

ANALYSE DE SÉQUENCES OCULOMÉTRIQUES ET D'ÉLECTROENCÉPHALOGRAMMES PAR MODÈLES MARKOVIENS CACHÉS

Jean-Baptiste Durand¹, Anne Guérin-Dugué & Sophie Achard²

¹ *Univ. Grenoble Alpes, Laboratoire Jean Kuntzmann et Inria, Mistis, 51 rue des Mathématiques, B.P. 53, F-38041 Grenoble cedex 9, France ; jean-baptiste.durand@imag.fr*

² *Univ. Grenoble Alpes, CNRS, GIPSA-lab, 11 rue des Mathématiques, B.P. 46, F-38040 Grenoble, France ; [anne.guerin|sophie.achard]@gipsa-lab.grenoble-inp.fr*

Résumé. Cette étude vise à analyser des séquences de mouvements oculaires collectées au cours de tâches de lecture visant à acquérir de l'information à des fins décisionnelles. L'analyse, basée sur des semi-chaînes de Markov cachées, met en évidence différentes phases d'acquisition, qui sont alors reliées à des caractéristiques de signaux électroencéphalographiques multicanaux collectés concomitamment à la lecture. Cette analyse permet de révéler des changements de variance et de corrélation entre canaux suivant les phases et les bandes de fréquence.

Mots-clés. Oculométrie, électroencéphalogrammes, séquences multimodales, semi-chaînes de Markov cachées.

Abstract. This work aims at analysing sequences of eye movements. These sequences were measured during reading tasks involving information acquisition so as to take decisions. Their analysis, based on hidden semi-Markov chains, highlights different phases of acquisition, which can be related to particular characteristics in multichannel electroencephalograms measured synchronously with eye movements during the reading tasks. This analysis reveals changes associated with the different phases of information acquisition, occurring in variances and correlations between channels at specific frequencies.

Keywords. Eye movements, electroencephalograms, multimodal sequences, hidden semi-Markov chains.

1 Introduction

Cette étude consiste en l'analyse jointe de signaux oculométriques (suivi du mouvement des yeux au cours du temps) et d'électroencéphalogrammes (EEG) collectés sur des participants lors de tâches de lecture visant à acquérir de l'information afin de prendre une décision.

1.1 Objectifs de l'étude

Simola *et al.* (2008) ont montré par l'analyse oculométrique d'expériences d'acquisition et de recherche d'information textuelle que la lecture n'est pas un processus homogène dans le temps, ni d'un point de vue cognitif, ni du point de vue du mouvement des yeux. Au contraire, ces types de processus mettent en œuvre plusieurs étapes ou phases, telles que la lecture rapide, normale, approfondie – en termes d'oculométrie – et la formulation, l'élimination d'hypothèses, la confirmation et la décision – en termes cognitifs. Sur la base d'un modèle markovien caché, les auteurs ont mis en évidence la possibilité d'identifier des étapes à partir des seuls mouvements oculaires. Il est alors possible d'interpréter la succession de ces étapes lors d'une séquence observée particulière comme une stratégie de lecture traduisant une progression par étapes cognitives distinctes.

Notre objectif est d'une part d'identifier ces étapes dans un nouveau contexte expérimental et d'autre part, d'arriver à les caractériser du point de vue des EEG. À terme, nous souhaiterions également quantifier la diversité des stratégies de lecture suivant les textes et suivant les participants.

1.2 Protocole et données

Les données ont été obtenues à partir d'expériences au cours desquelles des participants ont à accomplir une sorte de revue de presse. Un thème est présenté au participant, puis un texte qui traite soit de ce thème, soit d'un thème complètement différent, soit d'un thème proche. Le participant doit décider le plus rapidement possible si le texte traite ou non du thème présenté. Le protocole est intégralement décrit dans Frey *et al.* (2013). Notre étude se limite aux textes traitant du thème présenté préalablement. Le corpus se compose de 60 textes, chacun d'entre eux étant lu (dans un ordre randomisé) par les 15 participants. Lors de la lecture et jusqu'à la prise de décision, la position des yeux sur l'écran et les EEG sur 30 canaux sont enregistrés avec un pas de temps de 1 ms. La lecture est composée d'une succession de périodes d'environ 180 ms, appelées fixations, durant lesquelles le regard se pose sur une portion du texte comprenant typiquement de 5 à 10 caractères, et de périodes d'environ 40 ms, appelées saccades, qui permettent aux yeux de se fixer sur une autre portion de texte. La fixation est le pas de temps naturel (bien que de durée variable) pour l'analyse oculométrique, car on peut associer à chaque saccade la lecture de quelques syllabes ou mots très courts.

