à gauche), image de classification temps-fréquence moyenne brute des mots (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les visages familiers qu'avec les mots, et un * blanc lorsque la différence significative est en direction opposée. À noter ; l'étendue des scores Z pour les images de classification est de -50 à 80 afin de mieux représenter la différence visuellement.

Finalement, la comparaison entre les objets non familiers et les mots, réalisée à titre de vérification, dévoile 33 cellules temps-fréquence significativement différentes ($dl = 18$, $p < 0,0009$, avec correction de Bonferroni) à travers les 55 tests t réalisés (Figure 13). Parmi ces différences, trois ensembles de cellules temps-fréquence peuvent être dissociés, mais représentent tous une efficacité de traitement visuel plus élevée en reconnaissance de mots qu'en reconnaissance d'objets non familiers. D'abord, nous avons un grand ensemble de cellules correspondant à des fréquences basses à intermédiaires (5 à 35 Hz) s'étalant sur toute la durée de présentation du stimulus pour les fréquences de 20 Hz et moins et progressant vers des périodes de temps débutant de plus en plus tardivement pour les plus hautes fréquences. Ensuite, au niveau des hautes fréquences, nous avons une cellule isolée (40 Hz) de 0 à 67 ms et un ensemble de cellules (45 à 55 Hz) dans la période de temps de 33 à 100 ms.

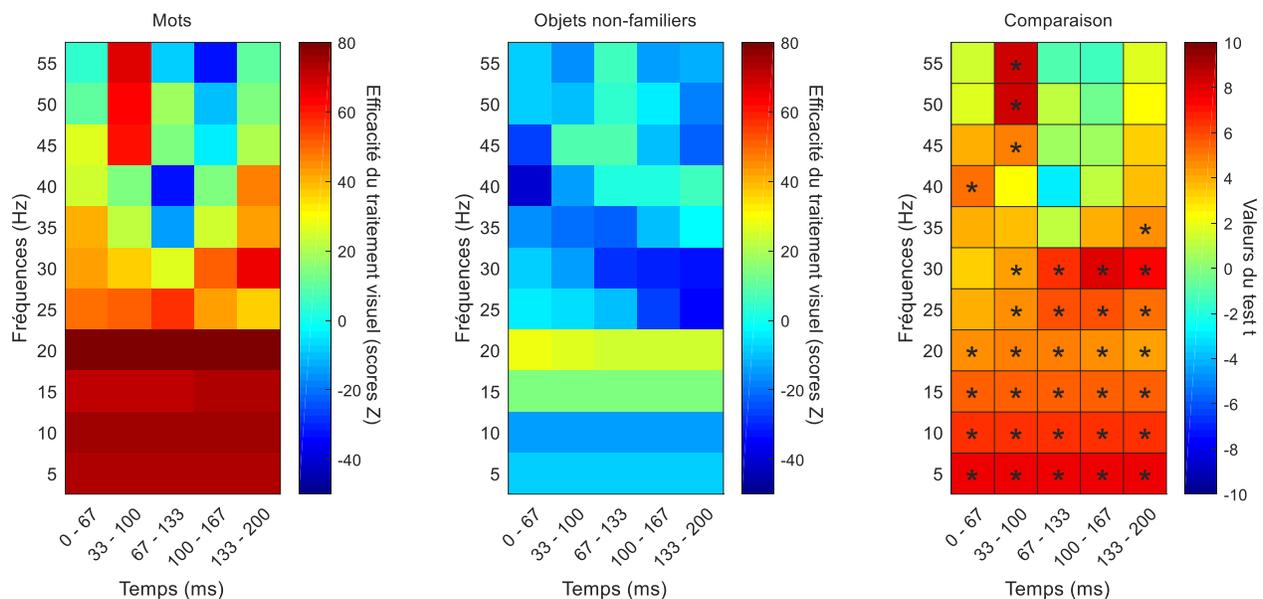


Figure 13. Comparaison statistique entre les profils temps-fréquence des mots et des objets non familiers; image de classification temps-fréquence moyenne brute des mots (à gauche), image de classification temps-fréquence moyenne brute des objets non familiers (au milieu) et résultats des tests t (à droite). Les différences significatives ($p < 0,0009$) sont identifiées par un * noir lorsque la différence significative indique une efficacité de traitement visuel supérieure avec les mots qu'avec les objets non familiers,

et un * blanc lorsque la différence significative est en direction opposée. À noter ; l'étendue des scores Z pour les images de classification est de -50 à 80 afin de mieux représenter la différence visuellement.

Images de classification phase-amplitude. Étant donné que les images de classification dans le domaine phase-amplitude sont uniformément non-significative ($p < 0,05$) pour tous les types de stimuli, aucune comparaison à travers les différentes tâches ou classes de stimuli n'a été effectuée.

Discussion

En ce qui concerne le domaine temporel, nous avons vu qu'il n'y a que le vecteur de classification obtenu avec les mots qui se distingue des autres types de stimuli à quelques moments seulement. Cela est compréhensible puisque la cohérence inter-sujets obtenue pour les vecteurs de classification des objets familiers, des visages familiers et des objets non familiers est nulle. Ainsi, cela augmente la variance d'erreur, ce qui nous amène à des contrastes non significatifs. Dans le domaine temporel, l'indice de cohérence inter-sujets obtenu avec les mots (0,18) est le seul à être significativement différent de 0 ($t(27) = 5.57$, $p = .001$). Malgré tout, le fait que les mots sont construits d'un nombre limité de traits relativement simples peut aussi amener une réelle différence, c'est-à-dire une efficacité de traitement visuel plus élevée en reconnaissance de mots qu'en reconnaissance d'objets ou de visages, puisque ces derniers n'ont pas de contrainte en termes de nombre et de complexité des traits qui les composent. De plus, la plus grande efficacité de traitement visuel rapportée pour les mots pourrait être expliquée par l'implication de processus cognitifs différents. En effet, il s'agit ici de lecture de mots comparativement à de la reconnaissance d'images comme telle.

Bien sûr, dans tous les cas, nous devons opérer une reconnaissance de forme, mais les lettres sont des formes très faciles à reconnaître étant donné toute la pratique que nous avons avec ces stimuli, un phénomène connu depuis longtemps dans la littérature sur l'automatisme de la lecture (De Soto et De Soto, 1985; Horn et Manis, 1987; Stroop, 1935).

