

e-FRAN > PLATEFORME

e-FRAN > DES TERRITOIRES ÉDUCATIFS
D'INNOVATION NUMÉRIQUE

Mission Monteil > POUR LE NUMÉRIQUE
DANS L'ÉDUCATION

ProFAN > DES COMPÉTENCES
POUR LES EMPLOIS DU FUTUR



Mieux comprendre les bases cognitives de la lecture pour en faciliter l'apprentissage

Ali SAGHIRAN

Mots-clés – Niveaux et Public concernés

Mots-clés : lecture à voix haute, attention visuelle, irrégularités de la langue

Niveaux : primaire

Public : enseignant-e-s d'élémentaire

À quelles questions cette étude tente-t-elle de répondre ?

Mon travail de recherche avait pour objectif de mieux comprendre quelles sont les connaissances et opérations cognitives impliquées dans la lecture à travers la modélisation computationnelle de cette activité. Un modèle computationnel est un modèle mathématique implémenté sur ordinateur qui permet de simuler le comportement de lecture. Autrement dit, on définit mathématiquement quels sont les composants du modèle et comment ils interagissent et on peut ensuite tester si le modèle parvient à lire des mots qu'il connaît ou qu'il n'a jamais rencontrés comme les lecteurs humains. Le fait que le modèle reproduise fidèlement le comportement humain permet de valider la conception théorique de la lecture qu'il propose. La plupart des modèles computationnels de lecture (Phénix *et al.*, 2016) s'accordent sur le fait que la lecture repose sur deux procédures : une procédure sérielle qui consiste à mettre en relation des unités orthographiques sous-lexicales avec les unités phonologiques correspondantes (par exemple « AU » correspond à /o/) et une procédure lexicale qui permet d'associer directement le mot entier à sa forme phonologique (par exemple, « BATEAU » associé à /bato/). Cependant, la question de savoir quelle est la taille des unités orthographiques sous-lexicales qui sont prises en compte reste entière. Traite-t-on toujours des graphèmes, ou plutôt des syllabes, ou la taille des unités peut-elle varier ? De la même façon, reconnaître et traiter des unités orthographiques à l'intérieur du mot écrit suppose un mécanisme de segmentation en sous-unités. La question relative à la nature des mécanismes cognitifs impliqués dans cette segmentation n'est actuellement pas résolue. Enfin, doit-on concevoir la mise en relation des unités sous-lexicales orthographiques et phonologiques comme définissant un système de traitement indépendant des connaissances lexicales, ou au contraire, comme s'effectuant par analogie, sur la base des connaissances lexicales mémorisées ? C'est à l'ensemble de ces trois questions que mon travail tente d'apporter des éléments de réponse.

Pourquoi ces questions sont-elles pertinentes ?

Question 1 : Quelle est la taille des unités orthographiques sous-lexicales traitées lors d'une lecture sérielle du mot ?

Tout le monde est d'accord sur le fait qu'un apprentissage systématique des relations graphèmes-phonèmes est indispensable pour l'apprentissage de la lecture. Un enfant lecteur débutant traite en général les mots qu'il rencontre pour la première fois, en associant chaque lettre au phonème correspondant, très systématiquement et de façon sérielle de gauche à droite (Valdois, 2020). Par exemple, le mot « LAC » sera lu « L-/l/, A-/a/, C - /k/ ». Un adulte bon lecteur traite aussi les mots qu'il voit pour la première fois de façon sérielle; mais quand un adulte lit le mot « ORNITHORYNQUE » pour la première fois, est-ce qu'il va utiliser un système de conversion graphème-phonème (O-/o/, R -/ʀ/, N-/n/, etc.) ou est-il capable de traiter le mot écrit en le segmentant en unités plus grandes (par exemple en syllabes)? De la même façon, quand on mesure la vitesse de lecture des enfants sur des mots inventés (comme « VERDULIN » ou « SCROPALE ») qui ne peuvent être lus que de façon sérielle, on observe que cette vitesse augmente au cours du primaire (Sprenger-Charolles *et al.*, 2005). Faut-il interpréter cette vitesse accrue comme démontrant une mise en relation plus rapide des graphèmes et des phonèmes? ou se pourrait-il que les enfants deviennent de plus en plus sensibles aux régularités statistiques de la langue (Chetail, 2017), c'est-à-dire aux combinaisons de lettres qui sont fréquentes, que celles-ci correspondent à des graphèmes en début d'apprentissage ou, plus tard, à des unités plus grandes comme les syllabes ou même les morphèmes? Cette question peut devenir primordiale si l'on souhaite améliorer la vitesse de lecture, prévenir les difficultés d'apprentissage ou remédier aux troubles dyslexiques.

