Modélisation bayésienne et étude expérimentale du rôle de l'attention visuelle dans l'acquisition des connaissances lexicales orthographiques

Emilie Carail

Communauté UNIVERSITÉ Grenoble Alpes

THÈSE

Pour obtenir le grade de

DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : PCN - Sciences cognitives, psychologie et neurocognition

Arrêté ministériel : 25 mai 2016

Présentée par

Emilie CARAIL

Thèse dirigée par **Sylviane VALDOIS**, chercheur, Communauté Université Grenoble Alpes
et codirigée par **Julien DIARD**, CNRS

préparée au sein du **Laboratoire Laboratoire de Psychologie et Neuro Cognition**
dans **l'École Doctorale Ingénierie pour la santé la Cognition et l'Environnement**

Modélisation bayésienne et étude expérimentale du rôle de l'attention visuelle dans l'acquisition des connaissances lexicales orthographiques

Bayesian modeling and experimental study of the role of visual attention in the acquisition of orthographic lexical knowledge

Thèse soutenue publiquement le **11 décembre 2019**,
devant le jury composé de :

Madame Sylviane VALDOIS
DR1, Communauté Université Grenoble Alpes, Directeur de thèse
Madame Fabienne CHETAIL
Professeur associé, Université Libre de Bruxelles, Rapporteur
Monsieur Eric CASTET
Directeur de Recherche, Aix-Marseille Université, Rapporteur
Madame Maryse BIANCO
Professeur Emérite, Communauté Université Grenoble Alpes, Examinateur
Monsieur Jean-Luc SCHWARTZ
Directeur de Recherche, Communauté Université Grenoble Alpes, Examinateur
Monsieur Ludovic FERRAND
Directeur de Recherche, Université Clermont Auvergne, Examinateur

Table des Matières

Liste des Illustrations

Liste des Tableaux

Chapitre 1

Introduction

Alors que la communication écrite connaît un véritable essor depuis quelques années, aussi bien dans la vie personnelle via les réseaux sociaux, les forums, les SMS, que dans la vie professionnelle via l'intensification des échanges par courriel (Lacroux, 2015), les performances en orthographe décroissent (Andreu & Steinmetz, 2016). Ainsi, écrire « sans faute » paraît être réservé à l'élite de la société, et pourrait générer un sentiment de honte chez ceux qui ne maîtrisent pas l'ensemble des contraintes et règles orthographiques du français. Dans le cadre de l'insertion professionnelle, certains recruteurs utilisent même les compétences orthographiques des candidats comme filtre sélectif (Lacroux, 2015) faisant de la qualité de l'orthographe un indicateur de confiance, de sérieux et de professionnalisme.

En réalité, il n'existe pas de preuve qu'une mauvaise maîtrise de l'orthographe corrèle avec, ou cause, un quelconque déficit cognitif, de raisonnement, ou de sérieux professionnel. Et pourtant, l'acquisition de l'orthographe, qui parait évidente pour certains, peut réellement être déficitaire pour d'autres. Leur stigmatisation sociale paraît alors injustifiée et évitable. De fait, la compréhension des mécanismes cognitifs impliqués dans l'apprentissage orthographique constitue une question de recherche majeure et d'actualité.

Dans ce travail de thèse, nous nous limitons à l'apprentissage de l'orthographe lexicale – c'est-à-dire l'apprentissage de l'orthographe canonique des mots de la langue –, sans évoquer le développement des compétences en orthographe grammaticale – soit les compétences qui régissent les accords de genre et nombre entre mots et les conjugaisons des verbes. En effet, l'étude de l'orthographe lexicale représente en soi un challenge, notamment dans les langues opaques comme le français où la plupart des mots ne s'écrivent pas comme ils se prononcent, c'est-à-dire, par application des règles de conversion graphème-phonème les plus fréquentes. Ainsi, dans la suite, le terme « orthographe » fait référence à l'orthographe lexicale.

Quels sont les mécanismes cognitifs impliqués dans l'apprentissage orthographique ?

Si nous devions définir simplement l'apprentissage orthographique d'un nouveau mot (par exemple *siampoie*), nous pourrions dire qu'il consiste à décoder la séquence de lettres qui le compose et à la mémoriser. Le décodage, dans les modèles conceptuels de l'apprentissage ou dans les modèles implémentés de transcodage phonologique, est constitué de trois étapes : l'analyse visuelle du mot, le découpage du mot en sous-unités orthographiques et l'encodage phonologique. Ainsi, l'apprentissage orthographique met en jeu de nombreux processus cognitifs : des mécanismes visuels pour identifier les lettres qui composent le mot et encoder leur position relative, des mécanismes moteurs pour réaliser les mouvements oculaires lors de l'exploration du mot écrit, des mécanismes phonologiques pour générer la forme phonologique de la

forme écrite et activer la représentation phonologique correspondante en mémoire, mais aussi, des mécanismes mémoriels pour la création et le renforcement des représentations lexicales correspondantes.

Il est admis que l'apprentissage orthographique se produit le plus souvent implicitement au cours de la lecture (Castles & Nation, 2006). Autrement dit, le lecteur, n'ayant pas conscience d'apprendre un nouveau mot, n'adopte pas de stratégie particulière ; en tout cas, pas volontairement.

Malgré le rôle essentiel de l'apprentissage orthographique dans la lecture – pour la reconnaissance visuelle des mots écrits – et l'écriture – pour l'orthographe spécifique aux mots –, les recherches étudiant les mécanismes liés à l'apprentissage orthographique restent moins nombreuses au regard de celles portant sur la lecture.

A l'heure actuelle, l'hypothèse d'auto-apprentissage formulée par Share (1995, 1999, 2004) demeure le modèle le plus reconnu. Cette théorie fait l'hypothèse que chaque décodage phonologique réussi permet l'apprentissage orthographique du nouveau mot. De fait, le décodage phonologique est vu comme le principal mécanisme impliqué dans l'apprentissage orthographique, suggérant que l'analyse visuelle des mots et les mécanismes associés ne joueraient qu'un rôle mineur dans la mémorisation de la représentation orthographique des mots.

Cependant, quelques études récentes remettent en question cette hypothèse en observant qu'un décodage réussi ne suffit pas à mémoriser l'orthographe des nouveaux mots et que réciproquement, l'orthographe peut être mémorisée malgré un décodage erroné (Castles & Nation, 2006, 2008; Nation, Angell, & Castles, 2007; Tucker, Castles, Laroche, & Deacon, 2016). Ces résultats suggèrent que d'autres mécanismes cognitifs sont impliqués dans l'apprentissage orthographique, ce qui a conduit les chercheurs à adopter de nouvelles méthodes expérimentales permettant de mieux cerner les mécanismes en jeu au cours de l'apprentissage.

C'est dans cette optique que Nation and Castles (2017) proposent d'étudier l'apprentissage orthographique en temps réel grâce à l'analyse des mouvements oculaires pendant la lecture de nouveaux mots. Même s'il n'existe actuellement que peu de données comportementales (notamment en français) de ce type, les quelques études ayant utilisé le suivi oculométrique ont montré que la lecture répétée d'un nouveau mot affecte le comportement oculomoteur (H. Joseph & Nation, 2018; H. Joseph, Wonnacott, Forbes, & Nation, 2014). Ces données suggèrent qu'un traitement visuel spécifique s'opère lors de la première rencontre avec un nouveau mot et que les variations observées sur les mesures oculaires au cours des expositions successives résultent du processus de mémorisation orthographique.

Comment modéliser l'apprentissage de l'orthographe lexicale ?

Les modèles conceptuels de l'apprentissage orthographique posent un cadre théorique en identifiant les mécanismes cognitifs impliqués. Ces modèles, construits à partir des résultats d'expériences comportementales, postulent l'existence de relations entre les mécanismes impliqués sans pour autant spécifier la manière dont ces mécanismes interagissent. Par exemple, le modèle théorique d'auto-apprentissage de Share (1995, 1999, 2004) repose sur l'hypothèse que l'apprentissage orthographique dépend principalement du traitement phonologique ; mais, comme tout modèle conceptuel, le modèle de Share ne décrit pas le fonctionnement exact de cette relation de dépendance.

A l'inverse, les méthodes computationnelles – modèles mathématiques implémentés sur ordinateur – ont pour but de modéliser les fonctions cognitives en réalisant une séquence de calculs. Ainsi, ces modèles, dont l'organisation structurelle découle directement des hypothèses théoriques formulées, nécessitent une grande clarté dans la spécification des hypothèses et par conséquent un degré de précision supérieur aux modèles conceptuels. La validité théorique du

modèle (c'est-à-dire, des mécanismes supposés impliqués dans le processus étudié) dépend de sa capacité à simuler – reproduire fidèlement – les effets comportementaux classiquement observés dans la littérature. Parallèlement, l'implémentation de relations mathématiques claires entre les mécanismes permet aux modèles computationnels de simuler, dans notre cas, l'apprentissage en condition normale, mais aussi, en condition détériorée, afin d'identifier les compétences déficitaires responsables des troubles de l'apprentissage. Finalement, ils permettent de prédire des effets comportementaux non encore observés. Par extension, ils peuvent permettre d'élaborer des outils améliorant l'apprentissage normal ou permettant la rééducation des troubles d'apprentissage.

Les très nombreuses recherches comportementales sur l'activité de lecture chez le lecteur expert ou débutant ont conduit à proposer des modèles de lecture, des modèles de reconnaissance de mots ou de contrôle des mouvements oculaires en lecture de texte (Phénix, Diard, & Valdois, 2016), mais peu de travaux ont eu pour but de modéliser l'apprentissage orthographique. Pourtant, modéliser l'acquisition de l'orthographe lexicale semble pertinent au regard de l'influence des connaissances lexicales orthographiques sur la lecture. En effet, comme le soulignent Chaves, Totereau, and Bosse (2012), de nombreuses études ont fait le lien entre niveau de lecture et compétences orthographiques, et suggèrent qu'un enfant avec de bonnes connaissances lexicales orthographiques lira plus rapidement et fera moins de fautes d'orthographe en production écrite de mots qu'un enfant dont la mémoire orthographique serait moins développée. Castles, Rastle, and Nation (2018) concluent notamment que la lecture ne se résume pas au décodage phonologique et que l'apprentissage orthographique est la clé du passage de la lecture sérielle laborieuse de l'apprenti lecteur à la lecture fluente qui caractérise le lecteur expert.

A l'heure actuelle, les modèles de lecture *Connectionist Dual Process* (CDP, Perry et al., 2007; Perry, Ziegler, & Zorzi, 2010, 2013) et *Dual Route Cascade* (DRC, Coltheart et al., 2001) ont été étendus pour modéliser l'apprentissage orthographique (respectivement, Perry, Zorzi, & Ziegler, 2019; Ziegler et al., 2014 et Pritchard et al., 2018). Parce qu'ils sont basés sur la théorie de l'auto-apprentissage formulée par Share (1995, 1999, 2004), ces modèles implémentent un rôle fondamental des traitements phonologiques dans l'acquisition des connaissances orthographiques lexicales. En parallèle, ils font l'hypothèse que l'information orthographique est parfaitement acquise (c'est-à-dire, sans erreur) dès la première exposition au nouveau mot. Cela induit un rôle négligeable de l'analyse visuelle des lettres au cours de l'apprentissage orthographique (Ziegler et al., 2014). Ces modèles ne permettent notamment pas de rendre compte des spécificités du traitement visuel opéré lors de l'acquisition de l'orthographe lexicale de nouveaux mots, tel que révélé par les études en oculomotricité.

Notre contribution

L'objectif principal de cette thèse est de proposer un modèle d'acquisition des connaissances lexicales orthographiques. Ce modèle, nommé *BRAID-Learn*, est une extension du modèle probabiliste de reconnaissance visuelle de mots, le modèle BRAID (pour *Bayesian word Recognition using Attention, Interference and Dynamics*, Phénix, 2018; Phénix, Valdois, & Diard, 2018). Contrairement au cadre théorique de l'auto-apprentissage et aux modèles computationnels existants, notre modèle d'apprentissage met l'emphase sur les mécanismes visuels opérés pendant la lecture d'un nouveau mot. Il intègre un mécanisme de contrôle de mouvements oculaires, permettant ainsi de simuler les déplacements oculaires et attentionnels pendant la lecture de stimuli, nouveaux ou familiers. Dans notre modèle, nous faisons l'hypothèse que les mouvements oculaires et attentionnels sont guidés par un unique principe : optimiser la vitesse d'accumulation d'information perceptive. A travers une série de simulations, nous montrons que cette hypothèse permet de rendre compte notamment des données oculomotrices observées lors de

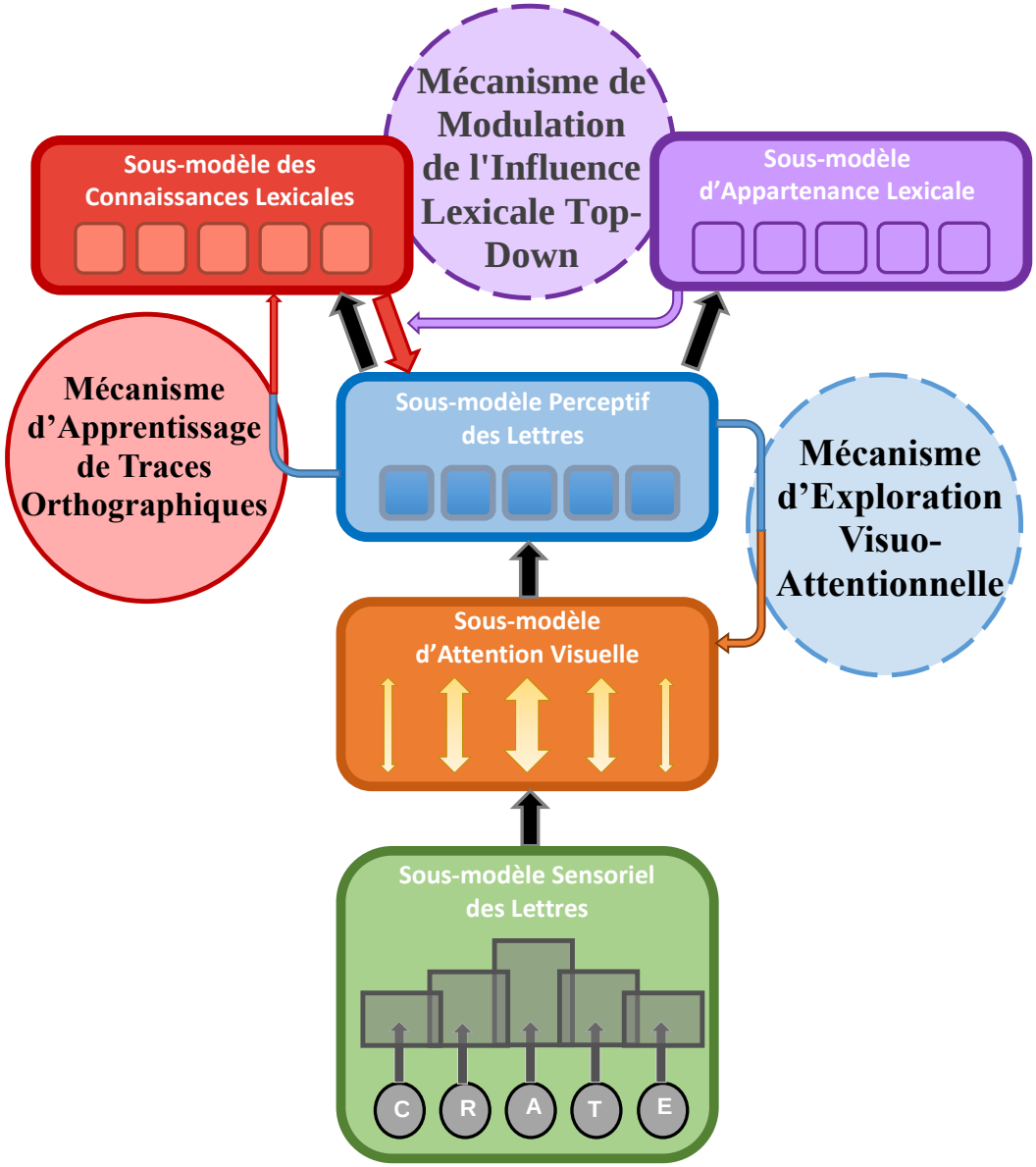

FIGURE 1.1 – Schéma du modèle conceptuel *BRAID-Learn*. Les cinq sous-modèles qui composent l'architecture du modèle BRAID sont représentés par des rectangles colorés, connectés par des flèches qui représentent les flux d'échange d'information. Le modèle *BRAID-Learn* repose sur trois mécanismes ajoutés à cette architecture, représentés par des ovales colorés.

l'apprentissage incident de nouveaux mots. La Figure 1.1 schématise la structure générale du modèle *BRAID-Learn* ainsi que les interactions entre les différents composants du modèle.

Le développement d'un modèle computationnel tel que nous le proposons avec le modèle *BRAID-Learn* nécessite l'utilisation de données comportementales pour, d'une part, calibrer le modèle et d'autre part, évaluer sa pertinence en comparant les résultats obtenus lors des simulations aux données comportementales observées.

Dans une première étude, nous simulons l'effet de longueur – l'augmentation des temps de réponse avec la longueur des mots – observé pour les mots en décision lexicale, et décrits dans une méga-étude classique (*French Lexicon Project, FLP ;* Ferrand et al., 2010). La comparaison

des données simulées avec les données comportementales nous permet d'une part, de calibrer une partie des paramètres essentiels au modèle *BRAID-Learn*, et d'autre part, de simuler l'effet de longueur en faisant varier les capacités visuo-attentionnelles (paramètres liés au sous-modèle d'attention visuelle, bloc orange dans la Figure 1.1). Les résultats obtenus montrent un effet de longueur accentué lorsque le modèle traite toutes les lettres en parallèle (c'est-à-dire, avec une seule fixation centrale) mais que le modèle reproduit fidèlement l'effet de longueur lorsqu'il réalise deux fixations sur les mots longs (c'est-à-dire, à partir de 7 lettres). Ces résultats suggèrent que les déplacements oculaires et attentionnels modulent la vitesse d'accumulation d'information perceptive et lexicale sur le mot, et résultent en un comportement efficace, résolvant le compromis entre précision et vitesse du traitement.

Dans *BRAID-Learn*, l'apprentissage orthographique découle du traitement visuel (c'est-à-dire, du comportement oculomoteur) opéré sur le nouveau mot au cours de la lecture. L'absence de données dans la littérature nous a amené à réaliser une seconde étude dans laquelle sont analysés les déplacements oculomoteurs associés à la lecture répétée de nouveaux mots en condition d'apprentissage implicite. Les données comportementales recueillies nous permettent d'une part, d'observer le lien entre les variations du comportement oculomoteur et l'apprentissage orthographique, et d'autre part, d'observer l'effet de chaque nouvelle lecture sur les mesures oculaires. Les résultats de cette étude suggèrent un lien entre apprentissage orthographique incident et capacités d'attention visuelle.

Munis de ces données, nous développons le modèle *BRAID-Learn* dont le fonctionnement résulte d'une hypothèse principale : l'apprentissage orthographique repose sur l'optimisation de l'accumulation d'information perceptive quant à l'identité des lettres qui composent le mot. Mathématiquement défini par une maximisation du gain d'entropie, ce principe d'optimisation permet de déterminer et contrôler les déplacements attentionnels (et donc oculaires) qui s'effectuent sur la séquence du mot lors de l'apprentissage.

Finalement, une série de simulations nous permet de tester notre hypothèse théorique. En évaluant la capacité du modèle *BRAID-Learn* à reproduire l'évolution des patterns oculomoteurs observée dans l'étude expérimentale précédente, occurrence par occurrence, nous montrons que ce mécanisme rend possible l'apprentissage orthographique. Les résultats simulés reproduisent un temps de traitement plus élevé pour les pseudo-mots que pour les mots ainsi qu'une diminution du temps de traitement à chaque nouvelle présentation, caractéristique du renforcement des traces orthographiques mémorisées.

Plan de lecture

Ce document de thèse est composé de cinq chapitres – numérotés de 2 à 6 – dont les Chapitres 3, 4 et 5 présentent les travaux de recherche répondant à notre problématique. Ces trois chapitres correspondent chacun à un article, présenté en langue anglaise, et précédé d'un résumé en français.

Nous donnons ci-dessous une courte description du corps de ce document, chapitre par chapitre.

Chapitre 2, Cadre théorique et hypothèses de recherche

Dans ce chapitre, nous définissons le cadre théorique et de recherche dans lequel nous nous plaçons pour répondre à notre problématique. Pour cela, nous faisons une revue critique des modèles d'apprentissage orthographique existants. Puis nous étendons ce chapitre par une présentation non exhaustive des modèles de contrôle de mouvements oculaires et finissons par une présentation de nos hypothèses de recherche et de nos contributions au développement du modèle *BRAID-Learn*.

Chapitre 3, Modélisation de l'effet de longueur des mots en décision lexicale : le rôle de l'attention visuelle

Ce chapitre est centré sur la simulation des effets de longueur observés comportementalement en tâche de décision lexicale. Notre objectif est double. Grâce à une première série de simulations, nous calibrons finement trois paramètres qui jouent un rôle essentiel dans le modèle *BRAID-Learn* : deux sont dédiés au sous-modèle d'appartenance lexicale (bloc violet dans la Figure 1.1), le troisième détermine les caractéristiques du sous-modèle d'attention visuelle (bloc orange dans la Figure 1.1). Finalement, dans une deuxième série de simulations, nous étudions le rôle de l'attention visuelle sur l'effet de longueur. Ces simulations ont été présentées dans un article publié dans *Vision Research* (Ginestet, Phénix, Diard, & Valdois, 2019).

Chapitre 4, Apprentissage implicite de mots nouveaux chez l'adulte : effets du nombre d'occurrences et de l'attention visuelle sur les mouvements oculaires

Dans cette étude comportementale, nous étudions la relation entre apprentissage orthographique et mouvements oculaires en analysant les variations du comportement oculomoteur en fonction du type d'item (mot nouveau ou mot connu), du nombre d'occurrences et des capacités visuo-attentionnelles de lecteurs adultes (experts) francophones au cours de la lecture répétée de nouveaux mots. Une fois encore, notre objectif est double puisque, en plus de constituer un jeu d'observations original dans la littérature, les données obtenues sont par la suite utilisées pour évaluer la validité du modèle *BRAID-Learn*. Les données comportementales sont présentées et discutées dans un article actuellement en révision (Ginestet, Valdois, Diard, & Bosse, submitted).

Chapitre 5, *BRAID-Learn* : un modèle computationnel d'apprentissage de l'orthographe lexicale

Ce chapitre, dédié au modèle d'apprentissage, présente sa description mathématique ainsi que son évaluation. En nous basant sur l'expérience comportementale présentée au chapitre précédent, nous simulons l'apprentissage de nouveaux mots et observons l'effet du nombre d'occurrences sur les mesures oculomotrices. Une comparaison entre les données simulées et les données comportementales nous permet d'évaluer la capacité du modèle *BRAID-Learn* à reproduire l'effet de la répétition sur les mouvements oculaires associé à la mémorisation orthographique.

Chapitre 6, Discussion

Dans ce dernier chapitre, après un résumé des principaux résultats obtenus au cours de cette thèse, nous discutons les apports du modèle *BRAID-Learn*, et la plausibilité des hypothèses théoriques qu'il implémente. Puis, nous concluons sur les conséquences de notre contribution sur le panorama des modèles théoriques et computationnels actuels, avant d'évoquer les améliorations ou extensions qui sont envisagées. Finalement, nous discutons des perspectives qu'offrent les travaux de recherche présentés dans cette thèse sur les mécanismes impliqués dans l'apprentissage de la lecture et de l'orthographe lexicale chez l'enfant ainsi que sur leurs conséquences sur les méthodes d'apprentissage.

Bonus

Comment s'écrit /sjãpwa/ ?

Chapitre 2

Cadre théorique et hypothèses de recherche

Note

Une partie de ce Chapitre reprend un article publié dans *ANAE – Approche Neuropsychologique des Apprentissages chez l'Enfant* (Ginestet, Valdois, Diard, & Bosse, in press).

Notre objectif principal dans le cadre de cette thèse est de proposer un nouveau modèle d'apprentissage de l'orthographe lexicale et d'en évaluer la plausibilité à travers sa capacité à simuler les données oculomotrices observées online en situation d'apprentissage incident de mots nouveaux.

Dans ce contexte, le chapitre 2 propose une revue de questions portant d'une part sur les modèles d'apprentissage orthographique et d'autre part sur les modèles oculomoteurs qui ont été proposés jusqu'ici dans le contexte de la lecture. Une première originalité de notre travail est de proposer un modèle à l'interface de ces deux courants de recherche. D'une part, les modèles d'apprentissage orthographique existants, qu'ils soient conceptuels ou computationnels, minimisent l'impact des traitements visuels et visuo-attentionnels dans l'apprentissage orthographique et n'ont donc pas l'ambition de rendre compte des patterns de mouvements oculaires. D'autre part, les modèles qui simulent les mouvements oculaires lors de la lecture n'ont pas l'ambition de rendre compte de l'apprentissage orthographique.

Le modèle *BRAID-Learn* que nous développons se situe à l'interface de ces deux cadres de recherche. Dans la mesure où il a été très récemment démontré que l'apprentissage orthographique affecte les mouvements oculomoteurs, il convient de développer de nouveaux cadres théoriques permettant à la fois d'expliquer comment s'acquiert la représentation orthographique d'un nouveau mot lors de rencontres successives et d'évaluer dans quelle mesure les mécanismes proposés rendent compte des mouvements oculaires observés comportementalement en situation d'apprentissage incident. Cette revue de questions sera l'occasion de discuter du rôle de l'attention visuelle dans l'apprentissage de la lecture. L'attention visuelle est très largement absente des modèles existants d'apprentissage orthographique mais c'est un composant central des modèles de mouvements oculaires. *BRAID-Learn* inclut un sous-modèle d'attention visuelle et ce composant joue un rôle majeur dans l'apprentissage orthographique par le modèle et la simulation des mouvements oculaires lors de cet apprentissage.

