

PROJET SCIENTIFIQUE COLLECTIF

Neural Representation of Temporal Sequences

24 Avril 2019

Groupe HSS 11

Tuteurs : PALLIER Christophe - Van KERKOERLE Timo

Coordinateurs : DULPHY Anne - GRENIER Jean-Yves



TABLE DES MATIÈRES

Remerciements	4
Introduction	5
1 Élaboration du protocole global et de l'hypothèse à tester	7
1.1 Choix global de la forme des stimuli en jeu et du protocole à suivre	7
1.2 Proposition d'une hypothèse à tester : construction d'un langage	8
1.2.1 Description du langage construit	9
1.2.2 Définition de la complexité du langage	10
1.2.3 But précis de l'expérience	11
2 Évolution concrète du développement de l'expérience et de la récupération des données	12
2.1 Démarche expérimentale	12
2.2 Création et mise en place de l'expérience	13
2.2.1 Un premier prototype sur ordinateur	13
2.2.2 Développement mobile	16
2.3 Récupération des données de l'expérience	18
2.3.1 Récupération informatique des données	18
2.3.2 Publication de l'application	20
2.3.3 Techniques de communication pour inciter au jeu	21
3 Résultats et interprétations	23
3.1 Premiers résultats et interprétation	23
3.1.1 Corrélation temps-complexité	24
3.1.2 Mesure quantitative de l'erreur	26
3.1.3 Analyse des proportions d'erreur	28
3.2 Perspectives et regard critique sur les résultats	32
3.2.1 Nettoyage des données, problème du format application de l'expérience .	32
3.2.2 La non pertinence du miroir comme primitive	33
Conclusion	34

Annexe	35
3.3 Annexe 1 - Architecture de l'application	35
3.4 Annexe 2 - Différentes configurations	35
Références bibliographiques	37

REMERCIEMENTS

Parvenus au terme de ce projet, nous prenons conscience du chemin parcouru et de l'accompagnement qui nous a été apporté tout au long de l'année écoulée. Nous tenons premièrement à remercier nos coordinateurs, Anne Dulphy et Jean-Yves Grenier qui nous ont encadrés, en particulier en nous conseillant de nous recentrer sur une problématique plus précise à la suite des premières formes que prenait le projet.

Mais ce projet dépasse le cadre de l'Ecole. En effet, il est né d'une visite chez Neurospin, centre de recherche pour l'innovation en imagerie cérébrale soutenu par le CEA (centre pour l'énergie atomique). Là-bas, nous avons rencontré notre tuteur, Christophe Pallier, directeur de l'équipe "Neuroimagerie du langage" de l'unité de neuroimagerie cognitive INSERM-CEA implantée au centre Neurospin. Il nous a étroitement aidés à choisir un sujet viable parmi les idées qui nous venaient à l'esprit au fil des discussions et des lectures de ses travaux. Nous avons également beaucoup collaboré avec Timo Van Kerkoerle, directeur de la section Deep Imaging à Neurospin. Il nous a permis de trouver les informations nécessaires à l'avancement du projet et nous a guidés tout au long de l'année. Nous le remercions fortement pour tout le temps qu'il nous a consacré, pour nous aiguiller sur le choix des paramètres de l'expérience, pour nous expliquer le traitement des données et pour répondre à toutes nos questions.

Nous tenons également à remercier tous les chercheurs de Neurospin qui nous ont apporté leur expertise. Nous remercions chaleureusement Joël Fagot pour tout l'intérêt qu'il porte à notre expérience dans le cadre d'un projet sur des singes ; Samuel Planton pour nous avoir présenté ses travaux sur les stimuli auditifs et la façon d'interpréter les résultats. Nous remercions également Mathias Sablé Meyer et Yair Lakretz pour l'intérêt qu'ils ont porté à nos travaux. Nous remercions enfin Stanislas Dehaene pour ses conseils avisés et ses encouragements.

Nous tenons de plus à remercier Isabelle Denghien pour la mise à disposition des serveurs de Neurospin et Hadrien Renaud, vice-président du Binet Réseau, pour son expertise en récupération de données.

Un grand merci à tous ceux qui ont contribué, de près ou de loin, à ce projet. Merci notamment à tous ceux qui se sont prêtés à notre expérience.

INTRODUCTION

On pense que l'homme possède des capacités remarquables par rapport à d'autres espèces (langage oral et écrit, mathématiques, musique...). Cela s'appuie sur les capacités de traitement des séquences : de l'apprentissage des probabilités de transition entre les événements jusqu'à l'extraction de structures imbriquées complexes, Dehaene et al., 2015.

La question du séquencement étant au coeur de ces concepts considérés comme spécifiques à l'Homme, elle a intéressé de nombreux chercheurs. Notamment, la distinction entre les hommes et les singes dans ce domaine a déjà fait l'objet de nombreuses expériences.

Dans les dernières décennies, les études d'apprentissage de séquences ont décrit un mécanisme possible par lequel les représentations mentales complexes sont construites de primitives plus simples. La capacité humaine à extraire le simple du complexe réside dans les séquences [1]. Tandis que les primates non-humains échouent à reconnaître une systématique dans l'apprentissage des langues [2], les humains semblent être naturellement dotés d'une capacité à saisir rapidement des règles complexes implicites. À 8 mois déjà, des enfants à qui on présente une courte séquence de syllabes extraient aisément des mots trisyllabiques récurrents [3, 4] et à 12 mois ils comprennent comment ces mots se combinent pour former de plus grandes structures [5]. Une capacité semblable à grouper des éléments consécutifs selon des régularités implicites a aussi été démontrée pendant l'apprentissage de séquences visuo-motrices par des adultes [6,7]. Les enfants et les adultes sont également en mesure d'apprendre des règles algébriques plus abstraites telles que « AAB » (une répétition de deux éléments suivis d'un troisième) [8,9]. Cette capacité d'apprentissage abstrait des règles semble être renforcée chez les humains et s'appuyer sur le cortex préfrontal inférieur (« zone de Broca ») [10,11,12]. En outre, des secteurs du cortex préfrontal inférieur semblent être utilisés pour les règles linguistiques et mathématiques [13]. La question se pose donc de savoir si une capacité de représentation interne et de manipulation des séquences imbriquées différencie l'Homme des grands singes. Bien qu'il y ait eu plusieurs études sur l'apprentissage artificiel du langage ([14,15,16]) et sur l'existence d'une faculté pour reconnaître les séquences, il y a eu comparativement moins d'études sur l'encodage à proprement parler des séquences. Les humains et les singes utilisent-ils spontanément les mêmes mécanismes pour « comprimer » l'information de séquence, c'est-à-dire utilisent-ils les mêmes primitives pour coder le modèle (p. ex., AAABBB est « deux triplets ») ?

Curieux d'en savoir plus sur cette distinction, nous avons voulu explorer le séquencement

chez l'être humain afin de le comparer à celui chez les singes. Réunis par l'intérêt commun que nous portons aux neurosciences, le sujet de notre projet scientifique s'est construit lors de nos premiers entretiens avec notre tuteur Christophe Pallier. Notre motivation principale était d'appliquer une approche quantitative et pluridisciplinaire aux sciences du langage.

Notre objectif ici est d'introduire un nouveau paradigme expérimental pour étudier l'acquisition de structures élémentaires. Nous souhaitons que ce nouveau paradigme fonctionne de la manière suivante : proposition d'une hypothèse sur l'encodage des séquences, élaboration d'un protocole permettant de vérifier cette hypothèse, en faisant en sorte que ce protocole puisse être reproduit sur des singes, et interprétation de nos résultats.

Pour notre travail, nous nous sommes séparés les tâches en trois pôles, en interaction constante, pour proposer l'expérience la plus pertinente et cohérente. Nous avons aussi continuellement échangé sur les différents choix avec nos tuteurs, dans l'optique de proposer une expérience poussée qui permet de vérifier au mieux la véracité de notre hypothèse.

1

ÉLABORATION DU PROTOCOLE GLOBAL ET DE L'HYPOTHÈSE À TESTER

1.1 CHOIX GLOBAL DE LA FORME DES STIMULI EN JEU ET DU PROTOCOLE À SUIVRE

La première question qui se pose est la forme des stimuli utilisés pour notre expérience : doit-on travailler avec des séquences de sons ? Des séquences d'images ? Des séquences de couleurs ?

En discutant avec des chercheurs qui travaillent avec des singes, nous nous sommes rendu compte que peu d'expériences faites sur les humains pouvaient être reproduites sur des singes. Il est pourtant intéressant de pouvoir faire la même expérience sur les deux populations pour mieux les comparer (en simplifiant éventuellement celle sur les singes). Nous avons donc essayé de chercher un stimulus neutre pour les Hommes et les singes.

Le papier [17] (Amalric M, Wang L, Pica P, FigueiraS, Sigman M, Dehaene S (2017), Language of Geometry) explore une approche des séquences spatiales chez l'Homme. Il montre une grande capacité des humains à découper en séquence une représentation spatiale. Mais cette expérience fait aussi appel à une autre capacité très prononcée chez l'Homme qu'est la reconnaissance spatiale. La preuve en est notre sensibilité à la symétrie. Nous avons alors décidé de nous orienter vers une expérience construite sur le même principe que celle-ci mais qui fait intervenir le moins possible l'organisation de l'espace.

D'autres expériences menées avec des stimuli auditifs [1] ont mis en évidence les capacités de séquençage de l'Homme. Cependant, les singes sont moins performants sur des tests d'audition car leur cortex préfrontal ventrolatéral n'est pas relié de la même manière aux zones cérébrales impliquées dans l'audition que chez l'Homme. De plus, d'après [18], le gène ayant la pression de sélection positive la plus significative chez l'homme, et qui est peu significatif chez les primates, code pour l' α -tectorine, une protéine qui joue un rôle vital dans la membrane tectoriale de l'oreille interne et dont les anomalies sont responsables de surdité d'origine génétique. Si nous

voulons comprendre s'il existe une capacité abstraite à reconnaître les séquences à la fois chez les singes et chez les Hommes nous devrons donc nous abstenir d'utiliser un stimulus sonore pour ne pas être influencés par la différence d'acuité auditive.

Après plusieurs discussions avec Timo Van Koerkerle, qui a beaucoup travaillé avec des singes, nous avons retenu un élément important : les singes sont sensibles aux couleurs. **Nous avons donc décidé de travailler sur des séquences temporelles de couleurs différentes. À l'image de l'article *Language of geometry* [17] le but de notre PSC est alors de proposer une hypothèse de langage du cerveau pour ces séquences de couleurs puis de la tester.**

1.2 PROPOSITION D'UNE HYPOTHÈSE À TESTER : CONSTRUCTION D'UN LANGAGE

Notre hypothèse globale est de dire que le cerveau est capable, de manière abstraite, de reconnaître des patterns, des régularités, dans une séquence. Par exemple, nous sommes capables de percevoir un doublet ou un triplet de couleur dans une séquence.