Question 2 : Comment s'effectue la segmentation en sous-unités lors d'un traitement sériel du mot écrit ?

Les modèles de lecture les plus largement reconnus aujourd'hui sont les modèles double-voie (Coltheart *et al.*, 2001; Perry *et al.*, 2007). Ils font l'hypothèse que le traitement sériel repose sur une segmentation en graphèmes (par exemple, P-OU-SS-IN) mais même les modélisations computationnelles récentes de ces modèles n'explicitent pas quels sont les mécanismes cognitifs qui permettent cette segmentation graphémique (Perry *et al.*, 2013). Cette question est très importante parce que la segmentation en graphèmes est loin d'être aisée pour tous les enfants et elle peut s'avérer particulièrement difficile dans une langue comme le français. D'une part, notre langue possède des graphèmes dont la longueur est très variable allant d'une lettre (comme pour 'I' ou 'B') à des séquences longues (comme 'EAU', 'OIN', 'ILLE' ou encore 'AIENT'). D'autre part, une même séquence de lettres peut correspondre à un ou deux graphèmes. Par exemple, la séquence 'AN' correspond au graphème 'AN' dans « BANC » mais à deux graphèmes dans « CANE ». On peut aussi remarquer que dans certains mots la segmentation graphémique n'est pas évidente même pour des adultes (par exemple, comment segmenter « VIEILLESSE »?). Il n'est donc pas étonnant que les débutants lecteurs fassent des erreurs de segmentation ou que celles-ci persistent en contexte dyslexique (Zoubrinetzky *et al.*, 2014). Identifier les mécanismes qui permettent la segmentation du mot en sous-unités (graphèmes ou autre) pourrait conduire à apporter une aide plus efficace à ces enfants, voire à proposer des entraînements systématiques en classe visant à prévenir les difficultés de segmentation.

Question 3 : La lecture sérielle, repose-t-elle sur des connaissances explicites ou se fait-elle par analogie aux connaissances lexicales ?

Depuis de nombreuses décennies se pose la question de savoir si lire un mot que l'on voit pour la première fois implique de recourir à un système de conversion graphème-phonème spécifique ou si la lecture des mots nouveaux repose sur les connaissances lexicales. Cette question est étroitement dépendante de la taille des unités et des mécanismes de segmentation. Les modèles double-voie supposent que le mot nouveau est lu en mobilisant un système de conversion qui intervient au sein de la voie sous-lexicale, indépendamment des connaissances mémorisées sur les mots de la langue (Coltheart *et al.*, 2001; Perry *et al.*, 2007). Cette conception conduit à faire l'hypothèse d'un système de

conversion limitée aux seuls graphèmes et à supposer que les mécanismes qui permettent la segmentation graphémique sont spécifiques à la voie sous-lexicale. D'autres modèles, dits de traitement par analogie (Ans *et al.*, 1998; Glushko, 1979), font l'hypothèse que les mots nouveaux activent plusieurs mots existants mémorisés, et que ce sont les sous-unités des mots activés qui contribuent à construire la prononciation du mot nouveau. Cette conception par analogie a deux conséquences : (a) la taille des sous-unités dépend des mots qui sont activés et du nombre de lettres qu'ils partagent, elle n'est donc pas fixée a priori, et, (b) si le traitement des mots connus et des mots nouveaux repose sur l'activation des connaissances lexicales, alors le mécanisme de segmentation impliqué dans le traitement des mots nouveaux est nécessairement également impliqué dans le traitement des mots connus. Ces différences théoriques ont des conséquences directes sur l'apprentissage. Dans le premier cas, par exemple, améliorer la lecture des mots nouveaux nécessiterait d'entraîner les seules conversions graphème-phonème; dans le second, cela demanderait d'entraîner également les enfants sur de vrais mots renfermant des unités communes.

