

Modélisation probabiliste de l'apprentissage de la lecture : simulation de l'acquisition du lexique dans un modèle de reconnaissance visuelle de mots

Cadre théorique

Les modèles mathématiques sont cruciaux pour comprendre les processus mis en œuvre pendant la lecture et son apprentissage. Cependant, bien qu'il existe déjà de nombreux modèles de lecture à haute voix (modèles double-voie (Coltheart et al., 2001), en triangle (Seidenberg & McClelland, 1999), modèle multi-traces mémorielles (Ans, Carbonnel & Valdois, 1998)) ou de reconnaissance visuelle de mots (modèles Bayesian Reader (Norris, 2006), Seriol (Whitney, 2001), Overlap (Gomez, 2008), Spatial coding model (Davis, 2010), Letters in Time and Retinotopic Space (Adelman, 2011), etc.), la littérature est encore limitée en ce qui concerne l'apprentissage des nouveaux mots par les adultes et l'acquisition du lexique pendant l'apprentissage de la lecture.

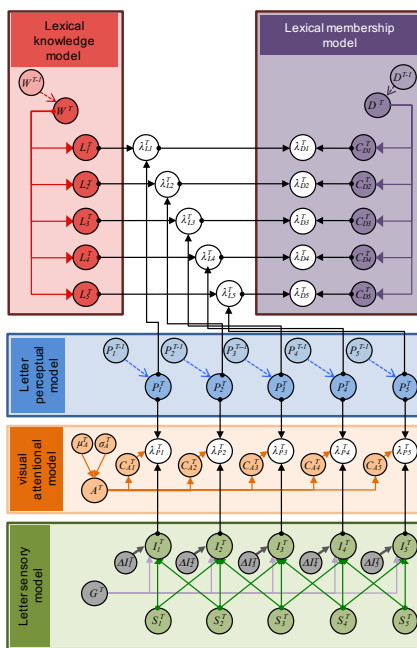
Dans ce contexte, la théorie de l'auto-apprentissage (Share, 1995) suppose que l'apprentissage de nouveaux mots reposerait essentiellement sur le décodage phonologique : c'est prononcer, ou inférer la forme prononcée d'un nouveau mot, qui permettrait une mémorisation de ce nouveau mot. Une conséquence de cette théorie est que l'apprentissage de nouveaux mots repose sur la voie auditive, et ne peut procéder uniquement dans la voie visuelle.

Nous avons développé un modèle d'identification visuelle de mots, le modèle BRAID (pour *Bayesian model of word Recognition with Attention, Interference and Dynamics*). C'est un modèle probabiliste structuré, dont les ingrédients originaux principaux sont un modèle d'interférence visuelle entre lettres voisines, un modèle visuo-attentionnel qui limite la capacité de traitement parallèle perceptif, et un modèle de la dynamique temporelle d'accumulation d'indices perceptifs sur l'identité des lettres et des mots reconnus. Le modèle BRAID rend compte d'effets classiques du domaine (par exemple, effet de supériorité de reconnaissance des lettres dans un mot, effet du voisinage lexical, effet de fréquence du mot), mais aussi d'effets jamais modélisés jusqu'alors, comme la position optimale du regard (OVP pour Optimal Viewing Position).

Nous avons récemment étendu le modèle BRAID pour l'apprentissage de nouveaux mots : un mécanisme de détection de nouveauté déclenche une stratégie d'exploration visuo-attentionnelle qui a pour but de maximiser la prise d'information visuelle sur le nouveau mot. Nous avons montré que ce mécanisme simulait de manière réaliste l'apprentissage d'un nouveau mot, sans composante

phonologique, contrairement à la théorie de l'auto-apprentissage. Nous avons simulé l'apprentissage de nouveaux mots par un « lecteur expert », c'est-à-dire disposant préalablement d'un lexique complet.

L'objectif de ce stage est d'explorer la capacité de ce mécanisme à décrire également l'acquisition du lexique complet pendant l'apprentissage de la lecture par un enfant. Nous étudierons les propriétés du mécanisme proposé lorsque le lexique initial est petit, et l'influence de facteurs comme la fréquence des mots ou la densité de leur voisinage sur la dynamique temporelle de l'apprentissage.



Travail proposé

Pour réaliser cet objectif, il s'agira de mettre en œuvre des expériences de simulation du modèle. Un premier jeu de simulations explorera les propriétés générales du mécanisme d'apprentissage de nouveaux mots lorsqu'il est employé pour l'acquisition du lexique : vitesse d'apprentissage, croissance du lexique, erreurs commises, etc. Un second jeu de simulations aura pour but de vérifier la capacité du modèle à reproduire des effets expérimentaux connus du domaine. Par exemple, un mot est plus rapide à apprendre s'il ressemble à beaucoup de mots déjà connus (effet facilitateur du voisinage pour l'apprentissage), mais son identification est moins rapide (effet inhibiteur du voisinage pour la reconnaissance).

D'un point de vue pratique, les points d'entrée du travail demandé consistent en une revue ciblée de la littérature sur les effets à simuler, la prise en main du modèle probabiliste (programme Mathematica), l'implantation informatique des tâches à simuler et des outils de visualisation nécessaires. Une étude des propriétés de l'espace des paramètres du modèle pourra être poursuivie.

Compétences demandées

Le candidat devra idéalement avoir une formation en programmation et simulation informatique, et une affinité pour la modélisation mathématique en Sciences Cognitives (pour les étudiants d'autres profils, n'hésitez pas à nous contacter). Des connaissances préalables en probabilités sont un plus, mais ne sont pas indispensables. Une préférence sera donnée aux candidats souhaitant poursuivre en doctorat.

Contacts

Julien Diard (LPNC) CR CNRS, julien.diard@univ-grenoble-alpes.fr, Tél. : 04 76 82 78 07
<http://diard.wordpress.com>

Sylviane Valdois (LPNC) DR CNRS, sylviane.valdois@univ-grenoble-alpes.fr

Emilie Ginestet (LPNC) doctorante, emilie.ginestet@univ-grenoble-alpes.fr

Financement

Indemnités de stage assurées par le projet FLUENCE

Références

- Adelman, JS. (2011). Letters in time and retinotopic space. *Psych. Rev.*, 118, 570-582.
- Ans, B., Carbonnel, S. & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psych. Rev.*, 105, 678-723.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psych. Rev.*, 108: 204-256.
- Davis, C.J. (2010). The spatial coding model of visual word identification. *Psych. Rev.*, 117, 713-758.
- Gomez, P., Ratcliff, R. & Perea, M. (2008). The overlap model: a model of letter position coding. *Psych. Rev.*, 115, 577-600.
- Norris, D. (2006). The Bayesian reader: Word recognition as an optimal Bayesian Decision Process. *Psych. Rev.*, 113, 327-357.
- Seidenberg MS, and McClelland JL. (1999). A distributed developmental model of word recognition and naming. *Psych. Rev.*, 96, 523-568.
- Share, DL. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*, 55:151-218.
- Whitney, C. (2001). How the brain encodes the order of letters: The Serial model. *Psych. Bulletin & Rev.*, 8, 221-243.