L'hypothèse principale que nous faisons alors est que la présence de régularité, si elle est détectée, facilite la mémorisation d'une séquence. Le principe de notre expérience se dessine maintenant : le sujet devra mémoriser une séquence puis la reproduire. Une séquence régulière pour le cerveau humain ou simiesque devrait être reproduite avec moins d'erreurs et plus rapidement qu'une séquence quelconque.

Toutefois le but de notre expérience est plus précis que cela. Nous cherchons à déterminer la nature de ces régularités repérées et avec quelle facilité elles sont perçues.

Nous avons alors cherché à établir une hypothèse précise sur l'encodage des séquences temporelles de couleurs.

Nous avons mené en parallèle de cette réflexion une série de tests pour trouver les modalités pertinentes pour notre expérience. Nous nous sommes fixés sur des séquences de 6 éléments avec 3 couleurs possibles : jaune, rouge et bleu. (Ce choix est détaillé dans la section "Un premier prototype sur ordinateur ") Nous avons donc cherché plus particulièrement une hypothèse précise sur l'encodage des séquences de 6 éléments à 3 couleurs.

1.2.1 DESCRIPTION DU LANGAGE CONSTRUIT

Nous avons d'abord conçu un langage formel capable de décrire, de manière compacte, toutes les séquences de 6 éléments avec 3 couleurs différentes. L'ensemble des instructions primitives est constitué des 3 couleurs élémentaires : rouge, jaune et bleu. Nous les désignons comme R (apparition du rouge), B (apparition du bleu) et J (apparition du jaune). A partir de ces primitives, une séquence peut alors être générée par simple concaténation.

Par exemple l'expression  peut s'écrire : JJBBRR.

Bien que n'importe quelle séquence puisse être encodée de cette façon basique, nous fournissons la preuve que les humains détectent et encodent des séquences régulières sous une forme souvent bien plus comprimée. On suppose donc que le cerveau humain reconnaît des patterns et non uniquement une liste de six couleurs. Ainsi, nous faisons également l'hypothèse que le « langage de la pensée » comprend des instructions pour répéter les opérations. Et nous noterons (*expression*)^k la répétition k fois de l'expression entre parenthèses. Par exemple, la séquence  peut être encodée comme $J^2B^2R^2$.

Le langage que nous construisons permet également une autre opération : le miroir. Et on notera $M(s_1s_2s_3)=s_1s_2s_3s_3s_2s_1$ le miroir de l'expression entre parenthèses.

Par exemple, la séquence  peut être encodée sous la forme $M(JJB)$, correspondant à JJBBJJ.

Pourquoi avoir choisi les fonctions "miroir" et "répétition" ? Nous avons testé certaines séquences sur nous 6 et nous avons remarqué que l'on retenait les séquences en repérant ces deux effets. (On se dit par exemple dans nos têtes « Alors c'est bleu jaune rouge puis la même chose dans l'autre sens » ou « C'est donc 4 bleus puis 2 jaunes » pour retenir les séquences). D'ailleurs, ce n'est pas par hasard si l'anaphore ou le chiasme sont des figures stylistiques si marquantes ! La compression comme nous l'avons formalisée peut aussi se comprendre de façon plus pratique avec la manière de prononcer les séquences, plus spécifiquement le ton et les groupements de lettres que l'on réalise naturellement. Ainsi une manière de prononcer la séquence 'ababab' est 'ab' pause 'ab' pause 'ab'. Cela correspond à la compression ' ab^3 ' selon notre système expliqué plus tôt. Une autre, un peu moins naturelle est 'a' pause 'ba' pause 'ba' pause 'b'. Nous pourrons avec nos résultats voir laquelle est effectivement utilisée pour mémoriser cette séquence. Pour ce qui est du miroir, on peut le voir avec le ton utilisé. Ainsi pour la séquence 'babbac', on pourra dire de manière assez naturelle : baBBac (où "B" est plus aigu et fort, surtout en comparaison au "a" le précédent ou le suivant).

Ensuite, dans la plupart des langues, de nombreuses expressions équivalentes fournissent le même résultat. C'est aussi le cas ici. Par exemple, l'expression  peut être capturée comme $J^2B^2J^2$ ou comme M(JJB) ou comme JJBBJJ ou etc. Nous supposons alors que les sujets appliquent le rasoir d'Occam, c'est-à-dire qu'ils tentent de sélectionner l'expression la plus parcimonieuse qui tient compte de la séquence observée.

1.2.2 DÉFINITION DE LA COMPLEXITÉ DU LANGAGE

Mais comment définir proprement ce qu'est l'expression la plus parcimonieuse ?

Le concept de complexité Kolmogorov, une notion de la théorie de l'information algorithmique, fournit un cadre mathématique naturel pour ces idées [19,20]. Ce cadre définit la complexité d'une séquence donnée comme la longueur de l'expression la plus courte capable de la produire dans un langage Turing-complet, (c'est-à-dire n'importe quel langage de programmation raisonnable). Malheureusement, un résultat classique de la théorie algorithmique de l'information est que, pour n'importe quel langage Turing-complet, la complexité de Kolmogorov n'est pas calculable. Même pour des langages simples, la complexité de Kolmogorov est souvent difficile à calculer en pratique, parce qu'elle implique l'examen, pour chaque séquence, de tous les programmes qui la calculent. Cette recherche grandit typiquement exponentiellement avec la taille de la séquence. De plus, quand on teste un programme on ne peut savoir s'il va s'arrêter ou pas (c'est un problème indécidable) et on doit donc mettre un temps de calcul limite. Différentes méthodes ont été développées pour approcher la complexité de Kolmogorov [19]. Une idée est de l'approcher en utilisant des compresseurs de fichiers standards comme Lempel-Ziv. Une telle approche a été utilisée par exemple dans [21] pour des groupes de grands documents via une définition de distance universelle. Les compresseurs de fichiers se comportent bien dans les textes relativement volumineux, mais ne fournissent pas de compression significative lorsque l'entrée est un très petit texte comportant peu de répétitions, comme c'est le cas pour les séquences de 6 couleurs que nous avons utilisées ici.

Dans notre cas, pour quantifier la complexité d'une séquence, nous utilisons alors plutôt la notion de « longueur minimale de description », qui est étroitement liée à la complexité de Kolmogorov [22]. Désormais, nous appellerons « complexité » d'une séquence de couleurs X, et notée K(X), la longueur de l'expression la plus courte qui lui correspond dans notre langage. L'hypothèse psychologique correspondante, que nous avons mise à l'épreuve dans nos expériences, est que les participants humains tentent de « comprimer » mentalement la séquence, c'est-à-dire de minimiser le coût en mémoire en identifiant la plus simple (la plus courte) expression mentale qui leur permet de stocker la séquence. Nous faisons les hypothèses les plus simples possibles concernant la longueur d'expression.

En substance,

- (1) chaque instruction primitive supplémentaire (R, J ou B) ajoute un coût fixe de 1
- (2) répéter un ensemble d'instructions ajoute un coût p_1 pour la répétition d'un élément, p_2 pour la répétition de 2 éléments et p_3 pour 3 éléments
- (3) le miroir a de même un coût m_2 ou m_3 selon le nombre d'éléments

On souhaite garder p_1 p_2 p_3 m_2 et m_3 comme des paramètres modifiables. Ainsi en fonction des données récoltées nous pourrons en testant plusieurs valeurs de ces paramètres estimer le poids relatif du miroir par rapport à la répétition.

Ainsi par exemple, l'expression  comprimée en $JJBRR$ a un coût de 6 et comprimée en $J^2B^2R^2$ a un coût de $3(1 + p_1)$ avec $p_1 < 1$. La complexité de cette séquence est donc $3(1 + p_1)$ (son coût minimal).

L'expression  quant à elle a un coût :

- Pour $J^2B^2J^2$ de $3(p_1 + 1)$
- Pour $M(J^2B)$ un coût de $m_3 + p_1 + 2$. Selon les valeurs de m_3 et p_1 la complexité de l'expression ne sera pas la même.

1.2.3 BUT PRÉCIS DE L'EXPÉRIENCE

L'hypothèse que nous faisons dans notre PSC est donc la suivante : le cerveau humain retient les séquences avec le langage décrit ci-dessus en codant chaque séquence par son expression de plus faible complexité.

Le but de notre PSC sera donc le suivant :

- * Tester notre hypothèse sur l'encodage des séquences grâce à un large jeu de données fournies par une application
- * Pour tester les différences entre humains et singes dans des expériences futures, fournir un protocole aisément transférable
- * Estimer les poids relatifs de répétition et miroir si les données le permettent

2

ÉVOLUTION CONCRÈTE DU DÉVELOPPEMENT DE L'EXPÉRIENCE ET DE LA RÉCUPÉRATION DES DONNÉES

2.1 DÉMARCHE EXPÉIMENTALE

Nous avions à notre disposition plusieurs méthodes pour réaliser l'expérience : payer un petit nombre de sujets qui réalisent une série de tests restreinte, ou réaliser une application ou une interface web. Dans notre cas, la dernière méthode présente de nombreux avantages :

- La collecte d'un grand nombre de données : comme notre expérience porte sur un grand nombre de séquences (729), il faut récupérer beaucoup de données pour être capable de faire de réelles statistiques et de déterminer les paramètres p et m dont on cherche l'importance.
- Une collecte qui peut se faire en continu et perdurer même après la fin de notre projet sous la supervision des chercheurs de NeuroSpin.

Afin de récupérer ces données de la façon la plus efficace possible, nous avons créé une application pour smartphone qui récupère les données de l'expérience à chaque essai.

En effet, alors que la puissance de calcul des smartphones ne cesse de croître et que les applications mobiles, aussi appelées “app”, continuent de dominer le milieu du numérique (c'est par exemple la plateforme qui représente le plus gros marché des jeux vidéos), les applications sont devenues une nouvelle frontière pour faire progresser la méthodologie des expériences de terrain. L'utilisation d'applications permet de recueillir une variété de données objectives et autodéclarées au fil du temps, de contrôler avec précision la randomisation et le support d'expérimentation, et surtout d'élargir l'échelle à laquelle l'expérience peut-être déployée. Une démarche intéressante que nous avons écartée du projet car difficilement réalisable aurait été d'analyser l'impact des langues natives de chaque individu sur les données. La raison en est simple : l'immense majorité de notre public a été le public polytechnicien, francophone. La démarche de trouver un public qui ne parle par exemple aucune langue indo-européenne, ou au moins pas le français est de fait très complexe pour un projet étudiant basé à l'X et avec de faibles moyens.