Après élimination de données aberrantes, le jeu analysé comprend 897 séquences, appelées traces oculométriques, dont la longueur cumulée est de 13 797 fixations. Pour une fixation t d'un participant s sur un document d , la donnée est une variable catégorielle $X_{s,d,t}$ caractérisant la saccade suivant cette fixation. En effet l'amplitude de la saccade «sortante» (relativement à une fixation donnée) caractérise localement la stratégie de lecture. En lecture usuelle, les mots sont fixés en séquence avec une probabilité plus importante de les sauter, s'ils sont courts, fréquents et prédictibles dans le contexte du texte.

Pour une lecture plus rapide, le nombre de mots sautés entre deux fixations sera plus important. Enfin, un mot plus long, moins fréquent pourra être refixé et on peut également observer des retours en arrière de relecture de portions de texte. Aussi ces différentes situations ont été catégorisées et codées par une variable pouvant prendre cinq valeurs possibles : 0 si la saccade a une amplitude de 2 mots ou plus dans le sens de lecture, 1 si l’amplitude est d’un mot, 2 si le même mot est relu, 3 si le mot précédent est lu et 4 si la saccade a une amplitude de 2 mots ou plus dans le sens opposé à la lecture. Les durées associées à chaque fixation sont aussi disponibles. La longueur de la séquence (s, d) est notée $T_{s,d}$. Un exemple de trace oculométrique est représenté Figure 1.

2 Modélisation des traces oculométriques

Afin d’identifier des phases dans le processus de lecture qui puissent être interprétées comme des étapes dans les processus cognitifs, des modèles de semi-chaîne de Markov cachée ont été utilisés. Des modèles de Markov cachés discriminatifs avaient été utilisés par Simola *et al.* (2008) mais dans un contexte un peu différent où l’objectif était la discrimination (supervisée) entre plusieurs types de tâches de lecture (trouver un mot, répondre à une question, trouver des textes intéressants). L’hypothèse sous-jacente à ces modèles est que des changements de stratégie au cours de la lecture reflètent des étapes cognitives, traduites indirectement à travers des changements de dynamique oculométrique (et donc des changements de régime des $(X_{s,d,t})_{1 \leq t \leq T_{s,d}}$).

Le modèle considéré est markovien caché à durée explicite, suivant la terminologie de Yu (2010). Les paramètres sont estimés par maximum de vraisemblance, et le choix de la famille des lois d’observation et de temps de séjour est effectué en même temps que l’estimation des paramètres, en utilisant le module `sequence_analysis` de la plate-forme OpenAlea (description détaillée par Durand, 2011). Les 897 séquences sont considérées comme mutuellement indépendantes, et de même loi. Le nombre d’états cachés est estimé à l’aide du critère ICL-BIC de McLachlan et Peel (2000, chap. 6) afin d’obtenir des états aux caractéristiques contrastées, et donc interprétables. Un modèle à 5 états a été sélectionné, avec deux classes d’états transitoires $\{0, 1, 2\}$ et $\{3\}$ et un état absorbant 4. Les états 0 et 1 ont tendance à alterner, et leur durée de séjour est très courte ; ils peuvent être interprétés comme une seule phase de lecture normale (phase 0) avec principalement $X_{s,d,t} \in \{0, 1, 2\}$. L’état 2 (phase 1) peut s’interpréter comme une phase de lecture rapide avec une plus forte proportion de $X_{s,d,t} = 0$, et l’état 3 (phase 2) comme une phase de lecture approfondie avec une plus faible proportion de $X_{s,d,t} = 0$. L’état absorbant 4 (phase 3) est caractérisé par une plus forte proportion de $X_{s,d,t} = 4$. Les lois des durées de fixation n’étant pas différenciées suivant ces états, cette variable a été ignorée.

La restauration des états cachés par principe du maximum a posteriori (séquence d’états la plus probable sachant les observations, calculée par l’algorithme de Viterbi) a permis de résumer les traces oculométriques à une séquence de phases (voir exemple Figure 1).