1 Les modèles d'apprentissage orthographique

Si l'automatisation du processus d'identification de mots joue un rôle majeur dans l'apprentissage de la lecture, le passage d'une lecture sérielle lente chez le lecteur débutant, à une lecture globale plus rapide chez le lecteur expert, est un processus long. Cette transition est rendue possible grâce au développement des connaissances lexicales orthographiques. Un lecteur apprend l'orthographe d'un mot lors d'expositions répétées à la forme écrite de ce mot, le plus souvent en contexte de lecture. Cet apprentissage est principalement incident : il se fait sans que le lecteur en ait conscience et sans passer par une stratégie particulière. Ainsi, plus le mot est lu, meilleure est la représentation orthographique mémorisée, ce qui induit une reconnaissance rapide du mot. Ainsi, l'apprentissage orthographique joue un rôle un essentiel dans le développement des compétences en lecture (Mancheva et al., 2015).

L'utilisation des méthodes computationnelles a permis la formulation d'hypothèses claires sur les mécanismes impliqués dans la lecture à « deux vitesses ». En fonction de leur structure et des hypothèses qu'ils implémentent, les modèles computationnels de lecture simulent une lecture sérielle (ralentie) ou globale (rapide) de différentes façons. D'un côté, les modèles double-voie (Coltheart et al., 2001; Perry et al., 2007) implémentent l'hypothèse de deux voies de lecture. La voie sous-lexicale (ou analytique) inclut un système de transcodage des unités orthographiques en leurs correspondants phonologiques (conversion graphème-phonème). Dédiée à la lecture de pseudo-mots ou de nouveaux mots réguliers, la voie sous-lexicale simule une lecture sérielle associée au décodage phonologique. La voie lexicale (ou globale) simule la lecture de mots connus, réguliers ou irréguliers, grâce à la reconnaissance directe – c'est-à-dire sans décodage phonologique – du mot dont les informations sur la forme orthographique et phonologique sont mémorisées au sein de lexiques.

A ces modèles double-voie s'oppose le modèle en triangle (modèle PDP, pour *Parallel Distributed Processing*, Seidenberg & McClelland, 1989) et le modèle multitraces (MTM, pour *MultiTrace Memory model*, Ans, Carbonnel, & Valdois, 1998). Ces modèles n'implémentent pas de mécanisme de décodage phonologique (conversion graphème-phonème). Le modèle PDP fait l'hypothèse qu'une procédure unique permet de lire aussi bien les mots connus que nouveaux, qu'ils soient réguliers ou non. Plus le mot a été traité, plus la lecture de ce mot est rapide. C'est donc le nombre de rencontres avec un mot qui détermine sa vitesse de reconnaissance. Le modèle MTM, quant à lui, postule l'existence d'un composant visuo-attentionnel déterminant dans le choix de la procédure de lecture (analytique vs. globale) à utiliser. Lorsqu'une séquence de lettres est présentée au modèle, l'attention visuelle, centrée sur l'ensemble des lettres du stimulus, génère l'activation d'un pattern orthographique. Si ce pattern n'est pas identique au stimulus, le modèle procède à une lecture analytique. Dans le cas contraire, le modèle procède à une lecture globale du mot.

L'hypothèse de deux voies de lecture a été émise dans les années 70-80, ce qui fait de ce cadre théorique l'un des plus anciens. Elle reste encore aujourd'hui très influente dans le domaine. En effet, seule cette hypothèse a été reprise pour modéliser l'apprentissage orthographique, dans un premier temps, conceptuellement (Share, 1995), puis computationnellement, à travers deux modèles récents (Pritchard et al., 2018; Ziegler et al., 2014).

L'objectif de cette section est de présenter chacun de ces modèles d'apprentissage orthographique en faisant une description sommaire des hypothèses et mécanismes qu'ils implémentent.

1.1 Le modèle d'auto-apprentissage de Share (1995)

Initialement proposée par Jorm and Share (1983), l'hypothèse de l'auto-apprentissage est conceptualisée par Share (1995). Ce modèle théorique repose sur l'architecture des deux voies de lecture postulée par les modèles double-voie. Il fait de plus l'hypothèse principale que « tout décodage réussi d'une séquence orthographique nouvelle donne l'occasion de mémoriser sa forme orthographique ». A travers ce modèle, Share (1995) propose un modèle d'apprentissage orthographique dans lequel la construction de traces orthographiques en mémoire est progressive, incidente et dépendante des traitements phonologiques.

Hypothèses postulées par le modèle d'auto-apprentissage Share (1995) décrit trois caractéristiques majeures du modèle d'auto-apprentissage associées au décodage phonologique.

Premièrement, le processus de reconnaissance des mots dépend du nombre de fois où le mot a été correctement décodé. En d'autres termes, lors de la première rencontre avec un mot, le lecteur, débutant ou expert, va procéder à une reconnaissance du mot par décodage phonologique via la voie sous-lexicale. Chaque nouvelle rencontre avec ce même mot est une occasion de renforcer la trace orthographique mémorisée. Plus la fréquence d'exposition au mot est élevée, moins le décodage phonologique est requis, privilégiant une reconnaissance rapide par la voie lexicale. On parle d'apprentissage item-dépendant ou *item-based*. Cette première hypothèse induit le développement simultané des deux voies de lecture ce qui contraste avec les modèles développementaux par stades – tel que le modèle de Frith (1985) – qui postulent un apprentissage orthographique par développement successif des voies sous-lexicales et lexicales (on parle alors d'apprentissage étape-dépendant ou *stage-based*). En résumé, le décodage phonologique joue un rôle majeur lors des premières expositions à un mot, justifiant une lecture ralentie des nouveaux mots et des mots rares. Le rôle du décodage phonologique décline progressivement, reflétant un apprentissage orthographique graduel.

Une seconde caractéristique principale du modèle d'auto-apprentissage concerne le développement de la voie sous-lexicale. Plus précisément, Share (1995) suggère que le processus de décodage phonologique devient plus performant au cours de l'apprentissage de la lecture. Initialement, le lecteur débutant connaît les règles basiques de conversion lettre-son (c'est-à-dire, à une lettre correspond un son). Selon Share (1995), la connaissance d'un minimum de règles de conversion graphème-phonème suffit à démarrer le mécanisme d'auto-apprentissage. Ainsi, l'accroissement du lexique orthographique découlant de décodages réussis permet au lecteur débutant de développer sa connaissance des règles de conversion grapho-phonémique. Par conséquent, l'auto-apprentissage de ces règles permet l'apprentissage d'un plus grand nombre de nouveaux mots aboutissant à l'augmentation du nombre de représentations orthographiques mémorisées.

Enfin, si Share (1995) fait l'hypothèse que les traitements phonologiques constituent la pierre angulaire du processus d'apprentissage orthographique – et donc de l'acquisition de la lecture –, le modèle d'auto-apprentissage tient compte des traitements visuo-orthographiques. Ces deux processus seraient à l'origine de la variabilité inter-individuelle observée expérimentalement dans les compétences en lecture. Selon Share (1995), les compétences phonologiques représentent la « condition *sine qua non* de l'acquisition de la lecture ». Par conséquent, elles permettraient d'expliquer une grande partie de la variabilité inter-individuelle. Les traitements visuels sont associés à la vitesse et la qualité de la mémorisation des représentations orthographiques mais restent dépendants d'un décodage phonologique réussi. Pour cette raison, les traitements visuels sont qualifiés de secondaires et n'expliqueraient qu'une part négligeable de la variabilité inter-individuelle observée dans le développement de la lecture.

Ces trois caractéristiques principales du modèle d'auto-apprentissage mettent en avant le rôle majeur et fondamental du décodage phonologique – réussi – dans 1) le développement du lexique orthographique et donc, de la voie lexicale (orthographique), 2) dans le développement des règles de conversion graphème-phonème et donc, de la voie sous-lexicale (phonologique) et 3) dans la vitesse d'acquisition de la lecture.

Si le modèle d'auto-apprentissage postule un rôle secondaire des traitements visuels dans l'apprentissage orthographique, le modèle conceptuel de traitement de l'information en lecture de LaBerge and Samuels (1974) propose une alternative intéressante. Ce modèle de traitement de l'information en lecture (illustré Figure 2.1), inclut quatre blocs mémoriels représentant les mémoires visuelle (VM), phonologique (PM), sémantique (SM) et épisodique (EM). Selon LaBerge and Samuels (1974), les ressources attentionnelles (notées A) sont allouées à un seul composant mémoriel à la fois, en fonction du degré d'automatisation des processus de reconnaissance, de décodage et de compréhension. Nous ne décrirons pas en détail le fonctionnement complet de ce modèle, préférant nous concentrer sur le processus d'apprentissage orthographique.

La Figure 2.2 illustre le processus d'apprentissage à divers niveaux. Considérons dans cet exemple uniquement l'apprentissage du mot $v(w_2)$ composé des éléments sp_2 et sp_3, des lettres l_2 à l_5 et des traits caractéristiques f_4 à f_8. Au tout début de l'apprentissage de $v(w_2)$, les traits caractéristiques de lettres non connues (représentées par f_7 et f_8) sont associés pour former, sous contrôle attentionnel, la représentation de la lettre inconnue l_5. Supposons maintenant que la reconnaissance des lettres est acquise et qu'un lecteur, débutant ou expert, rencontre un nouveau mot. Les lettres l_2, l_3 et l_4 sont identifiées automatiquement, c'est-à-dire sans contrôle attentionnel, puis associées pour former les sous-unités orthographiques sp_2 et sp_3. L'activation de l'unité non familière sp_3 induit un déplacement du point de focalisation attentionnelle (centre de l'attention) sur cette unité. Finalement, les sous-unités orthographiques sont associées pour former la représentation orthographique du mot (notée $v(w_2)$). La répétition du processus d'identification de ce nouveau mot conduit à l'automatisation du traitement de l'information visuelle des sous-unités puis des représentations orthographiques permettant une reconnaissance du mot plus rapide et une allocation des ressources attentionnelles en mémoire phonologique ou sémantique. Ainsi, l'apprentissage orthographique de nouveaux mots est graduelle et n'est rendu possible que par un mécanisme attentionnel permettant un traitement efficace de l'information visuelle. Autrement dit, l'attention visuelle est impliquée à chaque niveau en début de traitement pour permettre la formation d'une unité plus large. Une fois créée cette unité est automatiquement activée et ne requiert plus de ressource attentionnelle ou minimalement.

Le paradigme expérimental de l'auto-apprentissage Share (1999) a également proposé une série de tâches pour évaluer les hypothèses découlant du modèle d'auto-apprentissage. Cet ensemble de tâches forme un paradigme expérimental, logiquement nommé paradigme d'auto-apprentissage, qui se compose de deux phases. Nous décrivons ce paradigme par l'exemple, en décrivant son utilisation dans l'Expérience 1 de Share (1999).

La première phase est une phase de lecture permettant de mettre le sujet en situation d'apprentissage incident de nouveaux mots. Pour cela, quarante élèves de deuxième grade (équivalent à la classe de CE1 en France) de langue maternelle hébraïque sont exposés à dix pseudo-mots. Ces nouveaux mots, d'une longueur comprise entre trois et cinq lettres et composés de deux à quatre syllabes, comportent deux graphèmes ambigus. Chaque pseudo-mot est associé à une catégorie sémantique différente représentée par un nom fictif de ville (par exemple, *DURP*), d'animal ou encore de fleur (dix catégories au total). Lors de la phase de lecture, il est demandé aux élèves de lire à haute voix dix courtes histoires et de désigner celle qu'ils ont préférée.

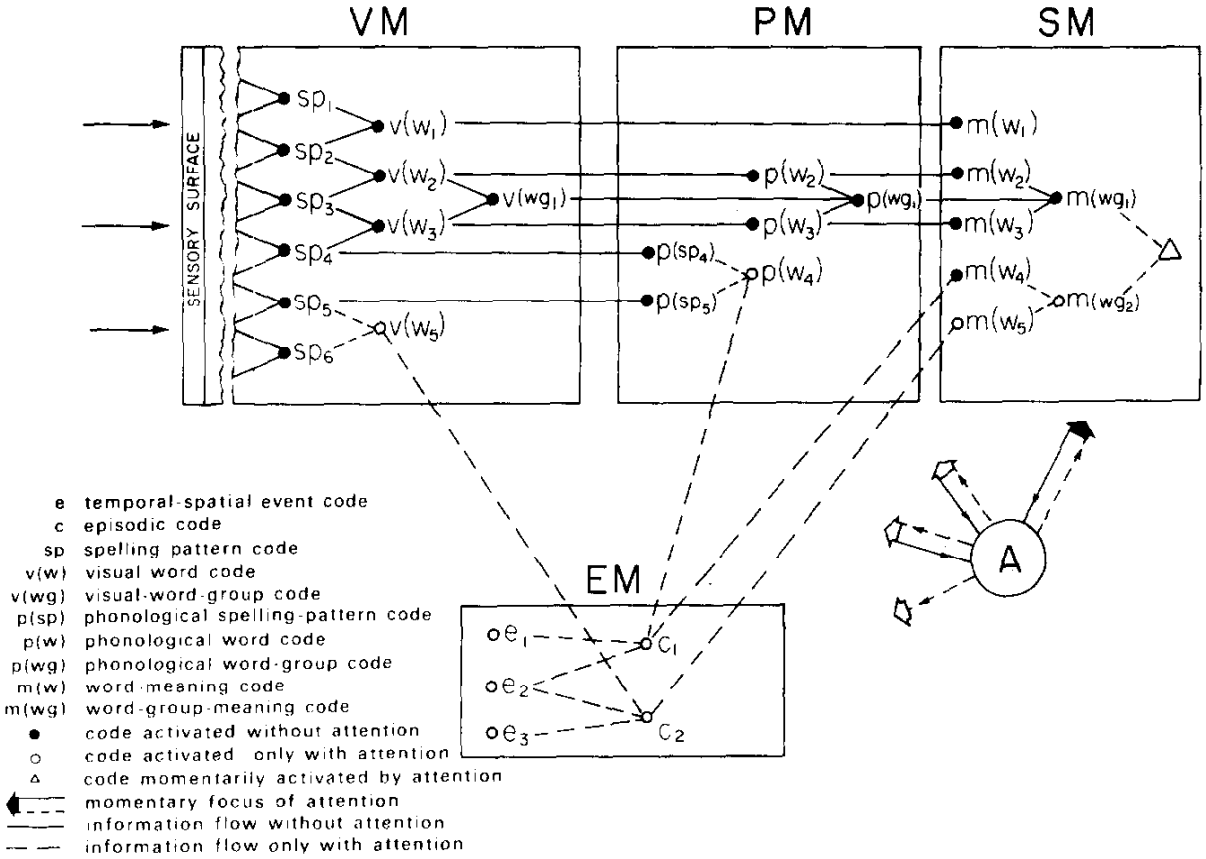

FIGURE 2.1 – Schéma complet du modèle de LaBerge and Samuels (1974). Les rectangles notés VM, PM, SM et EM représentent les blocs de mémoire, respectivement, visuelle, phonologique, sémantique et épisodique. Les ressources attentionnelles, notées A sur le schéma, sont allouées à un seul bloc mémoriel à la fois (dans l'exemple, au bloc de mémoire sémantique). Les lignes en pointillés caractérisent la propagation de l'information sous contrôle attentionnel, activant les unités représentées par un cercle vide à travers tout le système. Figure tirée de (LaBerge & Samuels, 1974).

Chaque histoire contient un pseudo-mot répété quatre ou six fois dans le texte. Un questionnaire de compréhension composé de cinq questions suit la lecture de chaque texte. Durant cette phase, le score de décodage du pseudo-mot (prononciation correcte ou non) est enregistré. L'élève n'est pas informé des tâches suivantes : on parle donc d'apprentissage incident.

La seconde phase est une phase de post-tests permettant de mesurer la qualité de l'apprentissage orthographique. Dans l'Expérience 1 de Share (1999), elle est mesurée trois jours plus tard et est composée de trois tâches. La première est une tâche de choix orthographique dans laquelle quatre items sont présentés simultanément : le mot cible tel que lu pendant la phase d'apprentissage (exemple : *DURP*), un pseudo-homophone du mot cible (exemple : *DERP*), un distracteur créé par substitution d'une des lettres du mot cible (exemple : *DURB*) et un distracteur créé, à partir du mot cible, par transposition de deux lettres adjacentes (exemple : *DRUP*). L'élève doit choisir l'item correspondant à un pseudo-mot qu'il a lu trois jours auparavant. Le taux de bonnes réponses est enregistré. La tâche de dénomination suit la tâche de choix orthographique. Au cours de cette tâche, les mots cibles, leur homophone et quarante mots connus sont affichés l'un après l'autre à l'écran. L'élève doit lire (à haute voix) le plus vite

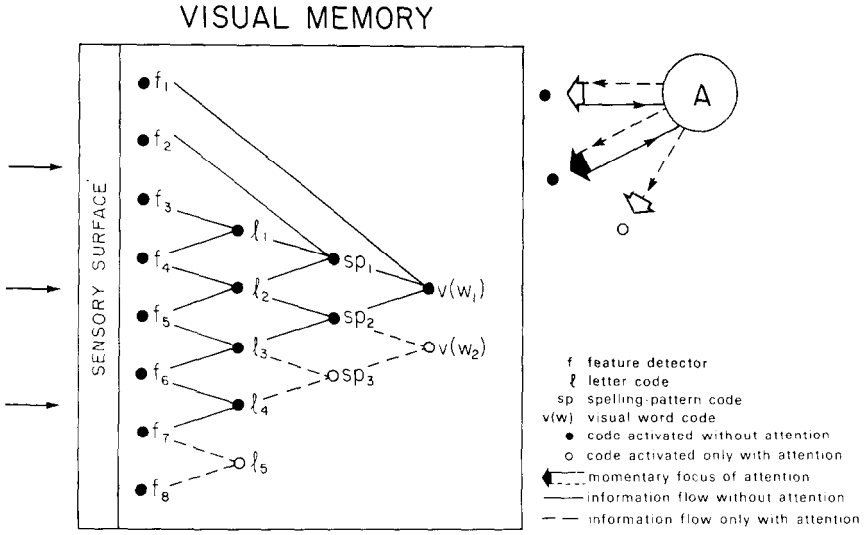

FIGURE 2.2 – Schéma du bloc de mémoire visuelle du modèle de LaBerge and Samuels (1974). Ce schéma illustre la propagation de l'information sensorielle venant du stimulus (flèches à gauche du schéma) par activation progressive des traits caractéristiques f_i des lettres l_i qui composent les sous-unités orthographiques sp_i des mots $v(w_i)$. Chaque unité peut-être activée avec ou sans contrôle attentionnel (représenté respectivement par une ligne en pointillés ou un trait plein). Figure tirée de (LaBerge & Samuels, 1974).

et le plus correctement possible les items présentés à l'écran. Les temps de réponse ainsi que le score en lecture (prononciation correcte ou non) sont enregistrés. Finalement, la phase de post-test se termine par une dictée. Après un rappel du thème de l'histoire associée au pseudo-mot à écrire, l'expérimentateur demande à l'élève d'écrire le nom de la ville, de l'animal ou de la fleur. Le taux de pseudo-mots correctement écrits (mot entier ou lettres des graphèmes ambigus) sont enregistrés.

Ce paradigme a inspiré de nombreuses études expérimentales dans le domaine, faisant du paradigme d'auto-apprentissage l'un des plus utilisés actuellement. Ces études évaluent l'apprentissage orthographique de nouveaux mots intégrés dans des textes (Bowey & Muller, 2005; Nation et al., 2007; Share, 2004; Tamura, Castles, & Nation, 2017; Tucker et al., 2016; Wang, Castles, Nickels, & Nation, 2011) mais aussi de nouveaux mots présentés isolément (Bosse, Chaves, Largy, & Valdois, 2015; Nation et al., 2007). Bien que ces études aient adapté leur protocole expérimental relativement à celui décrit par Share (1999), en fonction des participants (débutant ou expert) et en fonction du contexte d'apprentissage (mot isolé ou intégré dans un texte), toutes conservent un protocole comportant deux phases dédiées à l'apprentissage incident via la lecture pour la première et à la mesure de l'apprentissage orthographique effectif via des épreuves dédiées pour la seconde.

1.2 Les modèles computationnels d'apprentissage orthographique

L'utilisation des méthodes computationnelles a permis récemment le développement de deux modèles d'apprentissage de l'orthographe lexicale, tous deux issus de l'extension de modèles de lecture à haute voix, respectivement, le modèle CDP (pour *Connectionist Dual Process*, Perry et al., 2007, 2010, 2013) et le modèle DRC (pour *Dual Route Cascade*, Coltheart et al., 2001).

Leur structure commune de type « double-voie » fait de ces modèles des outils particulière-

ment adaptés à l'implémentation de l'hypothèse d'auto-apprentissage avancée par Share (1995, 1999, 2004). En effet, les modèles double-voie ont pour principal objectif de simuler la lecture de mots isolés, c'est-à-dire, la production orale d'une séquence de lettres écrites. De ce fait, ils mettent principalement l'emphase sur les processus phonologiques et implémentent une version minimaliste des mécanismes visuels associés aux traitements visuo-orthographiques.

Ainsi, les modèles d'apprentissage découlant des modèles double-voie postulent, par définition, un rôle majeur des traitements phonologiques pendant l'apprentissage orthographique de nouveaux mots. Plus précisément, ces modèles font l'hypothèse que la connaissance des relations graphème-phonème suffit à associer la représentation orthographique d'un nouveau mot à sa forme phonologique, synonyme d'un apprentissage orthographique réussi.

Si ces modèles implémentent des hypothèses théoriques similaires – auto-apprentissage grâce à deux voies de lecture –, ils diffèrent sur de nombreux points. L'objectif de cette section est donc de proposer une revue critique de ces modèles en présentant succinctement leur fonctionnement ainsi que les données comportementales qu'ils permettent de simuler.

1.2.1 Le modèle de Ziegler et al. (2014)

Le modèle computationnel d'apprentissage orthographique proposé par Ziegler et al. (2014) constitue, à notre connaissance, le premier modèle permettant de simuler l'acquisition implicite des connaissances orthographiques lexicales. Il implémente l'hypothèse d'auto-apprentissage défendue par Share (Share, 1995, 1999, 2004), modélisant ainsi l'acquisition des connaissances lexicales orthographiques chez le lecteur débutant.

Structure et fonctionnement du modèle Son architecture de base est similaire à celle du modèle connexionniste de lecture à haute voix CDP (Perry et al., 2007, 2010, 2013), schématisée en Figure 2.3. Le modèle CDP retient l'hypothèse de deux voies de lecture. La voie lexicale est constituée des lexiques orthographique et phonologique, associés aux connaissances sémantiques (représentées dans le modèle conceptuel mais non implémentées). Elle permet, grâce à un accès direct au lexique mental, de retrouver la prononciation d'un mot connu. La voie sous-lexicale comprend un premier module, le « buffer graphémique », qui a pour rôle de découper le stimulus en séquence grapho-syllabique et parcourir cette séquence. Chaque élément de cette séquence alimente le second module, nommé TLA (pour Two-Layer network of phonological Assembly), pouvant être résumé par la connaissance et l'application des associations entre graphème et phonème (après entraînement) permettant de générer une représentation phonologique du stimulus écrit. La voie sous-lexicale permet la lecture de mots ou pseudo-mots réguliers et est particulièrement impliquée lors de l'apprentissage des représentations orthographiques de nouveaux mots.

L'hypothèse d'auto-apprentissage implémentée dans le modèle de Ziegler et al. (2014) se compose de trois étapes, illustrées dans la Figure 2.4.

Avant apprentissage, la voie lexicale caractérise les connaissances présentes à l'état initial. Ziegler et al. (2014) postulent des connaissances lexicales orthographiques initiales nulles (c'est-à-dire, un lexique orthographique vide), mais supposent que des connaissances phonologiques sont déjà disponibles, et induites par l'exposition au langage oral. Le réseau TLA de la voie sous-lexicale est explicitement entraîné sur un petit set de mots, aux règles simples de conversion graphème-phonème (Hutzler, Ziegler, Perry, Wimmer, & Zorzi, 2004). La Figure 2.4(b)(i) illustre cet état initial, préalable à l'apprentissage orthographique.

Lorsqu'un nouveau mot écrit est « présenté » au modèle, le traitement visuel du stimulus est assuré par le module situé en amont des voies de lecture (rectangle « letters » dans la

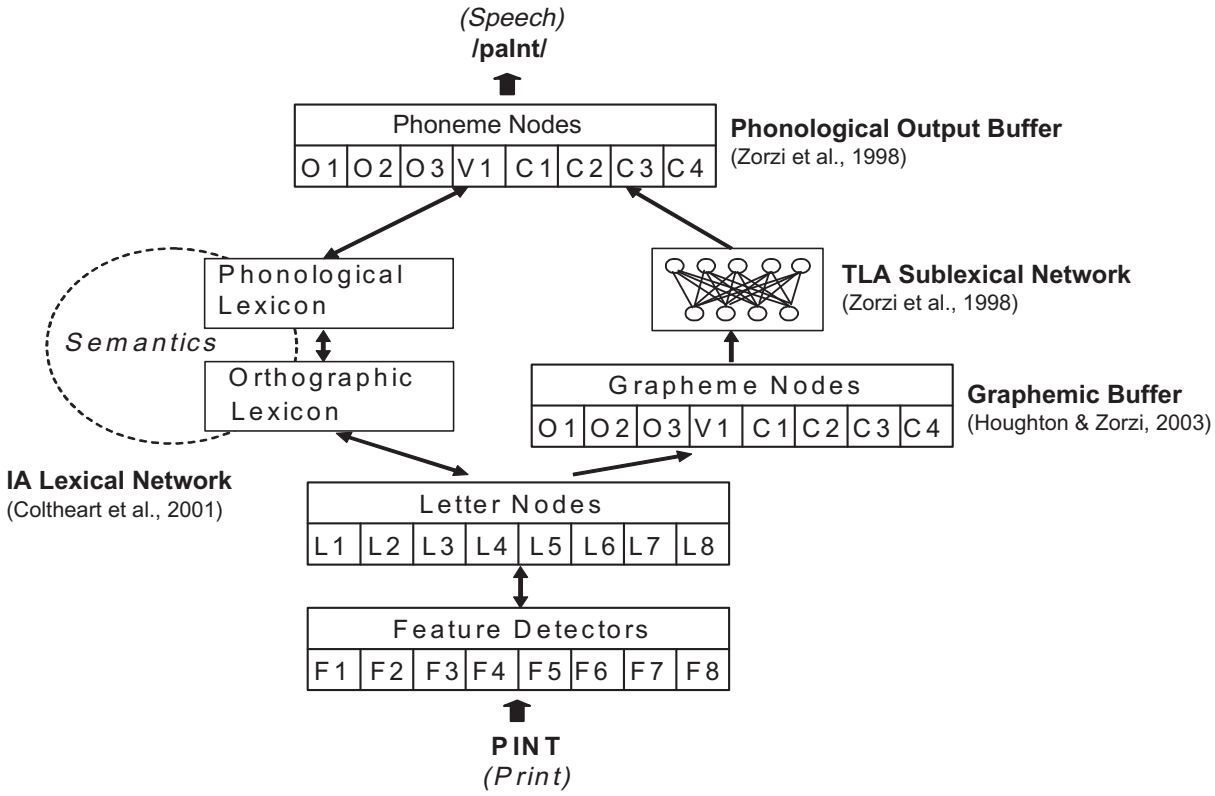

FIGURE 2.3 – Schéma du modèle CDP+ de Perry et al. (2007). Chaque module implémenté dans le modèle est représenté par un bloc, et les flèches représentent les voies de transfert de l'information. Le traitement visuel de la séquence de lettres formant le stimulus présenté à l'entrée du modèle (ici, le mot anglais *PINT*) aboutit à sa lecture à haute voix en sortie, soit par décodage phonologique (voie de droite sur le schéma), soit par lecture directe (voie de gauche sur le schéma). Figure tirée de (Perry et al., 2010).