Par ailleurs, la collecte des données sur smartphone est le plus souvent passive, c'est-à-dire qu'elle ne nécessite pas que l'utilisateur réalise une action : par exemple Google Maps qui récupère la localisation de l'utilisateur, Youtube qui récupère les vidéos regardées, etc. Toutefois, à l'heure actuelle, très peu d'expériences, où l'utilisateur devrait être actif, sont réalisées sur smartphone. C'est aussi pour nous une façon de nous démarquer et de montrer qu'un tel usage du smartphone peut se révéler très fructueux et permettre aux expériences de neuroscience d'une part d'être plus accessibles et d'autres part d'être plus pertinentes grâce au grand nombre de données collectées.

2.2 CRÉATION ET MISE EN PLACE DE L'EXPÉRIENCE

Le développement de l'expérience s'est réalisé en deux temps : une première partie d'expérimentation sur ordinateur précédée de discussions pour savoir quelle forme allait prendre l'expérience, une deuxième partie de création de l'application pour smartphone.

2.2.1 UN PREMIER PROTOTYPE SUR ORDINATEUR

Afin de nous décider sur la forme que prendrait notre expérience, nous avons commencé par développer un premier prototype sur ordinateur. L'avantage de ce mode de développement, alors que nous souhaitions déjà produire une application mobile, est sa rapidité de développement. Ainsi, nous avons pu produire en quelques heures un premier prototype.

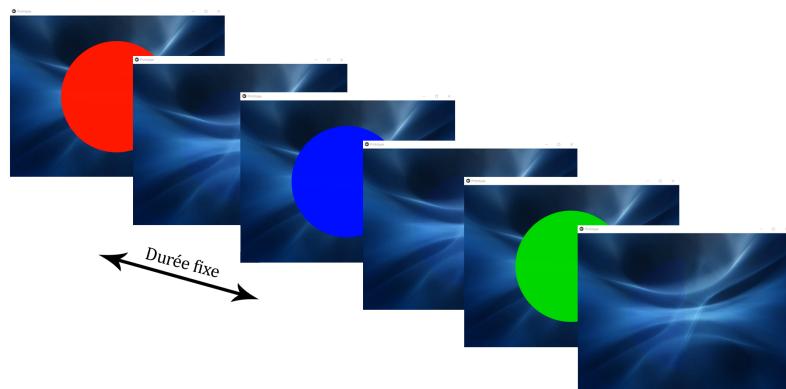


FIGURE 1 – Premier prototype d'expérience

Ceci nous a permis d'établir différents paramètres de l'expérience :

◊ **Forme générale de l'expérience -** Nous nous sommes accordés sur la forme de l'expérience. Celle-ci présenterait rapidement des disques de couleurs différentes. En effet souvent, en science cognitive, les stimuli sont des lettres, des chiffres, des formes, des couleurs comme nous l'avons lu à travers *The mind of an ape* de David Premack [14]. Aussi, dans les expériences sur les séquences, un élément très employé est le son (comme on l'a vu dans les expériences menées sur les souris par Brice Bathelier - chercheur que nous avons contacté en premier dans la mise en place de notre PSC) : ici, nous l'avons vite écarté car au-delà des problèmes évoqués en sous-section 1.2, cela n'aurait pas été adapté à une application smartphone (où les sons ne sont pas écoutables dans le métro par exemple), et car les sons sont trop complexes et codés de façon différente par le cerveau (ce que nous avons appris de Christophe Pallier). Lors de nos différentes discussions, nous avons aussi mentionné le fait que les stimuli puissent être des lettres, voire des mots, de la forme ABAB par exemple. Le problème rencontré se situait au niveau de l'adaptation de l'expérience pour les singes : les lettres et les mots sont des notions qui nous paraissent familières de par l'apprentissage que nous en avons eu. Chez les singes, les mots et les lettres n'ont aucune signification, et il est extrêmement difficile pour eux de les distinguer ; à la différence des formes et surtout des couleurs qui leur sont assez intuitives. D'autant plus que chez le cerveau humain et simiesque, les données de forme et de couleur sont traitées par la même aire corticale (appelée V2), tandis qu'une autre aire traite l'information spatiale et l'aspect saccadé de la séquence (l'aire V3).

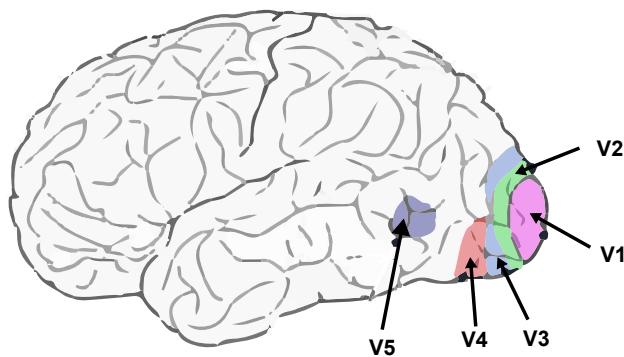


FIGURE 2 – Représentation schématique des aires cérébrales visuelles

La couleur qui occupe une place plus importante de notre cerveau, interfère peu avec la zone chargée du caractère séquentiel, et est donc plus facile à distinguer, elle serait donc

le paramètre constitutif de la séquence. Pour le support de la couleur nous avons choisi assez naturellement le rond. Ce choix évoluera un peu par la suite (cf "Ajout de texture").

- ◊ **Nombre de stimuli -** Nous avions initialement conçu l'expérience comme composée d'une séquence de 8 stimuli de 4 couleurs différentes. Lors des tests que nous avons menés sur nous-même nous nous sommes rendu compte de leur difficulté élevée : sur une séquence donnée nous étions environ deux sur six dans le groupe à réussir à la reproduire. Par ailleurs, après discussion avec nos tuteurs il est remonté que cette forme d'expérience serait trop difficile à mettre en place sur des singes car cela prendrait trop de temps pour leur faire prendre en main l'expérience. Par conséquent nous avons abaissé la difficulté en passant à une expérience composée de 6 stimuli en 3 couleurs. La complexité des séquences étant ainsi réduite, nous avons eu à réfléchir à l'augmentation de la vitesse de flash des bulles afin de rendre le jeu plus difficile et intéressant pour les utilisateurs, pour qu'ils continuent à participer à l'expérience et à fournir des données.
- ◊ **Choix des couleurs -** Une fois le type de stimuli établi, nous avons réfléchi avec nos tuteurs au meilleur choix de couleurs à présenter. L'idée principale étant de choisir des couleurs dont le nom est monosyllabique. En effet, lorsque l'utilisateur reproduit la séquence, il a tendance à répéter le nom des couleurs dans sa tête. Il était donc nécessaire que tous les noms de couleurs aient la même longueur afin de ne pas impacter le temps de réaction que l'on cherche à mesurer.

Pour faire office de contre-exemple, on peut imaginer que l'on ait choisi comme couleurs "violet" (deux syllabes) et "bleu" (une syllabe). On peut imaginer que la séquence composée de six "violet" soit plus longue à reproduire que la séquence composée de six "bleu" pour les raisons évoquées précédemment alors qu'elles ont toutes les deux la même complexité. Nous nous sommes donc arrêtés sur le "bleu", le "jaune", et le "rouge" qui ont de plus l'avantage d'être parfaitement distinguables par tout le monde, même les daltoniens.



FIGURE 3 – Deux choix de couleurs possibles

2.2.2 DEVELOPPEMENT MOBILE

Pour le développement de l'expérience en tant que telle, nous avons d'abord dû choisir l'environnement de développement. Le choix s'est porté sur ReactNative pour deux raisons principales. D'abord, cela permet de développer sur les deux plateformes portables principales (Android et iOS) simultanément ce qui réduit grandement la difficulté et le temps de développement tout en touchant un maximum de personnes. De plus ReactNative est réputé comme d'une difficulté d'apprentissage raisonnable en plus d'être fortement documenté. L'architecture de l'application que nous avons codée est simplifiée dans "Annexe 1 - Architecture de l'application". Puis, au cours des réunions et de nos discussions nous avons fixé un certain nombre de paramètres.

Ecran réponse - L'une des problématiques majeures que nous avions à traiter était de trouver la manière optimale pour l'utilisateur de reproduire la séquence sur un écran tactile. Il fallait trouver une solution qui limite un maximum la présence de biais moteurs : le temps pris par l'utilisateur pour reproduire une séquence devait être uniquement lié à la complexité d'encodage de celle-ci et non à la façon dont elle est tapée sur l'écran. Ainsi, une entrée sous forme de trois boutons (Figure 4.a) sur lesquels on peut cliquer plusieurs fois n'était pas envisageable, puisque l'utilisateur serait beaucoup plus rapide lorsqu'une couleur se répète dans la séquence, sans pour autant que la séquence soit réellement moins complexe. Nous en sommes arrivé à deux possibilités :

- Des boutons disposés en triangle, de manière à ce qu'après chaque clic la position des boutons permutent aléatoirement (Figure 4.b).
- Une matrice où chaque couleur est présente six fois, dont la disposition est aléatoire, et où chaque bouton disparaît une fois qu'on a cliqué dessus (Figure 4.c).

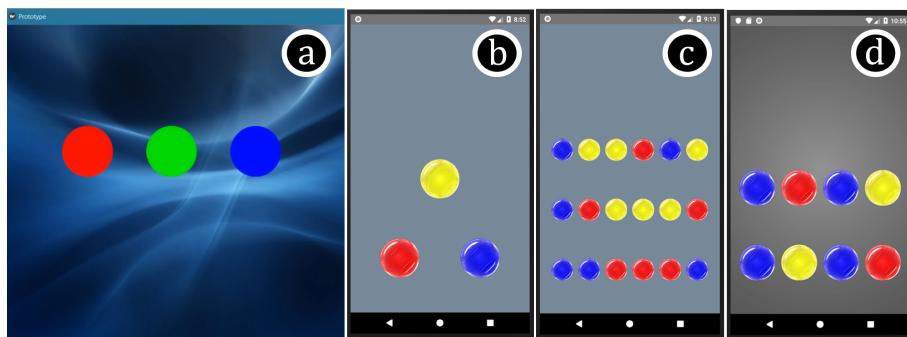


FIGURE 4 – Quatre types d'écran réponses envisagés

Finalement, une dernière modification a été faite à cette représentation matricielle. La ma-

trice ne comporte plus que huit boutons (Figure 4.d) : les six couleurs présentes dans la séquence à reproduire et deux distracteurs. Ce choix a été motivé par le fait qu'avec un grand nombre de boutons, ceux-ci pouvaient être trop petits pour certains doigts. Par ailleurs, l'expérience étant vouée à être menée par des singes, on nous a expliqué qu'il était préférable de réduire le nombre de boutons pour faciliter leur compréhension de l'expérience et ainsi réduire le temps d'apprentissage des animaux testés.

