

## Proposition de Master Recherche

### “Utiliser des modèles de Markov cachés (HMM) pour comprendre comment les bébés apprennent à parler”

Joan Birulés (post-doc Babylab LPNC), Mathilde Fort (LPNC), Julien Diard (LPNC)

Ce projet vise à utiliser des modèles de Markov cachés (HMM) pour analyser les données d'un ensemble d'études d'oculométrie sur l'attention sélective aux visages parlants lors de l'apprentissage de mots. L'objectif est de révéler les stratégies attentionnelles distinctes employées par les nourrissons, les enfants et les adultes lorsqu'ils explorent un visage parlant dans diverses conditions linguistiques. Ce projet exigera de l'étudiant de

- Se familiariser avec le domaine théorique de l'attention visuelle, de la perception audiovisuelle de la parole d'un point de vue développemental (par exemple, Birulés, et al, 2018 ; 2020 ; Fort et al. 2018).
- Apprendre à traiter les données d'oculométrie (en Matlab, R, etc.), et à utiliser le package SMAC Matlab d'A. Coutrot pour l'analyse basée sur les HMM (Coutrot, Hsiao, & Chan, 2018), package disponible à l'adresse : <http://antoinecoutrot.magix.net/public/code.html>.
- Interpréter, intégrer et discuter les résultats avec le contexte théorique dans un compte rendu final.

#### Informations pratiques :

Les candidats peuvent soit venir d'une formation en science des données ou en informatique, soit d'une formation en sciences cognitives, soit d'une formation en psychologie expérimentale, à condition qu'ils aient de bonnes compétences en analyse des données.

**Contact :** Joan Birulés : [joan.birules@univ-grenoble-alpes.fr](mailto:joan.birules@univ-grenoble-alpes.fr) et Mathilde Fort : [mathilde.fort@univ-grenoble-alpes.fr](mailto:mathilde.fort@univ-grenoble-alpes.fr)

#### Contexte théorique :

Le visage d'un interlocuteur fournit un grand nombre d'informations ; les yeux peuvent aider à identifier l'interlocuteur et, grâce à ses mouvements, nous pouvons déduire les états d'esprit, les attitudes et les intentions potentielles. D'autre part, le fait de regarder la bouche fournit des indices de parole auditifs et visuels spatio-temporels et acoustiquement congruents qui renforcent la saillance du signal de parole. L'exploration des stratégies attentionnelles distinctes employées par les nourrissons, les enfants et les adultes lors de la perception de la parole audiovisuelle nous permet de comprendre quelles sont les informations qu'ils privilégient à un moment donné, et offre donc une fenêtre sur les mécanismes cognitifs qui sous-tendent l'apprentissage du langage.

Les recherches antérieures sur le sujet ont principalement enregistré et analysé les données relatives au regard des percepteurs sur les visages parlants en calculant les proportions de temps total de regard (PTLT) sur deux zones d'intérêt (AOI) définies a priori, respectivement autour des yeux et de la bouche des visages parlants. Il reste à explorer si des analyses plus fines - à la fois spatiales et temporelles, au-delà des différences de temps de regard moyen - pourraient aider à mieux caractériser les stratégies d'exploration visuelle à différents stades du développement. Le projet actuel vise à explorer cette question en réanalysant un vaste ensemble de données d'études de perception audiovisuelle de la parole - des données de nourrissons de 4 mois à celles d'enfants et d'adultes (Birulés, Bosch, Brieke, Pons, & Lewkowicz, 2018 ; Birulés et al., 2020 ; Fort, Ayneto-Gimeno, Escrichs, & Sebastián-Gallés, 2018 ; Pons, Bosch, & Lewkowicz, 2015) - à l'aide de modèles de Markov cachés (HMM, exemple dans la figure 1). Cette méthode est axée sur les données et est capable d'encoder non seulement les aspects spatiaux mais aussi temporels des données, fournissant ainsi une description plus complète des schémas attentionnels dans le regard des perceurs. De cette façon, nous visons à fournir une explication plus détaillée et plus complète des différentes stratégies exploratoires employées pour traiter la parole audiovisuelle à différents stades de développement et dans différentes conditions linguistiques.