Figure 2.4(a)). Ce module garantit une reconnaissance immédiate et parfaite de la position et de l'identité des lettres composant le stimulus. Le décodage phonologique résultant de l'application des associations graphème-phonème via le réseau TLA aboutit à l'activation de l'ensemble des phonèmes correspondant aux unités graphémiques du mot. Les phonèmes générés par la voie sous-lexicale entraînent l'activation des représentations phonologiques correspondantes au sein du lexique phonologique (Figure 2.4(b)(ii). La représentation phonologique la plus activée est alors associée à la représentation orthographique lexicale nouvellement créée.

La boucle d'auto-apprentissage se solde par la mise en mémoire lexicale de la trace orthographique. Cette dernière étape, illustrée Figure 2.4(b)(iii), correspond à la création d'une connexion directe entre le mot le plus activé au sein du lexique phonologique et la représentation orthographique du mot présenté en entrée. Parallèlement, chaque décodage réussi permet l'amélioration du mécanisme de décodage implémenté dans le réseau TLA grâce à un ajustement des poids de connexion. Les auteurs parlent d'entraînement du réseau TLA. Finalement, chaque nouvelle présentation du mot permet une mise à jour de la fréquence de ce mot.

Principaux résultats Une première série de simulations permet d'évaluer la capacité du modèle à apprendre des connaissances lexicales orthographiques. Après la phase de pré-entraînement

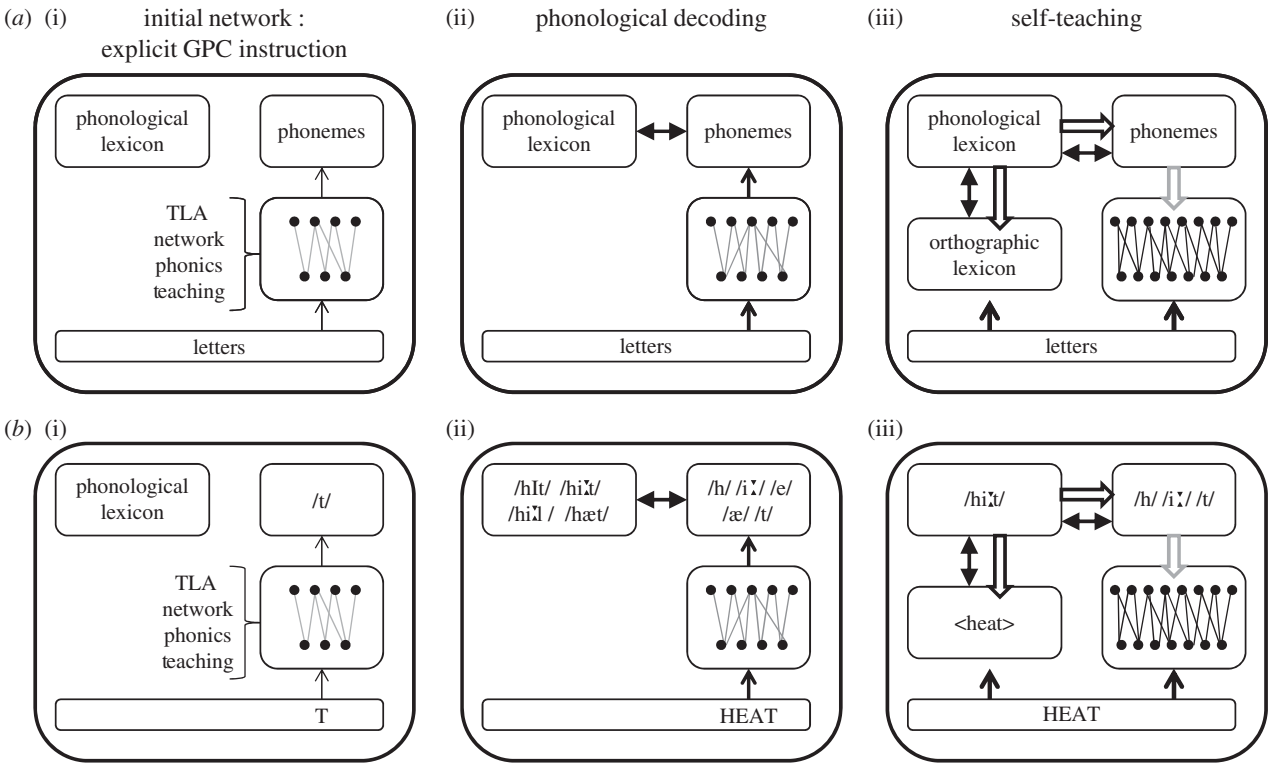

FIGURE 2.4 – Schéma du processus d'apprentissage orthographique implémenté dans le modèle de Ziegler et al. (2014). Les figures notées (*a*) illustrent les trois étapes du processus d'apprentissage : le pré-entraînement (explicite) du réseau TLA (i), le décodage phonologique (ii) et, la mémorisation de la représentation orthographique en mémoire lexicale suivie de la mise à jour du réseau TLA par auto-apprentissage (iii). Les figures notées (*b*) illustrent, par l'exemple, les différentes étapes du processus d'apprentissage du mot anglais *HEAT*. Figure tirée de (Ziegler et al., 2014).

du réseau TLA, le modèle est utilisé pour simuler l'apprentissage orthographique de 32 735 mots anglais (voir Perry et al., 2013 pour une liste complète) présentés 500 000 fois. Les résultats reportés indiquent qu'environ 80 % des mots sont correctement appris lorsque le seuil d'activation des mots stockés dans le lexique phonologique est de 0,05 et que ce score n'est que de 45 % lorsque le seuil est fixé à 0,45. En d'autres termes, plus le nombre de mots activés dans le lexique phonologique est élevé, plus le nombre de mots appris est élevé. Baisser le seuil d'activation revient à réduire l'inhibition entre les phonèmes activés dans le buffer phonologique et les représentations phonologiques contenues dans le lexique phonologique, favorisant ainsi la reconnaissance des mots irréguliers. Ainsi, les auteurs concluent que l'absence d'apprentissage observé pour certains mots pourrait être liée à l'opacité de la langue anglaise, induisant une correspondance lettres-sons ambiguë chez un nombre important de mots « à apprendre ». Cette observation justifierait l'exposition à des nouveaux mots à travers des contextes variés afin de lever toute ambiguïté, permettant ainsi une mémorisation correcte de la trace orthographique de ces nouveaux mots.

Dans une version plus récente du modèle, Perry et al. (2019) ajoutent un mécanisme permettant au modèle d'apprendre les mots considérés comme trop irréguliers pour être appris par décodage phonologique (c'est-à-dire les 20 % de mots non appris avec un seuil de 0,05). Ce mécanisme modélise l'apprentissage par instruction directe ou apprentissage explicite obtenu, par exemple, grâce à l'utilisation de *flash cards* (Perry et al., 2019). Ainsi, lorsque la lexicali-

sation par décodage phonologique échoue, ce mécanisme permet au modèle de déclencher un apprentissage explicite par la voie lexicale.

Selon l'hypothèse d'auto-apprentissage implémentée dans ce modèle, le décodage phonologique joue un rôle essentiel dans l'apprentissage orthographique. A travers la Simulation 3, Ziegler et al. (2014) évaluent les conséquences d'un apprentissage incorrect, lié à une erreur de décodage en début d'apprentissage, sur les performances générales du modèle. Pour cela, deux cas de figure sont envisagés. Dans le premier cas de figure, le modèle simule un apprentissage inexistant, lorsqu'aucun mot du lexique phonologique n'est sélectionné par activation des phonèmes de l'entrée, signifiant l'absence de création de connexion entre entrées phonologique et orthographique, et donc, aucune modification des connexions du réseau TLA. Dans le second cas de figure, le modèle simule un apprentissage erroné, induisant l'association d'une représentation phonologique incohérente – choisie aléatoirement parmi les mots du lexique phonologique activés – avec la séquence orthographique présentée et par conséquent, une mise à jour du réseau TLA incorrecte.

Les résultats obtenus montrent un faible impact de l'absence totale de lexicalisation puisque, à la fin des 500 000 présentations, le taux d'apprentissage est similaire quelle que soit la probabilité de non-lexicalisation. A l'inverse, un décodage erroné génère un taux d'apprentissage d'autant plus faible que la probabilité d'erreur de lexicalisation est élevée mais le taux d'apprentissage reste néanmoins globalement élevé (supérieur à 60 %). Selon les auteurs, ces résultats étayent la conclusion précédente suggérant que la diversité et la multiplicité des contextes de lecture augmentent la probabilité qu'un lecteur débutant décode avec succès une séquence de lettres, engendrant une lexicalisation correcte du mot.

Enfin, Ziegler et al. (2014) (Simulation 4) évaluent l'importance relative des mécanismes visuel et phonologique implémentés sur l'apprentissage orthographique. Pour cela, ils réalisent deux séries de simulations, chacune d'elles permettant de simuler le principal processus déficitaire – visuel ou phonologique – associé à la dyslexie – respectivement de surface ou phonologique. Un traitement visuel déficitaire est simulé par la probabilité (comprise entre 0,02 et 0,10) que chaque lettre composant le mot soit échangée avec la lettre suivante (exemple : CAT devenant ACT). Chaque fois qu'un mot est activé dans le lexique phonologique, chaque phonème généré en sortie du réseau TLA (c'est-a-dire dans le buffer phonologique, voir Figure 2.3) est modifié suivant une probabilité comprise entre 0,05 et 0,45. En conséquence, les poids de connexion du réseau TLA sont incorrectement ajustés, simulant une connaissance des relations graphème-phonème partiellement correctes, par conséquent, un traitement phonologique déficitaire. Les résultats obtenus montrent que seulement 8 % des mots sont appris correctement après 500 000 présentations lorsque le modèle simule un apprentissage perturbé phonologiquement (probabilité de 0,45). Par contre, le modèle apprend encore environ 70 % des mots présentés lorsqu'un déficit visuel jugé important est simulé (probabilité de 0,10). Ziegler et al. (2014) en concluent que seul le dysfonctionnement phonologique a un effet majeur sur la performance et que les capacités phonologiques sont indispensables à l'apprentissage orthographique.

Dans une étude récente, Perry et al. (2019) évaluent la capacité du modèle à rendre compte des différences inter-individuelles dans les performances en lecture de mots réguliers, irréguliers et pseudo-mots. Pour cela, ils utilisent les résultats comportementaux de 622 enfants anglophones – dont 388 dyslexiques – décrits par Peterson, Pennington, and Olson (2013). Dans un premier temps, les performances de chaque enfant reportées lors de trois tâches comportementales (choix orthographique, suppression phonémique, vocabulaire) sont utilisées, avant apprentissage, pour créer un modèle unique, par ajustement des paramètres du modèle. Ce sont donc autant de modèles que d'enfants qui sont générés. Les performances du traitement visuel, associées aux résultats obtenus dans la tâche de choix orthographique, sont modulées par

l'ajout de bruit dans le lexique orthographique et par la probabilité de lexicalisation du mot. Les performances du traitement phonologique, associées aux résultats obtenus dans la tâche de suppression phonémique, sont modulées par l'ajout de bruit dans le réseau TLA durant la phase d'entraînement explicite (avant apprentissage). Enfin, la taille du lexique phonologique est fonction des scores en vocabulaire. Des modèles multi-déficits sont considérés, c'est-à-dire que chaque modèle peut présenter un déficit sur chacune ou plusieurs de ces dimensions.

Après ajustement des paramètres du modèle, une phase complète d'apprentissage est réalisée pour ce modèle. Enfin, après cet apprentissage, les simulations de lecture de mots réguliers, irréguliers et non-mots sont comparées aux données comportementales. Les résultats indiquent que le modèle simule avec succès les performances des normo-lecteurs et des dyslexiques, rendant compte des différences inter-individuelles observées comportementalement. Les résultats simulés par les modèles multi-déficits pour les dyslexiques sont comparés à ceux obtenus par trois autres modèles alternatifs. Le premier présente un déficit purement visuel induit par l'inversion de lettres composant le stimulus, le second, un déficit purement phonologique caractérisé par des erreurs de conversion graphème-phonème, et le troisième est un modèle global bruité représentant un traitement général déficitaire (ajout de bruit dans chaque composant du modèle). Perry et al. (2019) concluent que la dyslexie pourrait être associée à de multiples déficits, montrant que le modèle multi-déficits permet une meilleure simulation des performances individuelles en lecture.

1.2.2 Le modèle de Pritchard et al. (2018)

Le modèle d'apprentissage orthographique développé par Pritchard et al. (2018) partage de nombreux points communs avec celui proposé par Ziegler et al. (2014) et décrit dans la section précédente. En effet, le modèle de Pritchard et al. (2018) constitue une extension d'un modèle connexionniste de lecture à haute voix de type double-voie. De plus, il implémente l'hypothèse d'auto-apprentissage défendue par Share (1995, 1999, 2004), permettant ainsi de simuler le développement des connaissances lexicales orthographiques chez le lecteur débutant.

Structure et fonctionnement du modèle Comme son nom l'indique, le modèle de Pritchard et al. (2018), nommé ST-DRC (pour *Self-Teaching Dual Route Cascade*) est issu du modèle DRC (Coltheart et al., 2001). Son architecture générale est schématisée en Figure 2.5. Cette dernière étant composée de deux voies de lecture (lexicale et sous-lexicale) similaires à celles du modèle CDP décrit dans la Section 1.2.1, nous ne décrivons pas chacune des voies en détail, préférant présenter les spécificités du modèle ST-DRC à travers la description du mécanisme d'auto-apprentissage implémenté.

A l'état initial représentant les connaissances préalables à l'apprentissage de la lecture, le lexique phonologique contient l'ensemble des représentations phonologiques des mots monosyllabiques anglais alors que le lexique orthographique est vide, caractérisant l'absence de connaissances lexicales orthographiques. Dans le modèle ST-DRC, l'ensemble des règles de conversion graphème-phonème impliquées dans le décodage phonologique (voie sous-lexicale) sont supposées connues.

L'apprentissage est conditionné par la reconnaissance « phonologique » du mot. Ce processus est similaire à celui décrit dans le modèle de Ziegler et al. (2014). En entrée, le module de reconnaissance de lettres implémenté dans le modèle ST-DRC simule une reconnaissance parfaite de la position et de l'identité des lettres qui composent le stimulus. Grâce à l'application des règles de conversion graphème-phonème, le décodage phonologique (réalisé dans la voie sous-lexicale) permet la reconnaissance des phonèmes constituant le stimulus, activant les re-

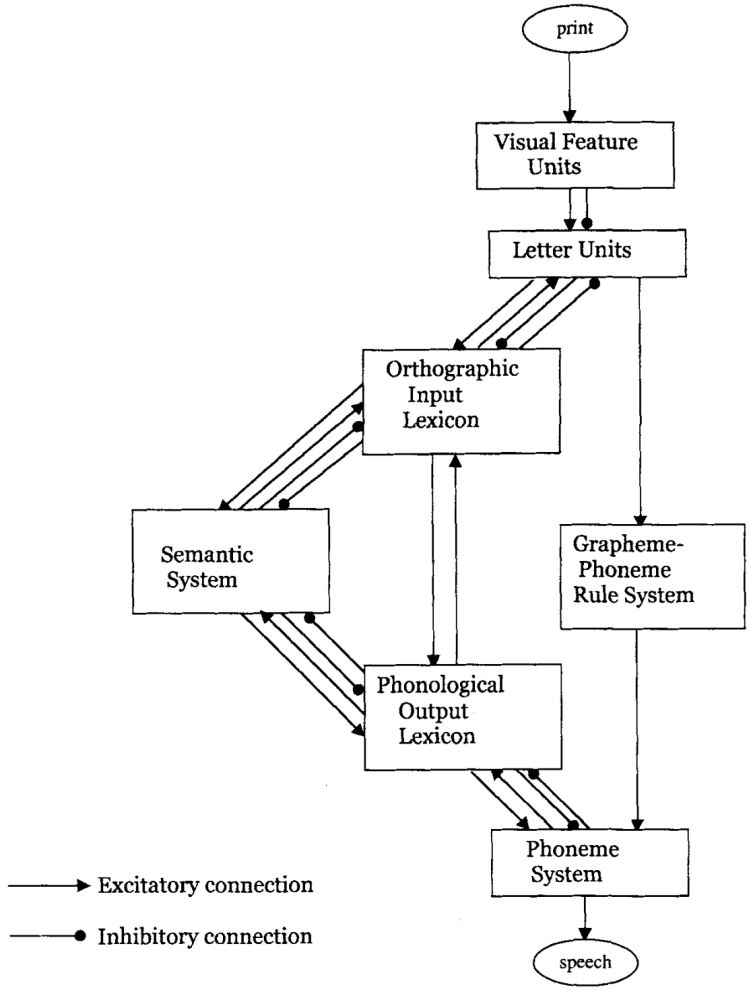

FIGURE 2.5 – Schéma du modèle DRC de Coltheart et al. (2001). Ce schéma représente chaque module implémenté dans le modèle, représenté par un bloc, et les flèches indiquent les influences, soit excitatrices soit inhibitrices. Le traitement visuel d'un stimulus présenté en entrée aboutit à sa lecture à haute voix en sortie, soit par décodage phonologique (voie de droite sur le schéma), soit par lecture directe (voie de gauche sur le schéma). Figure tirée de (Coltheart et al., 2001).

présentations phonologiques stockées dans le lexique phonologique. La trace phonologique dont l'activation dépasse le seuil de reconnaissance « phonologique » fixé signifie que le stimulus est connu, déclenchant le processus d'apprentissage.

Deux cas de figure sont alors possibles. Le premier cas, nommé apprentissage *type-based*, correspond à l'apprentissage lors de la première exposition à la forme écrite de ce mot. Par définition, dans ce cas, aucune représentation orthographique du mot n'est présente dans le lexique orthographique. Ainsi, l'apprentissage correspond à la mise en mémoire orthographique de la séquence de lettres présentée au modèle. Ce processus conduit à la création de connexions bi-directionnelles excitatrices entre la représentation orthographique du stimulus nouvellement ajoutée au lexique orthographique et la représentation phonologique reconnue. Parallèlement, des connexions inhibitrices sont créées entre la nouvelle entrée orthographique et, d'une part, les autres mot du vocabulaire oral (connexions bi-directionnelles) et d'autre part, les autres mots du lexique orthographique. Finalement, à chaque nouvelle entrée orthographique est associée une fréquence initiale égale à 10 (valeur fixée pour les simulations).

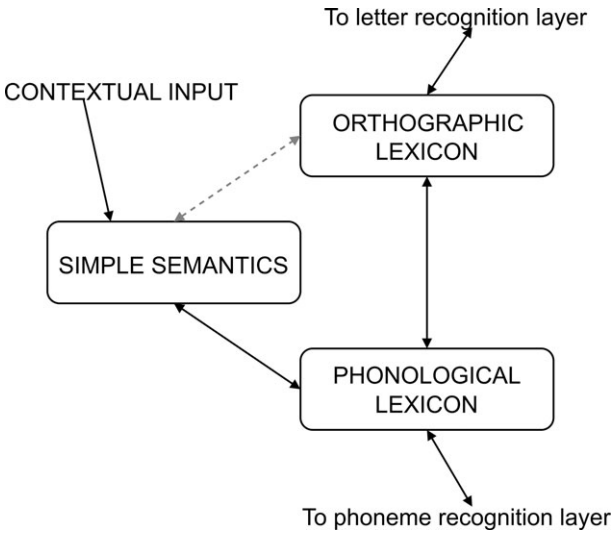

FIGURE 2.6 – Schéma de la voie lexicale du modèle ST-DRC de Pritchard et al. (2018). Ce schéma représente les modules implémentés dans la voie sous-lexicale incluant la quantité d'informations sémantiques disponibles (notée *contextual input*). La connexion bi-directionnelle entre la représentation sémantique et la représentation orthographique d'un mot n'est pas implémentée dans le modèle (flèche pointillée). Figure tirée de (Pritchard et al., 2018).

Le second cas, nommé apprentissage *token-based*, correspond à l'apprentissage lorsque la représentation orthographique du mot présenté au modèle est déjà contenue dans le lexique orthographique. Ce second cas de figure caractérise le renforcement des connaissances lexicales orthographiques suite à la lecture répétée du nouveau mot. Autrement dit, l'apprentissage *token-based* succède à l'apprentissage *type-based*. Son mécanisme est simple. Si le mot est lu par la voie lexicale (c'est-à-dire, reconnu grâce à l'activation d'une représentation orthographique au-dessus du seuil de reconnaissance de mots écrits fixé), le modèle procède à une mise à jour de ses connaissances sur le mot reconnu. La représentation orthographique ne comportant pas d'erreur dès la première présentation, seule la fréquence du mot est mise à jour. Ainsi, à chaque nouvelle lecture d'un mot présent dans le lexique orthographique, la fréquence de ce dernier est augmentée de 10.

Enfin, Pritchard et al. (2018) ont implémenté une première version de la représentation sémantique des mots. Ce module, nommé « Simple Semantics » dans le modèle ST-DRC est intégré à la voie lexicale (Figure 2.6). Il est connecté au lexique phonologique et au lexique orthographique (cette connexion n'est pas implémentée dans le version actuelle du modèle) permettant ainsi l'association des traces orthographiques et phonologiques mémorisées à leur représentation sémantique. Ce module est systématiquement activé, quel que soit le stimulus présenté. L'activation de la représentation sémantique du mot est fonction des informations fournies par le contexte. Bien que le modèle ST-DRC simule la lecture de mots isolés, Pritchard et al. (2018) intègrent un paramètre permettant au modèle de simuler la quantité d'informations sémantiques disponibles, issues, par exemple, de la lecture d'un texte. La valeur de ce paramètre peut être assimilée à la prédictibilité du mot présenté. Ainsi, plus l'activation de la représentation sémantique est élevée, plus le mot est prédictible, favorisant sa reconnaissance.

Principaux résultats Pendant la phase d'apprentissage, le modèle ST-DRC est exposé à 30 220 items formant le corpus. Il s'agit de 8 017 nouveaux mots anglais monosyllabiques

(6 658 mots réguliers et 1 359 mots irréguliers) dont 1 620 apparaissent plus d'une fois dans le corpus. Cela permet la simulation de la lecture répétée de certains mots, représentative de la fréquence d'apparition des mots dans les textes. Après la phase d'apprentissage, le modèle évalue la qualité de la mémorisation orthographique en simulant la lecture à haute voix (post-test) des 8 017 nouveaux mots. Pour que l'apprentissage soit considéré comme correct, il doit aboutir à la création d'une connexion cohérente entre la représentation orthographique mémorisée et la représentation phonologique du mot reconnu et le mot nouvellement appris doit être correctement prononcé lors du post-test.

Dans une première série de simulations (Simulation 1 dans Pritchard et al. (2018)), les auteurs examinent l'influence du contexte sur l'apprentissage. Pour cela, ils font varier, pendant la phase d'apprentissage, 1/ le nombre de représentations sémantiques activées de 0 à 100 (0 correspond à l'absence de contexte simulant l'apprentissage d'un mot isolé, 1 représente un contexte très informatif prédisant un unique mot, et 100 représente un contexte très ambigu activant des mots sémantiquement différents) et 2/ l'effet « excitateur » du module sémantique de 0,02 à 1. Les résultats obtenus montrent un effet (qualitatif) massif du contexte sur l'apprentissage orthographique des mots irréguliers. Plus le contexte sémantique est précis, plus le nombre de mots irréguliers appris est élevé. En effet, alors qu'en l'absence de contexte moins d'1 % des mots irréguliers sont appris, un contexte précisé – une seule représentation sémantique activée et effet « excitateur » du module sémantique supérieure à 0,02 – permet au modèle d'en apprendre plus de 92 %. A l'inverse, aucun effet du contexte n'est observé lors de l'apprentissage de mots réguliers.

Ces résultats sont aisément interprétables au vu du mécanisme d'apprentissage implémenté dans le modèle ST-DRC. En effet, l'apprentissage est conditionné par l'activation d'un mot du lexique phonologique suite au décodage. Or, plus le mot contient de graphèmes ambigus, plus la reconnaissance du mot par la voie sous-lexicale est compromise. Ainsi, en l'absence de connaissances sémantiques, la lecture de nouveaux mots irréguliers aboutit rarement au déclenchement du processus d'apprentissage. Pour les auteurs, ces résultats sont en accord avec l'hypothèse d'auto-apprentissage qui suggère que le décodage partiel de nouveaux mots irréguliers dans un contexte sémantique précis permet leur reconnaissance et par conséquent, leur apprentissage.