Vitesse d'affichage - La deuxième question qui se posait était l'affichage de la séquence au sujet. En effet, nous nous sommes rendu compte durant le codage de l'application que l'on pouvait jouer aisément sur 4 durées : la durée d'affichage, la durée de persistance pendant laquelle la bulle reste sur l'écran, la durée de disparition et la durée qui sépare l'apparition de deux bulles. Nous avons donc réalisé un écran permettant de changer ces paramètres, pour pouvoir les tester et incorporer les bons paramètres dans l'expérience finale. Finalement, nous avons fixé la durée de persistance à 100 millisecondes et la durée de séparation de deux stimuli à 300 millisecondes (Les deux autres durées ont été réduites à 5 millisecondes ce qui est équivalent à une durée nulle en considération du fait que l'oeil traite seulement 40 images par seconde).

Ces durées forment un compromis : un flash trop court entraîne un mélange des différentes couleurs lié à la fréquence d'affichage des téléphones, tandis qu'un flash trop long entraîne la mémorisation des noms des couleurs créant ainsi un biais supplémentaire, lié au fait que l'on ne retienne non plus une séquence temporelle mais un seul long mot. De plus, en science cognitive, la durée de 400 millisecondes est très utilisée car elle correspond à la durée que le cerveau met pour traiter un stimulus. Cette durée a été mise en évidence par Kutas et Hillyard qui ont découvert un saut dans le potentiel électrique neuronal survenant au bout de 400 millisecondes lorsque qu'ils présentaient un stimulus (un mot en l'occurrence) anormal dans une séquence [23]. Ainsi, lorsqu'une bulle arrive sur l'écran, celle qui la précède aura déjà été traitée par le cerveau avec les paramètres que l'on a fixés.

Ajout de texture - Nous avons assez vite eu l'idée, pour rendre l'expérience plus ludique de modifier les ronds, patchs de couleur uniforme, en une structure 3D de bulle ; en ayant ainsi l'impression d'éclater des bulles et non de cliquer sur une couleur. Mais l'ajout de la texture sur le rond a un autre avantage : le cortex visuel n'encode pas très bien les stimuli homogènes. On s'affranchit donc de ce dernier problème. Cet ajout de texture se fait à travers un calque de luminosité ajouté sur



FIGURE 5 – Texture de bulle

le patch rond de couleur homogène : la présence seule de zones de lumière et de zones d'ombre sur le patch suffit à tromper le cerveau et donner l'impression d'une structure 3D.

Choix du fond de l'application - Le fond présente deux caractéristiques utiles :

- La couleur grise qui est parfaitement neutre vis-à-vis de toutes les autres couleurs de l'expérience et ne favorise donc pas la perception d'une couleur en particulier
- Le dégradé qui attire le regard vers le centre de l'image et améliore la perception du stimulus

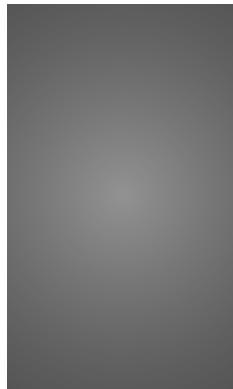


FIGURE 6 – Fond convergent vers le point d'attention

Implémentation d'un tutoriel - Enfin, afin de s'assurer que les sujets aient bien compris l'expérience avant de commencer, nous avons mis en place un court tutoriel au cours duquel ils doivent reproduire des séquences simples de quatre bulles seulement.

2.3 RÉCUPÉRATION DES DONNÉES DE L'EXPÉRIENCE

2.3.1 RÉCUPÉRATION INFORMATIQUE DES DONNÉES

Afin d'analyser les données, il était nécessaire de pouvoir les récupérer et les stocker. S'est alors posée la question de quelles données nous seraient utiles pour l'analyse ultérieure (cf "Résultats et interprétations"). Sur la première version du jeu (Kivy sur ordinateur), la récupération des données se faisait en extrayant du jeu la séquence d'entrée, la réponse de l'utilisateur, le temps de réponse, et un numéro aléatoire permettant d'identifier la partie. Ces données étaient transmises dans un fichier .csv avant d'être envoyées vers un tableau Excell. Dans le cas de version smartphone de l'expérience, nous avons dû essayer plusieurs méthodes afin de récupérer les données. Pour cela, nous avons exploré plusieurs pistes, certaines ne conduisant pas à une

solution satisfaisante. Une des pistes que nous avons suivies est d'utiliser une API (ou interface de programmation applicative), c'est-à-dire un ensemble de fonctionnalités logicielles, déjà établies, fournies par un des sites d'une grande entreprise connue (Microsoft ou IBM). Cependant, pour cette solution, le coût de stockage des informations aurait été très important, et la mise en ligne de l'application aurait pris plus de temps. Nous avons alors décidé d'utiliser des serveurs disponibles pour stocker nos données. Nous disposions des serveurs de NeuroSpin, ainsi que de ceux du BR (binet réseau, association d'informatique de l'École polytechnique).

Le codage de la base de données a eu lieu en plusieurs étapes :

- la création de la base de données,
- la mise en place du routage des informations,
- la mise en ligne sur les serveurs du BR,
- la mise en ligne sur les serveurs de NeuroSpin.

La création de la base de données ne posa pas de problèmes. Nous avons pour cela utilisé Mysql, afin de créer une base de données qui pouvait être transférée sur n'importe quel serveur. Le routage fut la partie la plus compliquée, parce qu'elle nécessitait la connaissance des différents langages indispensables à cela. L'intégration de la base de données dans le code java de l'application a été réalisé en collaboration avec l'équipe qui s'occupait du codage de l'application. Le principal avantage de cette méthode est qu'elle permettait de changer de serveur facilement (passer le routage depuis la base de données stockée sur les serveurs du BR à celui du NeuroSpin). Enfin, il a fallu déposer les différents fichiers sur les serveurs. Pour des raisons de sécurité, celui du BR ne pouvait accueillir que des données provenant du réseau Eduroam, c'est-à-dire d'étudiants jouant sur le serveur de l'École polytechnique.

Le type de données récupérées fut aussi une des principales questions que nous nous sommes posées. Nous avons décidé de récupérer la séquence d'entrée, la séquence de sortie, ainsi que les temps de réponse. A partir de ces données, nous pouvons représenter les différentes informations qui nous intéressent, c'est-à-dire le taux de réussite global, le taux de réussite par séquence, le temps entre deux appuis successifs, en reliant chaque séquence à sa complexité. Nous avons décidé d'écartier l'identifiant utilisateur, puisque notre but n'est pas de suivre une seule personne qui joue de nombreuses fois, mais plutôt que chaque séquence soit jouée un maximum de fois.

2.3.2 PUBLICATION DE L'APPLICATION

Le 9 février, l'application était publiée. Il a donc fallu réfléchir à quelques aspects auxquels nous n'avions pas porté notre attention jusqu'ici :

- ◊ **Le nom -** Il devait à la fois être attractif tout en faisant apparaître le fait que c'était une expérience de neurosciences. Nous avons également pensé à jouer sur le fait qu'il y ait plusieurs couleurs et des bulles pour l'inclure dans le nom. Finalement, le choix s'est arrêté sur NeuroSequence.
- ◊ **Le design de l'application -** Il devait de la même façon concilier l'aspect "jeu" et "expérience". Le fond neutre permet de créer une ambiance sérieuse et d'éviter les distractions lors de l'expérience. L'allure du fond gris a déjà été évoquée dans "Choix du fond". Au contraire l'aspect du titre et du personnage de la page d'accueil donne un caractère ludique à l'expérience.

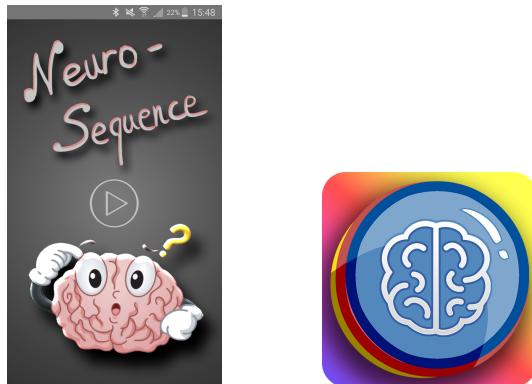


FIGURE 7 – Design général et logo de l'application

- ◊ **Le design de la page PlayStore -** Il reprend les thèmes de l'application en ajoutant le logo de NeuroSpin et celui de l'IPP.



FIGURE 8 – Photo de couverture Google Play Store

Comme l'application n'était pas achevée lors de la publication, les changements progressifs se sont faits par ajout de mises à jour. Ces dernières se font sur le PlayStore dans lequel il faut charger le nouveau fichier apk (fichier qui permet l'installation de l'expérience sur l'appareil). La version telle qu'elle se présente au moment de la rédaction du rapport est la 15ème mise à jour de l'application, chaque version permettant de corriger certains bugs, de changer des paramètres ou de les créer.

2.3.3 TECHNIQUES DE COMMUNICATION POUR INCITER AU JEU

Pour pouvoir tirer des conclusions pertinentes par rapport à notre analyse de données, nous avions besoin qu'un nombre important de parties soient jouées. Nous avons dans un premier lieu eu recours à notre cercle d'amis et à la participation des membres de la famille, néanmoins le nombre de parties jouées ne répondait toujours pas à nos attentes. C'est ainsi que nous avons eu l'idée de créer un concours ouvert aux étudiants de notre école, mais pas seulement, puisque nous avons également eu la participation d'étudiants de l'ENSTA Paritech et de l'ENSAE. Pour inciter les participants à jouer le plus possible, nous avons muni notre application d'un score qui récompense les séquences réussies successivement et puis nous avons promis des lots aux personnes justifiant du meilleur score. Le concours a été, entre autre, l'objet d'une publication sur les groupes d'élèves X 2017 et X 2018, dont l'image est présentée ci-dessous.



FIGURE 9 – Image de la publication Facebook

Notre stratégie de communication a été un réel succès puisqu'elle nous a permis de collecter pas moins de 5000 séquences jouées.

On peut aussi noter l'influence des deux publications sur les installations de l'application sur le PlayStore (Figure 10).