Avant d'entrer dans les comparaisons entre chaque paire de stimuli, il est impressionnant de voir que chaque image de classification temps-fréquence, à travers tous les types de stimuli, diffère des autres en plusieurs points, et ce considérant la cohérence inter-sujets assez élevée (0,93 pour objets familiers, 0,87 pour mots (Arguin, 2018), 0,67 pour objets non familiers (Arguin et Massé, 2019) et 0,57 pour visages familiers). On pourrait penser que l'indice de cohérence inter-sujets obtenu avec les visages familiers est plutôt faible relativement à celui obtenu pour les objets familiers, mais considérant la variabilité interindividuelle connue en reconnaissance de visages, cela n'est pas alarmant. En effet, les habiletés à reconnaître un visage varient d'un individu à l'autre, pouvant être représentés sur un spectre allant de personnes souffrant de prosopagnosie développementale jusqu'aux super-reconnaisseurs, et ce niveau d'habileté influence la manière dont les individus utilisent l'information visuelle (Tardif et al., 2019). Une autre étude a démontré, par exemple, que les prosopagnosiques font des fixations plus longues et à différents endroits du visage par rapport aux neurotypiques (De Luca et al., 2019). Stacchi, Ramon, Lao et Caldara (2019) ont aussi montré l'existence de patrons d'échantillonnage de l'information visuelle idiosyncratiques en reconnaissance de visage, corrélés avec les réponses neuronales différant aussi à travers les participants. Puis, il y a aussi de la variabilité en fonction de l'âge, où les aînés feraient davantage de fixations visuelles que les jeunes adultes en reconnaissance de visage (Liu, Shen, Olsen et Ryan, 2018), et en fonction de la culture, où les Occidentaux et les Orientaux se

différencieraient au niveau de leurs mouvements oculaires (Caldara, 2017; Senju et al., 2013). Donc, un indice de cohérence inter-sujets de 0,57 peut être considéré comme relativement élevé pour une tâche de reconnaissance de visages et peut-être même comparable à celui obtenu avec les objets familiers.

La comparaison entre l'image de classification temps-fréquence obtenue avec les objets familiers et celle obtenue avec les objets non familiers est intéressante dans le sens où elle nous permet d'évaluer l'effet de la familiarité du stimulus. Il est vrai que les différences observées entre ces deux images de classification pourraient être attribuables aux différences importantes entre les caractéristiques visuelles des objets présentés. Dans le cas des objets familiers, nous avons présenté des objets de la vie quotidienne tels que des outils, des légumes, des jouets et autres, alors que dans l'étude sur les objets non familiers, ce sont des objets synthétiques 3D de formes abstraites qui ont été utilisés (voir Annexe A et C pour exemples de stimuli). Nous savons que les caractéristiques de base du stimulus peuvent modifier nos processus neuronaux. Entre autres, la forme du stimulus (Ito, Sugata et Kuwabara, 1998), le nombre de caractéristiques (Fournier, Scheffers, Coles, Adamson et Abad, 2000) et la complexité visuelle de l'objet (Gerlach et Marques, 2014) ont un impact sur notre performance et notre activité cérébrale. Par contre, la familiarité des objets pourrait aussi expliquer les différences observées entre les deux images de classification temps-fréquence. Selon l'effet de supériorité de l'objet, les attributs visuels sont encodés plus efficacement lorsqu'on a affaire à des contextes familiers que lorsque ce sont des contextes non familiers. Cet effet a d'abord été rapporté dans une tâche de reconnaissance d'orientation et de position de lignes où les performances étaient meilleures lorsque la ligne faisait partie d'une configuration paraissant unitaire et tridimensionnelle que lorsqu'elle faisait partie d'une configuration plus abstraite et

compliquée (Weisstein et Harris, 1974). Puis, dans une tâche similaire, une autre étude a montré que l'émergence de caractéristiques dans la figure est aussi importante, autrement dit, si la ligne cible est positionnée de manière à créer une forme dans la figure, telle qu'un triangle dans le coin d'un rectangle, elle sera plus facilement encodée que si elle est incluse dans la figure sans y être reliée, telle qu'une ligne au milieu d'un rectangle (Lanze, Maguire et Weisstein, 1985). Ainsi, il est possible que les attributs visuels constituant les objets familiers soient plus facilement encodés que ceux constituant les objets non familiers. Cela expliquerait pourquoi les images de classification de ces deux stimuli présentent plusieurs différences, mais aussi pourquoi ces différences sont, en grande majorité, à l'avantage des objets familiers, c'est-à-dire que l'efficacité du traitement visuel est globalement plus élevée en reconnaissance d'objets familiers relativement aux objets non familiers.

Toujours dans l'objectif d'évaluer l'impact de la familiarité sur les profils temps-fréquence d'efficacité du traitement visuel, voyons ensemble le contraste entre l'image de classification obtenue avec les visages familiers et les résultats obtenus dans l'étude de Blais, Arguin et Gosselin (2013) avec des visages non familiers. Ces chercheurs ont trouvé des corrélations positives entre la performance de reconnaissance et un ratio signal-bruit variant selon des fréquences de 10 et 15 Hz. Au contraire, dans la présente étude, nous avons trouvé qu'avec les visages familiers, les fréquences de 15 Hz et moins sont associées avec une efficacité de traitement visuel neutre ou moyenne sur toute la durée de présentation du stimulus. Ce sont plutôt les hautes fréquences d'environ 25 à 55 Hz qui facilitent le plus la reconnaissance de visages au tout début et à la fin de la présentation de ces derniers. Il faut tout de même rappeler que Blais, Arguin et Gosselin ont fait varier le ratio signal-bruit à des rythmes de 5 à 20 Hz (par pas de 5 Hz) uniquement, alors que nous avons utilisé des rythmes

de 5 à 55 Hz. Il est bien possible que cela explique la différence entre ces deux études. Par contre, il est aussi probable qu'une réelle différence existe et qu'elle puisse être expliquée par un processus similaire à celui sollicité dans l'effet de supériorité de l'objet décrit précédemment. Ainsi, les traits d'un visage connu pourraient potentiellement être encodés plus facilement que ceux d'un visage inconnu. La différence analysée ici n'est, cependant, pas appuyée par la littérature. Entre autres, dans l'étude de Sakihara, Gunji, Furushima et Inagaki (2012), seulement la synchronisation reliée aux événements à un rythme dans la bande de fréquences delta (1 à 4 Hz) était plus grande dans la région temporale gauche avec les visages familiers comparativement aux visages non familiers. Dans cette étude, aucune autre bande de fréquences ne montrait de différence entre les visages familiers et non familiers dans la fenêtre de temps 0-600 ms. Puis, dans une étude de Collins, Robinson et Behrmann (2018), où plusieurs visages étaient présentés à un rythme de 15 Hz, les visages familiers évoquaient un signal EEG plus élevé que les visages non familiers. Au final, l'effet de la familiarité des visages reste incertain et mériterait d'être investigué plus spécifiquement.