Deux types d'erreurs aboutissent à une erreur d'apprentissage : les erreurs de décodage et les erreurs de reconnaissance de mot. Les scores en lecture obtenus lors du post-test montrent que malgré la mise en mémoire lexicale des représentations orthographiques, le modèle fait des erreurs de lecture. Ces erreurs sont la conséquence d'une erreur de décodage. En d'autres termes, la forme phonologique activée ne correspond pas à la représentation orthographique présentée au modèle. Les résultats montrent que ce type d'erreur intervient majoritairement lors de l'apprentissage de mots irréguliers et lorsque le contexte sémantique est précisé (nombre de représentations sémantiques activées compris entre 1 et 5). Deux cas de figure sont reportés. Lorsque la lecture « régulière » d'un mot irrégulier aboutit à l'activation d'un mot régulier (exemple : *HEAD-HEED*), on parle d'erreur de régularisation. Parfois, une même représentation orthographique est associée à différentes formes phonologiques. On parle alors d'homographes hétérophones, qui résultent d'un décodage non consistant au cours de présentations répétées. Les erreurs de reconnaissance de mots apparaissent lors de l'apprentissage d'homophones hétérographes (exemple : *BALL-BAWL*). Ce type d'erreur s'observe plus spécifiquement lorsque le contexte sémantique est clair. L'apprentissage d'homophones aboutit, normalement, à l'association d'une même forme phonologique à différentes formes orthographiques. Selon Pritchard et al. (2018), la forte activation d'une forme phonologique résultant de l'association des connaissances sémantiques et du décodage phonologique conduit à l'activation de la représentation

orthographique de l'homophone déjà mémorisé. Par conséquent, la trace orthographique du mot présenté n'est pas mémorisée.

Dans une deuxième série de simulations (Simulation 3), Pritchard et al. (2018) évaluent l'effet de la lecture répétée sur l'apprentissage orthographique. Pour cela, le modèle est exposé quatre fois au même corpus. Ce sont donc 120 880 mots qui sont présentés au modèle. L'effet « excitateur » du module sémantique est fixé à 0,25 et le nombre de représentations sémantiques activées admet deux valeurs (4 et 100) permettant de comparer l'apprentissage dans un contexte sémantique précis ou ambigu. Comme précédemment, la phase d'apprentissage est suivie d'une phase de lecture (post-test). Les résultats montrent une augmentation du nombre de mots appris – qu'ils soient réguliers ou non – avec le nombre d'occurrences, uniquement lorsque le contexte est précis. Lorsque le contexte est ambigu, aucun effet du nombre d'occurrences n'est observé, montrant que 100 % des mots réguliers sont lus correctement et seulement 10 % des mots irréguliers le sont, quel que soit le nombre d'occurrences. La représentation orthographique ne comportant aucune erreur dès sa mémorisation, ces observations reflètent la capacité du modèle à associer à une même représentation orthographique différentes phonologies.

1.2.3 Discussion

Les modèles d'apprentissage orthographique de Ziegler et al. (2014) et de Pritchard et al. (2018) implémentent l'hypothèse d'auto-apprentissage soutenue par Share (1995, 1999, 2004). A travers plusieurs séries de simulations, ces auteurs ont évalué la capacité de leur modèle respectif à simuler l'acquisition de connaissances lexicales orthographiques chez le lecteur débutant.

Développés à partir des modèles double-voie CDP (Perry et al., 2007, 2010, 2013) et DRC (Coltheart et al., 2001), le mécanisme de décodage phonologique qu'ils implémentent constitue le point fort de ces modèles. Les représentations phonologiques pré-existantes et la connaissance préalable, même minimaliste, des relations graphème-phonème suffisent à mémoriser correctement la trace orthographique de nouveaux-mots réguliers. Par ailleurs, l'implémentation d'un mécanisme phonologique complet permet de simuler le comportement de lecteurs dont les compétences phonologiques seraient altérées (Perry et al., 2019; Ziegler et al., 2014).

Ces modèles présentent des différences qu'il nous paraît intéressant de relever. De par son architecture, seul le modèle de Ziegler et al. (2014) permet l'ajustement du mécanisme de décodage (réseau TLA de la voie sous-lexicale) au cours de l'apprentissage, ce que les auteurs apparentent au développement par auto-apprentissage de la connaissance des règles de conversion graphème-phonème. Ainsi, Ziegler et al. (2014) font l'hypothèse forte que les voies lexicale et sous-lexicale sont entraînées simultanément. Le modèle ST-DRC, quant à lui, intègre un module sémantique dont l'excitation dépend du degré d'ambiguïté du contexte de lecture. Grâce à cette implémentation, Pritchard et al. (2018) font l'hypothèse forte que l'apprentissage orthographique résulte de l'interaction entre le décodage (partiel dans le cas des mots irréguliers) et la précision du contexte d'apprentissage. Selon Pritchard et al. (2018), ces différences non contradictoires font de ces modèles des propositions complémentaires permettant de valider l'hypothèse d'auto-apprentissage soutenue par Share (1995, 1999, 2004).

Cependant, malgré les conclusions faites par les auteurs, plusieurs points nous paraissent discutables. Un premier point de discussion concerne la définition de l'apprentissage orthographique. A travers ces modèles, l'apprentissage orthographique résulte de la création d'une connexion entre une représentation orthographique et sa représentation phonologique correspondante (et la représentation sémantique dans le cas du modèle ST-DRC). Dans le modèle de Ziegler et al. (2014) – et en l'absence de contexte dans le modèle ST-DRC –, seul un décodage réussi permet l'activation de la représentation phonologique correspondant à la séquence des lettres présentées en entrée.

De cette définition découlent deux propriétés fortes. Premièrement, cela implique que pour être mémorisée, la représentation phonologique du mot lu (c'est-à-dire, présenté à l'écrit) doit être connue. Ces modèles sont donc incapables de mémoriser la trace orthographique d'un nouveau mot jamais traité oralement, ce qui correspond à la situation générale, dans laquelle le lecteur débutant connaît les mots sous leur forme orale avant d'en voir la forme écrite. Cependant, des données comportementales montrent que des lecteurs débutants (pour quelques exemples, voir Bosse et al., 2015; Nation et al., 2007; Share, 1999, 2004; Tucker et al., 2016) et experts sont capables de mémoriser correctement la représentation orthographique de pseudo-mots : ces deux modèles sont en l'état incapables de rendre compte de cette situation.

Deuxièmement, cela induit qu'un lecteur, débutant ou expert, mémorise l'orthographe d'un mot si et seulement si ce dernier est décodé correctement, suggérant un rôle majeur du traitement phonologique. En accord avec cette hypothèse, les simulations effectuées dans le modèle d'apprentissage (Ziegler et al., 2014) montrent la quasi-absence d'apprentissage orthographique lorsque le modèle fait face à un déficit phonologique majeur. Mais ces observations ne sont pas en accord avec les données comportementales observées chez l'humain ou chez l'animal. Si les traitements phonologiques sont à la base de l'apprentissage orthographique, on devrait s'attendre à ce que les enfants dyslexiques qui présentent un trouble phonologique sévère aient des connaissances orthographiques lexicales très limitées. Même si un trouble phonologique peut effectivement s'accompagner de troubles de l'orthographe lexicale (Zoubrinetzky, Bielle, & Valdois, 2014), les formes prototypiques de dyslexies phonologiques contredisent cette prédiction. En effet, les formes pures de dyslexie phonologique se caractérisent par une lecture déficitaire des mots nouveaux mais une lecture préservée des mots, réguliers et irréguliers (voir Colé and Valdois (2007) pour une revue de questions). Ces dyslexiques disposent donc de connaissances lexicales leur permettant de reconnaître les mots aussi efficacement que des sujets non dyslexiques. On trouve d'ailleurs des dissociations similaires en production écrite. Chez ces sujets, seule la dictée de pseudo-mots est déficitaire (Campbell & Butterworth, 1985). Il est donc possible de développer des connaissances lexicales normales malgré un trouble phonologique majeur (Howard, 1996).

Les résultats obtenus chez l'animal appuient ces observations. En effet, Grainger, Dufau, Montant, Ziegler, and Fagot (2012) et Scarf et al. (2016) ont réalisé une étude portant sur les compétences en décision lexicale orthographique respectivement chez le singe et chez le pigeon. Dans ces deux études, le taux de bonnes réponses obtenu au cours d'une tâche de discrimination (équivalente à la décision lexicale chez l'humain) a permis aux auteurs de conclure que ces animaux, pourtant dépourvus de capacités phonologiques (supposées spécifiquement « humaines »), sont capables d'apprendre l'orthographe des mots. En effet, en seulement quelques jours, les singes ont appris en moyenne 139 mots (sur les 500 présentés) et les pigeons, après 8 mois d'entraînement, arrivent à discriminer en moyenne 43 mots sur les 500 présentés. Il semble donc qu'un apprentissage orthographique purement visuel soit possible indépendamment de toute capacité de traitement phonologique.

Enfin, l'indépendance relative entre les capacités phonologiques et l'apprentissage orthographique a également été observée chez l'enfant normo-lecteur. De récentes études montrent non seulement qu'un décodage réussi ne garantit pas l'apprentissage orthographique mais qu'une mémorisation correcte peut être observée malgré un décodage erroné (Castles & Nation, 2006, 2008; Nation et al., 2007; Tucker et al., 2016) laissant supposer qu'un apprentissage purement orthographique – c'est-à-dire, sans influence phonologique – est possible.

Si le décodage phonologique permet de mémoriser la représentation orthographique d'un grand nombre de mots réguliers, il paraît évident que l'association lettre-son ne peut suffire à lier la représentation orthographique d'un mot irrégulier avec sa forme phonologique. Pour

cette raison, les résultats simulés reportés par Ziegler et al. (2014) et Pritchard et al. (2018) montrent un grand nombre d'erreurs d'apprentissage issues d'un décodage incorrect dû à la présence d'irrégularités graphémiques. Selon Ziegler et al. (2014), ces erreurs justifient la nécessité de multiplier et diversifier les contextes de lecture, pour permettre la reconnaissance de mots malgré leur irrégularité. Pour pallier les erreurs de décodage, Ziegler et al. (2014) proposent dans un premier temps d'augmenter les représentations phonologiques activées grâce au décodage. Cette solution ne semble pas cohérente avec leur conclusion puisque l'apprentissage en contexte devrait aboutir à une réduction du nombre de mots activés. Pour cette raison, Perry et al. (2019) ajoutent un mécanisme d'apprentissage par instruction directe, déclenché en cas d'absence de lexicalisation du mot traité par décodage phonologique. Pritchard et al. (2018) font l'hypothèse, plus plausible, que l'activation d'un mot résulte à la fois du décodage et d'informations fournies par le contexte. Le modèle ST-DRC intègre donc un module sémantique associant la représentation sémantique des mots à leur trace phonologique. Les résultats obtenus par Pritchard et al. (2018) sont en accord avec les données comportementales montrant 1/ que l'orthographe de nouveaux mots réguliers présentés isolément (c'est-à-dire sans contexte) peut être mémorisée efficacement (Bosse et al., 2015; Nation et al., 2007; Share, 1999) et 2/ qu'un contexte de lecture précis a un effet bénéfique uniquement sur l'apprentissage de mots irréguliers (Wang et al., 2011). Cependant, si le contexte sémantique peut contribuer à la mémorisation orthographique d'un mot irrégulier, cette condition ne semble pas suffisante. En effet, les données neuropsychologiques montrent, chez les dyslexiques de surface, des erreurs en lecture et orthographe de mots irréguliers malgré de très bonnes compétences sémantiques. Chez un cas prototypique, Castles, Coltheart, Savage, Bates, and Reid (1996) concluent à une bonne connaissance sémantique des mots irréguliers que le sujet dyslexique ne pouvait pas lire correctement. Globalement, à notre connaissance, aucune étude ne montre un déficit sémantique chez les enfants qui présentent une dyslexie de surface et donc, un déficit spécifique en lecture et orthographe de mots irréguliers.

Un deuxième point de discussion concerne le rôle des traitements visuels sur l'apprentissage orthographique. Ziegler et al. (2014) et Pritchard et al. (2018) adoptent un point de vue très différent sur cette question. En effet, alors que Pritchard et al. (2018) reconnaissent que plusieurs mécanismes cognitifs impliqués dans l'apprentissage orthographique ne sont pas implémentés dans le modèle ST-DRC, incluant les dimensions visuo-attentionnelles, les conclusions de Ziegler et al. (2014) suggèrent un rôle mineur des traitements visuels dans la mémorisation orthographique.

De notre point de vue, ces analyses ne sont pas en accord avec les données comportementales. Un rôle privilégié des traitements phonologiques dans l'apprentissage orthographique devrait garantir l'acquisition de bonnes connaissances lexicales chez les enfants ayant des formes de dyslexies sans trouble phonologique associé (c'est-à-dire, des dyslexies de surface). Or, ces enfants présentent un trouble majeur de la lecture des mots irréguliers et de l'orthographe lexicale (par exemple, « aquarium » écrit « acoiriome ») (Inserm, 2007; Valdois et al., 2003). Il est donc possible d'avoir des difficultés sévères à acquérir la forme orthographique des mots malgré de bonnes aptitudes phonologiques. Les études portant sur ces formes particulières de dyslexies ont montré que le déficit orthographique chez ces enfants dyslexiques était associé à des difficultés de traitement simultané de plusieurs éléments visuels, donc à un déficit de l'empan visuo-attentionnel (empan VA ; Bosse, 2005; Dubois et al., 2010; Valdois et al., 2003). Ce déficit parait donc concerner les traitements visuels et non phonologiques.

Par conséquent, les simulations de déficits visuels par le modèle de Ziegler et al. (2014), qui montrent un effet relativement mineur de ces déficits sur l'apprentissage, pourraient être la simple conséquence de la sous-spécification des traitements visuels dans le modèle. Par ailleurs,

la simulation d'un déficit visuel proposée par Ziegler et al. (2014) et plus récemment dans Perry et al. (2019) semble discutable. En effet, le déficit visuel est associé à un déficit d'encodage positionnel ; les lettres du mot présenté au modèle lors de l'apprentissage sont identiques à celles du nouveau mot mais l'ordre des lettres est modifié (transposition de lettres adjacentes, exemple : *CAT-ACT*). Or, les théories visuelles de la dyslexie suggèrent soit des déficits de bas niveau dans le cadre de la théorie magnocellulaire (Stein, 2014), soit un déficit de l'orientation spatiale (Facoetti & Molteni, 2001) soit un déficit du traitement simultané de plusieurs éléments ou trouble de l'empan VA (Bosse, Tainturier, & Valdois, 2007). Le trouble simulé ne correspond à aucune de ces théories.

En d'autres termes, la sous-spécification des niveaux de traitement visuel – commune aux modèles de Ziegler et al. (2014) et Pritchard et al. (2018) – s'accompagne d'un traitement simultané et de la reconnaissance parfaite, mais non réaliste, de la position et de l'identité des lettres. De fait, les limites du traitement parallèle des lettres dues à l'acuité visuelle, au phénomène d'interférence (ou *crowding*) et à l'attention visuelle ne sont pas prises en compte. Cela a deux conséquences majeures. Premièrement, et comme nous venons de le discuter, les troubles spécifiques de l'apprentissage orthographique liés à un déficit des capacités visuo-attentionnelles observés dans la dyslexie de surface ne peuvent pas être simulés par ce modèle. Deuxièmement, dès qu'elle est mémorisée, la trace orthographique ne contient aucune erreur, quelle que soit sa longueur ou sa structure orthographique (nombre de graphèmes ambigus). Cela ne signifie pas pour autant qu'un apprentissage est réussi dès la première occurrence. En effet, les résultats reportés par Ziegler et al. (2014) montrent une augmentation progressive du nombre de mots appris en fonction du nombre d'items présentés au modèle. Ainsi, pour qu'environ 80% des mots soient appris, le modèle doit être exposé à plus de 300 000 items. Ces résultats suggèrent que le développement progressif du réseau TLA va permettre un décodage réussi d'un plus grand nombre de mots, illustrant les difficultés d'apprentissage du modèle lorsque le décodage phonologique est compromis. En effet, l'accroissement des performances du traitement phonologique permettrait soit de corriger des erreurs d'apprentissage induites par un décodage phonologique erroné, soit de déclencher l'apprentissage si les décodages précédents n'ont pas permis l'activation d'une représentation phonologique. Bien que de nombreuses études ont montré qu'un petit nombre de présentations (entre 1 et 4) suffit à un bon apprentissage orthographique des mots (Nation et al., 2007; Share, 2004), l'acquisition parfaite *one-shot* – c'est-à-dire dès la première présentation – de l'information orthographique supposée à travers le modèle ST-DRC ne semble pas plausible. En effet, les données oculométriques reportées dans de récentes études (H. Joseph & Nation, 2018; H. Joseph et al., 2014) montrent une évolution du traitement visuo-orthographique au cours des présentations suggérant une acquisition graduelle de la trace orthographique lexicale d'un nouveau mot.

Pour conclure, nous conservons de cette analyse deux hypothèses pour l'élaboration du modèle *BRAID-Learn*. La première est qu'un apprentissage orthographique de nouveaux mots doit être possible, même en l'absence de connaissances phonologiques, donc y compris dans un modèle purement visuel tel que le modèle BRAID. La seconde concerne la dynamique et la temporalité de cet apprentissage : nous faisons l'hypothèse que l'apprentissage doit être graduel, et aboutir à une représentation orthographique stable en un petit nombre d'exposition au nouveau mot (moins d'une dizaine). Les hypothèses présentées ici, notamment la première, n'ont pour pas pour objectif de nier l'impact sur l'apprentissage orthographique des capacités de recodage phonologique ou de remettre en question l'effet positif de connaissances phonologiques préalables. Notre objectif est plutôt d'affirmer l'importance des traitements visuels et visuo-attentionnels dans le traitement de la forme écrite des mots et de mieux appréhender la dynamique d'acquisition de leur forme orthographique.

2 Les modèles de contrôle des mouvements oculaires en lecture de texte

De nombreuses études portant sur l'analyse des mouvements oculaires générés au cours de la lecture de texte ont conduit à proposer divers modèles computationnels de contrôle des mouvements oculaires. Si les mouvements oculaires générés lors de la lecture de nouveaux mots nous informent sur les mécanismes en jeu lors de l'apprentissage orthographique alors les mécanismes implémentés dans les modèles computationnels de contrôle des mouvements oculaires peuvent apporter des informations utiles à la modélisation de l'apprentissage orthographique de nouveaux mots.

Cette section n'a pas pour objectif de présenter de façon exhaustive les modèles de contrôle des mouvements oculaires en lecture de texte. A la place, nous retenons les deux modèles les plus influents du domaine : le modèle E-Z Reader (Pollatsek, Reichle, & Rayner, 2006; Reichle, Pollatsek, Fisher, & Rayner, 1998; Reichle, Rayner, & Pollatsek, 2003; Reichle et al., 2009) et le modèle SWIFT (pour *Saccade-generation With Inhibition by Foveal Targets*, Engbert, Longtin, & Kliegl, 2002; Engbert et al., 2005).

Ces modèles ont été développés dans le but de décrire et expliquer les processus cognitifs impliqués dans la génération des mouvements oculaires observés au cours de la lecture de texte. Ils ne permettent donc pas de simuler l'apprentissage orthographique de nouveaux mots. Ils font l'hypothèse que les mouvements oculaires – et leurs caractéristiques – sont étroitement liés à l'identification des mots. Pour atteindre cet objectif, ils postulent un rôle majeur de l'attention visuelle et des propriétés lexicales du mot pour contrôler *où* et *quand* le lecteur déplace son regard. Étonnamment, ces modèles ont été développés indépendamment des modèles de reconnaissance de mots et ne représentent pas en détail les processus cognitifs associés à l'identification des mots (Reichle et al., 2003). En effet, ces modèles sont spécialisés dans la description des mécanismes permettant de rendre compte des déplacements oculaires pour parcourir les mots d'un texte.

Pour chaque modèle, nous présentons succinctement son fonctionnement ainsi que son intérêt vis-à-vis de l'apprentissage orthographique, en attachant une importance particulière à la description des processus les plus pertinents pour ce travail de thèse.

2.1 Le modèle E-Z Reader (Reichle et al., 2009)

Une première version du modèle E-Z Reader est proposée par Reichle et al. (1998). Dans cette section, nous décrivons le fonctionnement de la dixième et, à notre connaissance, dernière version du modèle, E-Z Reader 10 (Reichle et al., 2009).

Structure et fonctionnement du modèle Inspiré du modèle qualitatif de Morrison (1984), le modèle E-Z Reader repose sur un principe simple (Pollatsek et al., 2006) : lorsqu'un mot est identifié, l'attention visuelle se déplace sur le mot suivant générant la programmation d'un mouvement oculaire et la réalisation d'une saccade.

Le modèle E-Z Reader fait deux hypothèses principales. Premièrement, l'attention est distribuée sur un seul mot à la fois (Figure 2.7) reflétant le caractère sériel du traitement visuel. On parle alors de modèle SAS (pour *Serial Attention Shift*). Deuxièmement, les déplacements attentionnel et oculaire sont découplés, c'est-à-dire, qu'ils sont effectués à des moments différents du traitement lexical.

La structure du modèle repose sur l'implémentation de cinq systèmes cognitifs et oculomoteurs majeurs : le traitement visuel, l'identification de mot, l'attention visuelle, le contrôle

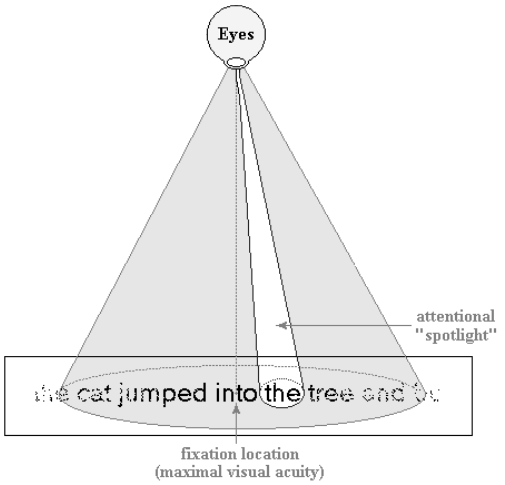

FIGURE 2.7 – Schématisation de la distribution de l'attention visuelle et de l'acuité visuelle dans le modèle E-Z Reader de Reichle et al. (2009). Le cône blanc représente l'allocation des ressources attentionnelles focalisées sur le mot « *the* ». L'acuité visuelle, représentée par le cône gris, est maximale sous la position du regard. Figure adaptée de (Reichle et al., 2003).

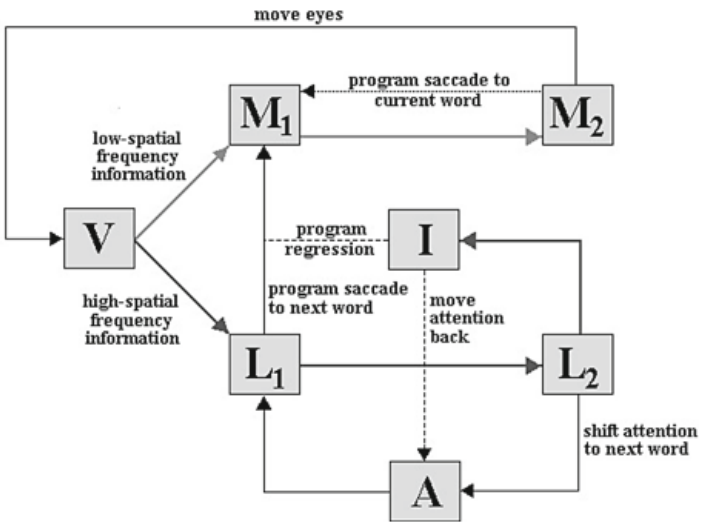

FIGURE 2.8 – Schéma du modèle E-Z Reader de Reichle et al. (2009). Les carrés représentent les différents systèmes implémentés dans le modèle : le traitement visuel (V), l'identification du mot (L_1 et L_2), l'attention visuelle (A), le contrôle oculomoteur (M_1 et M_2) et l'intégration post-lexicale (I). Les flèches épaisses représentent la circulation des informations entre les différents systèmes. Les flèches fines représentent les passages obligatoires entre les composants. Les flèches en pointillés représentent les passages probables mais non systématiques. Figure adaptée de (Reichle et al., 2009).

oculomoteur et l'intégration post-lexicale. Ces différents systèmes et leurs interactions sont représentés dans la Figure 2.8. L'intégration complète du mot (traitement visuel, identification et représentation sémantique globale) se compose de trois processus : le traitement visuel pré-attentif (ou pré-lexical, noté V), le traitement lexical (noté L) et le traitement post-lexical (noté I).

L'étape V représente le traitement de l'information visuelle précédant le traitement lexical.

Au cours de cette étape, deux types d'informations sont extraits. Premièrement, les informations basse-fréquence (par exemple, les espaces entre les mots) dont l'accès est limité par l'acuité visuelle, ralentissant la prise d'information loin du point de fixation (cône gris sur la Figure 2.7). Ces informations sont transmises au système oculomoteur et permettent de déterminer, lors de la programmation saccadique, la longueur de la saccade oculaire. Deuxièmement, les informations haute-fréquence (par exemple, les traits caractéristiques des lettres) dont l'accès est limité par l'attention visuelle (cône blanc sur la Figure 2.7). Répartie sur l'ensemble du mot quelle que soit sa longueur (répartition uniforme des ressources attentionnelles sur l'ensemble des lettres du mot), elle permet la prise d'informations visuelles sur le mot. Transmises au système pour le traitement lexical, ces informations constituent le point de départ de l'identification de mot.

L'hypothèse d'un découplage des déplacements attentionnel et oculaire se traduit par l'implémentation d'un processus de traitement lexical scindé en deux étapes : la vérification de la familiarité (notée L_1) et l'accès au lexique (noté L_2). Ces deux étapes forment ensemble le mécanisme d'identification de mot. Si l'attention joue un rôle majeur dans la récupération d'informations visuelles précoces au cours de l'étape V et de la preview parafovéale (voir paragraphe suivant **Principaux résultats**), elle semble tout aussi indispensable lors du traitement lexical. En effet, elle permet le maintien de l'information visuelle au cours du traitement lexical et se révèle donc nécessaire pour aboutir à l'identification du mot (Pollatsek et al., 2006).