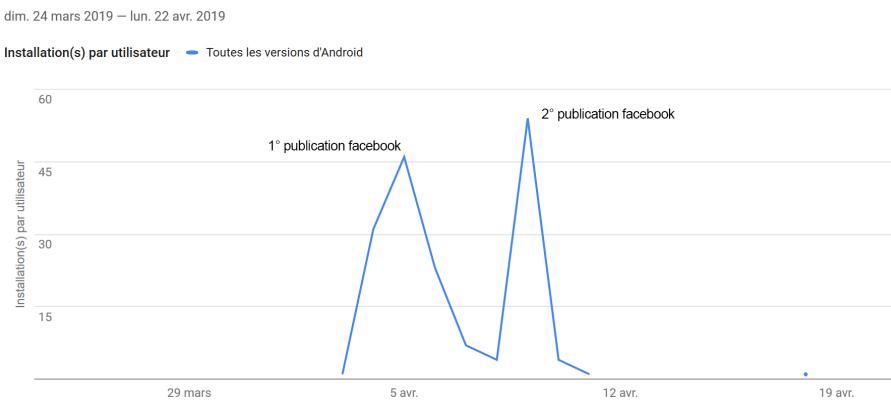


FIGURE 10 – Outil de statistiques de Google Play Console

Le nombre d'installations connaît deux pics :

- 45 installations le 5 avril, jour de la publication du post concours sur le groupe X2017
- 53 installations le 11 avril, jour de la publication du post concours sur le groupe X2018

L'application est par ailleurs encore disponible à cette adresse web sur le Play Store : <https://play.google.com/store/apps/details?id=com.neurosequence>

3

RÉSULTATS ET INTERPRÉTATIONS

Nous avions deux objectifs majeurs pour l'analyse des données. Premièrement valider le modèle d'un cerveau humain qui traite les séquences grâce aux deux opérations que sont la puissance et le miroir. Deuxièmement, tenter d'affiner ce modèle en cherchant le poids précis de ces deux méthodes dans le fonctionnement de la psyché humaine. Ce travail nécessite un grand jeu de données, ce que nous avons pu obtenir en ayant plus de 5 000 parties jouées. Nous en avons obtenus plusieurs résultats qui sont présentés dans la suite de cette partie.

Tout d'abord, nous avons procédé à un nettoyage et à une remise en forme des données. Ici il a évidemment fallu supprimer les séquences de longueur quatre qui servaient d'entraînement. Mais nous avons aussi enlevé les données absurdes ou incohérentes comme celles dont le temps de réponse dépassait les dix secondes ce qui correspond le plus souvent à un utilisateur non concentré ou qui a simplement oublié de répondre. Ainsi, après avoir procédé au nettoyage, nous avons transformé les trois types de données que nous avions récoltées (séquence demandée, séquence répondue, temps de réaction) en d'autres grandeurs utiles pour l'analyse : justesse ou non de la réponse, nombre de bulles bien placées, temps moyen entre les clics, et beaucoup d'autres...

3.1 PREMIERS RÉSULTATS ET INTERPRÉTATION

Notre objectif étant d'obtenir un maximum de résultats sur la base de données que nous avons construite, nous avons procédé à une démarche heuristique, exploratoire en choisissant de corrélérer des grandeurs qui nous semblaient pouvoir l'être. Ainsi, la première étape de notre travail a été la représentation graphique des données, afin de nous faire une première idée de la forme des résultats que nous avions récoltés.

On observe sur la Figure 11 un premier résultat positif et fondateur : les séquences de faible complexité sont en proportion plus souvent réussies, et avec des temps de réponse plus réduits, alors qu'en inversement, plus la complexité augmente plus la proportion d'erreurs et les temps de réponse sont élevés. On note cependant que cela ne représente qu'une tendance et qu'aucun groupe de points ne se détache réellement. On ne retrouve pas de manière évidente nos groupes

de complexités. Il faut donc affiner notre analyse, ce que nous ferons dans la suite.

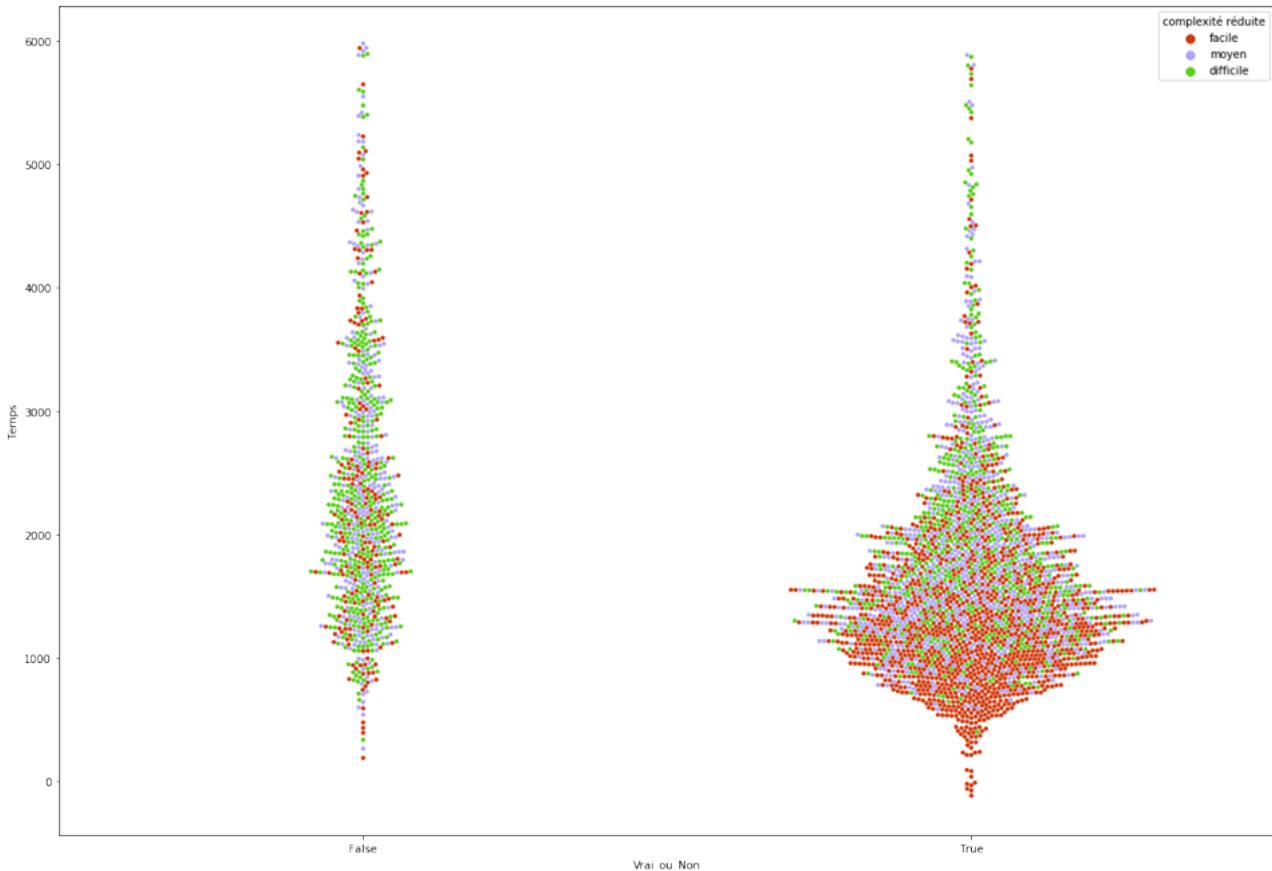


FIGURE 11 – Temps de réponse et réussite selon la complexité réduite

Nous avons ici séparé les séquences en trois groupes de complexité similaire. Celles de complexité 1, 2 et 3 sont “faciles” (représentés par les points rouges), 4 et 5 “moyennes” (points bleus) et 6 “difficiles” (points verts).

3.1.1 CORRÉLATION TEMPS-COMPLEXITÉ

Nous cherchons dans un premier temps à établir une corrélation entre la complexité théorique des séquences et une grandeur hybride qui prend en compte à la fois le temps de réponse et le taux d’erreur. Le problème de corrélation double étant assez courant en neurosciences, plusieurs grandeurs de ce type sont utilisées dans la littérature, nous avons choisi d’utiliser le temps LISAS (linear integrated speed-accuracy score), considéré comme plus intéressant sur de nombreux points [29]

Pour chaque type de séquence de couleurs, le temps LISAS est défini de la façon suivante :

$$LISAS = RT_j + \frac{S_{RT}}{S_{PE}} \times PE_j$$

Où :

- RT_j est le temps de réaction moyen sur l'ensemble des sujets sur cette séquence, lorsqu'elle était réussie
- PE_j est le taux d'erreur sur cette séquence
- S_{RT} est l'écart-type du temps de réaction
- S_{PE} est l'écart-type du taux d'erreurs

On observe en Figure 12.(a) un coefficient de corrélation très faible, en dessous du dixième. La faible corrélation est due à l'étalement (variance élevée) du LISAS au sein d'un groupe de même complexité. Il en ressort que le temps de réaction ne semble pas être un critère réellement discriminant selon les complexités des séquences. Cependant, en choisissant pour chaque complexité k , le LISAS moyen des séquences de séquences k , on observe sur la figure 12.(b) une croissance de la fonction de complexité. Pour comprendre pourquoi le LISAS n'est pas aussi pertinent qu'on ne le pensait, nous avons alors étudié une de ses composantes : le temps de réaction. Nous avons proposé l'hypothèse suivante : soit la séquence est bien retenue avant de commencer à donner la réponse, soit la séquence n'est pas retenue et le sujet se trompe. Dans tous les cas, avec cette hypothèse, le temps de réaction n'est plus un paramètre pertinent, car le délai pris pour transcrire la réponse est du même ordre de grandeur quelle que soit sa complexité.

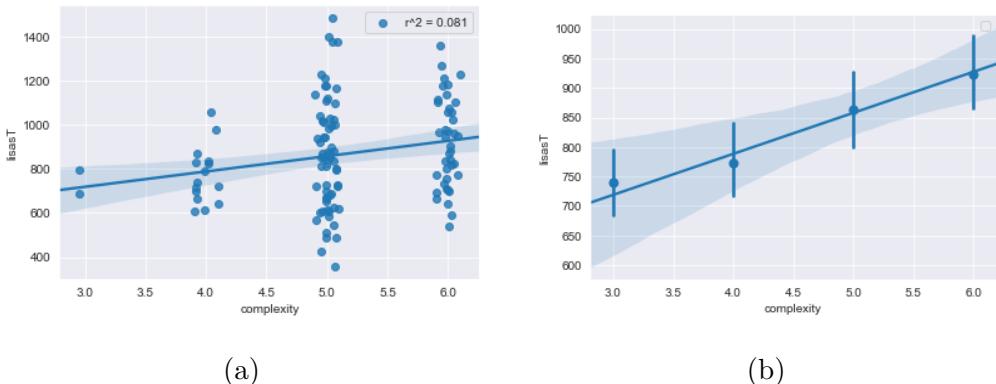


FIGURE 12 – Corrélation LISAS/complexité

Pour tester cette hypothèse, nous avons alors sélectionné les séquences qui avaient été réussies et avons calculé le temps moyen entre chaque clic selon les complexités. Le résultat est affiché en Figure 13.

complexité	Temps moyen entre click	Ecart type	IC bas	IC haut
2	267.88	99.49	238.81	296.95
3	342.37	163.11	319.05	365.69
4	369.57	161.65	360.32	378.82
5	443.52	202.34	431.17	455.87
6	457.67	183.88	446.34	469.01

FIGURE 13 – Temps moyen, écart type et intervalle de confiance selon la complexité

Ainsi, même en ne considérant que les séquences réussies, on retrouve la tendance que l'on avait en figure 12.(b). Il apparaît effectivement que les temps moyens entre chaque clic sur un bouton croissent avec la complexité de la séquence. Cependant, les écarts-types et les variances sont très élevés. Ici on a déterminé la moyenne sur toutes les séquences d'une même complexité de la moyenne des temps de réponses. Or la moyenne est un estimateur pratique mais assez instable, il suffit d'une réponse aberrante pour déplacer énormément l'estimation. Ainsi, on peut tout de même conclure une corrélation entre le temps de réponse et la complexité d'une séquence, peut-être le sujet réfléchit-il à la présence d'une structure lorsqu'il reproduit la séquence ? Cependant le critère n'apparaît pas comme assez discriminant et on préférera en utiliser d'autres par la suite. D'autant que ces résultats souffrent des biais moteurs liés à l'appui sur les boutons, même si nous avons cherché à les réduire autant que possible. Cela correspond par ailleurs à la conclusion à laquelle nous avions abouti a priori lors d'une réunion avec nos coordinateurs.