Maintenant, penchons-nous sur le contraste entre les objets familiers et les visages familiers. Cette comparaison est particulière, nous permettant d'investiguer l'effet de la classe du stimulus en contrôlant pour la familiarité, puisque nous avons deux stimuli reconnaissables par les participants. Les nombreuses différences obtenues entre ces deux images de classification semblent ainsi attribuables au fait que les systèmes neuronaux impliqués dans la reconnaissance des objets sont partiellement distincts de ceux utilisés dans la reconnaissance de visages. En effet, il est bien connu qu'une spécialisation fonctionnelle cérébrale s'opère dans le cortex occipito-temporal où différentes régions traitent les visages, les scènes, les objets (Grill-Spector, Golarai et Gabrieli, 2008), les parties du corps (Schwarzlose, Swisher,

Dang et Kanwisher, 2008; Spiridon, Fischl et Kanwisher, 2006) et même les mots (Grill-Spector et Weiner, 2014). Ainsi, l'aire fusiforme des visages (Coggan, Baker et Andrews, 2019; Gao, Conte, Richards, Xie et Hanayik, 2019; Grill-Spector et al., 2008; Grill-Spector, Knouf et Kanwisher, 2004; Grill-Spector et Weiner, 2014; Idaka, 2014; Kanwisher, McDermott et Chun, 1997; Schwarzlose et al., 2008; Spiridon et al., 2006; Zhang, Li, Song et Liu, 2012) dans le gyrus fusiforme, l'aire occipitale des visages (Ambrus, Windel, Burton et Kovács, 2017; Grill-Spector et al., 2008; Grill-Spector et Weiner, 2014; Pitcher, Walsh et Duchaine, 2011; Schwarzlose et al., 2008; Spiridon et al., 2006) dans le gyrus occipital inférieur et une partie du sulcus temporal supérieur postérieur (Grill-Spector et al., 2008; Idaka, 2014; Schobert, Corradi-Dell'Acqua, Frühholz, van der Zwaag et Vuilleumier, 2018) sont des régions qui répondent préférentiellement aux visages. Puis, une région plus étendue, s'étalant du cortex occipital inférieur jusqu'au cortex temporal, appelée le complexe occipital latéral, s'active davantage lorsque des objets sont visibles comparativement à d'autres types de stimuli (Grill-Spector et al., 2008; Grill-Spector, Kourtzi et Kanwisher, 2001; Grill-Spector et Weiner, 2014, Malach et al., 1995; Schwarzlose et al., 2008; Spiridon et al., 2006; Vinberg et Grill-Spector, 2008). De plus, malgré la forte association entre la prosopagnosie congénitale et l'agnosie visuelle d'objet, qui sont rapportées chez 80 % des individus figurant dans les études (de 1976 à 2016) ayant mesuré les habiletés de reconnaissance de visages et d'objets en considérant l'exactitude et le temps de réponse, il reste tout de même un 20% d'individus prosopagnosiques n'ayant aucun trouble en reconnaissance d'objets (Geskin et Behrmann, 2018). Ainsi, indépendamment de s'il s'agit d'un système neuronal plus général s'occupant du traitement visuel des objets, incluant un sous-système pour le traitement des visages, ou s'il

s'agit de deux systèmes séparés, ce résultat atteste du fait qu'il y a au moins une partie du traitement neuronal des objets qui se distingue de celui des visages.

Cette interprétation s'appliquerait aussi pour les autres comparaisons effectuées entre les objets, les visages et les mots. En effet, dans tous les cas, plusieurs différences ont été rapportées au niveau de l'efficacité du traitement visuel à travers le temps et les fréquences de présentation de la cible. Tel que mentionné précédemment, la présentation de mots provoque aussi l'activation préférentielle d'une région cérébrale spécifique. En fait, il s'agirait d'une section du gyrus fusiforme gauche, nommée l'aire des formes de mots visuels (ou « visual word form area » en anglais), qui répond de manière significativement plus élevée pour des suites de lettres et des symboles que pour d'autres types de stimuli (Centanni, King, Eddy, Whitfield-Gabrieli et Gabrieli, 2017; Cohen et al., 2000; Dehaene, Clec'H, Poline, Bihan et Cohen, 2002; Grill-Spector et Weiner, 2014; Reinke, Fernandes, Schwindt, O'Craven et Grady, 2008). Ainsi, puisque les objets, les visages et les mots sont traités dans des régions cérébrales partiellement distinctes, il n'est pas étonnant de retrouver des différences dans les profils d'échantillonnage de l'information visuelle fait par le cerveau lors du traitement de ces derniers. De plus, dans tous les contrastes réalisés avec l'image de classification des mots, la grande majorité, sinon toutes les différences sont à l'avantage des mots, c'est-à-dire que l'efficacité du traitement visuel est plus grande en reconnaissance de mots qu'en reconnaissance d'objets ou de visages. Comme pour le domaine temporel, le fait que les mots sont construits d'un nombre limité de traits relativement simples pourrait expliquer cette plus grande efficacité de traitement en reconnaissance de mots, considérant que les objets et les visages n'ont pas de contrainte en termes de nombre et de complexité des traits qui les composent. De plus, on pourrait penser que l'automaticité de la lecture (De Soto et De Soto,

1985; Horn et Manis, 1987; Stroop, 1935) amène les participants à performer mieux, ce qui se traduirait, dans ce cas-ci, par une forte capacité de reconnaissance de mots à travers un plus grand registre de fréquence et sur de plus longues périodes de temps relativement aux autres types de stimuli.

Au final, il est vraiment intéressant de constater que la méthodologie appliquée nous permet de faire ressortir des profils, ou plus spécifiquement des images de classification temps-fréquence, capable de discriminer à travers les tâches effectuées par les participants. D'autant plus, nous avons vu que ces images de classification sont sensibles à la classe de stimuli présentés, mais aussi à la familiarité de ces derniers.