La durée de l'étape L_1 est fonction de la fréquence du mot, de sa prédictibilité, de sa longueur et de la position relative de la lettre centrale du mot par rapport à la position du regard. Cela permet au modèle E-Z Reader de rendre compte des principaux effets comportementaux observés sur les temps de traitement (Pollatsek et al., 2006). En résumé, plus le mot est difficile à traiter (mot long et/ou peu fréquent et/ou peu prévisible et/ou éloigné du point de fixation), plus la durée allouée à la vérification de la familiarité du mot est élevée. Bien qu'elle ne soit pas explicitement définie comme telle, cette étape pourrait s'assimiler à la récupération des informations orthographiques et phonologiques du mot. L'achèvement de cette étape signifie que l'identification complète du mot est imminente ; le modèle débute la programmation saccadique, c'est-à-dire, la détermination de la position du prochain point de fixation (étape notée M_1 dans la Figure 2.8) aboutissant à la réalisation de la saccade, c'est-à-dire, au déplacement de l'œil (étape notée M_2 dans la Figure 2.8). Nous reviendrons sur ces étapes dans le paragraphe suivant (**Principaux résultats**).

L'étape L_2 (accès au lexique) clôt le processus de traitement lexical – et donc l'identification du mot. Elle correspond à l'accès au sens du mot, utilisant des représentations plus abstraites. Par conséquent et contrairement à l'étape L_1, elle n'implique pas de traitement visuel. Sa durée n'est fonction ni de la longueur du mot ni de l'acuité visuelle et demeure inférieure à la durée de l'étape L_1. A la fin de cette étape, le mot est identifié entraînant le déplacement attentionnel sur le mot suivant et simultanément, le début du traitement post-lexical (respectivement, étapes A et I dans la Figure 2.8).

Le traitement post-lexical (étape I) met en jeu l'ensemble des représentations haut-niveau (par exemple, représentation sémantique du mot liée au contexte ou analyse syntaxique) permettant au lecteur de construire une représentation sémantique globale du texte associée à la compréhension de texte. Reichle et al. (2009) parlent d'intégration post-lexicale du mot. Une intégration post-lexicale trop lente ou inexistante est associée à une difficulté de compréhension. Ainsi, si l'intégration post-lexicale du mot n échoue ou survient après l'identification du mot $n+1$, le regard et l'attention sont redirigés sur le mot n.

Principaux résultats A travers cette section, nous présentons deux effets simulés par le modèle E-Z Reader (Pollatsek et al., 2006) : l'effet de débordement et l'effet de longueur des

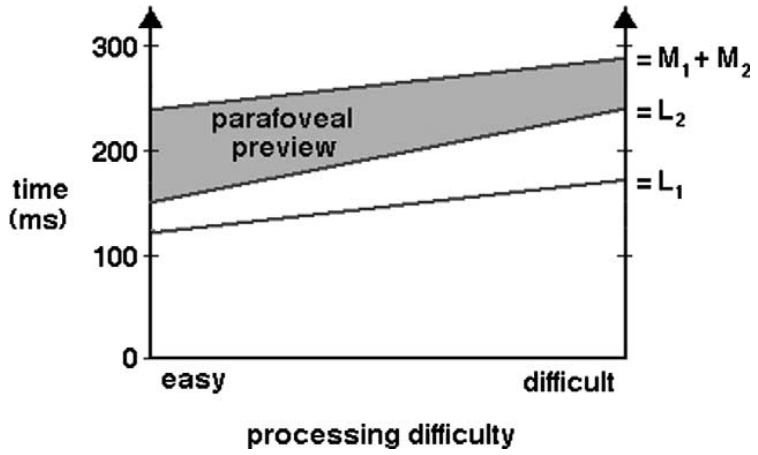

FIGURE 2.9 – Détermination du temps de traitement parafovéal dans le modèle E-Z Reader de Reichle et al. (2009). Ce graphe représente le durée allouée à la programmation/réalisation de la saccade oculaire (étapes M_1 et M_2) en fonction de la durée d'identification du mot n (étapes L_1 et L_2) et comment cette relation module le temps de traitement parafovéal sur le mot $n+1$. Figure tirée de (Pollatsek et al., 2006).

mots. Ces effets influencent le comportement oculomoteur intra-mot – nombre de fixations et temps de traitement –, apparaissant de fait comme particulièrement pertinents dans ce travail de thèse.

L'implémentation de deux étapes de traitement lexical permet de dissocier le déplacement attentionnel du déplacement du regard. D'après Pollatsek et al. (2006), ce découplage permet d'expliquer l'effet de débordement, c'est-à-dire, l'effet du temps de traitement du mot n sur le temps de traitement du mot $n+1$ (pour un exemple, voir White, Rayner, & Liversedge, 2005). Compte-tenu des valeurs de paramètres fixées dans le modèle, le déplacement de l'attention visuelle se produit toujours avant le déplacement du regard. La durée écoulée entre le déplacement attentionnel et le déplacement du regard représente le temps de traitement parafovéal (effet de *preview*). La durée du traitement lexical sur le mot n dépend des caractéristiques intrinsèques du mot, comme sa fréquence et sa prédictibilité, et des limites du traitement visuel dues à l'acuité et l'attention visuelle. Plus le mot est difficile à traiter, plus le temps alloué au traitement lexical (étapes L_1 et L_2) est élevé, induisant une durée de traitement visuel parafovéal sur le mot $n+1$ plus faible, comme illustré Figure 2.9. Au cours du traitement parafovéal, le mot $n+1$ est sous le focus attentionnel. Par conséquent, plus la durée de la *preview* est élevée, plus la quantité d'informations visuelles « précoces » est élevée, réduisant la durée nécessaire à l'identification du mot $n+1$ (après déplacement du regard).

Enfin, les résultats reportés par Pollatsek et al. (2006) indiquent que le modèle simule avec succès l'effet de longueur des mots sur le nombre de fixations – et donc sur les temps de traitement tels que la durée du regard et la durée de première fixation. Observé expérimentalement, cet effet associe un nombre de fixations et des temps de traitement d'autant plus élevés que le mot est long (pour un exemple, voir Lowell & Morris, 2014). Selon Pollatsek et al. (2006), cet effet résulte de la correction d'une erreur de programmation saccadique. Lors de l'étape M_1, le modèle détermine la prochaine position du point de fixation. Plus précisément, considérant que le modèle traite le mot n, cette étape permet de définir la longueur de la saccade à réaliser pour traiter le mot $n+1$. Toutes les saccades sont dirigées vers du centre du mot suivant. En raison des limites du traitement visuel dues à l'acuité visuelle, cette position permet

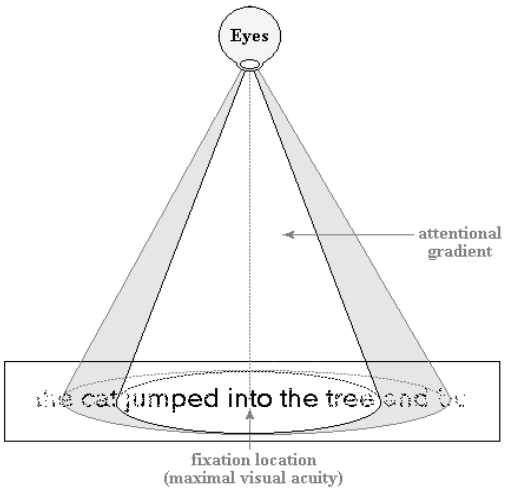

FIGURE 2.10 – Schématisation de la distribution de l'attention visuelle et de l'acuité visuelle dans le modèle SWIFT de Engbert et al. (2005). Le cône blanc représente la distribution parallèle des ressources attentionnelles autour de la position du regard – dont l'asymétrie n'est pas représentée sur la figure. L'acuité visuelle (représentée par le cône gris) et l'attention visuelle sont maximales sous la position du regard. Figure adaptée de (Reichle et al., 2003).

un traitement lexical optimal, c'est-à-dire, sur l'ensemble des lettres du mot. Les informations visuelles basse-fréquence transmises au système oculomoteur permettent de prédire la position du centre du prochain mot. Modulée par un calcul d'erreur, cette dernière définit la longueur de la saccade à réaliser. Dans le modèle E-Z Reader, la probabilité de générer une refixation (intra-mot) est fonction de la différence entre le centre du mot et la position réelle du regard. Cette différence augmentant avec la longueur du mot, la probabilité de refixation est d'autant plus grande que le mot est plus long, générant ainsi une augmentation du temps de traitement. En conclusion, dans E-Z Reader, l'effet de longueur est généré par la correction d'une erreur de programmation saccadique elle-même induite par les limites du traitement de l'information visuelle basse-fréquence, c'est-à-dire, l'acuité visuelle.

2.2 Le modèle SWIFT (Engbert et al., 2005)

Le modèle SWIFT (Engbert et al., 2002, 2005) se distingue principalement du modèle E-Z Reader sur deux aspects : l'allocation de l'attention visuelle et les processus impliqués dans la programmation saccadique. Dans cette section, nous résumons le fonctionnement du modèle et les processus qu'il suppose impliqués dans le contrôle des mouvements oculaires.

Structure et fonctionnement du modèle Le modèle SWIFT postule deux hypothèses principales. Premièrement, l'attention visuelle est distribuée parallèlement sur plusieurs mots de manière graduelle (Figure 2.10). On parle alors de modèle GAG (pour *Gradient by Attention Guidance*). Deuxièmement, les processus calculant *où* et *quand* déplacer son regard sont découplés, impliquant l'implémentation de deux mécanismes distincts pour modéliser le contrôle spatial (*où*) et temporel (*quand*) de la programmation saccadique.

La structure du modèle est représentée Figure 2.11. Elle est composée de deux sous-modèles principaux : le sous-modèle d'activations lexicales qui va permettre, via deux voies distinctes, d'activer le sous-modèle de programmation saccadique qui détermine *où* et *quand* déplacer le

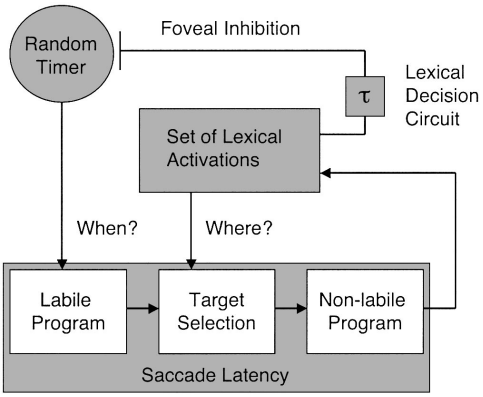

FIGURE 2.11 – Schéma du modèle SWIFT de Engbert et al. (2005). Les sous-modèles d'activations lexicales et de programmation saccadique sont représentés. Le bloc « *Labile Program* » débute la programmation saccadique, le bloc « *Target Selection* » détermine quel sera le mot fixé après la saccade et le bloc « *Non-labile Program* » réalise la saccade oculaire. Les flèches représentent la circulation des informations entre les différentes composantes du modèle. Figure tirée de (Engbert et al., 2005).

regard, et le sous-modèle de latence saccadique qui permet la programmation et la réalisation de la saccade (programme moteur).

Dans le modèle SWIFT, la position du regard (fovéa) et le point de focalisation attentionnelle coïncident et leurs déplacements sont couplés (contrairement au modèle E-Z Reader). La distribution de l'attention est asymétrique : considérant le mot n fixé, les mots $n-1$, $n+1$ et $n+2$ seront situés sous le gradient attentionnel (cône blanc sur la Figure 2.10). Ces mots constituent le champ d'activation.

Au cours du traitement lexical, tous les mots du champ d'activation font l'objet d'une analyse lexicale. Ces mots sont donc activés en parallèle. Le mécanisme d'identification de mot (traitement lexical) se décompose en deux étapes. C'est un processus dynamique : l'état du système évolue au cours du temps. Au cours de la première étape, appelée pré-traitement, l'activation lexicale augmente progressivement au cours du temps jusqu'à atteindre sa valeur maximale. Cette dernière, calculée pour chacun des mots du champ d'activation, est fonction de la fréquence et de la prédictibilité du mot. Au cours de la seconde étape, appelée achèvement lexical, l'activation diminue, jusqu'à atteindre zéro. Cette étape modélise le déclin mémoriel. A la fin de cette étape, le mot est supposé parfaitement identifié.

Les vitesses d'accumulation d'information et de déclin mémoriel sont modulées par deux facteurs : la quantité d'attention visuelle allouée au mot et sa prédictibilité. Modélisée mathématiquement par une distribution gaussienne asymétrique (implémentée par deux demi-gaussiennes de variances différentes), la quantité d'attention est maximale sous le point de focalisation attentionnelle. Elle diminue avec la distance entre le mot et la fovéa (excentricité). Par conséquent, la vitesse d'accumulation d'information lexicale est plus élevée pour le mot situé sous la fovéa. Finalement, l'asymétrie de la distribution permet de favoriser les mots situés à droite du mot fixé. L'impact de la prédicitibilité du mot sur les temps de traitement dépend de l'étape. Lors du pré-traitement, plus le mot est prédictible, plus le temps de traitement sera élevé, augmentant la probabilité de sauter les mots hautement prédictibles. Cependant, au cours de l'achèvement lexical, plus le mot est prédictible, plus la quantité attentionnelle allouée sera élevée, générant une reconnaissance plus rapide pour les mots hautement prédictibles.

Intéressons-nous maintenant aux commandes oculomotrices. Considérons le mot n traité en

fovéa. Le *quand* représente le moment auquel la programmation saccadique débute (programme noté « *Labile Program* » dans la Figure 2.11). Dans le modèle SWIFT, la programmation saccadique débute après un intervalle de temps aléatoire (noté « random timer » dans la Figure 2.11). Afin de tenir compte de la difficulté du traitement lexical du mot n, le déclenchement de la programmation saccadique est retardé par l'activation du mot, représentant l'influence lexicale top-down du traitement lexical sur la durée de fixation. Modélisé par le circuit de décision lexicale τ (Figure 2.11), le retard est d'autant plus élevé que le mot est difficile à traiter. Les auteurs parlent d'inhibition fovéale.

Le « *où* » représente la position du regard – et donc du point de focalisation attentionnelle – après réalisation de la saccade. Il est opérationnalisé par la probabilité que le mot k (autre que n mais appartenant au champ d'activation) soit, à l'instant t, sélectionné par le système comme mot cible. Mathématiquement, cette probabilité est égale à l'activation relative du mot k, à l'instant t. Une fois que le mot cible est sélectionné, le programme moteur (noté « *Non-labile Program* » dans la Figure 2.11) déclenche la saccade oculaire et donc le déplacement simultané du regard et de l'attention visuelle.

Principaux résultats Le modèle SWIFT simule de nombreux effets classiques observés dans la littérature tel que l'effet de longueur des mots sur le nombre de fixations intra-mot ou encore l'effet de fréquence sur la durée de première fixation.

Il nous parait particulièrement intéressant de présenter deux effets produits par le modèle SWIFT : l'effet de la position optimale du regard (noté *OVP* pour *Optimal Viewing Position*) et l'effet inversé de la position optimale du regard (noté *IOVP* pour *Inverted Optimal Viewing Position*). Ces effets, jamais simulés à notre connaissance, sont induits par deux mécanismes spécifiques et indépendants, respectivement, l'attention visuelle et la correction d'erreur saccadique. L'analyse de ces effets nous apparaît pertinente dans ce travail de thèse tant elle permet de comprendre comment le comportement oculomoteur intra-mot – nombre de fixations et durée de première fixation – est influencé, dans le modèle SWIFT, par l'attention visuelle et la longueur des mots.

L'*OVP* est définie comme la position du regard (c'est-à-dire, de la fovéa) permettant la reconnaissance d'un mot isolé la plus rapide possible (pour un exemple, voir O'Regan & Jacobs, 1992). L'OVP se situe légèrement à la gauche du milieu d'un mot. Ainsi, plus la position de la fovéa est éloignée de l'*OVP*, moins les mots sont reconnus. On parle alors d'effet de l'*OVP*. Dans le modèle SWIFT, cet effet est induit par le mécanisme d'attention visuelle implémenté. L'attention est conceptualisée par une distribution gaussienne asymétrique. Ainsi, la quantité d'attention décroît avec l'excentricité et l'asymétrie de la distribution permet une allocation attentionnelle plus importante pour les lettres situées en début de mot. Finalement, cette forme mathématique choisie pour l'attention visuelle induit donc directement l'effet de l'*OVP*. Ce résultat est donc induit par les propriétés du modèle, et non pas par analyse des régularités observées dans des simulations.

Comme nous l'avons vu, Reichle et al. (2009) simulent l'effet de longueur grâce à la correction de la position du regard lors de la première fixation, générant ainsi une refixation et donc un temps de traitement plus élevé. Dans le modèle SWIFT, le processus permettant de générer des refixations est identique à celui implémenté dans le modèle E-Z Reader. En se basant sur des données neurophysiologiques qui montrent que la détection d'erreur a lieu pendant la réalisation de la saccade, Engbert et al. (2005) font l'hypothèse que pour corriger une erreur de position de première fixation sur le mot n, la programmation saccadique s'active immédiatement après la fin de la saccade en provenance du mot $n-1$. Autrement dit, si la position de la fovéa (et donc du point de focalisation attentionnelle) ne coïncide pas avec la position optimale du regard sur

le mot n (c'est-à-dire, légèrement à gauche du centre du mot), une nouvelle programmation saccadique débute. Cela entraîne une diminution de la durée de première fixation sur le mot n. Pour Engbert et al. (2005), ce mécanisme permet d'expliquer l'effet d'*IOVP* observé expérimentalement (Vitu, McConkie, Kerr, & O'Regan, 2001). Dans cette étude, Vitu et al. (2001) observent que, lors de la lecture de texte, lorsque la position de la fixation est proche des limites d'un mot (début ou fin), la durée de première fixation (que la fixation soit unique ou non) est moins élevée que celle obtenue lorsque la fixation est centrée sur le mot. Les résultats simulés reportés par Engbert et al. (2005) montrent que le modèle SWIFT reproduit avec succès l'effet d'*IOVP*. Pour Engbert et al. (2005), ces résultats suggèrent que la correction d'erreur de programmation saccadique dans le choix de la position de première fixation est responsable de cet effet.

2.3 Discussion

Les modèles de contrôle des mouvements oculaires de Reichle et al. (2009) et Engbert et al. (2005) ont pour objectif de simuler le comportement oculomoteur au cours de la lecture de texte. Ces modèles simulent avec succès de nombreux effets observés comportementalement tels que la probabilité de refixation, la probabilité qu'un mot ne soit pas fixé (saut de mot), les effets de prédictibilité, de longueur du mot sur les durées de fixation ou encore l'*OVP* et l'*IOVP*. Nous présentons dans cette section quelques points de discussion qui nous paraissent particulièrement intéressants pour ce travail de thèse.

L'implémentation d'un mécanisme d'attention visuelle constitue le point fort de ces modèles. En effet, les modèles E-Z Reader et SWIFT postulent un rôle majeur de l'attention visuelle dans la simulation des mouvements oculaires pendant la lecture de texte. Bien que les effets induits par l'attention visuelle soient relativement similaires pour les deux modèles, son implémentation et la manière dont elle est déplacée constituent les points principaux de divergence entre ces modèles. Nous ne débattrons pas ici de la meilleure implémentation de l'attention visuelle préférant résumer son rôle au sein de chaque modèle.

Dans le modèle E-Z Reader, Reichle et al. (2009) postulent que l'attention visuelle n'est allouée qu'à un seul mot à la fois, et qu'une quantité d'attention identique est allouée à chacune des lettres du mot (distribution uniforme). Par ailleurs, le déplacement du point de focalisation attentionnelle et du regard sont découplés ; le déplacement attentionnel a lieu à la fin du traitement lexical (étape L_2) alors que le déplacement du regard, plus tardif, s'opère à la fin de la programmation saccadique (étape M_2). Cette hypothèse permet de rendre compte des effets de preview (traitement parafovéal) observés expérimentalement et par extension, de la probabilité qu'un mot hautement prédicitible ne soit pas fixé. Dans le modèle SWIFT, l'attention visuelle est distribuée sur plusieurs mots (quatre) à la fois. Modélisée par une distribution gaussienne asymétrique, elle permet, associée au gradient d'acuité visuelle, une répartition des ressources attentionnelles sur chaque lettre – et par extension sur chaque mot – fonction de la distance au point de focalisation attentionnelle. Indirectement, cette implémentation permet de rendre compte de la probabilité plus élevée de ne pas fixer un mot hautement prédictible, assimilée à du pré-traitement parafavoéal par Engbert et al. (2005). Enfin, dans le modèle SWIFT, le déplacement de l'attention et du regard coïncident. Pour conclure, les principaux effets observés expérimentalement sur les mouvements oculaires en lecture de texte – temps de traitement, saut de mot, effet de longueur, effet de preview – sont expliqués dans ces modèles par les propriétés du modèle d'attention visuelle.

Deux points nous semblent cependant discutables. Premièrement, ces modèles, dédiés à la simulation du comportement oculomoteur au cours de la lecture de texte, mettent l'emphase

sur les mouvements oculaires inter-mots. En effet et bien qu'ils tiennent compte de propriétés intrinsèques des mots telles que leur fréquence, leur longueur et leur prédictibilité, ils n'implémentent pas de mécanisme de reconnaissance de mots permettant de rendre compte d'effets liés à la structure orthographique du mot à traiter tels que, par exemple, l'effet du voisinage sur l'identification du mot. Deuxièmement, certains des effets simulés semblent être obtenus grâce à l'apport de programmes spécifiques, ajoutés aux modèles dans le but de rendre compte des données comportementales non simulées par les mécanismes principaux. Ainsi, par exemple, dans le modèle E-Z Reader, un programme de génération de refixations (Reichle et al., 1998) est implémenté, de manière indépendante du mécanisme attentionnel, permettant d'expliquer les saccades intra-mots. Des fixations additionnelles (troisième, quatrième, etc.) sont également possibles, mais sans que les durées et la structure spatiale de ces fixations soient évidentes. En particulier, l'ensemble des fixations sur un mot ne semble pas, dans ces modèles, pouvoir décrire un besoin d'exploration du mot pour sa reconnaissance, typique des lecteurs apprenants ou du contexte d'apprentissage orthographique.

3 Les modèles d'apprentissage et de contrôle de mouvements oculaires : Discussion générale

Dans les Sections 1.2 et 2, nous avons présenté le fonctionnement de deux modèles d'apprentissage orthographique développés récemment (Pritchard et al., 2018; Ziegler et al., 2014) ainsi que quelques-uns de leurs résultats principaux, et deux modèles classiques de contrôle de mouvements oculaires dans la lecture de texte.

Dans cette section, nous résumons, en trois parties, les raisons pour lesquelles ces modèles ne permettent pas, selon nous, de rendre compte de l'ensemble des processus impliqués dans l'apprentissage orthographique. Puis, nous proposons une réponse à chacune des limites exposées, afin de modéliser l'apprentissage orthographique, objectif principal de ce travail de thèse.

3.1 *BRAID-Learn* : un modèle hybride

Les modèles d'apprentissage orthographique sont conçus spécifiquement pour implémenter le système d'auto-apprentissage de Share (1995). Ils ne simulent pas les mouvements oculaires observés lors de la lecture d'un mot ou d'un pseudo-mot. De la même façon, les modèles de contrôle des mouvements oculaires ont pour seul objectif de simuler le comportement oculomoteur au cours de la lecture de texte. Ils ne simulent pas l'apprentissage orthographique de nouveaux mots. Ce sont des modèles dédiés. A notre connaissance, il n'existe pas à l'heure actuelle de modèle computationnel qui rende compte à la fois de l'apprentissage de l'orthographe lexicale et des mouvements oculaires en lecture de mots et de pseudo-mots.

Notre objectif est donc de développer un modèle « hybride », *BRAID-Learn*, qui rende compte de ces deux dimensions.

3.2 Et la reconnaissance de mots ?

Dans le cadre de l'hypothèse principale qu'ils implémentent (l'hypothèse d'auto-apprentissage), les modèles d'apprentissage mettent l'emphase sur les traitements phonologiques, considérant les mécanismes d'analyse visuelle comme secondaires. Les modèles proposés par Ziegler et al. (2014) et Pritchard et al. (2018) sont basés sur les modèles double-voie (respectivement, CDP et DRC) qui, eux-mêmes, n'incluent pas de sous-modèle de reconnaissance de mots pleine-

ment spécifié (Norris, 2013). De leur côté, les modèles de contrôle des mouvements oculaires sont également conçus indépendamment des avancées de la recherche dans le domaine de la reconnaissance de mots. Ils simulent principalement le comportement oculomoteur inter-mots. L'absence d'implémentation d'un mécanisme de reconnaissance de mots ne leur permet de tenir compte ni des effets liés à la similarité visuelle entre les lettres ni des influences du voisinage orthographique au cours de l'identification de mot.

Afin de pallier cette limite, le modèle *BRAID-Learn* est conçu comme une extension du modèle de reconnaissance de mots BRAID.

3.3 Et l'attention visuelle ?

Les modèles computationnels d'apprentissage orthographique et les modèles de contrôle des mouvements oculaires se distinguent fortement par la place qu'ils accordent au rôle de l'attention visuelle en lecture. Alors qu'aucun mécanisme d'attention visuelle n'est implémenté dans les premiers (à l'exception du modèle de LaBerge & Samuels, 1974), l'attention visuelle est un facteur crucial pour rendre compte des mouvements oculaires dans les seconds. Plus généralement, l'attention visuelle est cruellement absente de la plupart des modèles de lecture (à l'exception du modèle MTM, Ans et al., 1998) et des modèles de reconnaissance de mots (à l'exception du modèle MORSEL, Mozer & Behrmann, 1990, voir Phénix et al., 2016 pour une revue complète). Or, de très nombreux arguments plaident pour un rôle de l'attention visuelle en lecture et en reconnaissance de mots (Ginestet et al., 2019; Lachter, Forster, & Ruthruff, 2004).