3.1.2 MESURE QUANTITATIVE DE L'ERREUR

Jusqu'à maintenant nous avons considéré que l'utilisateur pouvait avoir juste ou avoir faux, mais nous n'avons pas analysé l'erreur que celui-ci réalise. S'il se trompe sur une seule des bulles ce n'est évidemment pas la même erreur que s'il s'est trompé sur leur quasi totalité, qu'il a échangé les couleurs, etc. Pour cette analyse nous avons fait appel à une mesure utilisée en linguistique et en biologie sur l'ADN, la distance de Damerau-Levenshtein, qui permet de quantifier la distance entre deux séquences (ici l'entrée et la réponse) et donc "à quel point l'utilisateur avait juste". En résumé cette distance calcule le nombre d'éditions des caractères

(insertion, suppression, substitution et transposition) nécessaires pour passer de l'une à l'autre (Figure 14).

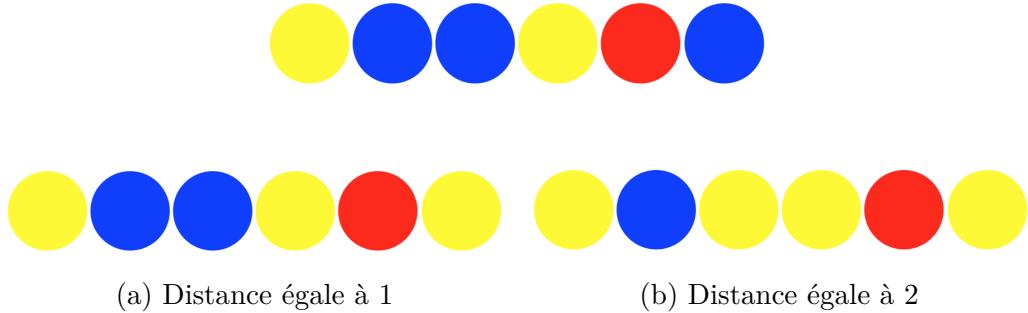


FIGURE 14 – Séquence initiale et exemple d'erreur et distance associée

Dans cette optique nous avons analysé, parmi les réponses fausses, la distance de Damerau-Levenshtein entre la séquence répondue par l'utilisateur et la réponse attendue. Nous avons regroupé les séquences par complexité. On obtient alors le jeu de courbes de la figure 4, tracé uniquement pour des complexités strictement supérieures à 3. Les courbes pour les complexités plus faibles n'ont pas été tracées car le nombre d'erreurs pour ces séquences était si faible - moins de dix - que les résultats étaient incohérents.

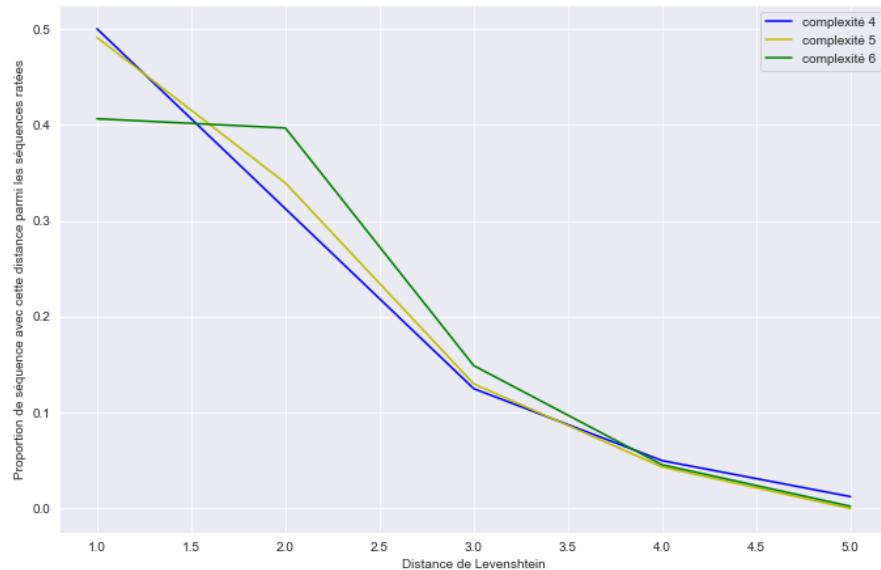


FIGURE 15

On observe sur la figure 15 que parmi les séquences ratées, celles de complexités élevées sont généralement plus éloignées de la bonne réponse que les séquences de complexité plus faible. Les différences peuvent sembler faibles sur le graphe étant donné que les échelles sont petites mais elles sont en fait de l'ordre de 7 à 20% .

Remarque : On notera que l'ensemble des études proposées ci-dessus ont été réalisées pour des complexités avec un choix de paramètres $m_2 = m_3 = 1$ et $p_2 = p_3 = 1$. Cela correspond au choix arbitraire fait pour le développement de l'application. On verra cependant par la suite que ce choix n'est pas incohérent.

3.1.3 ANALYSE DES PROPORTIONS D'ERREUR

Ainsi, lors d'une réunion avec nos coordinateurs, nous avons conclu que le critère discriminant majeur pour les complexités serait le taux de réussite au sein de chaque catégorie. La démarche de notre dernière analyse a alors été la suivante : pour plusieurs configurations de paramètres¹ numérotées de 1 à 5, nous avons cherché à établir des régressions linéaires entre les complexités qui en ressortent et les taux d'erreurs observés dans l'expérimentation.

Séquence	Configuration 1	Configuration 2	Configuration 4	Configuration 5
BBRRRR	3	3	3.5	3.5
JJB ¹ BJJ	5	4	4.5	4.5
JRJJRJ	4	3.5	3.75	3.75
BBRRBB	4	4.5	5	5.25

FIGURE 16 – Exemple de complexité selon les configurations

1. Les détails sur les différentes configurations sont disponibles en annexe

Les graphes représentant les taux d'erreurs obtenus pour chaque configuration sont les suivants :