Discussion générale

L'objectif premier de la présente étude était d'établir les profils temporel et temps-fréquence de l'efficacité du traitement visuel en reconnaissance d'objets et de visages familiers. Cela nous a permis de confirmer l'existence de variation temporelle rapide dans l'efficacité de l'encodage visuel durant les 200 premières ms d'exposition au stimulus. Les profils temporels sont, par contre, très variables d'un participant à l'autre, que ce soit en reconnaissance d'objets ou de visages familiers. D'un autre côté, les profils temps-fréquence sont remarquablement similaires à travers les participants. Cela suggère qu'à travers le temps, les meilleurs moments pour traiter l'information visuelle ne sont pas les mêmes entre les individus, mais que la rythmicité de l'accès à l'information visuelle la plus adéquate pour bien reconnaître un stimulus à différents intervalles de temps est pratiquement la même pour tous. Cela appuie aussi l'hypothèse voulant que l'efficacité du traitement visuel varie en fonction des oscillations cérébrales, qui suivent aussi différents rythmes, et que nos images de classification temps-fréquence arrivent à capter ces oscillations de façon indirecte à travers la performance perceptuelle.

Cependant, la fonction cognitive étant évaluée reste encore à confirmer. Dans la présente étude, nous parlons d'encodage visuel ou plus généralement du traitement visuel, mais il est aussi possible que ce soit la rythmicité de l'attention qui, par conséquent, module notre performance en reconnaissance visuelle. Dans la discussion de l'expérience 1 avec les objets familiers, nous avons rapporté plusieurs études, ayant trouvé des rythmes de basses fréquences dans la variation des performances sur différents types de tâches visuelles, qui expliquent leurs résultats par un échantillonnage attentionnel de l'information ou un déploiement variable de l'attention (Busch et VanRullen, 2010; Dugué, Marque et VanRullen,

2015; Dugué, McLelland, Lajous et VanRullen, 2015; Dugué et VanRullen, 2014; Fiebelkorn, Saalman et Kastner, 2013; Helfrich et al., 2018; Senoussi, Moreland, Busch et Dugué, 2019). Considérant alors que l'attention soit modulée par un rythme ou des rythmes de basses fréquences, nous aurions dû retrouver une haute efficacité de traitement visuel dans les basses fréquences pour les visages familiers, et non seulement pour les objets familiers. D'un autre côté, comme nous avons expliqué dans la discussion de l'expérience 2 sur les visages familiers, il y a une grande variabilité interindividuelle dans les habiletés à reconnaître des visages. Ainsi, il se pourrait qu'en excluant les individus ayant davantage de difficulté à reconnaître des visages, on obtienne une image de classification temps-fréquence où les basses fréquences seraient associées à une haute efficacité de traitement visuel. Il serait intéressant d'augmenter l'échantillon et d'évaluer les habiletés de reconnaissance de visages de chaque participant pour investiguer cet aspect.

En tant que deuxième objectif, nous désirions savoir si le profil phase-amplitude est uniforme pour chaque stimulus. Comme ce qui était attendu, nous avons trouvé des profils phase-amplitude uniformément non significatifs pour les objets et les visages familiers, confirmant que ce code de données n'est pas pertinent à l'évaluation de l'efficacité du traitement visuel.

Comme troisième objectif, nous voulions savoir si les profils temporel et temps-fréquence sont différents d'un stimulus à l'autre. En nous basant sur les études précédentes analysant la reconnaissance de visages non familiers (Blais, Arguin et Gosselin, 2013), de mots (Arguin, 2018) et d'objets non familiers (Arguin et Massé, 2019), nous avons émis l'hypothèse qu'il y aurait des différences. Pour ce qui est du domaine temporel, nous avons vu qu'au contraire, il n'y avait pas de différence entre les profils obtenus avec les objets familiers,

les visages familiers et les objets non familiers, mais que seulement le profil obtenu avec les mots se démarquait des autres. Par contre, pour le domaine temps-fréquence, nous avons confirmé notre hypothèse en retrouvant plusieurs différences significatives entre chaque type de stimulus. En présumant que nos images de classification temps-fréquence sont le reflet des oscillations cérébrales survenant durant le traitement visuel, cela pourrait vouloir dire que le rythme de ces oscillations s'adapte, par des processus descendants, au stimulus analysé. Sinon, il est aussi possible que les différentes régions du cortex utilisées dans le traitement des différents stimuli s'activent selon différents rythmes ou communiquent avec d'autres régions cérébrales selon différents rythmes.

Finalement, notre quatrième objectif était d'évaluer l'impact de la familiarité d'un stimulus sur la façon dont ce dernier est perçu en comparant nos profils obtenus avec les objets et visages familiers avec ceux obtenus avec les objets (Arguin et Massé, 2019) et visages non familiers (Blais, Arguin et Gosselin, 2013). Nous n'avons pas trouvé de différence entre les profils temporels des objets familiers et non familiers, mais nous avons trouvé plusieurs différences significatives entre les images de classification temps-fréquence de ces deux stimuli. Cela suggère, encore une fois, que le rythme des oscillations cérébrales s'adapte, par des processus descendants, au stimulus analysé selon qu'il soit familier ou non. Par contre, il n'est pas possible d'exclure, pour le moment, la possibilité que ces stimuli soient suffisamment distincts dans leurs caractéristiques visuelles et que le cerveau les traite comme deux types de stimuli aussi différents qu'un objet et un visage puissent l'être. Le fait que la comparaison entre les profils temps-fréquence des visages familiers et non familiers montre aussi des différences nous pousse à croire que la première explication, quant à la familiarité du stimulus influençant le rythme des oscillations cérébrales, est la bonne. Par contre, les

différentes méthodologies appliquées pour obtenir le profil des visages familiers et celui des visages non familiers nous empêchent d'effectuer une réelle comparaison statistique.

Cette étude présente tout de même quelques limites. D'abord, puisqu'il s'agit de tâches de reconnaissance, nous avons supposé que tous les participants connaissaient le nom, ou du moins la fonctionnalité, des objets présentés et le nom ou une information pertinente sur toutes les célébrités présentées. Par contre, chaque participant possède un niveau de familiarité différent pour chaque objet et chaque célébrité. Par exemple, si une personne fait souvent la cuisine, elle sera probablement plus habile à reconnaître un chaudron qu'une personne qui ne cuisine jamais. Le principe est le même pour les célébrités où une personne qui est fan d'Angelina Jolie pourrait la reconnaître plus facilement qu'une personne qui ne l'a vue que dans un seul film par exemple. De plus, certains objets ou visages utilisés se ressemblent plus que d'autres et il peut être difficile, voire impossible, de les différencier une fois superposés de bruit visuel. Cela peut amener de « fausses » mauvaises reconnaissances, par exemple le participant, sachant qu'il a eu une bonne réponse la dernière fois qu'il a répondu « pomme » en voyant un objet rond, aura tendance à répondre « pomme » à nouveau lorsque l'image d'une balle est présentée. Ainsi, cela peut interférer avec les « vraies » mauvaises reconnaissances, normalement associées à un ratio signal-bruit plus petit ou une gamme de fréquences moins propices à l'encodage visuel. Finalement, au cours de l'expérimentation, certains participants ont dit qu'ils ne savaient pas quel était l'objet ou qui était la célébrité présentée, mais une fois l'erreur enregistrée par le chercheur, ils retrouvaient le nom de l'objet ou de la célébrité. Malheureusement, une fois la réponse entrée dans le programme, il n'y a pas de façon de revenir en arrière, donc le participant entend un feedback inapproprié et le

programme ajuste le bruit visuel en conséquence. Ces erreurs, pourtant rares, peuvent avoir eu une légère influence sur les résultats.