Par ailleurs, l'implémentation d'autres phénomènes influençant le processus de reconnaissance d'un mot tels que l'acuité visuelle (C. Whitney, 2001) et l'encombrement visuel (*crowding* ; D. Whitney & Levi, 2011) est également indispensable. Si les modèles de contrôle de mouvements oculaires en lecture de texte E-Z Reader et SWIFT tiennent compte – en plus de l'attention visuelle – de l'acuité visuelle, aucun des modèles présentés dans ce chapitre n'implémente de mécanisme de *crowding*.

Plus généralement, aucun modèle de reconnaissance de mots ou de lecture existant n'implémente l'ensemble des trois mécanismes de traitement que sont l'acuité visuelle, le *crowding* et l'attention visuelle (voir Phénix, 2018; Phénix et al., 2016 pour une revue complète). Le modèle IA (pour *Interactive Activation Model* ; McClelland & Rumelhart, 1981) n'en intègre aucun, les modèles *Overlap* (Gomez, Ratcliff, & Perea, 2008) et SCM (pour *Spatial Coding Model* ; Davis, 2010) ne possèdent qu'un mécanisme de *crowding* et, les modèles MORSEL (pour *Multiple Object Recognition and attention SELection* ; Mozer & Behrmann, 1990) et MTM (pour *Multiple Trace Model* ; Ans et al., 1998), n'implémentent qu'un mécanisme d'attention visuelle. Plus récemment, Bernard and Castet (2019) proposent un modèle de reconnaissance de mots qui implémente à la fois un mécanisme de *crowding* et un gradient d'acuité visuelle.

Afin de tenir compte de ces trois facteurs limitant le traitement visuel, le modèle *BRAID-Learn* est développé à partir du modèle BRAID, premier modèle de reconnaissance de mots incluant à la fois un gradient d'acuité visuelle, un mécanisme d'interférences latérales entre lettres voisines (*crowding*) et un sous-modèle d'attention visuelle.

4 Contribution : développement d'un nouveau modèle d'apprentissage orthographique

En conséquence de l'analyse critique qui précède, notre objectif principal est le développement d'un modèle computationnel de traitement des mots et de leur apprentissage orthographique, qui puisse prédire et simuler le contrôle des mouvements oculaires (et donc les déplacements attentionnels), en prenant en compte les propriétés des traitements visuels et de l'attention visuelle. Le modèle *BRAID-Learn*, que nous définirons, a pour ambition de répondre à cet objectif.

Dans le cadre de cette thèse, nous utiliserons le modèle comme un outil pour étudier le rôle fondamental que joue l'attention visuelle à la fois pour expliquer les effets de longueur en décision lexicale (Chapitre 3), mais également pour décrire les mouvements oculaires et attentionnels dans l'apprentissage de nouveaux mots (Chapitre5), ce qui sera corroboré par des données expérimentales présentées dans le Chapitre 4.

Notre contribution au développement d'un nouveau modèle d'apprentissage orthographique se décompose en trois sous-parties.

- Dans le ***Chapitre 3***, nous utilisons le modèle BRAID dans sa version originale pour simuler les effets de longueur en décision lexicale de mots. Cette étude est triplement motivée :

    1. Les modèles de contrôle des mouvements oculaires postulent l'existence d'un module de « *familiarity check* » qui correspond à un certain niveau d'activation lexicale sans que le mot soit totalement reconnu. Cela s'apparente fortement aux mécanismes simulant la décision lexicale dans les modèles de reconnaissance de mots. Nous faisons l'hypothèse que l'apprentissage d'un mot nouveau résulte de l'adaptation du comportement oculomoteur en fonction de l'évaluation de la « non familiarité » du stimulus présenté. Le mécanisme de décision lexicale est étudié dans ce premier article en tant que processus impliqué à la fois dans l'apprentissage et dans le contrôle des mouvements oculaires.
    2. Les modèles de contrôle des mouvements oculaires proposent une simulation des effets de longueur sur les mots essentiellement grâce à une probabilité de refixation qui augmente avec la longueur du mot. Par ailleurs, les effets de longueur en lecture de mots nouveaux sont bien simulés par les modèles de lecture mais ces modèles ont beaucoup de mal à rendre compte des effets de longueur sur les mots en décision lexicale. Démontrer la capacité d'un modèle de reconnaissance de mots à rendre compte des effets de longueur sur les mots en décision lexicale est une étape nécessaire pour faire le lien avec les modèles de contrôle oculomoteur.
    3. Les modèles oculomoteurs accordent un rôle essentiel à l'attention. Le modèle BRAID est le premier modèle de reconnaissance de mots à implémenter un module attentionnel – en plus d'un gradient d'acuité et d'un mécanisme d'interférences latérales entre lettres voisines. Notre objectif sera donc plus particulièrement de montrer le rôle de ce module en décision lexicale de façon à justifier qu'il apparaisse comme le chaînon indispensable pour qu'un modèle de reconnaissance de mots puisse également rendre compte des patterns de mouvements oculaires sur les mots.

- Dans le ***Chapitre 4***, nous présentons les données d'une étude comportementale d'apprentissage orthographique incident de mots nouveaux chez de jeunes adultes normo-lecteurs. Trois objectifs principaux sont associés à cette étude :

1. Peu d'études comportementales se sont intéressées aux mouvements oculaires générés pendant la lecture répétée de nouveaux mots. A travers cette étude, nous étudierons l'effet de répétition en lecture de nouveaux mots chez le lecteur adulte expert francophone. Cet effet sera évalué aussi bien sur les mesures oculométriques que sur la qualité de la mémorisation orthographique afin d'observer le lien entre variations du comportement oculomoteur et apprentissage orthographique. L'utilisation de l'oculométrie nous permettra d'observer l'apprentissage orthographique « en train de se faire », apportant de nouvelles informations sur la vitesse d'acquisition des représentations orthographiques de nouveaux mots.
2. Quelques études suggèrent un rôle de l'attention visuelle dans l'apprentissage orthographique. Par ailleurs, il a été montré que le temps de traitement d'un mot connu est influencé par les capacités visuo-attentionnelles. Le modèle *BRAID-Learn* inclut un mécanisme d'attention visuelle déterminant dans le contrôle des mouvements attentionnels générés au cours de la lecture de mots nouveaux ou connus. Notre objectif sera donc d'explorer le rôle de l'attention visuelle sur la qualité de la mémorisation orthographique et sur le comportement oculomoteur afin de justifier son implication dans le processus d'apprentissage.
3. Le développement d'un modèle d'apprentissage orthographique tel que nous le proposons avec le modèle *BRAID-Learn* nécessite l'utilisation de données comportementales pour, d'une part, calibrer le modèle et d'autre part, évaluer sa pertinence en comparant les résultats obtenus lors des simulations aux données comportementales observées. Les données recueillies à travers cette expérience comportementale serviront de référence pour l'évaluation du modèle *BRAID-Learn*.

- Dans le ***Chapitre 5***, nous présentons le modèle *BRAID-Learn*, pour répondre à deux objectifs :

  1. Notre premier objectif est de définir formellement le modèle *BRAID-Learn*. Pour cela, nous préciserons tout d'abord les hypothèses constitutives du modèle, nous présenterons sa structure, comme une extension du modèle de reconnaissance de mots BRAID, et nous décrirons la définition mathématique du modèle et la manière dont il simule l'apprentissage orthographique.
  2. Notre second objectif est d'évaluer la capacité du modèle *BRAID-Learn* à reproduire les données comportementales. Ces données, issues de l'étude présentée au Chapitre 4, seront comparées aux données simulées par le modèle. La comparaison entre les données simulées et comportementales nous permettra d'évaluer la pertinence des hypothèses implémentées dans le modèle, et de discuter les limites du modèle.

Chapitre 3

Modélisation de l'effet de longueur des mots en décision lexicale : le rôle de l'attention visuelle

Note

Ce Chapitre reprend un article publié dans *Vision Research* (Ginestet et al., 2019).

Résumé

L'effet de longueur de mots en Décision Lexicale a été mis en évidence dans de nombreuses expériences comportementales mais simuler cet effet représente un défi pour la plupart des modèles computationnels existants. Les effets de longueur dans ces modèles sont attribués au mode de traitement sériel engagé lors du décodage phonologique de la séquence à lire, ce qui permet très facilement de rendre compte d'effets de longueur en lecture de pseudo-mots mais plus difficilement d'expliquer les effets de longueur sur les mots, notamment en décision lexicale, puisque le traitement est alors spécifiquement visuo-orthographique.

Notre premier objectif dans le cadre de cet article est d'évaluer la capacité de BRAID, un nouveau modèle Bayésien de reconnaissance visuelle de mots récemment développé par des membres de notre équipe (Phénix, Valdois, & Diard, soumis), à rendre compte de l'effet de longueur observé comportementalement sur les mots en décision lexicale. BRAID est un modèle très sophistiqué dans la mesure où il intègre non seulement un gradient d'acuité visuelle et un mécanisme d'interférences latérales entre lettres adjacentes (ou *crowding*) mais également, un composant visuo-attentionnel modélisé par une distribution de probabilité Gaussienne. Le second objectif de l'étude est de mieux comprendre l'impact de l'attention visuelle sur l'effet de longueur.

Nous avons utilisé les données comportementales issues du French Lexicon Project (FLP ; Ferrand et al., 2010) comme données de référence auxquelles comparer les performances du modèle. Au total, ce sont 1 200 mots français de 4 à 11 lettres (150 par longueur) appariés en fréquence et voisinage orthographique qui ont été utilisés dans chacune des cinq simulations que nous avons réalisées. Une première série de deux simulations évalue la capacité du modèle à rendre compte des données dans deux conditions de focalisation de l'attention, suite à une focalisation unique ou lorsque deux focalisations sont possibles pour les mots les plus longs.

Les autres simulations évaluent les variations de l'effet de longueur selon que la distribution de l'attention visuelle sur les lettres du mot soit homogène ou piquée.

Le FLP nous donne accès au temps de réaction moyen recueilli auprès de plusieurs centaines de participants pour chacun des mots proposés. L'effet de longueur sur les mots en décision lexicale est estimé à 7.25 ms par lettre supplémentaire. Dans les simulations, c'est le nombre d'itérations requis pour décider que l'item est un mot qui est calculé. Des calibrations ont été effectuées au cours des recherches précédentes pour qu'une itération équivale à environ une milliseconde. Nous calculons l'augmentation moyenne du temps de traitement avec la longueur, ce qui est exprimé en nombre d'itérations par lettre supplémentaire.

Les simulations montrent qu'un effet de longueur massif (44,6 ms par lettre supplémentaire) est obtenu en condition de focalisation attentionnelle unique pour toutes les longueurs de mot. C'est seulement lorsque deux focalisations attentionnelles sont effectuées dans le cas des mots les plus longs (à partir de 7 lettres) que l'effet de longueur simulé (8 itérations par lettre) correspond aux données comportementales.

Ces résultats sont doublement intéressants et contre-intuitifs. D'une part, ils montrent qu'un effet de longueur massif est obtenu en condition de traitement parallèle, ce qui s'oppose à la conception majoritaire selon laquelle un effet de longueur est nécessairement le reflet d'un traitement sériel. D'autre part, les résultats des simulations montrent que faire une seconde focalisation attentionnelle permet une lecture « économique » en diminuant les temps de traitement des mots longs. En d'autres termes, le fait de réaliser un déplacement attentionnel (et du regard) accélère le traitement du mot. Ce second résultat s'oppose là encore au dogme selon lequel analyser un mot sériellement en plusieurs fixations et donc plusieurs déplacements de l'attention se traduirait nécessairement par un temps de traitement plus long.

Les autres simulations montrent qu'une modification de la distribution de l'attention (uniforme ou « piquée ») permet, quant à elle, de moduler l'amplitude et le sens de l'effet de longueur. c'est d'abord l'impact d'une distribution attentionnelle homogène sur la performance en décision lexicale qui est évalué. Deux conditions sont testées. Dans la première, la quantité de ressources attentionnelles reste fixe et similaire pour toutes les longueurs de mot et l'attention est répartie de façon équivalente sur toutes les lettres. Cela revient à allouer d'autant moins d'attention à chaque lettre que le mot est plus long, ce qui conduit à un effet de longueur plus fort que l'effet observé comportementalement. La seconde condition de distribution homogène consiste à allouer la même quantité d'attention visuelle à toutes les lettres du mot quelle que soit sa longueur. Cette seconde condition de répartition uniforme de l'attention permet d'inverser l'effet de longueur, les temps de réaction sont alors d'autant plus courts que le mot est plus long. Finalement, la dernière simulation évalue l'effet d'une distribution « piquée » de l'attention (c'est-à-dire, lorsque l'attention se distribue sur une lettre à la fois). Les résultats montrent un effet de longueur massif (32,6 itérations par lettre) dans cette condition.

L'ensemble des résultats montre que les effets de longueur simulés varient en fonction des paramètres d'attention visuelle (nombre de focalisations et pattern de distribution de l'attention). L'attention visuelle semble donc un composant crucial pour expliquer les effets de longueur observés en décision lexicale chez l'humain. Un rétrécissement de la distribution de l'attention pourrait être responsable des effets de longueur exagéré observé dans certaines pathologies de la lecture.

Abstract

The word length effect in Lexical Decision (LD) has been studied in many behavioral experiments but no computational models has yet simulated this effect. We use a new Bayesian model of visual word recognition, the BRAID model, that simulates expert readers' performance. BRAID integrates an attentional component modeled by a Gaussian probability distribution, a mechanism of lateral interference between adjacent letters and an acuity gradient, but no phonological component. We explored the role of visual attention on the word length effect using 1,200 French words from 4 to 11 letters. A series of five simulations was carried out to assess (a) the impact of a single attentional focus versus multiple shifts of attention on the word length effect and (b) how this effect is modulated by variations in the distribution of attention. Results show that the model successfully simulates the word length effect reported for humans in the French Lexicon Project when allowing multiple shifts of attention for longer words. The magnitude and direction of the effect can be modulated depending on the use of a uniform or narrow distribution of attention. The present study provides evidence that visual attention is critical for the recognition of single words and that a narrowing of the attention distribution might account for the exaggerated length effect reported in some reading disorders.

1 Introduction

The effect of word length on visual word processing has been examined with a variety of techniques, including the lexical decision (LD) task in which the participant has to decide as quickly and as accurately as possible if the stimulus presented on the screen is a word or not. Initial research in this field yielded inconsistent results, the effect being in turn reported as inhibitory – longer words yield longer latencies (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Hudson & Bergman, 1985; O'Regan & Jacobs, 1992) – or null (Acha & Perea, 2008; Frederiksen & Kroll, 1976; Richardson, 1976). In measuring LD reaction times (RT) for a large sample of words of different lengths, megastudies helped understand these apparent inconsistencies. Indeed, it appears that the effect of word length is not linear: reaction times are constant for words between 5 to 8 letters, but they increase with length for words longer than 8 letters (Ferrand et al., 2011, 2017, 2010; New, Ferrand, Pallier, & Brysbaert, 2006).

Simulating the length effect in LD is challenging for all classes of word processing models, namely the dual-route model, the triangle model or the multi-trace memory model. These models typically attribute the length effect to serial processing and can account for this effect in reading aloud but not in LD. According to dual route models (DRC: Coltheart et al., 2001; CDP, CDP+, Perry et al., 2007, 2010), an effect of length on naming latencies reflects serial processing within the non-lexical route, due to left-to-right grapheme parsing (Perry & Ziegler, 2002; Perry et al., 2007, 2010) or serial letter-sound mapping (Coltheart et al., 2001). Serial processing within the non-lexical route can straightforwardly account for the strong length effect typically reported in pseudo-word naming (Ferrand, 2000; Ferrand & New, 2003). The interaction between parallel processing by the lexical route and serial processing by the non-lexical route accounts for smaller length effects in word naming. Perry et al. (2007) showed that no length effect occurred when the non-lexical route of the CDP+ model was turned off, suggesting that serial processing is critical to explain length effects on words in naming. In a comparison of the length effect in English and German, Perry and Ziegler (2002) reported an effect of length on naming latency in German but not in English for words from 3 to 6 letters. They simulated this differential effect within the DRC framework by changing the balance between lexical and non-lexical processing. Decreasing the strength of the lexical route

while increasing that of the non-lexical route resulted in the observed length effect on German words, suggesting again that length effect would be inherently related to serial processing for phonological recoding. It follows that dual route models cannot account for the word length effect in LD, except by explaining it as following from serial phonological recoding, that is to say, assuming that the phonology of the word is generated in LD, as in naming, and that decision would be based on the phonological output.

The triangle model of reading (Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989), that postulates parallel processing and does not include any serial mechanism, failed to simulate any length effect in word naming (Perry et al., 2007; Seidenberg & McClelland, 1989; Seidenberg & Plaut, 1998). In an attempt to accommodate the length effect on naming latency within this framework, Plaut (1999) implemented a serial processing mechanism that was orthographic and not phonological. The network initially fixated the first letter of the input string and tried to generate the appropriate phonological output of the whole word. When unable to generate the appropriate output based on this first fixation, the model had the ability to refixate the input string. Using the number of fixations as a proxy for naming latency, an overall length effect was obtained for words. This effect was null for 3-to-4 letter words that could be accurately identified within a single fixation, but inhibitory for 4-to-6 letter words that required more than one fixation. This study again emphasizes the critical role of serial processing for an account of length effects in word naming but provides no explanation on such effects in LD. Indeed, the number of refixations reflects the degree of difficulty that the network experiences in constructing word pronunciation without postulating any orthographic word recognition system that would be critical to simulate the LD task.

While the two previous classes of models assume involvement of phonological recoding in length effects, the multitrace memory (MTM) model of reading (Ans et al., 1998) postulates that such effects follow from visual attention processing at the orthographic level. The model postulates an attentional device, the visual attentional window, that delineates the amount of orthographic information under processing. Most familiar words are processed as a whole following single visual attention focusing, but shifts of focused attention are required to process unfamiliar letter strings. As a direct consequence, simulations showed a strong length effect for pseudo-words in naming (Ans et al., 1998; Valdois et al., 2006). However, no length effect was simulated in LD for either words or pseudo-words, since decision was taken following parallel processing (i.e., single attention focusing) of the input letter-string (Valdois et al., 2006). A length effect in LD was simulated within this framework by assuming a reduced visual attentional window. The simulations carried out to account for impaired reading due to limited visual attention capacity showed significant length effect in both reading and LD and a stronger length effect for pseudo-words than for words (Juphard, Carbonnel, & Valdois, 2004). Thus, the MTM model offers an account of the length effect as deriving from visual processing at the orthographic level, constrained by attention, but failed to simulate the length effect on words exhibited by typical readers in LD.

Overall, there is a relative consensus that the length effect, when observed, would reflect some kind of serial processing. What remains controversial is whether this serial mechanism relates to phonological decoding or to orthographic processing. Some behavioral data rather support an orthographic processing interpretation. Indeed, evidence that length effect can be seen in different word recognition tasks and that a larger effect is reported in tasks of progressive demasking, that more tap into visual factors than in lexical decision or naming tasks (Ferrand et al., 2011), suggests that this effect may reflect visual-orthographic encoding processes rather than orthography-to-phonology mapping. Furthermore, strong evidence for a visual-orthographic account comes from pathological data on brain-damaged patients and

dyslexic children (Barton, Hanif, Björnström, & Hills, 2014). The reports of an increased length effect in word reading following brain-damage in letter-by-letter readers (Arguin & Bub, 2005; Rayner & Johnson, 2005) or after surgical intervention at the level of visual-orthographic brain regions in patients with preserved oral language and phonological skills (Gaillard et al., 2006) support a visual-orthographic origin of the word length effect. In the same way, dyslexic children with a visual attention span disorder (i.e., a multiletter parallel processing deficit) show an abnormally large length effect in naming and lexical decision tasks, despite preserved phonological skills (Juphard et al., 2004; Valdois et al., 2011, 2003). These findings place important constraints on recognition models in suggesting that a visual attention mechanism may contribute to the word length effect.

The main contribution of the current paper is to use an original Bayesian model of visual word recognition, called BRAID (Phénix, 2018; Phénix et al., 2018) to study and simulate the word length effect in LD. The BRAID model is a fully probabilistic model that includes a visual attention layer, an interference mechanism between adjacent letters and an acuity gradient. It was previously shown (Phénix, 2018) that it could account – as other models do – for classical effects in letter perception (e.g., word and pseudo-word superiority effects, context effects) and in word recognition and LD (e.g., frequency and neighborhood effects). We further showed that its visuo-attentional layer allowed to modulate some of these effects, accounting for more subtle effects such as letter spacing or cueing in letter perception. We also showed how its temporal nature allowed studying the dynamics of evidence accumulation about letters and words, allowing to reconcile seemingly contradictory effects of word neighborhood in LD (Phénix et al., 2018). Here, we use this model to simulate the length effect on LD latency for a large subset of words from the French Lexicon Project (FLP; Ferrand et al., 2010). We will specifically show the critical role of visual attention in the word length effect on LD latency.

The rest of this paper is structured in three main sections. In the first section, we describe the BRAID model and how LD is simulated by Bayesian inference. In the second section, we describe the stimuli used in the simulations and the method used to calibrate the simulated LD task. The third section reports five simulation experiments showing how the distribution of visual attention over the word letter string modulates word length effects in LD.

2 The BRAID model

The general structure of the BRAID model is illustrated in Figure 3.1. A full description of the model is provided elsewhere (Phénix, 2018), and beyond the scope of this paper. Instead, we propose a rapid summary of the salient points of the model, selected according to their relevance to the simulations carried out in the current study.

First, the overall architecture of the model is inspired from classical word recognition models (e.g., McClelland & Rumelhart, 1981), featuring three main representational levels. The first level, called, in our model, "the letter sensory submodel", implements low-level visual processing of letter stimuli. These are noted, considering time instant T, with variables S_1^T to S_N^T (see Figure 3.1), with N the length of the input string. Processing at this level aims at recognizing letter identity and position (Dehaene, Cohen, Sigman, & Vinckier, 2005; Grainger, Dufau, & Ziegler, 2016), which is also classical. Feature extraction is parallel and results in probability distribution over "internal" letter representations (variables I_1^T to I_N^T of Figure 3.1). The model implements several plausible components of low-level visual processing, such as an acuity gradient: information about letters decreases as distance from fixation position (variable G^T in Figure 3.1) increases. In probabilistic terms, this decrease of information is represented with increasing uncertainty in the corresponding probability distributions, of the form $P(I_n^T \mid S_n^T)$,

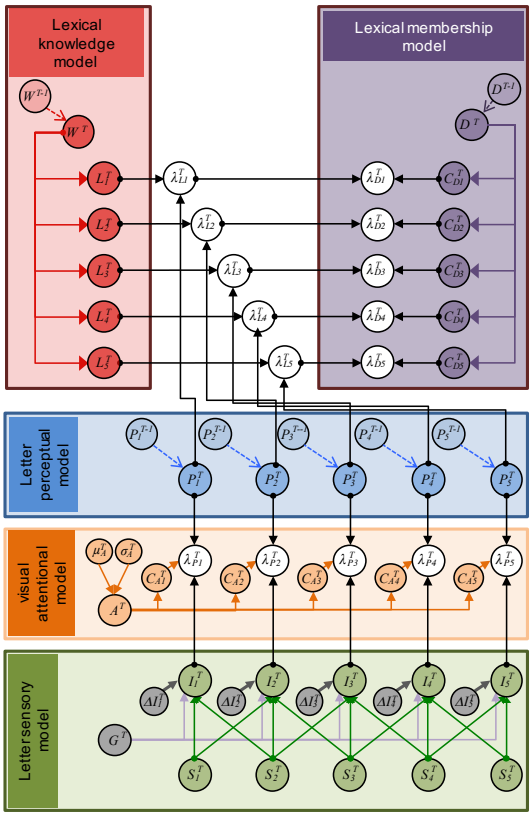

Figure 3.1 – Graphical representation of the structure of the BRAID model. Submodels are represented as colored blocks, and group together variables of the model (nodes). The dependency structure of the model is represented by arrows. This dependency structure corresponds to a 5-letter stimulus (e.g., note the 5 spatial positions of variables S_1^T to S_5^T); this structure also corresponds to time instant T (note that variables P_n^T, W^T and D^T, on which dynamical models are defined, depend on their previous iterations, P_n^{T-1}, W^{T-1} and D^{T-1}). See text for details.

which are identified from behavioral confusion matrices (Geyer, 1977). This acuity effect is symmetric around fixation (C. Whitney, 2001) and by default, we consider that gaze position G^T is located at word center. Concerning letter position identification, the model features a distributed position coding scheme (Davis, 2010; Gomez et al., 2008), such that information about a letter combines with neighboring letters. This mechanism is implemented by lateral interference between letters (represented with diagonal green edges in Figure 3.1).

The second submodel is the "letter perceptual submodel". It implements how probability distributions over variables I_1^T to I_N^T, at each time step, feed information to be accumulated into perceptual variables P_1^T to P_N^T (see Figure 3.1); this creates an internal representation of the input letter string. The third level is the "lexical knowledge submodel", implements knowledge about the spelling of 35,644 French words of the French Lexicon Project (Ferrand et al., 2010). For each word (variable W^T in Figure 3.1), its spelling is encoded as probability distributions over its letters (variables L_1^T to L_N^T in Figure 3.1). The probability to recognize a word is modulated by its frequency (prior probability distribution $P(W^0)$, not shown in Figure 3.1).

The letter perceptual and lexical knowledge submodels are linked by a layer of "coherence variables" (variables λ_{L1}^T to λ_{LN}^T in Figure 3.1), which allow, during word recognition or lexical decision, the comparison between the letter sequence currently perceived and those of known words. Coherence variables, here, can be interpreted as "Bayesian switches" (Bessière, Mazer,

Ahuactzin, & Mekhnacha, 2013; Gilet, Diard, & Bessière, 2011). Depending on their state (open/closed or unspecified), the coherence variables allow or do not allow propagation of information between adjacent submodels. In that sense, they allow to connect or disconnect portions of the model. In the BRAID model, propagation of information through variables λ_{Ln}^{T} is bidirectional between the lexical knowledge and the letter perceptual submodels.