#### References:

- Birulés, J., Bosch, L., Brieke, R., Pons, F., & Lewkowicz, D. J. (2018). Inside bilingualism: Language background modulates selective attention to a talker's mouth. *Developmental Science*, 22(3), 1–11. <https://doi.org/10.1111/desc.12755>
- Birulés, J., Bosch, L., Pons, F., & Lewkowicz, D. J. (2020). Highly proficient L2 speakers still need to attend to a talker's mouth when processing L2 speech. *Language, Cognition and Neuroscience*, 35(10), 1314–1325. <https://doi.org/https://doi.org/10.1080/23273798.2020.1762905>
- Coutrot, A., Hsiao, J. H., & Chan, A. B. (2018). Scanpath modeling and classification with hidden Markov models. *Behavior Research Methods*, 50(1), 362–379. <https://doi.org/10.3758/s13428-017-0876-8>
- Fort, M., Ayneto-Gimeno, A., Escrichs, A., & Sebastián-Gallés, N. (2018). Impact of Bilingualism on Infants' Ability to Learn From Talking and Nontalking Faces. *Language Learning*, 68, 31–57. <https://doi.org/10.1111/lang.12273>

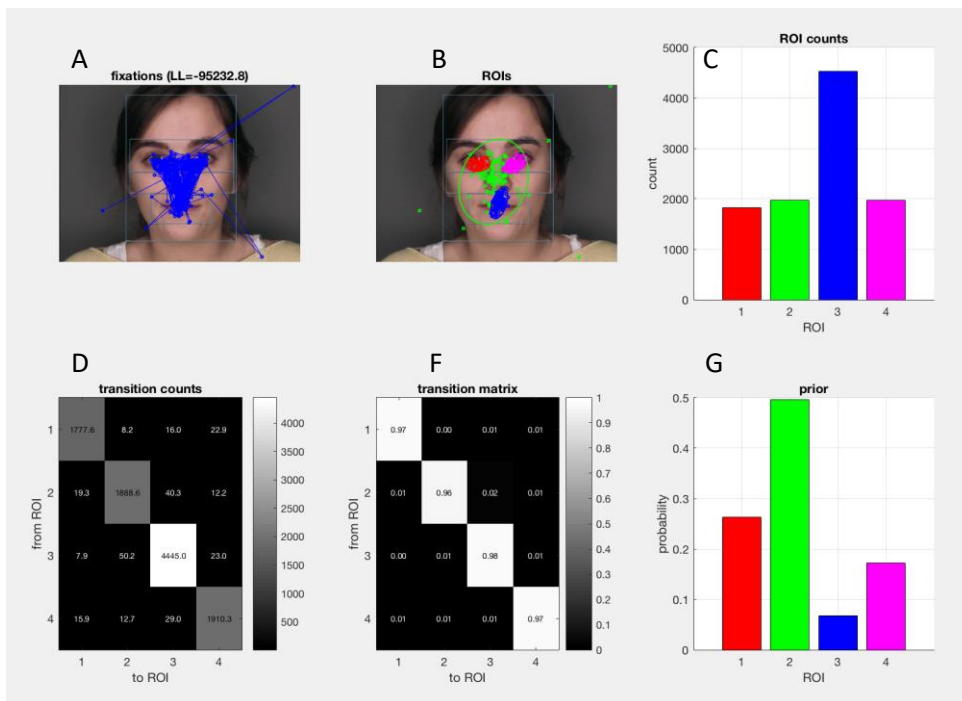


Figure 1. Exemple de sortie d'un HMM de groupe à 4 états. La figure montre A) les emplacements de fixation, B) les ROIs extraits par le HMM, C) le nombre de fixations dans chaque ROI, D) le nombre de transitions entre les différents ROIs, E) les probabilités de transitions entre les différents ROIs G) et la distribution préalable sur les ROIs. Cet exemple montre que l'analyse basée sur les HMM complète les états classiques des yeux et de la bouche par un nouvel état "reste du visage" (dans une analyse HMM à 4 états).