Conclusion

En conclusion, cette étude témoigne du fait qu'il y a des variations rapides dans l'efficacité du traitement visuel à travers le temps et que ces variations semblent être le mieux représentées par un code temps-fréquence en raison de la cohérence inter-sujets remarquablement élevée pour ce domaine. De plus, nous avons démontré qu'il y a des différences majeures entre chaque image de classification temps-fréquence moyenne de tous les types de stimuli analysés jusqu'à présent (mots, objets non familiers, objets familiers et visages familiers), que ce soit entre différentes classes de stimuli ou entre différents niveaux de familiarité d'une même classe de stimulus.

Nous pensons que ce sont les oscillations cérébrales qui modulent l'efficacité du traitement visuel. Cependant, cette étude n'inclut pas de mesure de l'activité cérébrale pour en avoir la preuve. Nous travaillons présentement sur un projet reprenant la technique d'échantillonnage temporel, mais comportant à la fois une mesure des oscillations cérébrales par l'entremise de l'EEG. Néanmoins, nos résultats suggèrent que les images de classification du domaine temps-fréquence reflètent un aspect hautement fondamental du traitement visuel. Il reste encore à déterminer s'il s'agit d'un processus attentionnel, d'un processus d'encodage visuel ou d'une autre fonction cognitive, mais dans tous les cas, le principe est le même, c'est-à-dire que notre cerveau échantillonne l'information visuelle temporellement et que les propriétés de cet échantillonnage sont affectées par la nature du stimulus à traiter. Pour la suite, il serait pertinent de reprendre la méthodologie appliquée dans cette étude et produire, à nouveau, des profils temporel et temps-fréquence en reconnaissance de visages non familiers afin de compléter le portrait de comparaisons que nous avons présenté. Considérant que les images de classification temps-fréquence font ressortir les patrons associés aux oscillations

cérébrales, il serait intéressant de tester des populations cliniques, telles que des aînés souffrant de démence légère, et de comparer les profils obtenus avec ceux de jeunes adultes pour voir si ces images de classification sont sensibles aux déficits neurologiques.

Références

- Alonso-Prieto, E., Belle, G. V., Liu-Shuang, J., Norcia, A. M. & Rossion, B. (2013). The 6Hz fundamental stimulation frequency rate for individual face discrimination in the right occipito-temporal cortex. *Neuropsychologia*, 51(13), 2863-2875. doi:10.1016/j.neuropsychologia.2013.08.018
- Ambrus, G. G., Windel, F., Burton, A. M. & Kovács, G. (2017). Causal evidence of the involvement of the right occipital face area in face-identity acquisition. *NeuroImage*, 148, 212-218. doi:10.1016/j.neuroimage.2017.01.043
- Arguin, M. (2018, mars). Dynamique temporelle de l'encodage visuel. Affiche présentée au 40e congrès annuel de la SQRP, Québec, Québec.
- Arguin, Martin. & Massé, Justine. (2019, mai). The time course of novel visual object recognition. Affiche présentée au 19e congrès annuel de la VSS, St. Pete Beach, Floride.
- Blais, C., Arguin, M. & Gosselin, F. (2013). Human visual processing oscillates: Evidence from a classification image technique. *Cognition*, 128(3), 353–362. <https://doi.org/10.1016/j.cognition.2013.04.009>
- Brodeur, M. B., Kehayia, E., Dion-Lessard, G., Chauret, M., Dionne-Dostie, E., Montreuil, T. & Lepage, M. (2012). The Bank of Standardized Stimuli (BOSS): Comparison between French and English norms. *Behavior Research Methods*, 44, 961-970.
- Busch, N. A., Dubois, J. & VanRullen, R. (2009). The Phase of Ongoing EEG Oscillations Predicts Visual Perception. *Journal of Neuroscience*, 29(24), 7869–7876. <https://doi.org/10.1523/JNEUROSCI.0113-09.2009>
- Busch, N. A. & VanRullen, R. (2010). Spontaneous EEG oscillations reveal periodic sampling of visual attention. *Proceedings of the National Academy of Sciences*, 107(37), 16048-16053. doi:10.1073/pnas.1004801107
- Butler, S., Blais, C., Gosselin, F., Bub, D. & Fiset, D. (2010). Recognizing famous people. *Attention, Perception, & Psychophysics*, 72(6), 1444-1449. doi:10.3758/APP.72.6.1444
- Caldara, R. (2017). Culture Reveals a Flexible System for Face Processing. *Current Directions in Psychological Science*, 26(3), 249-255. doi:10.1177/0963721417710036