- Résultat pour la première configuration

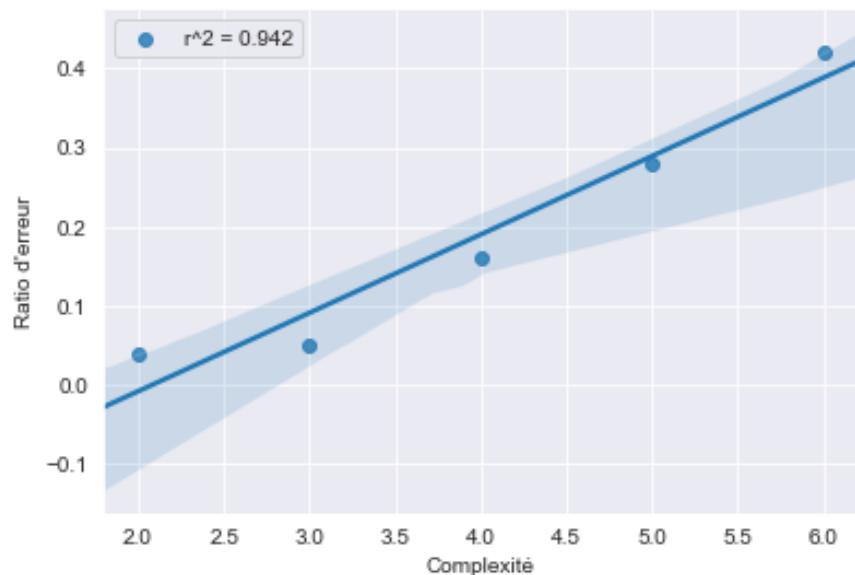


FIGURE 17 – 1^{ère} configuration

- Résultat pour la deuxième configuration

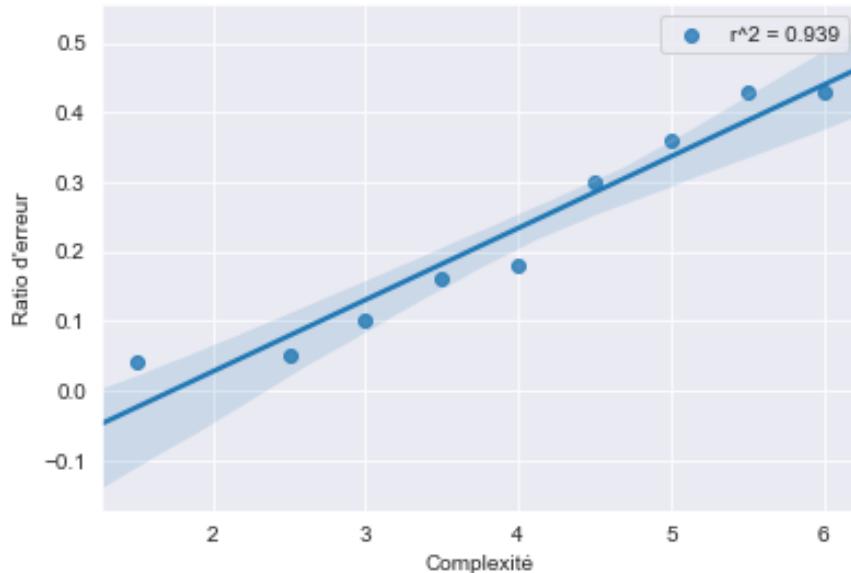


FIGURE 18 – 2^{ème} configuration

- Résultat pour la troisième configuration

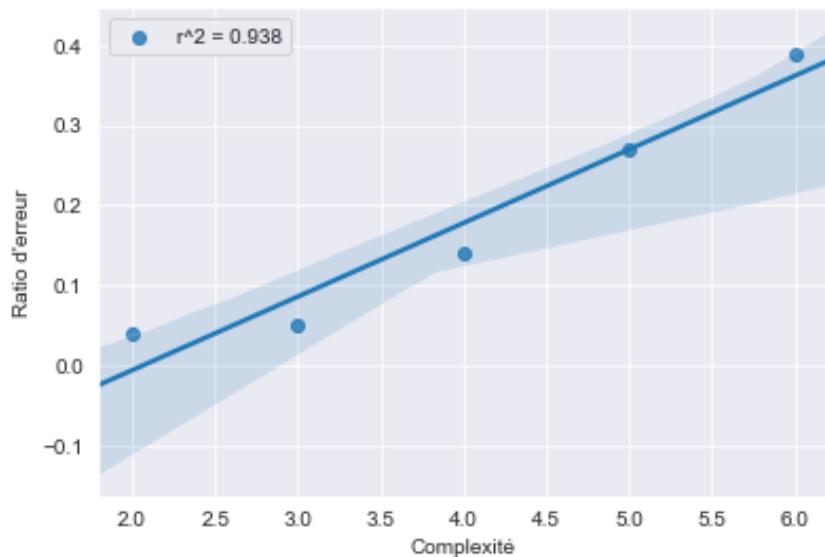
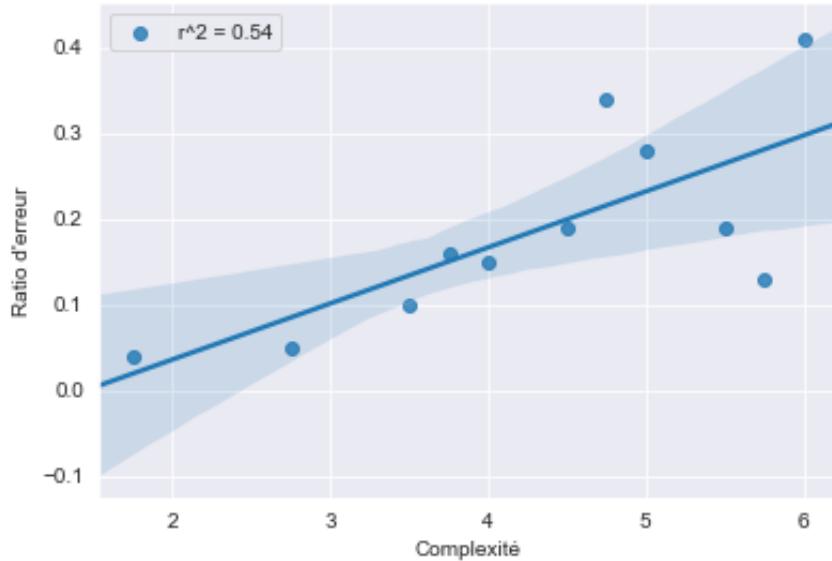
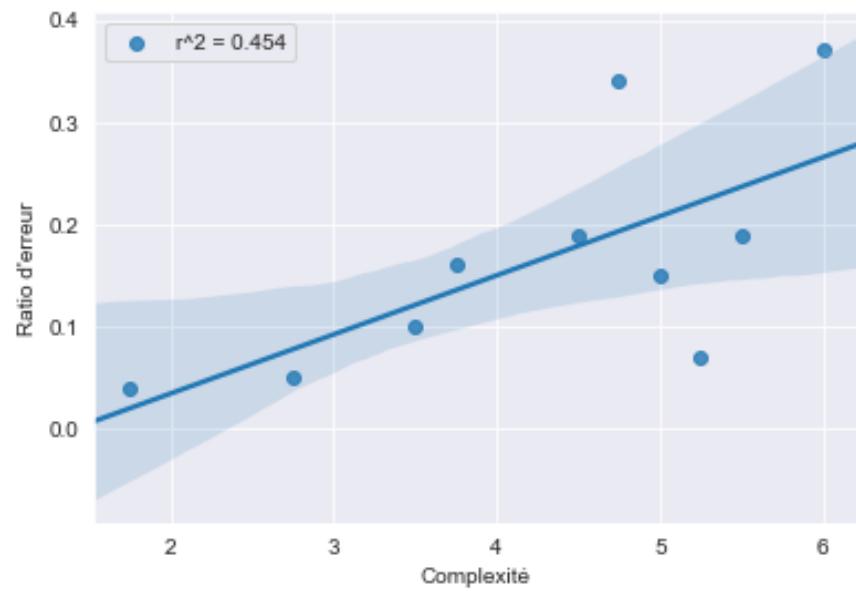


FIGURE 19 – 3^{ème} configuration

- Résultat pour la quatrième configuration

FIGURE 20 – 4^{ème} configuration

- Résultat pour la cinquième configuration

FIGURE 21 – 5^{ème} configuration

On remarque d'excellentes corrélations pour les trois premiers jeux de paramètres et des résultats bien plus décevants pour les deux suivants. On en conclut d'abord que le fait de prendre deux valeurs numériques différentes suivant si le miroir et la répétition concernent deux ou trois couleurs n'est pas pertinent. En effet, on peut interpréter cela par le fait que le nombre de couleurs dans la séquence à répéter, dont on prend le miroir n'influe pas sur sa détection par l'utilisateur.

Par ailleurs, en poussant l'analyse sur les trois premières configurations on observe que la valeur du paramètre de miroir semble n'avoir que très peu d'effets. En effet, les résultats sont quasiment les mêmes, que le miroir soit considéré aussi dur que la répétition, ou bien qu'il soit considéré comme indécelable (cas $m \geq 3$).

Quel jeu de paramètres choisir alors ? On peut supposer que la configuration $n^o 1$ profite de son faible nombre de points. Elle ne semble pas parfaitement adaptée dans le sens où elle ne valorise pas les répétitions doubles simples du type 'R,R' qui sont de complexité égale à 2 tout comme 'R,B' alors qu'on a ici une permutation qui est source de difficulté.

La configuration $n^o 3$ produit les mêmes complexités que la première quant aux répétitions. La seule différence est qu'elle donne une complexité plus élevée aux séquences miroir de type 'R,J,B,B,J,R'. Néanmoins, comme ces-dernières ne représentent qu'une minorité, les groupes de complexités restent en fait très proches de ceux de la configuration $n^o 1$.

Enfin, la configuration $n^o 2$ semble très intéressante. Elle détaille en fait les deux premières - on se rend compte que le groupe de complexité 3 de la configuration n1 se subdivise pour former les groupes 2.5, 3 et 3.5 de la configuration $n^o 2$. Cela est majoritairement dû au fait que l'on favorise maintenant les doublets de type 'R,R'. Obtenir une si belle corrélation pour un nombre de point deux fois plus élevé est un résultat très réjouissant. Ce choix de paramètres précise les groupes de complexité en préservant la tendance.

3.2 PERSPECTIVES ET REGARD CRITIQUE SUR LES RÉSULTATS

3.2.1 NETTOYAGE DES DONNÉES, PROBLÈME DU FORMAT APPLICATION DE L'EXPÉRIENCE

Contrairement à une expérience menée en laboratoire, nous obtenons certes énormément de résultats en un temps restreint, mais certaines données recueillies doivent être supprimées. Effectivement, les joueurs, sur leur smartphone, sont beaucoup moins concentrés que lors d'une

expérience en laboratoire. Ainsi, par exemple certains tests duraient six minutes pour une seule séquence car la personne avait parlé à son voisin et arrêté son jeu au milieu de la phase réponse. Nous avons donc dû nettoyer un certain nombre de données (i.e. enlever les résultats parasites). Il n'est pas exclu que des éléments nous aient échappés, nos analyses se font donc avec ces limitations.

Par ailleurs pour motiver les utilisateurs à rester concentrés et à faire de leur mieux, nous avions mis en place un jeu concours. Cela a cependant amené quelques individus à tenter de tricher. Nous avons alors, d'un côté, mis à jour l'application en prenant en compte les méthodes de triches utilisées et en les contrant. D'un autre côté, en ayant repéré certains joueurs tricheurs nous avons essayé au mieux d'effacer leurs données mais cela n'est pas aisément fait, bien que ce comportement demeure heureusement marginal, il a pu influencer nos analyses.

3.2.2 LA NON PERTINENCE DU MIROIR COMME PRIMITIVE

Le résultat négatif le plus fort de ce travail est que le miroir n'apparaît pas comme une primitive pertinente pour expliquer la manière de compresser des humains. Il semble que cela reflète le fait que les séquences sont flashées et qu'elles perdent ainsi leur caractère spatial qui est fondamental pour distinguer un miroir. Le miroir ne serait donc pas associé au codage abstrait des séquences ? Une autre explication serait en fait la longueur des séquences que nous traitons [6]. Celle-ci ne permet de toutes façons pas de montrer des miroirs convaincants car d'autres méthodes sont plus efficaces. Par exemple, $ABBA = M(AB)$ est certes un miroir, mais de manière assez claire, il peut aussi être vu comme AB^2A .

Il est à noter que de manière intéressante les six membres du PSC trouvent au contraire que les séquences présentant un miroir sont plus faciles à mémoriser. Naturellement nous avons tous passé un certain temps à travailler sur les séquences miroirs, notamment sur papier, et donc probablement à avoir lié différentes représentations (visuelles, auditives) entre elles. On pourrait imaginer qu'en expliquant à une personne les différentes méthodes de compressions qu'il peut utiliser, celui-ci les utilise effectivement pour aider sa mémoire. Cette capacité d'apprentissage très élevée est justement une des différences fondamentales entre singes et humains. Pour revenir au thème général dans lequel notre travail se situe, le langage, celui-ci est appris. Dans ce contexte un travail ultérieur sur une quantification de l'effet de l'apprentissage des compressions différentes de séquences serait intéressant. Entraînés, des sujets pourraient possiblement obtenir de meilleurs résultats. On pourrait par exemple comparer des sujets entraînés en autodidactes, à ceux ayant appris les primitives du langage développé pour ce travail.

CONCLUSION

A la frontière des sciences et des humanités notre travail de neurosciences s'est ancré dans une démarche surtout expérimentale. Nous avons formulé l'idée que l'homme utilise une compression passant par deux méthodes, l'exponentiation et le miroir.

La mise en place de toutes les modalités de l'expérience a été complexe et plus chronophage que prévue : développement des différentes versions de l'application, détermination et affinage des paramètres expérimentaux, génération automatique des séquences, modalités d'une publication d'application pour smartphone, etc. Les difficultés techniques ont été présentes, et nous ont poussés à revoir notre calendrier initial. Notre objectif, ambitieux, était d'être le plus complet possible : non pas étudier quelques séquences intéressantes quoique disparates, mais de réaliser une réelle base de données de résultats utilisables en vue de tester notre hypothèse. Grâce à plus de 16 heures de jeu de volontaires nous avons pu amasser une masse de données suffisante pour répondre à nos interrogations de façon pertinente.