The BRAID features a fourth submodel, which is more original: the "visual attentional submodel". It serves as an attentional filtering mechanism between the letter sensory submodel and the letter perceptual submodel. "Control variables" (variables C_1^T to C_N^T in Figure 3.1) pilot the states of coherence variables between the sensory and perceptual letter submodels (variables λ_{P1}^{T} to λ_{PN}^{T} in Figure 3.1): this allows explicitly controlling the transfer of information between these two submodels. We interpret this as an attentional model. We spatially constrain attention to be described by a probability distribution, so that sensory information cannot accumulate into perceptual variables in full, in all positions simultaneously: allocating attention to some positions is to the detriment of other positions. Mathematically, in the model, the distribution of attention over the letter string is Gaussian (variable A^T in Figure 3.1; the parameters of its distribution are its mean μ_A^T and standard-deviation σ_A^T). In this paper, we assume that the peak of attention is aligned with gaze position ($\mu_A^T = G^T$). This model affects perceptual accumulation of evidence as acuity does: the further the letter from the attention mean, the less attention it receives, the less information is transferred, and hence, the slower it is identified.

The fifth and final submodel is the "lexical membership submodel". It implements knowledge about whether a sequence of letters (variables C_{D1}^{T} to C_{DN}^{T} in Figure 3.1) corresponds (variable D^T in Figure 3.1 is *true*) or not (variable D^T is *false*) to a known word. The knowledge encoded here can be interpreted as an "error model": assuming the input string is a word, the perceived letters should match those of a known word in all positions; on the contrary, if the input is not a word, matching should fail in at least one position. This submodel is useful for simulating the lexical decision task.

2.1 Task simulation by Bayesian inference in BRAID

Variables appearing in Figure 3.1 form a state space with many dimensions. To mathematically define the BRAID model, the joint probability distribution over this state space is decomposed as a product of elementary probabilistic terms. Their definition and calibration is described in full elsewhere (Phénix, 2018). The model being defined, we then simulate tasks by computing a "question", that is to say, a probability distribution of interest. This question is solved using Bayesian inference, that is to say, applying the rules of probabilistic calculus to the model. Here, we show the questions that allow simulating letter recognition, word recognition and lexical decision. The corresponding mathematical derivations cannot be provided here in full, due to lack of space. Instead, we describe how these derivations can be interpreted, in terms of simulated processes. The lexical decision task is the task of interest here, but since it involves the two previous ones, we describe them in their nesting order, for clarity purpose.

2.1.1 Letter identification

Consider first letter identification, that is, the process of sensory evidence accumulation, from a given stimulus, to perceptual letter identity. We distinguish variables and their values by using uppercase and lowercase notations, respectively. Furthermore, we use a shorthand for denoting all positions 1 to N and all time-steps 1 to T of any variable X: $X_{1:N}^{1:T}$. As an example,

simulating that the model is given a sequence of letters as an input is setting variables $S_{1:N}^{1:T}$ to $s_{1:N}^{1:T}$.

Thus, to simulate letter identification, we set stimulus $s_{1:N}^{1:T}$, gaze position $g^{1:T}$ and attention parameters $\mu_A^{1:T}$ and $\sigma_A^{1:T}$, we allow information propagation by setting variables ${\lambda_P}_n^{1:T}$ to their "closed" state, and we compute the probability distribution, for a given time-step T and a given position n, over perceived letter P_n^T. Since variables λ_L are left unspecified, information does not propagate to lexical submodels, and is constrained in the letter sensory, visual attentional and letter perceptual submodels (see Figure 3.1).

As a result, we consider here letter identification without lexical influence. We note that a variant in which variables $\lambda_{L_{1:N}}^{1:T}$ are closed results in a bidirectional transfer of information, with both bottom-up perceptual influence of letters on words and top-down predictive influence of words on letters. This "letter identification with lexical influence" variant allows accounting for the classical word superiority effect (Phénix, 2018). We do not consider it further here; letter identification is therefore modeled by the question:

$$Q_{P_n}^T = P(P_n^T \mid s_{1:N}^{1:T}\ [{\lambda_P}_n^{1:T} = 1]\ \mu_A^{1:T}\ \sigma_A^{1:T}\ g^{1:T})\ . \tag{3.1}$$

$Q_{P_n}^T$ is a probability distribution over the perceived letter (P_n^T, at time T and position n). It is a discrete distribution, since P_n^T is a variable with 27 possible values (one for each letter plus one for missing or unknown characters).

Computing $Q_{P_n}^T$ involves two components, that are classical of inference in dynamical probabilistic models, such as Dynamic Bayesian Networks or Bayesian Filters (Bessière et al., 2013). The first component is a dynamical prediction computation, whereas the second describes sensory evidence processing. In the dynamical term, the knowledge about letters at previous time step is spread to the next time step. This involves information decay such that, if stimuli were absent, the probability distribution over letters would decay towards its initial state. In the BRAID model, this is a uniform distribution, representing lack of information: all letters are considered equally likely at the perceptual level (i.e., lexical information is restricted to be expressed at the lexical knowledge submodel).

The second component describes sensory evidence processing at current time-step and its accumulation into the dynamic state variable. Here, information is extracted from stimulus $s_{1:N}^T$, in the letter sensory submodel, to accumulate into variable P_n^T. Details are not provided here, but this feature processing involves interference effects from adjacent letters, if any, and loss of performance, due to the combined effects of acuity and attention functions, when gaze and attention are not located on the considered letter position n.

However, the output of the letter sensory submodel is modulated by the visual attention submodel before reaching perceptual variable P_n^T. More precisely, attention allocation affects the balance between information decay and sensory evidence accumulation. We note α_n the amount of attentional resources at position n (i.e., it is the probability value $P([A^T = n] \mid \mu_A^T\ \sigma_A^T)$). When α_n is high, sufficient attentional resources are allocated to position n, enough information from sensory processing accumulates into the perceptual variable so that temporal decay is counterbalanced and surpassed by sensory evidence. In this case, the probability distribution over variable P_n^T "acquires information" and gets more and more peaked at each time step. That peak, in the space of all possible letters, is always on the correct letter, provided enough attention (except for some pathological cases). In other words, letter identification, as simulated by the probability distribution $Q_{P_n}^T$, converges.

2.1.2 Word identification

Simulating word identification proceeds in a similar fashion as in isolated letter recognition above, except that information is allowed to propagate further into the model architecture, and more precisely to the lexical knowledge submodel, by setting $[\lambda_{L1:N}^{1:T} = 1]$ (see Figure 3.1). The probabilistic question is ${Q_W}^T$:

$$ {Q_W}^T = P(W^T \mid e^{1:T}\ [\lambda_{L1:N}^{1:T} = 1]\ \mu_A^{1:T}\ \sigma_A^{1:T}) \ , \tag{3.2} $$

with $e^t = s_{1:N}^t\ [\lambda_{P1:N}^t = 1]\ g^t$.

${Q_W}^T$ is a probability distribution over words of a given lexicon; at any time-step T, it encodes information about which of these words is likely to correspond to the letter sequence perceived from the stimulus. As in letter identification, the resulting Bayesian inference (not detailed here, see Phénix, 2018) involves a classical structure, combining a dynamical system simulation with perceptual evidence accumulation. First, information in the probability distribution ${Q_W}^T$ decays over time, so that, if the sensory stimuli were absent, it would converge back towards its resting and initial state. Here, this is the prior distribution over words $P(W^0)$, which encodes word frequency. Second, sensory evidence accumulation is based on the probabilistic comparison (through coherence variables λ_{L1}^T to λ_{LN}^T) between letter sequences associated to words w in lexical knowledge and the perceived identities of letters, as computed by the letter recognition question ${Q_{P_n}}^T$. This comparison, in BRAID, is influenced by the similarity between the letter sequences of the stimulus and of words of the lexicon, so that similar (neighbor) words compete with each other for recognition.

This results in a dynamical process of word recognition that depends on letter recognition. In the first few time-steps, ${Q_{P_n}}^T$ is still close to uniform, so that ${Q_W}^T$ is, too. Sensory evidence accumulation then proceeds; after some time, and even though all letters might not be perfectly identified yet (i.e., with high probability), the probability distributions over letters become diagnostic enough so that the input word is identified. The dynamics of this process are modulated, for instance, by the target word neighborhood density: recognition is faster when few perceptual evidence points toward a word with few or no competitors. Assuming that the input letter string corresponds to a known word, and with few pathological exceptions, word recognition as simulated by ${Q_W}^T$ converges toward the correct word.

2.1.3 Lexical decision

The final task we consider is lexical decision, our task of interest for the experiments that follow. It is modeled by question ${Q_D}^T$:

$$ {Q_D}^T = P(D^T \mid e^{1:T}\ [\lambda_{D1:N}^{1:T} = 1]\ \mu_A^{1:T}\ \sigma_A^{1:T}) \tag{3.3} $$

${Q_D}^T$ is a probability distribution, at each time-step, over the lexical membership variable D^T. It is a Boolean variable, that is to say, it is *true* when the stimulus is perceived to be a known word, and *false* otherwise. In question ${Q_D}^T$, the states of coherence variables allow information to propagate throughout the whole model, from the input letter-string to the lexical submodels (see Figure 3.1). Bayesian inference that derives from question ${Q_D}^T$, as before, involves a dynamical component (information decay towards a resting state) and perceptual evidence accumulation about lexical membership. However, what constitutes perceptual evidence, here, is less easily interpreted than in letter and word recognition. To explain, we consider, in turn, the two Boolean alternatives.

First, consider the hypothesis that the stimulus is a word (the $D^T = true$ case). Perceptual evidence of the LD process is the process of word recognition. Evidence about lexical membership is the probability that coherence variables λ_{L1}^T to λ_{LN}^T detect a match between perceived letters and those of a known word. In other words, lexical decision proceeds as if the lexical knowledge submodel was "observing" the probability of a match between the letter perceptual and lexical knowledge submodels. When a known word can be reliably identified from the stimulus by word recognition (or when a set of neighbor words is activated enough), then coherence variables $\lambda_{L1:N}^T$ have high probability, indicating a match, so that the probability that $D^T = true$ is high, indicating that the stimulus is recognized as being a known word.

Second, consider now the hypothesis that the stimulus is not a word (the $D^T = false$ case). Here, the expected states of coherence variables $\lambda_{L1:N}^T$ indicates that at least one letter of the stimulus should not match, when compared to known words. In other words, accumulating evidence from word recognition operates under the assumption that there would be one error in the stimulus, compared to known word forms. Technically, since it is unknown, all possible positions for the error have to be evaluated. For a given error position, word recognition ${Q_W}^T$ in computed with the input letter string in all other positions, and alternative letters in the considered position. In other words, the stimulus is probably not a word if changing at least a letter of the stimulus is required to match it to a known word. Perceptual evidence accumulation in ${Q_D}^T$ proceeds by pitting the two hypotheses against each other: lexical decision results from this competition. As in word recognition, which can recognize the input word even though its letter are not fully identified, lexical decision, in some cases, reaches high probability that the input is a known word before it is identified. The overall dynamics are further modulated by neighborhood density around the target word.

3 Method

Having described the general structure of the BRAID model and Bayesian inference involved in task simulation, we now describe the dataset that was used in the simulations and consider parameter calibration. The majority of the parameters of BRAID have been calibrated on independent data, and thus have default values that we used in previous simulations (Phénix, 2018). Here, we consider and calibrate the parameters that are specific to LD. For this purpose, we first explore the effect of these parameters on the simulated LD task to identify the parameter values that best fit the LD RTs reported for a large sample of French words in the FLP (Ferrand et al., 2010).

3.1 Material

The words used in these simulations were extracted from the FLP database (Ferrand et al., 2010). Given the great variability of the characteristics of the words of the FLP, we first chose to restrict the dataset to words from 4 to 11 letters. Then, we trimmed each list by removing words of very high frequency and those associated with high error rates. Finally, 150 words per length were randomly selected. Thus, our final word set consisted of 1,200 words from 4 to 11 letters, with a mean written frequency of 57.9 per million ranging from 20.3 to 104.2 per million and with a mean number of orthographic neighbors of 1.9 ranging from 0.4 to 5.1.

An ANOVA on mean RTs for the selected word set showed significant main effect of length ($F(7,1173) = 2.342$, $p = 0.022$) but no main effect of frequency ($F < 1, ns$) and no Length-by-Frequency interaction ($F < 1, ns$).

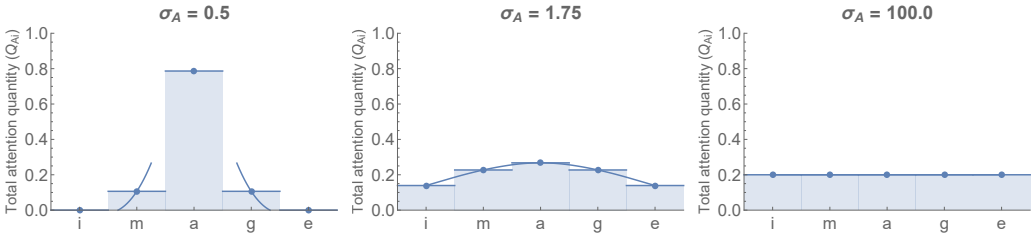

Figure 3.2 – Illustration of attention distribution over the letter string of the word IMAGE for different values of the σ_A parameter: Left: $\sigma_A = 0.5$; Middle: $\sigma_A = 1.75$; Right: $\sigma_A = 100.0$.

Table 3.1 – Definition domains of σ_A, τ_{YES} and τ_{NO} parameters for the calibration by grid search.

Parameters \ Word length		4 L	5 L	6 L	7 L	8 L	9 L	10 L	11 L
σ_A	**Min. value**	0.5	0.5	0.5	0.5	1.0	1.0	1.0	1.0
	Max. value	3.25	3.25	3.25	3.25	4.0	4.0	4.5	4.5
	Step	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
τ_{YES}	**Min. value**	0,6							
	Max. value	0,95							
	Step	0,05							
τ_{NO}	**Min. value**	0,6							
	Max. value	0,95							
	Step	0,05							

3.2 Calibration of parameters

We now present some of the parameters that have direct influence on simulating LD. First, we consider decision thresholds for the YES and NO answers, τ_{YES} and τ_{NO}. These parameters, with values between 0 and 1, are respectively linked to the probability $P([D^T = true] \mid e^{1:T}\ [\lambda_{D^{1:T}_{1:N}} = 1]\ \mu_A^{1:T}\ \sigma_A^{1:T})$ that the stimulus is a word and to the probability $P([D^T = false] \mid e^{1:T}\ [\lambda_{D^{1:T}_{1:N}} = 1]\ \mu_A^{1:T}\ \sigma_A^{1:T})$ that the stimulus is not a word. Thus, τ_{YES} and τ_{NO} set the probability to be reached by evidence accumulation to generate a decision. The relation between parameters τ_{YES} and τ_{NO}, as well as the values for the prior distribution over D^0, adapt the model to various types of non-words, according to their relation with real words (e.g., single-letter difference or full-consonant strings).

In the FLP experiment, whatever the word length, pseudowords were built in a single manner. Monosyllabic pseudowords were created by recombining onsets and rimes from the real words used. The same method was used to build polysyllabic pseudowords, but by recombining syllables instead of onsets and rimes. Despite controlling the same criteria as those used for the selection of real words, this method is questionable since it sometimes results in pseudowords that do not constitute plausible sequences in French. Thus, in our study, we only use the real words of the FLP experiment as stimuli. Therefore, we only consider simulations where words have to be identified as such (i.e., stimuli are always words), so that the parameter for deciding that a stimulus is not a word (τ_{NO}) should practically be irrelevant. To be precise, there could be fringe cases where a word would be incorrectly recognized as a non-word; such cases should be rare, especially for high values of τ_{NO}.

Other parameters are easily set, thanks to their physical interpretations. For instance, we

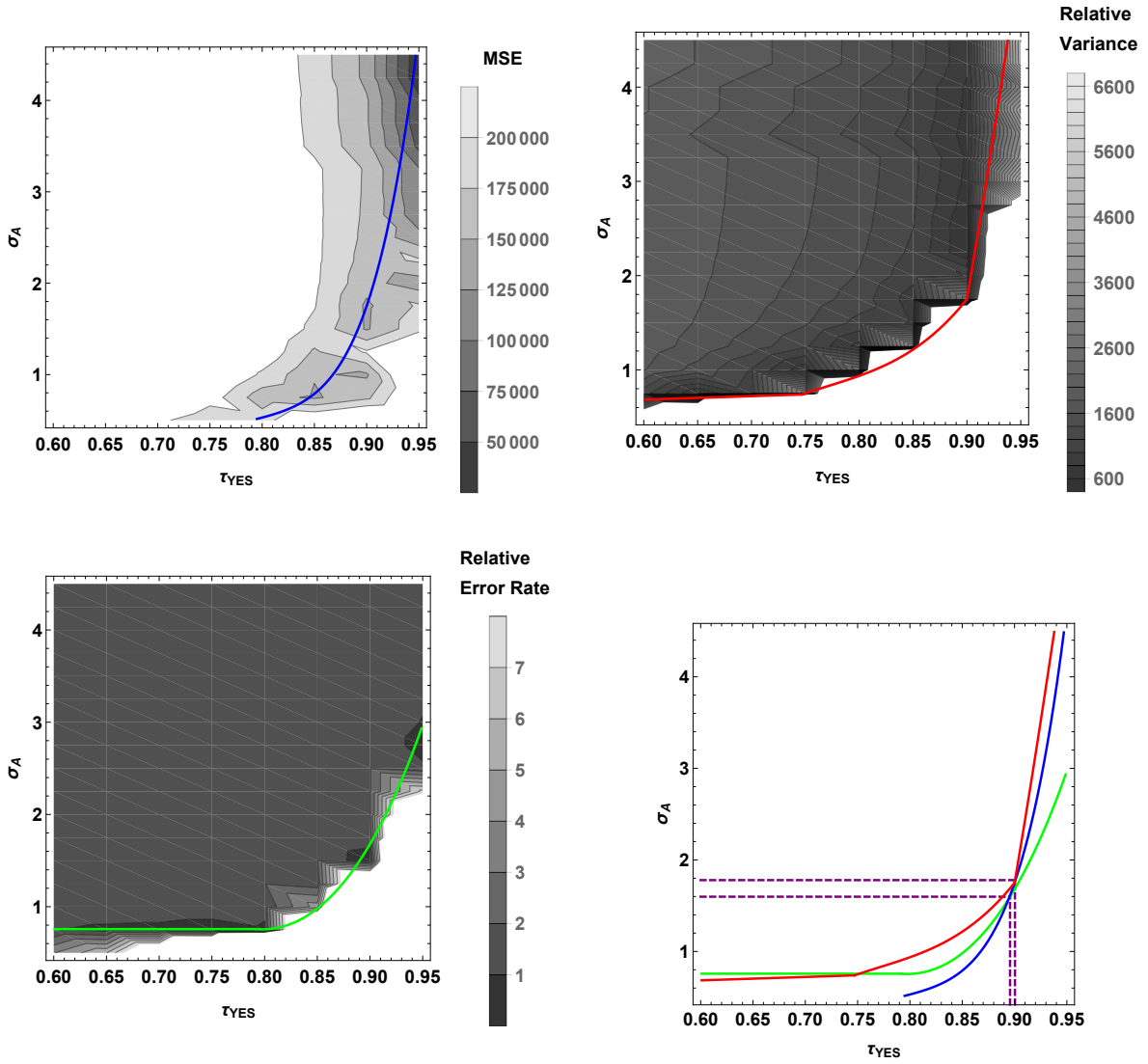

Figure 3.3 – Results of the grid search calibration of τ_{YES} and σ_A. Top left: MSE between simulated and observed RTs. Top right: absolute difference between simulated and observed variance of RT distributions. Bottom left: absolute difference between simulated and observed error rate. All measures are averaged over all considered word lengths (4 to 11 letters). Color gradients indicate measure values: the darker the color, the smaller the difference between the observed and the simulated data. Colored curves highlight parameter space regions where measure is close to optimal. Bottom right: superposition of the three curves of optimal parameter combinations, defining a small region where all measures are close to optimal (delineated by dashed lines).

classically assume that the gaze position (g) and the attentional focus (μ_A) coincide; thus, assuming a central fixation, we set $\mu_A = g = \frac{N+1}{2}$, with N the word length. The σ_A parameter characterizes the spread of attention, i.e., mathematically, the standard deviation of the Gaussian distribution of attention. Thus, the higher the value, the more spread out and uniform-like the distribution of attention. Since attention is modeled by a probability distribution, the sum of attention quantity Q_{Ai} over all letters is 1. Figure 3.2 illustrates various attention distributions for a given input. A reduced σ_A parameter value ($\sigma_A = 0.5$) induces

an attention distribution such that the central letter is efficiently processed, to the detriment of external letters. Conversely, a large σ_A parameter value ($\sigma_A = 100.0$) spreads attention uniformly over all letters, such that they are equally processed but potentially insufficiently. Indeed, uniformly distributing a constant amount of attention will spread it too thinly for long words. Finally, an intermediate σ_A parameter value ($\sigma_A = 1.75$) allows to distribute attention to favor some of the letters, but still provides enough processing resources to all the letters of a 5 letter word.

To specifically calibrate the LD related parameters σ_A, τ_{YES} and τ_{NO}, we applied a grid search method, that is, we explored a set of regularly spaced points in the domain of possible parameter value combinations (including τ_{NO} to verify that it is indeed irrelevant). We considered the grid search domain shown on Table 3.1.

The time unit of the BRAID model has been calibrated previously so that one iteration corresponds to around 1 ms. Thus, we set the maximum duration allocated for simulations to 2,000 iterations, which is enough time if we consider the mean RTs observed in the FLP (Ferrand et al., 2010).

Qualitative inspection of the results confirmed that the decision threshold τ_{NO} has no influence on the performance of the model. In the following, we therefore set τ_{NO} arbitrarily to 0.65, and only consider the values of couples (τ_{YES}, σ_A) for further analyses.

To calibrate σ_A and τ_{YES}, the model simulations were compared with the experimental observations of the FLP experiment. For this comparison, we used three measures: the Mean Square Error (MSE) between simulated and observed RTs for correct answers, the absolute difference between simulated and observed variances of RT distributions for correct answers, and, finally, the absolute difference between simulated and observed error rates. The MSE is calculated with:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(tc_i - tm_i)^2, \tag{3.4}$$

in which tc_i represents the observed RT (in ms) and tm_i the simulated RT (in number of iterations) for the i-th word, and n is the number of words correctly identified as words by the model. Results for each of these comparison measures, and for each explored (τ_{YES}, σ_A) combination, are shown in Figure 3.3, aggregated for all word lengths (with qualitative inspection confirming that results behave consistently across word lengths).

As results suggest, each measure individually does not provide a unique point where model simulations and observations maximally correspond. Instead, each measure suggests a one-dimensional curve, along which combinations of τ_{YES}, σ_A provide good results. More precisely, increasing decision threshold τ_{YES} can be compensated by also increasing σ_A, with a non-linear relationship in all three adequacy measures. This suggests that the model is robust, that is to say, a large number of parameter values allow the model to capture the experimental observations.

Combining the three adequacy measures was performed "geometrically" (Figure 3.3, bottom right), by considering the intersection of the three optimal curves of each measure. They intersect in a small region of the explored space, with σ_A between 1.6 and 1.8 and τ_{YES} around 0.9.

In summary, simulating the LD task on 1,200 words from 4 to 11 letter allowed to calibrate three parameters (σ_A, τ_{YES} and τ_{NO}) of the BRAID model. Therefore, in the remainder of this study, we consider the following values: $\tau_{NO} = 0.65$, $\tau_{YES} = 0.90$ and $\sigma_A = 1.75$.

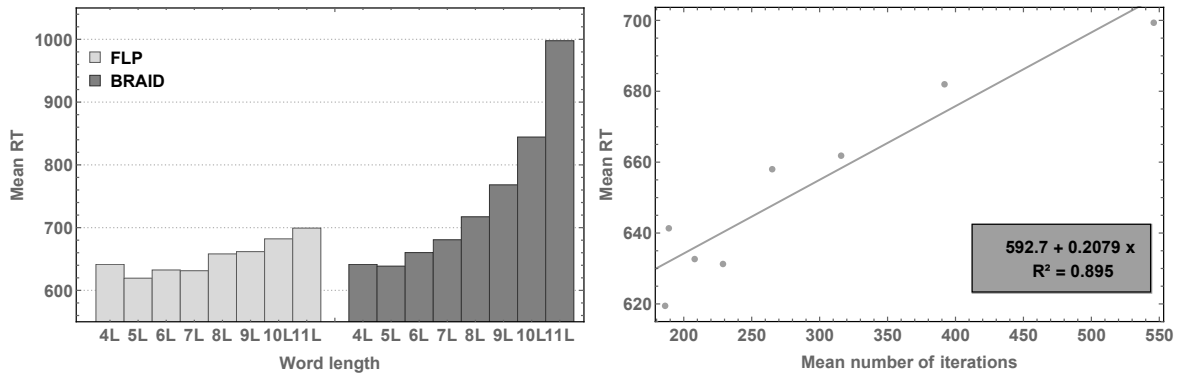

Figure 3.4 – Results of Simulation 1.A (simulation of LD with a unique centered attentional focus). Left: mean RTs reported in the FLP study (in ms; light gray) and simulated by the model (in number of iterations; dark gray), as a function of word length. Mean simulated RTs are scaled and adjusted by aligning them on the 4-letter word condition; other length conditions are thus model predictions. Right: linear regression between simulated and observed mean RTs ($R^2 = 0.895$).

4 Simulation 1: simulation of the word length effect

Above, in Section 3, we have aggregated simulation results over all word lengths, and shown that BRAID could account very well for the experimental observations, in terms of overall MSE, variance of response distributions, and accuracy. We now consider simulation results as a function of word length, to study the ability of the model to account for the word length effect. In this section, we study two simulations of the word length effect: in Simulation 1.A, as in parameter calibration, the model simulates LD with a single centered attentional focus, while in Simulation 1.B, the model performs several shifts of focused attention.

4.1 Simulation 1.A

4.1.1 Procedure

All parameter values used for this simulation are the same as those of parameter calibration (Section 3).

4.1.2 Results

For analysis, the errors made by the model were removed, which, for the chosen parameters, represent 5 words out of 1,200 (i.e., 0.42% of errors). Results are presented in Figure 3.4.

Figure 3.4 (left) shows mean behavioral and simulated RTs. To compare them, we scaled the simulated RTs by adding a constant, on the 4-letter word condition, so as to align them with behavioral RTs for this condition. Using this method, the simulation results for the other word lengths are predictions. Note that we will use this same method to present simulation results for all the following simulations.