- Centanni, T. M., King, L. W., Eddy, M. D., Whitfield-Gabrieli, S. & Gabrieli, J. D. E. (2017). Development of sensitivity versus specificity for print in the visual word form area. *Brain and Language*, 170, 62-70. doi:10.1016/j.bandl.2017.03.009
- Chakravarthi, R. & VanRullen, R. (2012). Conscious updating is a rhythmic process. *Proceedings of the National Academy of Sciences*, 109(26), 10599–10604. doi:10.1073/pnas.1121622109
- Chauvin, A., Worsley, K. J., Schyns, P. G., Arguin, M. & Gosselin, F. (2005). Accurate statistical tests for smooth classification images. *Journal of Vision*, 5(9), 659–667. <https://doi.org/10.1167/5.9.1>
- Coggan, D. D., Baker, D. H. & Andrews, T. J. (2019). Selectivity for mid-level properties of faces and places in the fusiform face area and parahippocampal place area. *European Journal of Neuroscience*, 49(12), 1587-1596. doi:10.1111/ejn.14327
- Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M.-A. & Michel, F. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, 123(2), 291-307. doi:10.1093/brain/123.2.291
- Cohen, M. X. (2015). *Cycles in mind: How brain rhythms control perception and action*. Kindle Edition: Sinc(x) press.
- Collins, E., Robinson, A. K. & Behrmann, M. (2018). Distinct neural processes for the perception of familiar versus unfamiliar faces along the visual hierarchy revealed by EEG. *NeuroImage*, 181, 120-131. doi:10.1016/j.neuroimage.2018.06.080
- Dehaene, S. (1993). Temporal Oscillations in Human Perception. *Psychological Science*, 4(4), 264–270. doi:10.1111/j.1467-9280.1993.tb00273.x
- Dehaene, S., Clec'H, G. L., Poline, J.-B., Bihan, D. L. & Cohen, L. (2002). The visual word form area: a prelexical representation of visual words in the fusiform gyrus. *Neuroreport*, 13(3), 321-325.
- De Luca, M., Pizzamiglio, M. R., Di Vita, A., Palermo, L., Tanzilli, A., Dacquino, C. & Piccardi, L. (2019). First the nose, last the eyes in congenital prosopagnosia: Look like your father looks. *Neuropsychology*, Advance online publication. doi:10.1037/neu0000556

- De Soto, J. L. & De Soto, C. B. (1985). Reading Achievement and Automatic Recognition of Words and Pseudowords. *Journal of Reading Behavior*, 17(2), 115-127. doi:10.1080/10862968509547534
- Dugué, L., Marque, P. & VanRullen, R. (2011). The Phase of Ongoing Oscillations Mediates the Causal Relation between Brain Excitation and Visual Perception. *Journal of Neuroscience*, 31(33), 11889–11893. <https://doi.org/10.1523/JNEUROSCI.1161-11.2011>
- Dugué, L., Marque, P. & VanRullen, R. (2015). Theta Oscillations Modulate Attentional Search Performance Periodically. *Journal of Cognitive Neuroscience*, 27(5), 945-958. doi:10.1162/jocn_a_00755
- Dugué, L., McLelland, D., Lajous, M. & VanRullen, R. (2015). Attention searches nonuniformly in space and in time. *Proceedings of the National Academy of Sciences of the United States of America*, 112(49), 15214-15219. doi:10.1073/pnas.1511331112
- Dugué, L. & VanRullen, R. (2014). The dynamics of attentional sampling during visual search revealed by Fourier analysis of periodic noise interference. *Journal of Vision*, 14(2), 11-11. doi:10.1167/14.2.11
- Ebaid, D. & Crewther, S. G. (2019). Visual Information Processing in Young and Older Adults. *Frontiers in Aging Neuroscience*, 11. doi:10.3389/fnagi.2019.00116
- Egley, R., Driver, J. & Rafal, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, 123(2), 161-177. doi:10.1037/0096-3445.123.2.161
- Fiebelkorn, I. C. & Kastner, S. (2019). A Rhythmic Theory of Attention. *Trends in Cognitive Sciences*, 23(2), 87-101. doi:10.1016/j.tics.2018.11.009
- Fiebelkorn, I. C., Saalman, Y. B. & Kastner, S. (2013). Rhythmic Sampling within and between Objects despite Sustained Attention at a Cued Location. *Current biology : CB*, 23(24). doi:10.1016/j.cub.2013.10.063
- Foley, E., Rippon, G. & Senior, C. (2017). Modulation of Neural Oscillatory Activity during Dynamic Face Processing. *Journal of Cognitive Neuroscience*, 30(3), 338-352. doi:10.1162/jocn_a_01209

- Fournier, L. R., Scheffers, M. K., Coles, M. G. H., Adamson, A. & Abad, E. V. (2000). When complexity helps: an electrophysiological analysis of multiple feature benefits in object perception. *Acta Psychologica*, *104*(1), 119-142. doi:10.1016/S0001-6918(00)00014-7
- Gao, C., Conte, S., Richards, J. E., Xie, W. & Hanayik, T. (2019). The neural sources of N170: Understanding timing of activation in face-selective areas. *Psychophysiology*, *56*(6), e13336. doi:10.1111/psyp.13336
- Gerlach, C. & Marques, J. F. (2014). Visual complexity exerts opposing effects on object categorization and identification. *Visual Cognition*, *22*(6), 751-769. doi:10.1080/13506285.2014.915908
- Geskin, J. & Behrmann, M. (2018). Congenital prosopagnosia without object agnosia? A literature review. *Cognitive Neuropsychology*, *35*(1-2), 4-54. doi:10.1080/02643294.2017.1392295
- Grill-Spector, K., Golarai, G. & Gabrieli, J. (2008). Developmental neuroimaging of the human ventral visual cortex. *Trends in Cognitive Sciences*, *12*(4), 152-162. doi:10.1016/j.tics.2008.01.009
- Grill-Spector, K., Knouf, N. & Kanwisher, N. (2004). The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience*, *7*(5), 555. doi:10.1038/nn1224
- Grill-Spector, K., Kourtzi, Z. & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*(10), 1409-1422. doi:10.1016/S0042-6989(01)00073-6
- Grill-Spector, K. & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, *15*(8), 536-548. doi:10.1038/nrn3747
- Gruber, T., Malinowski, P. & Müller, M. M. (2004). Modulation of oscillatory brain activity and evoked potentials in a repetition priming task in the human EEG. *European Journal of Neuroscience*, *19*(4), 1073-1082. doi:10.1111/j.0953-816X.2004.03176.x
- Helfrich, R. F., Fiebelkorn, I. C., Szczepanski, S. M., Lin, J. J., Parvizi, J., Knight, R. T. & Kastner, S. (2018). Neural Mechanisms of Sustained Attention Are Rhythmic. *Neuron*, *99*(4), 854-865.e5. doi:10.1016/j.neuron.2018.07.032