L'enseignement principal est que notre modèle originel, avec la primitive "miroir" ayant le même poids que celle "puissance", rend compte correctement des difficultés effectives vues par les utilisateurs pour mémoriser les séquences. Cependant, en affinant ce modèle, il apparaît que le miroir n'a pas d'influence, dans le cadre des séquences étudiées.

Qu'attendre alors de l'adaptation de notre expérience à des singes ? Tout au long de notre travail, nous avons réfléchi dans l'objectif que l'expérience mise en place puisse être effectuée sur des singes dans le laboratoire de Neurospin, afin de comparer les résultats avec ceux obtenus sur des humains. On peut postuler que les singes ne détectent pas les séquences "miroir", et sont capable de mémoriser des puissances, mais seulement si celles-ci ne sont pas complexes. En effet, des séquences de la forme a^4 ou $(ab)^3$ nécessitent une aisance avec les nombres que les singes ne possèdent à priori pas.

ANNEXE

3.3 ANNEXE 1 - ARCHITECTURE DE L'APPLICATION

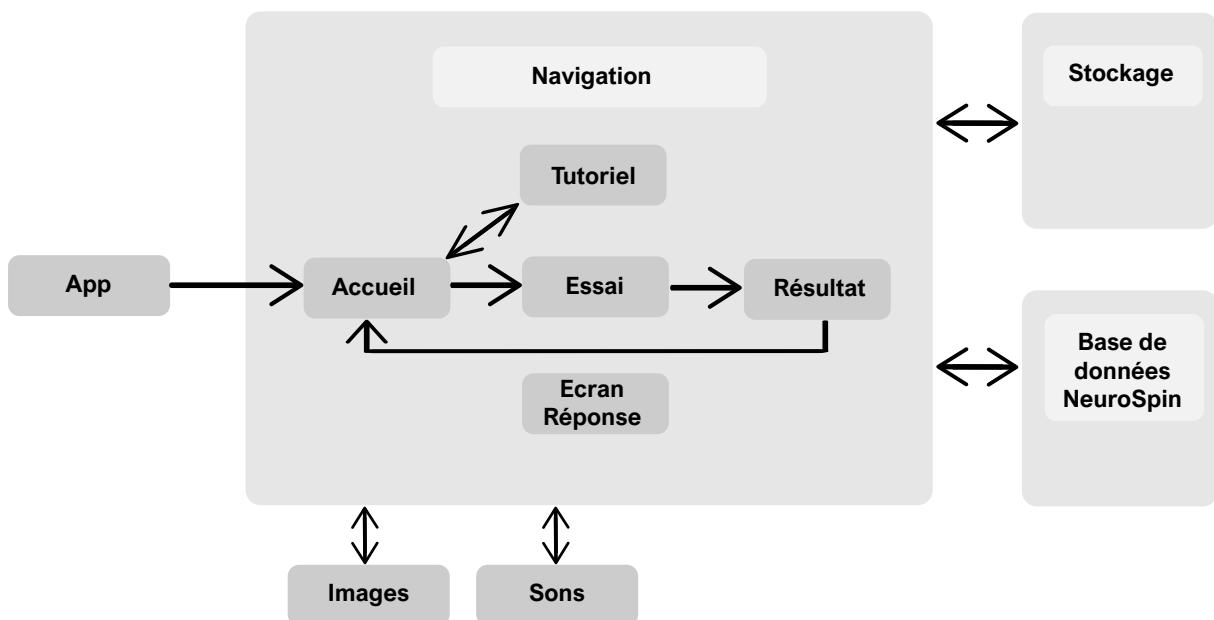


FIGURE 22 – Architecture de l'application

3.4 ANNEXE 2 - DIFFÉRENTES CONFIGURATIONS

- Configuration n° 1

Complexité miroir	1
Complexité répétition	1
Les différentes complexités des séquences	2, 3, 4, 5, 6

- Configuration n° 2

Complexité miroir	2
Complexité répétition	0.5
Les différentes complexités des séquences	1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6

- Configuration n° 3

Complexité miroir	6
Complexité répétition	1
Les différentes complexités des séquences	2, 3, 4, 5, 6

- Configuration n° 4

Complexité miroir	2
Complexité répétition	0.75
Complexité miroir expression de 3 couleurs	3
Complexité répétition expression de 3 couleurs	1
Les différentes complexités des séquences	1.75, 2.75, 3.5, 3.75, 4, 4.5, 4.75, 5, 5.5, 5.75, 6

- Configuration n° 5

Complexité miroir	6
Complexité répétition	0.75
Complexité miroir expression de 3 couleurs	6
Complexité répétition expression de 3 couleurs	2
Les différentes complexités des séquences	1.75, 2.75, 3.5, 3.75, 4.5, 4.75, 5, 5.25, 5.5, 6

RÉFÉRENCES BIBLIOGRAPHIQUES

- [1] DEHAENE, S., MEYNIEL, F., WACONGNE, C., WANG, L. AND PALLIER, C.. The neural representation of sequences : from transition probabilities to algebraic patterns and linguistic trees. *Neuron*, 88(1), 2015, pp.2-19.
- [2] YANG C.. Ontogeny and phylogeny of language. *Proc Natl Acad Sci*. 2013; 110 : 6324–6327. doi : 10.1073/pnas.1216803110 PMID : 23576720
- [3] SAFFRAN JR, ASLIN RN, NEWPORT EL. Statistical Learning by 8-Month-Old Infants. *Science*. 1996 ; 274 : 1926–1928. PMID : 8943209
- [4] KABDEBON C, PEÑA M, BUIATTI M, DEHAENE-LAMBERTZ G. Electrophysiological evidence of statistical learning of long-distance dependencies in 8-month-old preterm and full-term infants. *Brain Lang*. 2015 ; 148 : 25–36. doi : 10.1016/j.bandl.2015.03.005 PMID : 25865749
- [5] SAFFRAN JR, WILSON DP. From Syllables to Syntax : Multilevel Statistical Learning by 12-Month-Old Infants. *Infancy*. 2003 ; 4 : 273–284.
- [6] RESTLE F. Theory of serial pattern learning : structural trees. *Psychol Rev*. 1970 ; 77 : 481.
- [7] SAKAI K, KITAGUCHI K, HIKOSAKA O. Chunking during human visuomotor sequence learning. *Exp Brain Res*. 2003 ; 152 : 229–242. doi : 10.1007/s00221-003-1548-8 PMID : 12879170
- [8] PEÑA M., BONATTI L., NESPOR M., MEHLER J.. Signal-Driven Computations in Speech Processing. *Science*. 2002 ; 298 : 604–607. doi : 10.1126/science.1072901 PMID : 12202684
- [9] MARCUS GF, VIJAYAN S, RAO SB, VISHTON PM. Rule learning by seven-month-old infants. *Science*. 1999 ; 283 : 77–80. PMID : 9872745
- [10] WILSON, B., MARSLEN-WILSON, W.D. AND PETKOV, C.I.. Conserved sequence processing in primate frontal cortex. *Trends in neurosciences*, 40(2), 2017 pp.72-82.
- [11] WANG L, UHRIG L, JARRAYA B, DEHAENE S. Representation of Numerical and Sequential Patterns in Macaque and Human Brains. *Curr Biol*. 2015 ; 25 : 1966–1974. doi : 10.1016/j.cub.2015.06.035 PMID : 26212883
- [12] BAHLMANN J, SCHUBOTZ RI, FRIEDERICI AD. Hierarchical artificial grammar processing engages Broca's area. *NeuroImage*. 2008 ; 42 : 525–534.

- [13] MONTI MM, PARSONS LM, OSHERSON DN. Thought Beyond Language : Neural Dissociation of Algebra and Natural Language. *Psychol Sci.* 2012 ; 23 : 914–922. doi : 10.1177/0956797612437427 PMID : 22760883
- [14] PREMACK, D. AND PREMACK, A.. *The mind of an ape*. New York : Norton, 1983.
- [15] SAFFRAN JR, ASLIN RN, NEWPORT EL. Statistical Learning by 8-Month-Old Infants. *Science*. 1996 ; 274 : 1926–1928. PMID : 8943209
- [16] SAFFRAN JR, WILSON DP. From Syllables to Syntax : Multilevel Statistical Learning by 12-Month-Old Infants. *Infancy*. 2003 ; 4 : 273–284.
- [17] AMALRIC M, WANG L, PICA P, FIGUEIRAS, SIGMAN M, DEHAENE S (2017) The language of geometry :Fast comprehension of geometrical primitives and rules in human adults and preschoolers.
- [18] BARRIEL V., Ces 1,4 % qui nous séparent des chimpanzés!, Article de la revue M/S : médecine sciences Volume 20, Numéro 10, octobre 2004, p. 859–861
- [19] S. ROMANO, M. SIGMAN, S. FIGUEIRA LT2C2 : A language of thought with Turing-computable Kolmogorov complexity, *Papers in Physics*, vol. 5, art. 050001 (2013)
- [20] NICOLAS GAUVRIT·HENRIK SINGMANN· FERNANDO SOLER-TOSCANO·HECTOR ZENIL, Algorithmic complexity for psychology : a user-friendly implementation of the coding theorem method, *Psychonomic Society, Inc.* 2015
- [21] CILIBRASI R, VITA NYI PM. Clustering by compression. *IEEE Trans Inf Theory*. 2005 ; 51 : 1523–1545.
- [22] GRUNWALD P. A tutorial introduction to the minimum description length principle. *ArXiv Prepr Math0406077*. 2004 ; <http://arxiv.org/abs/math/0406077>
- [23] KUTAS, M. ; HILLYARD, S. A. (1980). "Reading senseless sentences : Brain potentials reflect semantic incongruity". *Science*. 207 (4427) : 203–208. doi :10.1126/science.7350657. PMID 7350657
- [24] AMALRIC M, WANG L, PICA P, FIGUEIRAS, SIGMAN M, DEHAENE S (2017), Language of Geometry – Syntax, Semantics and Complexity, May 4, 2016
- [25] MEYNIEL F, MAHEU M, DEHAENE S (2016) Human Inferences about Sequences : A Minimal Transition Probability Model.
- [26] DEHAENE, S., Origine du langage et singularité de l'espèce humaine, Cours de psychologie cognitive expérimentale, Collège de France, 2017-2018
- [27] JIANG, X., LONG, T., CAO, W., LI, J., DEHAENE, S. AND WANG, L. Production of supra-regular spatial sequences by macaque monkeys. *Current Biology*, 2018.

- [28] RESTLE, F. Theory of serial pattern learning : Structural trees. *Psychological Review*, 77(6), 1970, p.481.
- [29] A. VANDIERENDONCK A comparison of methods to combine speed and accuracy measures of performance : A rejoinder on the binning procedure Published online : 4 March 2016 in Psychonomic Society, Inc. 2016