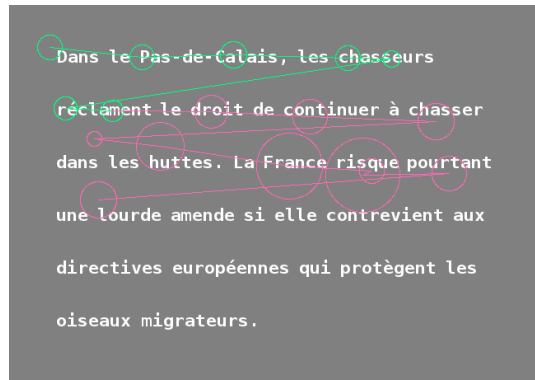


FIGURE 1 – Trace oculométrique avec phases restaurées. Chaque segment représente une saccade et chaque cercle, la durée de fixation (avec un rayon proportionnel à la durée). Les fixations de couleur verte sont associées à la phase restaurée 0 (lecture normale) et celles de couleur rose, à la phase 1 (lecture rapide).

3 Analyse des EEG

Par souci de synthèse nous nous focalisons sur 6 des 30 canaux choisis sur la base de l'étude préalable de Frey *et al.* (2013) pour leur rôle contrasté en terme d'analyse de processus cognitifs élémentaires mis en jeu dans l'analyse d'information textuelle : électrodes frontale gauche F7, frontale droite F8, frontale centrale Fz, centrale Cz, pariétale Pz et occipitale Oz (voir schéma des positions dans cette référence, Figure 5).

On s'attend à ce que l'activation d'une région du cerveau se traduise par l'augmentation de la variabilité des EEG dans certaines gammes de fréquences (voir Fitzgibbon *et al.*, 2004). Afin de conserver la dimension temporelle du signal et à la fois de distinguer suivant les fréquences, une représentation en ondelettes des EEG a été réalisée en utilisant une décomposition MODWT (Maximal Overlap Discret Transform) sur 7 échelles de résolution (la septième étant la moins résolue), les 5 dernières étant désignées par Gamma haute (62,5 à 125 Hz), Gamma basse (31,2 à 62,5 Hz), Bêta (15,6 à 62,5 Hz), Alpha (7,8 à 15,6 Hz) et Thêta (3,9 à 7,8 Hz).

La comparaison des variances par phase, par échelle et par électrode a mis en évidence certaines différences significatives (voir Figure 2a), si l'on considère que les coefficients d'ondelette sont indépendants et de même loi. Les différences les plus marquées sont associées à l'électrode Oz dans les bandes Thêta et Alpha, aux électrodes F7 et F8 dans la bande Thêta et à l'électrode F8 dans la bande Gamma haute. Cependant, la variabilité individuelle et liée au texte n'est pas prise en compte dans ces comparaisons (violation des hypothèses d'indépendance et de même loi) et pour certaines séquences, des inversions surviennent dans l'ordre des variances induit par les phases.

La comparaison des corrélations entre électrodes par phase fait apparaître, notamment dans la bande Thêta (Figure 2b), des différences marquées d'une phase à l'autre,

particulièrement entre les électrodes Pz et Cz, et également entre F7 et F8. De plus ces changements de corrélation sont bien plus reproductibles d’une séquence à l’autre que les changements de variance.