En résumé, les modèles théoriques de la lecture ont des conséquences directes sur nos conceptions de l'apprentissage de la lecture et des méthodes pédagogiques les plus adaptées pour favoriser cet apprentissage. Elles ont aussi des implications fortes sur la prise en charge et la remédiation des troubles dyslexiques. L'objectif de cette étude, était donc d'évaluer la plausibilité d'un nouveau modèle théorique qui s'inscrit dans le contexte des modèles de lecture par analogie et d'explicitier la nature et le fonctionnement des mécanismes impliqués dans la segmentation des unités sous-lexicales au sein de ce modèle.

Quelle méthodologie de recherche a-t-on utilisée ?

Nous avons développé un nouveau modèle computationnel de la lecture, appelé *BRAID-Phon* – extension du modèle *BRAID* (Phénix, 2018; Phénix *et al.*, soumis; Phénix *et al.*, 2018) –, qui implémente deux hypothèses principales : (1) la lecture se fait toujours par activation des connaissances lexicales, qu'il s'agisse de lire des mots existants ou des mots nouveaux; (2) le fait de traiter le mot dans sa totalité ou seulement des sous-parties de mots dépend de la distribution de l'attention visuelle lors du traitement. Pour une présentation détaillée du modèle, se référer à Saghiran (2021) et Saghiran *et al.* (2020); une illustration schématique du modèle est présentée dans la Figure 1.

1 Architecture du modèle

Le modèle *BRAID-Phon* (Saghiran, 2021) possède une architecture à 5 niveaux qui peut être décrite ainsi :

Le niveau sensoriel À ce premier niveau, le traitement vise à reconnaître les lettres qui composent le mot à lire. La reconnaissance de chaque lettre dépend de sa ressemblance avec les autres lettres de l'alphabet. Ainsi, 'M' est plus difficile à reconnaître du fait de sa proximité visuelle avec 'N'. La reconnaissance des lettres est aussi plus ou moins facile selon leur position par rapport au point de fixation. S'il fixe la lettre 'O', le modèle reconnaîtra mieux le 'P' initial du mot « POT », que celui de « PIROGUE » qui est plus éloigné de la lettre fixée. Ceci simule une baisse linéaire de l'acuité visuelle plus on s'éloigne du point de fixation. Enfin, les lettres qui sont à l'intérieur du mot sont un peu moins bien reconnues que la première ou la dernière lettre, ce qui correspond au phénomène de *crowding* bien connu au niveau comportemental. On a donc au niveau sensoriel, trois phénomènes (la ressemblance entre lettres, l'acuité et le *crowding*) qui dégradent la lisibilité des lettres.

Le niveau de traitement visuo-attentionnel Ce niveau explicite comment l'attention visuelle se déploie sur le mot pour moduler la quantité d'information qui transite du niveau sensoriel au niveau perceptif. Le modèle représente l'attention visuelle comme ayant la forme d'une distribution gaussienne :

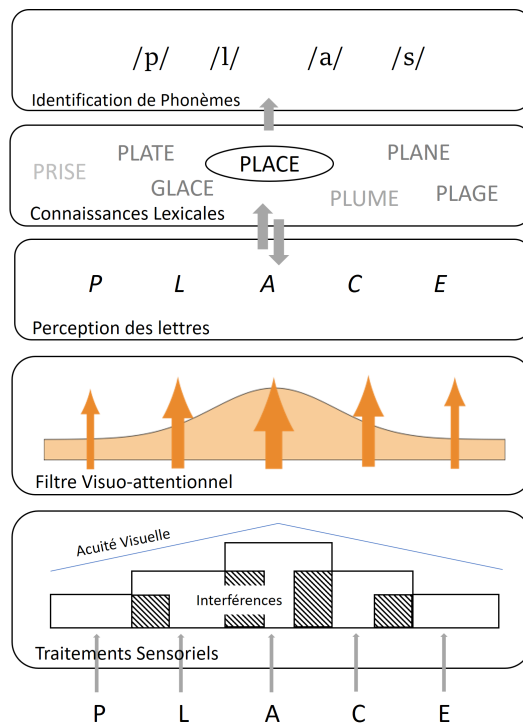


FIGURE 1 – Illustration schématique du modèle *BRAID-Phon*.