- Händel, B. F., Haarmeier, T. & Jensen, O. (2010). Alpha Oscillations Correlate with the Successful Inhibition of Unattended Stimuli. *Journal of Cognitive Neuroscience*, 23(9), 2494–2502. <https://doi.org/10.1162/jocn.2010.21557>
- Hanslmayr, S., Volberg, G., Wimber, M., Dalal, S. S. & Greenlee, M. W. (2013). Prestimulus Oscillatory Phase at 7 Hz Gates Cortical Information Flow and Visual Perception. *Current Biology*, 23(22), 2273–2278. doi:10.1016/j.cub.2013.09.020
- Horn, C. C. & Manis, F. R. (1987). Development of automatic and speeded reading of printed words. *Journal of Experimental Child Psychology*, 44(1), 92–108. doi:10.1016/0022-0965(87)90024-5
- Iidaka, T. (2014). Role of the fusiform gyrus and superior temporal sulcus in face perception and recognition: An empirical review. *Japanese Psychological Research*, 56(1), 33–45. doi:10.1111/jpr.12018
- Ito, M., Sugata, T. & Kuwabara, H. (1998). Effects of the sharp and round parts of geometric figures on visual evoked potentials. *Japanese Journal of Physiological Psychology and Psychophysiology*, 16(1), 33–40. doi:10.5674/jjppp1983.16.33
- Jaegle, A. & Ro, T. (2014). Direct control of visual perception with phase-specific modulation of posterior parietal cortex. *Journal of Cognitive Neuroscience*, 26(2), 422–432. https://doi.org/10.1162/jocn_a_00494
- Kanwisher, N., McDermott, J. & Chun, M. M. (1997). The Fusiform Face Area: A Module in Human Extrastriate Cortex Specialized for Face Perception. *Journal of Neuroscience*, 17(11), 4302–4311. doi:10.1523/JNEUROSCI.17-11-04302.1997
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R. & Broussard, C. (2007). What's new in psychtoolbox-3. *Perception*, 36(14), 1–16.
- Klimesch, W., Sauseng, P. & Hanslmayr, S. (2007). EEG alpha oscillations: The inhibition–timing hypothesis. *Brain Research Reviews*, 53(1), 63–88. <https://doi.org/10.1016/j.brainresrev.2006.06.003>
- Lachaux, J.-P., George, N., Tallon-Baudry, C., Martinerie, J., Hugueville, L., Minotti, L., ... Renault, B. (2005). The many faces of the gamma band response to complex visual stimuli. *NeuroImage*, 25(2), 491–501. doi:10.1016/j.neuroimage.2004.11.052

- Lakatos, P., Karmos, G., Mehta, A. D., Ulbert, I. & Schroeder, C. E. (2008). Entrainment of Neuronal Oscillations as a Mechanism of Attentional Selection. *Science*, 320(5872), 110–113. <https://doi.org/10.1126/science.1154735>
- Lange, J., Keil, J., Schnitzler, A., van Dijk, H. & Weisz, N. (2014). The role of alpha oscillations for illusory perception. *Behavioural Brain Research*, 271, 294–301. <https://doi.org/10.1016/j.bbr.2014.06.015>
- Lanze, M., Maguire, W. & Weisstein, N. (1985). Emergent features: A new factor in the object-superiority effect? *Perception & Psychophysics*, 38(5), 438-442. doi:10.3758/BF03207174
- Larousse, É. (s.d.). Définitions : oscillation - Dictionnaire de français Larousse. Repéré 7 juin 2019, à <https://www.larousse.fr/dictionnaires/francais/oscillation/56662>
- Latour, P. L. (1967). Evidence of internal clocks in the human operator. *Acta Psychologica*, 27, 341-348. doi:10.1016/0001-6918(67)90078-9
- Liu, M., Pei, G., Peng, Y., Wang, C., Yan, T. & Wu, J. (2018). Disordered high-frequency oscillation in face processing in schizophrenia patients. *Medicine*, 97(6). doi:10.1097/MD.00000000000009753
- Liu, Z.-X., Shen, K., Olsen, R. K. & Ryan, J. D. (2018). Age-related changes in the relationship between visual exploration and hippocampal activity. *Neuropsychologia*, 119, 81-91. doi:10.1016/j.neuropsychologia.2018.07.032
- Malach, R., Reppas, J. B., Benson, R. R., Kwong, K. K., Jiang, H., Kennedy, W. A., ... Tootell, R. B. (1995). Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences*, 92(18), 8135-8139. doi:10.1073/pnas.92.18.8135
- Marques, J. F. (2011). La nature des catégories superordonnées : caractéristiques, effets et explications. *L'Année psychologique*, 111(3), 533-548. <https://doi.org/10.4074/S0003503311003058>
- Martinovic, J., Gruber, T. & Müller, M. M. (2007). Induced Gamma Band Responses Predict Recognition Delays during Object Identification. *Journal of Cognitive Neuroscience*, 19(6), 921-934. doi:10.1162/jocn.2007.19.6.921

- Martinovic, J., Gruber, T., Hantsch, A. & Müller, M. M. (2008). Induced gamma-band activity is related to the time point of object identification. *Brain Research*, 1198, 93-106. doi:10.1016/j.brainres.2007.12.050
- Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M. & Ro, T. (2009). To See or Not to See: Prestimulus α Phase Predicts Visual Awareness. *Journal of Neuroscience*, 29(9), 2725–2732. <https://doi.org/10.1523/JNEUROSCI.3963-08.2009>
- Mathewson, K. E., Fabiani, M., Gratton, G., Beck, D. M. & Lleras, A. (2010). Rescuing stimuli from invisibility: Inducing a momentary release from visual masking with pre-target entrainment. *Cognition*, 115(1), 186–191. <https://doi.org/10.1016/j.cognition.2009.11.010>
- Mathewson, K. E., Prudhomme, C., Fabiani, M., Beck, D. M., Lleras, A. & Gratton, G. (2012). Making Waves in the Stream of Consciousness: Entraining Oscillations in EEG Alpha and Fluctuations in Visual Awareness with Rhythmic Visual Stimulation. *Journal of Cognitive Neuroscience*, 24(12), 2321–2333. https://doi.org/10.1162/jocn_a_00288
- McCabe, É., Blais, C. & Gosselin, F. (2005). Chapter 34 - CATEGORIZATION OF OBJECTS, SCENES, AND FACES THROUGH TIME. Dans H. Cohen et C. Lefebvre (dir.), *Handbook of Categorization in Cognitive Science* (p. 767 791). Oxford : Elsevier Science Ltd. doi:10.1016/B978-008044612-7/50089-5
- Pidoux, V. (2010). Expérimentation et clinique électroencéphalographiques entre physiologie, neurologie et psychiatrie (Suisse, 1935-1965). *Revue d'histoire des sciences*, 63(2), 439-472. doi:10.3917/rhs.632.0439
- Pitcher, D., Walsh, V. & Duchaine, B. (2011). The role of the occipital face area in the cortical face perception network. *Experimental Brain Research*, 209(4), 481-493. doi:10.1007/s00221-011-2579-1
- Rana, K. D. & Vaina, L. M. (2014). Functional Roles of 10 Hz Alpha-Band Power Modulating Engagement and Disengagement of Cortical Networks in a Complex Visual Motion Task. *PLoS ONE*, 9(10). <https://doi.org/10.1371/journal.pone.0107715>
- Reinke, K., Fernandes, M., Schwindt, G., O'Craven, K. & Grady, C. L. (2008). Functional specificity of the visual word form area: General activation for words and symbols but specific network activation for words. *Brain and Language*, 104(2), 180-189. doi:10.1016/j.bandl.2007.04.006