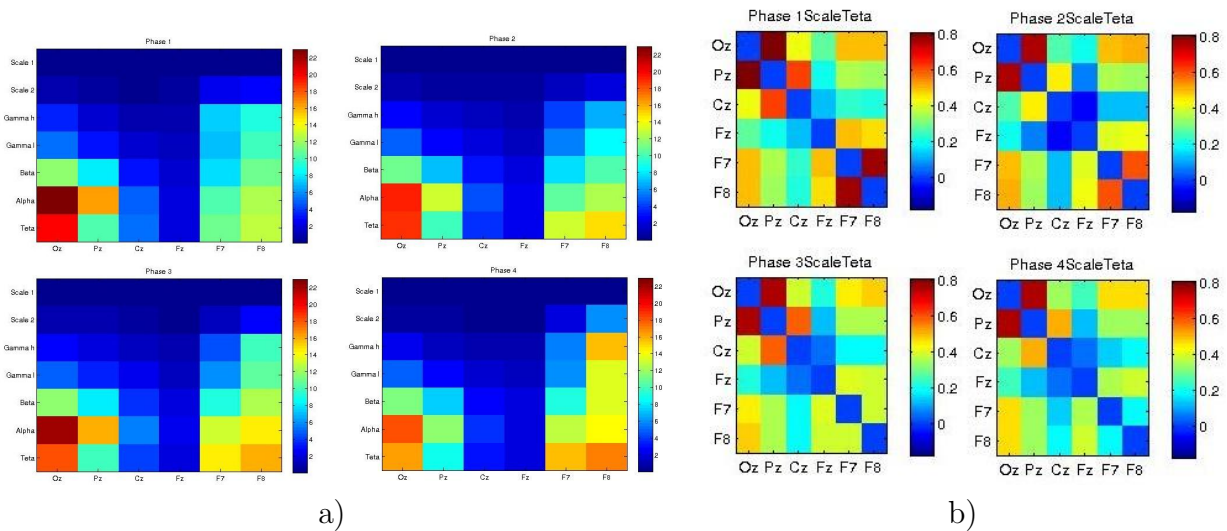


FIGURE 2 – Coefficients d’ondelette : a) variance empirique par phase, électrode et bande de fréquence ; b) corrélation empirique entre électrodes par phase dans la bande de fréquence Thêta.

4 Discussion et perspectives

La caractérisation des EEG par phase en section 3 n’est valide que si l’incertitude sur les états de la semi-chaîne de Markov cachée est faible, autrement dit si les caractéristiques des mouvements oculaires sont bien séparées d’un état à l’autre. Or des profils d’entropie (voir Durand et Guédon, 2016) mettent en évidence une incertitude importante dans les changements d’état. Il en résulte également une mauvaise identification des périodes de transition entre motifs de corrélation sur les coefficients d’ondelette.

Afin d’obtenir des phases aux caractéristiques plus contrastées, il est donc souhaitable d’intégrer les deux sources d’information, oculométrie et EEG, au sein d’un seul modèle composite qui étende par exemple les chaînes de Markov cachées couplées de Zhong et Ghosh (2001) au cas semi-Markovien avec des pas de temps différenciés pour les différentes sources, et un éventuel décalage temporel dans les transitions, reflétant une latence entre les processus de lecture et d’intégration de l’information.

Une fois réduite l’incertitude sur les phases, il deviendra alors possible pour les valider et affiner leur interprétation de les relier à des caractéristiques sémantiques du texte fixé à chaque instant, et de déterminer des clusters de participants et de textes pour

quantifier la variabilité des traces oculométriques liée à ces deux facteurs. Ces sources de variabilité pourront être intégrées au modèle proprement dit, plutôt qu'étudiées a posteriori et conditionnellement aux états restaurés.

Remerciements

Ce travail est en partie financé par le LabEx PERSYVAL-Lab (ANR-11-LABX-0025-01).

Bibliographie

- [1] Durand, J.-B. Statistical models of sequences and trees in OpenAlea. First OpenAlea workshop on Functional-Structural Plant Modelling. 14-15 Juin, Montpellier, France (2011)
- [2] Durand, J.-B et Guédon, Y. Localizing the latent structure canonical uncertainty : entropy profiles for hidden Markov models. *Statistics and Computing*, **26**(1), 549–567 (2016)
- [3] Fitzgibbon, S.P., Pope, K.J., Mackenzie, L., Clark, C.R. et Willoughby, J.O. Cognitive tasks augment gamma EEG power. *Clinical Neurophysiology*, **115**(8), 1802–1809 (2004).
- [4] Frey, A., Ionescu, B., Lemaire, B., Lopez Orozco, F. , Baccino, T. et Guérin-Dugué, A. Decision-making in information seeking on texts : an Eye-Fixation-Related Potentials investigation. *Frontiers in Systems Neuroscience* **7**, Article 39 (2013)
- [5] McLachlan, G., Peel, D. : Finite mixture models. Wiley Series in Probability and Statistics. John Wiley and Sons (2000)
- [6] Simola, J., Salojärvi, J. et Kojo, I. Using hidden Markov model to uncover processing states from eye movements in information search tasks. *Cognitive Systems Research* **9**(4), 237–251 (2008)
- [7] Yu, S.-Z. Hidden semi-Markov models. *Artificial Intelligence* **174**(2), 215–243 (2010)
- [8] Zhong, S. et Ghosh, J. A new formulation of coupled hidden Markov models. Rapport technique, Dept. of Electrical and Computer Eng., Univ. of Texas at Austin (2001)