les lettres qui sont sous focus attentionnel reçoivent un maximum d’attention, si bien que davantage d’information sur leur identité (représentée par flèche centrale la plus épaisse dans la Figure 1) transite vers le niveau perceptif; la quantité d’attention allouée à chaque lettre décroît ensuite de façon non linéaire plus on s’éloigne du focus attentionnel. La quantité d’attention visuelle allouée à chaque lettre du mot contribue à accroître sa lisibilité. La distribution de l’attention visuelle et la position du focus attentionnel sur le mot ne sont pas fixées a priori. Ce sont des processus dynamiques qui s’adaptent au cours du traitement de façon à optimiser l’identification des lettres.

Le niveau perceptif des lettres Ce niveau accumule l’information sur l’identité des lettres qui est transmise via le niveau sensoriel. L’accumulation d’information est modulée par la distribution de l’attention. L’information perceptive évolue de façon dynamique au cours du temps. À chaque pas de temps, de plus en plus d’information est accumulée permettant de se construire une image mentale transitoire de plus en plus précise de la séquence de lettres du mot.

Le niveau lexical Ce niveau correspond aux connaissances quant à la forme orthographique des mots de la langue. Le modèle possède une base lexicale de plusieurs dizaines de milliers de mots qui correspond à celle d’un lecteur adulte expert. Chaque mot est caractérisé par sa fréquence dans la langue. L’information accumulée sur l’identité des lettres au niveau perceptif permet d’activer les mots du lexique qui partagent des lettres communes avec la séquence en cours de traitement.

Le niveau phonologique À ce niveau, sont générés les phonèmes correspondant à la séquence orthographique du mot. En fait, l’activation des informations lexicales conduit à activer les phonèmes qui correspondent aux mots activés. L’activation des phonèmes est fonction du degré d’activation des mots, donc de leur proximité visuelle avec le mot présenté et de leur fréquence.

2 Fonctionnement du modèle

L'information transite du niveau sensoriel vers le niveau lexical orthographique et le niveau lexical phonologique. Comme nous l'avons dit précédemment, l'information sur l'identité des lettres qui s'accumule au niveau perceptif dépend de la qualité de l'information sensorielle modulée par l'attention visuelle. Le niveau perceptif est aussi influencé par les connaissances lexicales, si bien qu'une lettre sera d'autant mieux identifiée au niveau perceptif qu'elle fait partie d'un mot existant ou d'une séquence de lettres fréquente dans les mots de la langue.

Le modèle fait l'hypothèse d'une voie de traitement unique. Que la séquence de lettres à lire corresponde à un mot connu ou à un mot nouveau, le traitement sera toujours le même, impliquant les niveaux décrits ci-dessus. Comment dès lors, sera-t-il possible de traiter un mot connu très rapidement et un mot non connu de façon sérielle ?

Nous supposons que cela est possible grâce à l'attention visuelle. L'attention visuelle peut se déployer sur la séquence à lire et se déplacer de façon dynamique. À la première fixation sur le mot, la distribution de l'attention visuelle sur la séquence à lire est relativement large. Deux situations sont alors possibles : (a) soit cela permet d'accumuler de l'information assez rapidement sur l'ensemble des lettres de la séquence ; (b) soit l'accumulation d'information est lente et l'identité des lettres incertaine. Dans le premier cas, le modèle tendra à terminer le traitement en une seule distribution attentionnelle ; dans le second, il tendra à modifier la répartition de l'attention visuelle et la position du focus attentionnel, afin d'accumuler plus rapidement une information plus fiable sur l'identité des lettres du mot, quitte à ne traiter qu'une partie des lettres de la séquence à chaque capture attentionnelle. Le déplacement de l'attention visuelle lors de l'apprentissage d'un mot nouveau par le modèle *BRAID-Learn* (extension du modèle *BRAID*) a été étudié indépendamment (Ginestet, 2019 ; Ginestet *et al.*, 2022).