- Sakihara, K., Gunji, A., Furushima, W. & Inagaki, M. (2012). Event-related oscillations in structural and semantic encoding of faces. *Clinical Neurophysiology*, *123*(2), 270-277. doi:10.1016/j.clinph.2011.06.023
- Samaha, J., Bauer, P., Cimaroli, S. & Postle, B. R. (2015). Top-down control of the phase of alpha-band oscillations as a mechanism for temporal prediction. *Proceedings of the National Academy of Sciences*, *112*(27), 8439–8444. <https://doi.org/10.1073/pnas.1503686112>
- Samaha, J. & Postle, B. R. (2015). The Speed of Alpha-Band Oscillations Predicts the Temporal Resolution of Visual Perception. *Current Biology*, *25*(22), 2985–2990. <https://doi.org/10.1016/j.cub.2015.10.007>
- Schobert, A.-K., Corradi-Dell'Acqua, C., Frühholz, S., van der Zwaag, W. & Vuilleumier, P. (2018). Functional organization of face processing in the human superior temporal sulcus: a 7T high-resolution fMRI study. *Social Cognitive and Affective Neuroscience*, *13*(1), 102-113. doi:10.1093/scan/nsx119
- Schwarzlose, R. F., Swisher, J. D., Dang, S. & Kanwisher, N. (2008). The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences*, *105*(11), 4447-4452. doi:10.1073/pnas.0800431105
- Senju, A., Verneti, A., Kikuchi, Y., Akechi, H., Hasegawa, T. & Johnson, M. H. (2013). Cultural background modulates how we look at other persons' gaze. *International Journal of Behavioral Development*, *37*(2), 131-136. doi:10.1177/0165025412465360
- Senoussi, M., Moreland, J. C., Busch, N. A. & Dugué, L. (2019). Attention explores space periodically at the theta frequency. *Journal of Vision*, *19*(5), 22-22. doi:10.1167/19.5.22
- Spaak, E., Lange, F. P. de, & Jensen, O. (2014). Local Entrainment of Alpha Oscillations by Visual Stimuli Causes Cyclic Modulation of Perception. *Journal of Neuroscience*, *34*(10), 3536–3544. <https://doi.org/10.1523/JNEUROSCI.4385-13.2014>
- Spiridon, M., Fischl, B. & Kanwisher, N. (2006). Location and spatial profile of category-specific regions in human extrastriate cortex. *Human Brain Mapping*, *27*(1), 77-89. doi:10.1002/hbm.20169

- Stacchi, L., Ramon, M., Lao, J. & Caldara, R. (2019). Neural Representations of Faces Are Tuned to Eye Movements. *Journal of Neuroscience*, 39(21), 4113-4123. doi:10.1523/JNEUROSCI.2968-18.2019
- Stenner, M.-P., Bauer, M., Haggard, P., Heinze, H.-J. & Dolan, R. (2014). Enhanced alpha-oscillations in visual cortex during anticipation of self-generated visual stimulation. *Journal of Cognitive Neuroscience*, 26(11), 2540–2551. https://doi.org/10.1162/jocn_a_00658
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18(6), 643-662. doi:10.1037/h0054651
- Tardif, J., Morin Duchesne, X., Cohan, S., Royer, J., Blais, C., Fiset, D., ... Gosselin, F. (2019). Use of Face Information Varies Systematically From Developmental Prosopagnosics to Super-Recognizers. *Psychological Science*, 30(2), 300-308. doi:10.1177/0956797618811338
- VanRullen, R. & Koch, C. (2003). Is perception discrete or continuous? *Trends in Cognitive Sciences*, 7(5), 207-213. doi:10.1016/S1364-6613(03)00095-0
- VanRullen, R., Reddy, L. & Koch, C. (2005). Attention-driven discrete sampling of motion perception. *Proceedings of the National Academy of Sciences*, 102(14), 5291-5296. doi:10.1073/pnas.0409172102
- VanRullen, R., Reddy, L. & Koch, C. (2006). The Continuous Wagon Wheel Illusion Is Associated with Changes in Electroencephalogram Power at ~13 Hz. *Journal of Neuroscience*, 26(2), 502–507. <https://doi.org/10.1523/JNEUROSCI.4654-05.2006>
- Vanrullen, R., Reddy, L. & Koch, C. (2010). A motion illusion reveals the temporally discrete nature of visual awareness. Dans R. Nijhawan et B. Khurana (dir.), *Space and Time in Perception and Action* (p. 521-535). Cambridge University Press. Repéré à <https://hal.archives-ouvertes.fr/hal-00468119>
- Vinberg, J. & Grill-Spector, K. (2008). Representation of Shapes, Edges, and Surfaces Across Multiple Cues in the Human Visual Cortex. *Journal of Neurophysiology*, 99(3), 1380-1393. doi:10.1152/jn.01223.2007
- Vinette, C., Gosselin, F. & Schyns, P. G. (2004). Spatio-temporal dynamics of face recognition in a flash: it's in the eyes. *Cognitive Science*, 28(2), 289-301. doi:10.1207/s15516709cog2802_8

- Weisstein, N. & Harris, C. S. (1974). Visual Detection of Line Segments: An Object-Superiority Effect. *Science*, 186(4165), 752-755. doi:10.1126/science.186.4165.752
- Zhang, J., Li, X., Song, Y. & Liu, J. (2012). The Fusiform Face Area Is Engaged in Holistic, Not Parts-Based, Representation of Faces. *PLOS ONE*, 7(7), e40390. doi:10.1371/journal.pone.0040390

Annexe A

Exemples d'images d'objets familiers de la BOSS (Bordeur et al, 2012) utilisées dans l'expérience 1 :



Annexe B

Exemples d'images de visages familiers de la banque de visages célèbres construite par Butler, Blais, Gosselin, Bub et Fiset (2010) utilisées dans l'expérience 2 :



Annexe C

Exemples d'images d'objets non familiers utilisées dans l'étude d'Arguin et Massé
(2019) :