Nous avons donc fait deux hypothèses : la première est que l'attention visuelle est le mécanisme qui permet de passer d'un traitement du mot entier à un traitement par sous-unités ; la seconde stipule que la répartition de l'attention visuelle devrait déterminer la taille des unités sous-lexicales traitées et que cette répartition est modulée en fonction des influences lexicales qui participent à l'identification des lettres au niveau perceptif.

Quels résultats a-t-on obtenus ?

Pour vérifier ces hypothèses, nous avons effectué des simulations. Cela signifie que nous avons présenté au modèle des séquences de lettres correspondant à des mots connus ou à des mots nouveaux et nous avons étudié son comportement.

Nous avons tout d'abord vérifié que le modèle est capable de reproduire des effets comportementaux classiques, comme l'effet de fréquence ou de longueur. Le premier correspond au fait que les humains sont capables de lire un mot d'autant plus vite qu'il est plus fréquent. Le second au fait que la vitesse de lecture varie selon la longueur du mot. Nous avons montré que *BRAID-Phon* est parfaitement capable de simuler fidèlement les effets observés chez l'humain (Saghiran *et al.*, 2020). Ces simulations confirment que le modèle est capable de lire les mots connus à partir du traitement parallèle des lettres qui les composent.

Nous avons ensuite évalué la capacité du modèle à lire des mots qu'il ne connaît pas en utilisant la même procédure de traitement parallèle en une capture attentionnelle que pour les mots. Nous montrons que le traitement échoue. Le traitement des lettres du pseudo-mot présenté conduit à activer tous les mots lexicaux orthographiques qui partagent des lettres communes avec lui. Si l'on prend l'exemple du pseudo-mot « VIRDIN », c'est alors le mot qui est à la fois le plus proche visuellement de « VIRDIN » et le plus fréquent dans la base lexicale qui va être le plus activé. En l'occurrence, le mot « JARDIN » est le plus activé et la forme phonologique /ʒɑʁdɛ̃/ est produite par le modèle au lieu de /viʁdɛ̃/. Ce comportement du modèle est en accord avec l'observation que certains dyslexiques lisent

des pseudo-mots comme de vrais mots lorsque la procédure analytique sérielle est déficitaire.

Nous avons ensuite évalué la réponse du modèle lorsqu'on lui demande de traiter des mots inventés par segments et qu'on fait varier la taille des segments sous-lexicaux. Nous prendrons ici l'exemple du pseudo-mot « CHORAT ». Lorsque le segment se réduit à la première lettre 'C', le modèle active trois phonèmes, /k/ plus fortement que /s/ et /s/ plus fortement que /ʃ/. Son comportement change radicalement lorsque la capture attentionnelle porte sur la séquence 'CH', c'est alors uniquement le phonème /ʃ/ qui est activé. La séquence 'CHO' est associée sans ambiguïté à /ʃo/ du fait de l'activation forte de mots relativement fréquents (comme « CHOC », « CHOCOLAT », « CHOQUANT »). Le comportement du modèle change lorsque la séquence 'CHOR' est prise en compte. Dans ce cas, les phonèmes /ʃ/ et /k/ sont à peu près également activés du fait de l'appariement orthographique fort avec le mot « CHORALE » qui associé à « CHORUS », « CHORISTE » et d'autres renforce la prononciation /k/. Enfin, la seule prononciation générée est /k/ pour les séquences 'CHORA' et 'CHORAT' dans la mesure où les mots orthographiquement les plus proches (« CHORALE » et « CHORAL ») correspondent tous à la prononciation /k/. Cela témoigne du fait que la prononciation d'un graphème dans le modèle dépend largement du nombre de lettres qui sont simultanément traitées dans la séquence orthographique du pseudo-mot.

Le troisième exemple que nous détaillerons ici est l'exemple du pseudo-mot « VERDULIN ». Nous avons évalué le comportement du modèle selon que le traitement porte successivement sur chacune des trois syllabes ou sur des segments plus grands qui se chevauchent (comme VERD-RDUL-ULIN). Le modèle génère la prononciation attendue du pseudo-mot dans le cas des segments qui se chevauchent. La prise en compte de segments plus grands permet d'activer des mots lexicaux visuellement plus proches et dont la phonologie est davantage pertinente pour le pseudo-mot présenté.

Globalement les résultats des simulations montrent très clairement que :

1. La lecture d'un pseudo-mot requiert un traitement sériel, par analyse successive de groupes de lettres ou segments, plus petits que le mot.
2. C'est l'attention visuelle qui permet de se focaliser sur des sous-parties de mot pour en traiter les segments et c'est le déplacement de l'attention visuelle qui permet le traitement successif gauche-droite des segments orthographiques.
3. La prononciation privilégiée par le modèle dépend fortement de la taille des unités orthographiques traitées.
4. Dans certains cas, le fait de prendre en compte des segments orthographiques plus grands est nécessaire à une bonne prononciation et la taille des segments peut alors ne correspondre à aucune unité orthographique prédéfinie (ni graphème, ni syllabe).
5. Enfin, les simulations ont été effectuées à partir de la seule activation des informations lexicales, donc sans recours à un système explicite de conversion graphème-phonème. La lecture des pseudo-mots dans *BRAID-Phon* met en jeu exactement les mêmes connaissances et opérations mentales que la lecture des mots.

Que dois-je retenir de cette étude pour ma pratique ?

- La lecture des mots non connus repose sur le traitement sériel de segments orthographiques. Compte tenu du fait que l'enfant est confronté en tout début d'apprentissage à des mots écrits qu'il ne connaît qu'à l'oral, ces mots sont pour lui nouveaux à l'écrit. C'est donc bien la procédure sérielle de traitement qui va s'appliquer, ce qui justifie pleinement de privilégier un apprentissage explicite des relations graphème-phonème en CP.

- Si l'on accepte l'idée que les relations graphèmes-phonèmes ne sont pas mémorisées indépendamment des connaissances lexicales, alors l'introduction explicite de ces relations en classe doit nécessairement s'accompagner de l'exposition à des mots qui renferment la relation étudiée. Ceci est proposé dans la plupart des ouvrages pour l'apprentissage de la lecture en CP, mais pourrait disqualifier des méthodes qui privilégient la lecture de séquences ne correspondant pas à des mots.
- Les simulations suggèrent que l'attention visuelle joue un rôle important dans la segmentation de la séquence à lire. La taille des segments traités va donc en partie dépendre des ressources attentionnelles que l'enfant peut déployer sur le mot à lire. Moins de ressources sont nécessaires pour traiter des segments courts que des segments longs. Lors de l'apprentissage, cela pourrait justifier d'introduire d'abord les unités courtes correspondant à une lettre puis augmenter très progressivement la taille de ces unités.
- Le modèle nous montre que la plupart des graphèmes doivent être traités en contexte. Même la prononciation de 'A' ou 'O' est ambiguë tant qu'on ne sait pas quelle est la lettre qui suit. Mais tous les graphèmes d'une même longueur n'ont pas le même statut : la prononciation /u/ va fortement dominer pour 'OU' même si l'on ne connaît pas la lettre qui suit alors que cela ne sera pas du tout le cas pour 'ON'. Cela devrait conduire, à longueur de segment égale, à privilégier un apprentissage plus précoce des relations les moins ambiguës.
- Notre étude suggère que la taille des segments pris en compte lors du traitement dépend de la quantité de ressources attentionnelles disponibles et que le traitement de segments plus longs garantit en général une lecture avec moins d'erreurs et plus fluide. On devrait donc en conclure qu'entraîner l'attention visuelle en début d'apprentissage ou chez les pré-lecteurs devrait favoriser l'apprentissage (voir aussi Meyer, 2019; Meyer *et al.*, 2018).

Références

- Ans, B., Carbonnel, S. et Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, 105(4), 678-723. <https://doi.org/10.1037/0033-295x.105.4.678-723>
- Chetail, F. (2017). What do we do with what we learn? Statistical learning of orthographic regularities impacts written word processing. *Cognition*, 163, 103-120. <https://doi.org/10.1016/j.cognition.2017.02.015>
- Coltheart, M., Rastle, K., Perry, C., Langdon, R. et Ziegler, J. (2001). DRC : A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, 108(1), 204-256. <https://doi.org/10.1037/0033-295X.108.1.204>
- Ginestet, E. (2019). *Modélisation bayésienne et étude expérimentale du rôle de l'attention visuelle dans l'acquisition des connaissances lexicales orthographiques* [thèse de doctorat, Université Grenoble Alpes, Grenoble, France]. <https://tel.archives-ouvertes.fr/tel-02893469/document>
- Ginestet, E., Valdois, S. et Diard, J. (2022). Probabilistic modeling of orthographic learning based on visuo-attentional dynamics. *Psychonomic Bulletin and Review*, 1-24. <https://doi.org/10.3758/s13423-021-02042-4>
- Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology : Human Perception and Performance*, 5(4), 674-691. <https://doi.org/10.1037/0096-1523.5.4.674>
- Meyer, S. (2019). Conception et évaluation d'Evasion, un logiciel éducatif d'entraînement des capacités d'attention visuelle impliquées en lecture. Université Grenoble Alpes. <https://tel.archives-ouvertes.fr/tel-02402422>
- Meyer, S., Diard, J. et Valdois, S. (2018). Lecteurs, votre attention s'il vous plaît! Le rôle de l'attention visuelle en lecture. *ANAE - Approche Neuropsychologique des Apprentissages Chez L'enfant*. <https://hal.archives-ouvertes.fr/hal-02002545>

- Perry, C., Ziegler, J. C. et Zorzi, M. (2007). Nested incremental modeling in the development of computational theories : The CDP+ model of reading aloud. *Psychological Review*, 114(2), 273-315. <https://doi.org/10.1037/0033-295x.114.2.273>
- Perry, C., Ziegler, J. C. et Zorzi, M. (2013). A computational and empirical investigation of graphemes in reading. *Cognitive Science*, 37(5), 800-828. <https://doi.org/10.1111/cogs.12030>
- Phénix, T. (2018). *Modélisation bayésienne algorithmique de la reconnaissance visuelle de mots et de l'attention visuelle* [thèse de doctorat, Université Grenoble Alpes, Grenoble, France]. <https://tel.archives-ouvertes.fr/tel-02368168/document>
- Phénix, T., Diard, J. et Valdois, S. (2016). Les modèles computationnels de lecture. *Traité de neurolinguistique* (p. 167-182). De Boeck Supérieur.
- Phénix, T., Ginestet, E., Valdois, S. et Diard, J. (soumis). The role of attention in visual word recognition : A Bayesian programming approach.
- Phénix, T., Valdois, S. et Diard, J. (2018). Reconciling opposite neighborhood frequency effects in lexical decision : Evidence from a novel probabilistic model of visual word recognition. Dans T. Rogers, M. Rau, X. Zhu et C. W. Kalish (dir.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (p. 2238-2243). Cognitive Science Society. <https://hal.archives-ouvertes.fr/hal-01850020/document>
- Saghiran, A. (2021). *Modélisation bayésienne de la lecture* [thèse de doctorat, Université Grenoble Alpes, Grenoble, France]. <https://tel.archives-ouvertes.fr/tel-03364950>
- Saghiran, A., Valdois, S. et Diard, J. (2020). Simulating length et frequency effects across multiple tasks with the Bayesian model BRAID-Phon. Dans *Proceedings of the 42nd Annual Virtual Meeting of the Cognitive Science Society* (p. 3158-3163). <https://hal.archives-ouvertes.fr/hal-01850020/document>
- Sprenger-Charolles, L., Colé, P., Béchennec, D. et Kipffer-Piquard, A. (2005). French normative data on reading and related skills from EVALEC, a new computerized battery of tests. *Revue Européenne de Psychologie appliquée*, 55, 157-186. <https://doi.org/10.1016/j.cognition.2017.02.015>
- Valdois, S. (2020). L'apprentissage de la lecture. *Neurosciences Cognitives Développementales* (p. 129-151). De Boeck Supérieur. <https://hal.archives-ouvertes.fr/hal-02997281>
- Zoubrinetzky, R., Bielle, F. et Valdois, S. (2014). New Insights on Developmental Dyslexia Subtypes : Heterogeneity of Mixed Reading Profiles (E. Kroesbergen, dir.). *PLoS ONE*, 9(6), e99337. <https://doi.org/10.1371/journal.pone.0099337>