Figure 3.4 (left) shows that the word length effect is larger in the model than in the behavioral data. The model simulates a substantial word length effect of 44.6 iterations per letter on average, well above the 7.25 ms per letter found in the behavioral data. The linear regression computed between the simulated and behavioral data (Figure 3.4, right) yields a high correlation coefficient ($R^2 = 0.895$), which might suggest that the model, overall, adequately reproduces the variation of RTs as a function of word length. However, the linear regression

Table 3.2 – Fixation positions used for Simulation 1.B.

	4L	5L	6L	7L	8L	9L	10L	11L
$\mu_{A1} = g_1$	2.5	3.0	3.5	2.5	2.5	3.0	3.0	3.5
$\mu_{A2} = g_2$	–	–	–	5.5	6.5	7.0	8.0	8.5

parameters indicate that the relation between behavioral data and model simulations is mostly supported by the additional constant (592.7), thus decreasing strongly the weight of the model (0.2079).

4.2 Simulation 1.B

4.2.1 Procedure

The results of Simulation 1.A suggest that the model is able to reproduce the word length effect, but its magnitude is larger than observed experimentally. In Simulation 1.A, letters of long words are processed in a parallel manner, using a unique, central gaze and attentional focus. Attention distribution, in this case, allocates a small amount of processing resources to outer letters, basing word recognition essentially on a small number of inner letters; this makes it inefficient. Therefore, in Simulation 1.B, we explored a variant in which we assume that, for words of 7 letters or more, the model performs several shifts of attention.

Thus, we defined $\mu_{A1} = g_1$ and $\mu_{A2} = g_2$ as being two positions of the attention focus used by the model during the LD task. We adapted these values systematically for each word length (Table 3.2), following the assumption that a long word would be treated as two short words. For instance, we suppose that an 8-letter word is processed as two 4-letter words. Therefore, applying the same calculations as before for central focus (see Section 3.2) yields $\mu_{A1} = 2.5$ and $\mu_{A2} = 6.5$.

We set the maximum duration of the attention focus at each position to three conditions: 50, 100 and 200 iterations. Past this focus duration, the model shifts focused attention from this position to the other one, alternating either until the decision threshold or the time limit of 2,000 iterations is reached. All other parameter values are identical to those used in Simulation 1.A.

4.2.2 Results

No errors were made by the model in Simulation 1.B and the results were qualitatively similar, whatever the number of iterations allocated to each attention focus (50, 100 or 200). We here present the simulated length effects for the condition with 100 iterations.

Figure 3.5 presents a comparison of simulation results with experimental data. As previously, presented RTs are adjusted. Compared with Simulation 1.A (Figure 3.4), the main finding of Simulation 1.B is that allowing the model to perform several shifts of focused attention decreases RTs for words from 7 to 11 letters. The word length effect simulated by the model is characterized by a slope of 8 iterations per letter, very close to the experimentally observed slope in the FLP data (7.25 ms per letter). Linear regression of Simulation 1.B is also very satisfying ($R^2 = 0.868$). Finally, coefficients of the linear regression (Figure 3.5, right) show that the weight of the model is close to 1 (1.13), reducing the value of the additional constant (414.5). This time constant could reflect motor response time, measured in the FLP but not modeled in BRAID.

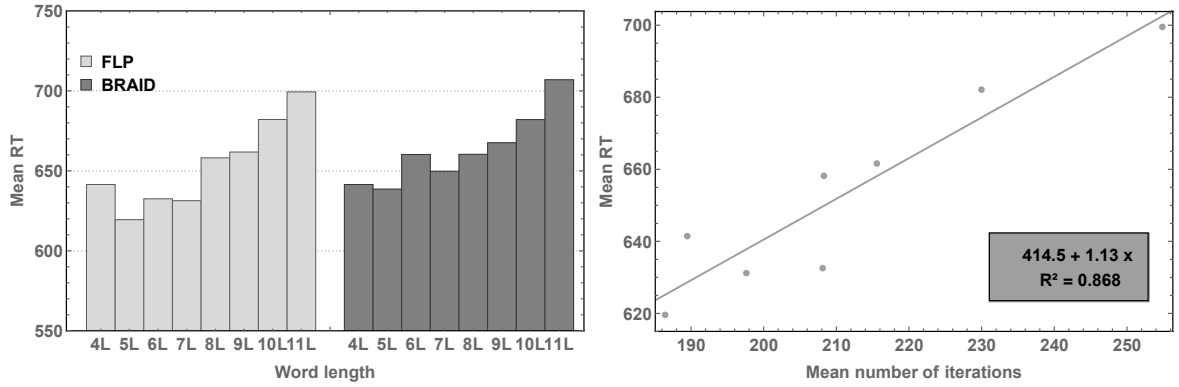

Figure 3.5 – Results of Simulation 1.B (simulation of LD in the condition with two attention focuses up to 7 letter words and a focus duration of 100 iterations). Left: mean RTs reported in the FLP study (in ms; light gray) and simulated by the model (in number of iterations; dark gray) as a function of word length. Mean simulated RTs are scaled and adjusted by aligning them on the 4-letter word condition; other length conditions are thus model predictions. Right: linear regression between simulated and observed mean RTs ($R^2 = 0.895$).

An additional simulation was carried out to explore whether a good fit of the data was only found when forcing multiple shifts of attention for words up to 7 letters. Results of the simulations showed that allowing multiple fixations for all word lengths resulted in a very similar length effect of 9 ms per letter well within the range of length effects reported in experimental studies for word lexical decision.

5 Simulation 2: Effect of attention distribution

In previous sections, we simulated the word length effect with either one (Simulation 1.A) or two (Simulation 1.B) shifts of attention, and with attention parameters set to default values. We will now investigate the effect of attention dispersion, that is, the σ_A parameter controlling the standard deviation of attention distribution, on the simulated word length effect. For this, we perform three simulations (Simulation 2.A, 2.B and 2.C), using the same experimental material as previously.

In Simulations 2.A and 2.B, the model simulates a single central focus of attention, and allocates identical attentional resources to each letter, using uniform-like distributions of attention. In Simulation 2.A, the total amount of attention is stable across lengths while in Simulation 2.B, a similar amount of attention is arbitrarily allocated to each letter whatever word length (thus increasing total attention with word length). On the contrary, Simulation 2.C explores the other extreme case, in which attention is fully allocated to each letter, in turn, in a serial manner.

5.1 Procedure

Using our Gaussian model of attention distribution with $\sigma_A = 100.0$ closely approximates a uniform distribution (see Figure 3.2, right). Recall that, in the model, the sum of the attentional quantity Q_{Ai} allocated to each letter is equal to 1. For this reason, the longer the word, the smaller Q_{Ai} for each letter. Thus, in Simulation 2.A, spreading attention uniformly gradually

decreases the amount of attention allocated to each letter. Considering words from 4 to 11 letters, this yields $0.09 \leq Q_{Ai} \leq 0.25$.

It follows that the uniform distribution of a fixed amount of total attention interacts with word length, to the detriment of longer words. This manipulation of attention distribution thus predicts an effect of word length in Simulation 2.A. Simulation 2.B was performed to explore the impact of a uniform distribution of attention without a priori interacting with word length. For this purpose, Q_{Ai} was arbitrarily set to 0.5 whatever letter position and word length.

In Simulation 2.C, σ_A was set to 0.5 to simulate an extreme case of narrow attention distribution, in which attention is mostly allocated to a single letter at a time (see Figure 3.2, left). After 100 iterations on the first letter, the focus of attention shifts to the next letter and so on, in a left-to-right manner (cycling back to the first position after the last, if required), until either the LD threshold is reached or 2,000 iterations have passed. All the other parameters are identical to those used in previous simulations.

5.2 Results

No errors were made by the model in either Simulation 2.A or Simulation 2.B. Around five percent errors (64 words out of 1,200; 5.33% errors) that were removed for subsequent analyses occurred in Simulation 2.C. Figure 3.6 presents a comparison of simulation results with experimental data, in which, as previously, simulation results are adjusted by reference to the 4-letter word condition.

As expected, Figure 3.6 (2.A, left) shows that the use of a uniform distribution of a stable amount of total attention across length resulted in a strong length effect on word LD RTs. An increase of 19 iterations per letter on average was obtained in the simulations against only 7.25 ms per letter in the behavioral data. The inhibitory length effect obtained on word LD RTs in Simulation 2.A contrasts with the facilitatory effect observed in Simulation 2.B. Indeed, as shown on Figure 3.6 (2.B), longer words were processed faster when the amount of attention allocated to each letter was constant across positions. Figure 3.6 (2.C, left) presents results obtained for Simulation 2.C. We observe that the magnitude of the simulated word length effect is larger than observed in the FLP study and this, starting from 6-letter words. The 32.6 iterations per letter found on average in the simulated data contrast with the 7.25 ms per letter reported in the behavioral data, showing that time of processing is highly and abnormally sensitive to word length.

The linear regressions presented on Figure 3.6 (2.A and 2.C, right) indicate that the relation between the behavioral and simulated data is almost linear (respectively $R^2 = 0.911$ and $R^2 = 0.865$). However, the parameters of the linear regressions, as in Simulation 1.A, show that these relations are mainly supported by the additional constant (respectively 541.7 and 558.1) decreasing strongly the weight of the model (respectively 0.4519 and 0.2669). Conversely, Figure 3.6 (2.B, right) shows that the relation between the behavioral and simulated data is not linear ($R^2 = 0.36$).

6 Discussion

Our main purpose in this study was to provide new insights on the role of visual attention on the word length effect in lexical decision. The BRAID model was used to perform a series of simulations and compare their results with LD RTs for 1,200 words from 4 to 11 letters taken from the FLP (Ferrand et al., 2010). A first series of two simulations (Simulation 1.A and 1.B) was performed to simulate the behavioral word length effect in LD. In Simulation 1.A, length

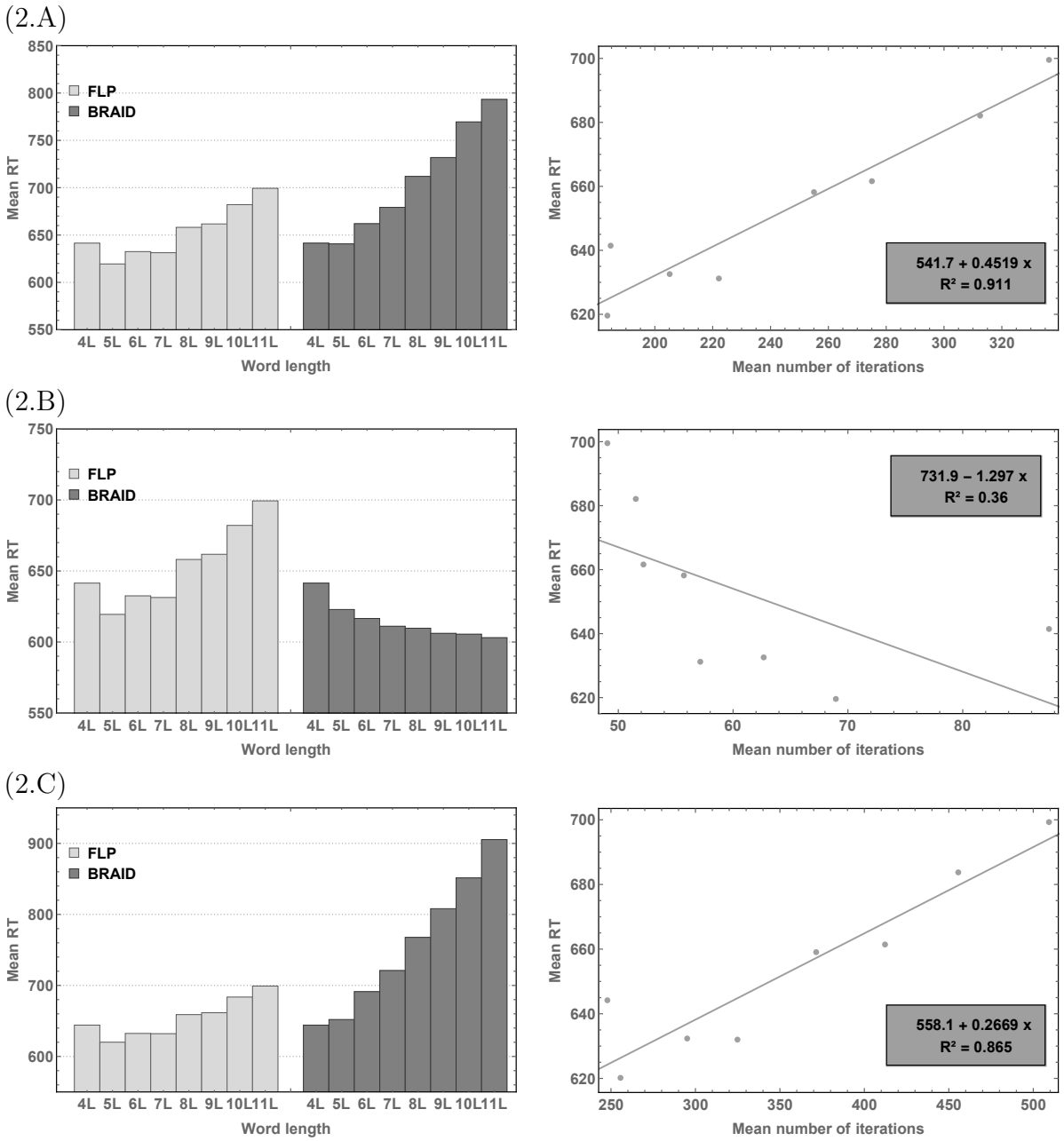

Figure 3.6 – Results of the simulations of LD RTs for different conditions of attention distribution. Simulation 2A: uniform distribution of attention, Q_{Ai} varies from 0.25 to 0.09 depending on word length; Simulations 2.B: uniform attention distribution, $Q_{Ai} = 0.5$ whatever word length; Simulation 2.C: narrow distribution of attention, $\sigma_A = 0.5$. Left: mean RTs obtained in the FLP study (in ms) and simulated by the model (in number of iterations) as a function of word length. Mean simulated RTs are scaled and adjusted by aligning them on the 4-letter word condition; other length conditions are model predictions. Right: linear regression between simulated (2.A, 2.B and 2.C) and observed mean RTs (respectively: $R^2 = 0.911$, $R^2 = 0.36$, $R^2 = 0.865$).

effect was simulated following a single central attention focus, which resulted in far stronger length effects than reported for humans in the FLP. In Simulation 1.B, a very good fit to the behavioral data was obtained using multiple shifts of attention during processing.

Another set of three simulations was then performed to manipulate the distribution of atten-

tion over the word letter-string and explore more in depth whether and how attention modulates the word length effect in LD. A uniform distribution of attention was used in Simulations 2.A and 2.B to explore processing in the absence of attention filtering. In Simulation 2.A, the total amount of attention remained constant across length so that each letter was allocated a lesser amount of attention as word length increased. In Simulation 2.B, the same amount of attention was allocated to each letter whatever word length. Results showed a stronger inhibitory word length effect than behaviorally reported, in Simulation 2.A, but a facilitatory effect in Simulation 2.B. Last, a Gaussian but narrow distribution of attention was used in Simulation 2.C that resulted in an exaggerated word length effect as compared to the behavioral data.

The current work shows that the BRAID model can successfully account for the length effect on word RTs in lexical decision, provided that multiple attention shifts across the letter-string are allowed for longer words. As previously claimed for length effects on naming latency (Perry & Ziegler, 2002; Plaut, 1999), this finding might suggest that length effects inherently relate to serial processing. Interestingly and rather counter-intuitively however, our simulations show that strong length effects can be generated with purely parallel processing (a single attention focus) and that, contrary to typical belief, additional shifts of attention lead to weaker rather than stronger length effects.

Indeed, comparing the results of the different simulations suggests that multiple shifts of attention fasten word processing, at least for longer words. Indeed, the most exaggerated word length effect on LD RTs was generated by BRAID following a single attention focus (Simulation 1.A). This effect was quite stronger than in conditions of multiple attention shifts, even as compared to the extreme case in which attention was allocated to each of the word letter successively (Simulation 2.C). Supportive evidence that length effect can follow from parallel processing comes from eye movement studies showing that the amount of time spent fixating a word increases with word length, even for single fixations (Rayner & Raney, 1996). Comparative studies of length effects in progressive demasking and lexical decision provide further support to the present findings. While visual target display duration is not experimentally constrained in LD tasks, display time in progressive demasking is short enough to prevent multiple fixations. Comparison of the two tasks showed that the word length effect was far stronger in progressive demasking than in LD (Ferrand et al., 2011), suggesting larger length effects in conditions of parallel processing. Thus, a first contribution of the current work is to show that length effects cannot a priori be interpreted as direct evidence for serial processing.

A reliable simulation of the word length effect reported for humans on word LD RTs was here obtained assuming multiple attention shifts, thus multiple fixations, for words of 7 letters or more. Evidence that refixation probability and gaze duration increase with word length in conditions of text reading (Juhasz & Rayner, 2003; Pollatsek, Juhasz, Reichle, Machacek, & Rayner, 2008; Rayner, 1998; Rayner & Raney, 1996) or isolated word reading (Vitu, O'Regan, & Mittau, 1990), and that gaze duration relates with LD RTs (Schilling, Rayner, & Chumbley, 1998), makes plausible the prediction of BRAID of multiple refixations for longer words. However, multiple fixations have been reported on even shorter words (5-to-6 letter long) in isolated word reading (Vitu et al., 1990), suggesting that multiple fixations might concern a wider range of word length than hypothesized in Simulation 1.B. Furthermore, the probability of a refixation is likely to gradually increase as a function of length. In contrast, we simulated a simple mechanism based on a length threshold: below it, refixations never happen; over it, refixations always occur. We also did not model the time delay induced by attention shifts. Modeling these mechanisms more precisely could help better account for the non-linearity of behavioral response times for words of lengths 5 to 7 (i.e., around our cutoff threshold for refixations in Simulation 1.B; see Figure 3.5). New experimental studies on eye movements during the lexical

decision task are required to address these issues, as precise data on these mechanisms appear to be lacking. Overall however, the current work already shows that extending the possibility of multiple attention shifts/fixations to the entire word-length range (from 4 to 11 letters) only marginally affects the simulated word length effect on LD RTs. Accordingly, the capacity of BRAID to successfully account for the word length effect on LD RTs is quite robust provided that processing allows multiple attention shifts (thus, refixations). Overall, the current findings lead to conclude that refixating is beneficial to the reader – processing time is improved and reading is more fluent – which might explain why refixating is the rule in all languages and all semantic contexts, independently of the readers' characteristics.

Another contribution of the current work is to show that length effects can be simulated in a word recognition model that does not include any phonological processing component. Previous computational accounts of the word length effect in isolated word naming attributed this effect to phonological recoding. For instance, in previous studies (Perry et al., 2007; Plaut, 1999), simulating the length effect was largely dependent on the capacity of the model to generate a plausible phonological output. As a direct consequence, such models cannot account for word length effects in lexical decision. Evidence in BRAID that mechanisms involved in visual word recognition can account for length effects independently of any phonological processing is well in line with evidence of word length effects in tasks such as progressive demasking, that primarily tap visual processing. This is also consistent with reports that the number of letters is a better predictor of LD latency than the number of phonemes (Balota et al., 2004). The fact that typical readers show exaggerated word length effects in conditions of visually degraded stimuli brings additional support for an early visual processing origin of word length effects in reading (Fiset, Gosselin, Blais, & Arguin, 2006). This is not to say that length effects in isolated word naming exclusively derive from visual orthographic encoding skills. The current findings suggest that, in reading as in LD, word length effects might reflect early orthographic processing skills without excluding that additional effects due to phonological recoding may further affect word length effects in reading aloud.

Importantly, the main contribution of our work is to demonstrate the causal role of visual attention in LD word length effects. The whole set of simulations shows that variations in the distribution of visual attention over the input word has direct impact on word length effect patterns. The uniform distribution of a fixed amount of attention resources over the letter string led to an abnormally strong word length effect, which resulted from the combination of lower attentional resources and increased visual acuity effects for longer words. In contrast and despite the acuity gradient, a large and similar amount of attention allocated to each letter whatever word length led to faster LD latency for longer words, thus a reversed length effect as compared to the behavioral data. As a sidenote, these simulations also clearly show that neither visual acuity nor lateral interference (i.e., crowding) — which remain constant across simulations — are critical to account for word length effect. More importantly, the inability of BRAID to generate human-like word length effect through simulations based on uniform distributions of visual attention is further evidence that attention distribution is critical for word processing. The role of attention is certainly not to make the processing of multiple stimuli equal as illustrated in these simulations but rather to act as a filter that, as in Simulations 1.A and 1.B, enhances letter visibility in some portions of the word to the detriment of others (Carrasco, 2011). The capacity of BRAID to account for word length effect in LD primarily relies on a Gaussian distribution of attention, thus modulating attention across letters within the input word. Our account of visual attention as modulating letter processing within single words is strongly supported by experimental findings that show an impact of visual attention in single word processing (Besner et al., 2016; Lien, Ruthruff, Kouchi, & Lachter, 2010) and lexical

decision (McCann, Folk, & Johnston, 1992), that support early prelexical involvement of visual attention (Risko, Stolz, & Besner, 2010; Stolz & Stevanovski, 2004) and that assume modulation through visual attention of the rate of feature uptake (Carrasco, 2011; Stolz & Stevanovski, 2004). In BRAID, a fully defined visual attention device is for the first time implemented in a word recognition model, showing how attention modulates sensory processing and what is the impact of this modulation on word processing.

Last, another very important issue for a word recognition model is to account not only for typical but also for atypical reading. Although simulating patterns of acquired or developmental dyslexia was beyond the scope of the current paper, evidence suggests that a reduction of visual attention distribution over the letter string results in exaggerated word length effects (Barton et al., 2014; Duncan et al., 1999; Juphard et al., 2004; Valdois et al., 2011, 2003). Our Simulation 2.C, with larger length effects deriving from reduced attention dispersion, suggests that BRAID could be able to account for a variety of reading disorders that show exaggerated length effects in the context of visual attention deficits. Interestingly, providing an account of atypical word recognition skills was beyond the scope of most previous computational models of reading but the few that were concerned with pathology and tried to simulate acquired disorders did postulate an attentional device (Ans et al., 1998; Mozer & Behrmann, 1990). Recognizing the role of visual attention in single word recognition and reading and implementing a visual attention device in a computational model of reading like BRAID is critical to offer a plausible and integrative account of human reading skills.

Acknowledgments

This work has been supported by a French Ministry of Research MESR grant for EG, and a Fondation de France research grant for TP.

Chapitre 4

Apprentissage implicite de mots nouveaux chez l'adulte : effets du nombre d'occurrences et de l'attention visuelle sur les mouvements oculaires

Note

Ce Chapitre reprend un article soumis dans une revue internationale.

Résumé

De manière surprenante, alors que de nombreuses études se sont penchées sur l'analyse des mouvements oculaires en lecture (Rayner, 1998, 2009), on ne dispose que de données très limitées sur les caractéristiques des mouvements oculaires lors du traitement d'un mot nouveau, et sur la manière dont les différentes mesures comportementales sont affectées par l'exposition répétée à un même mot. Une synthèse récente sur cette question suggère que l'étude de l'apprentissage « en train de se faire », par exemple par le biais des mesures oculométriques pendant la lecture, est une voie d'avenir pour mieux comprendre la construction de la mémorisation orthographique (Nation & Castles, 2017). Les résultats obtenus dans deux études récentes (H. Joseph & Nation, 2018; H. Joseph et al., 2014) ont permis d'étayer cette hypothèse, en montrant une diminution des temps de traitement (durée de première fixation et du temps total de fixation) au cours de la lecture répétée de nouveaux mots en situation d'apprentissage orthographique implicite.

Notre premier objectif est d'étudier l'effet de la lecture répétée de nouveaux mots sur les mouvements oculaires et sur l'apprentissage orthographique (c'est-à-dire, sur la qualité de la mémorisation orthographique). Parallèlement, à travers l'analyse du comportement oculomoteur au cours de l'apprentissage, nous proposons d'observer la dynamique temporelle du processus d'apprentissage orthographique. Finalement, notre troisième objectif est d'explorer le rôle de l'attention visuelle sur les mouvements oculaires et sur l'apprentissage orthographique.

Pour répondre à nos objectifs, 42 adultes normo-lecteurs de langue maternelle française sont mis en situation d'apprentissage incident par le biais de l'exposition répétée à 30 pseudo-mots (nouveaux mots) bisyllabiques, composés de 8 lettres et orthographiquement légaux. Trente mots connus bi- ou trisyllabiques de 8 lettres sont utilisés comme mots contrôles permettant de distinguer un effet d'apprentissage d'un simple effet de répétition sur les variations du

comportement oculomoteur. Chaque essai comporte la présentation d'une croix qui doit être fixée de façon à déclencher l'apparition d'un item (mot ou pseudo-mot) suivi d'un chiffre. Le participant doit lire l'item à haute voix puis le chiffre. Chaque item est présenté une, trois ou cinq fois. Les mots sont aléatoirement mélangés aux pseudo-mots et le nombre de répétitions varie aléatoirement. Les mouvements oculaires sont enregistrés tout au long de la tâche.

Immédiatement après la phase de lecture, l'apprentissage orthographique implicite des pseudo-mots est mesuré à travers deux tâches (phase test). Une première tâche de dictée permet de quantifier la mémorisation exacte de l'orthographe des nouveaux mots. Dans cette tâche, le participant doit écrire sous dictée le pseudo-mot tel qu'il était écrit dans la phase de lecture. Le score est calculé en considérant que la production est correcte lorsque l'orthographe produite sous dictée est strictement identique à celle du pseudo-mot cible lu. La seconde tâche est une tâche « originale » de décision orthographique qui permet de mesurer la capacité du sujet à reconnaître l'orthographe du pseudo-mot cible. Dans cette tâche, le participant est exposé à des pseudo-mots uniquement, incluant les pseudo-mots cibles et des pseudo-homophones des items cibles. Les items sont présentés un à un et pour chaque orthographe il faut décider le plus vite possible et en faisant le moins d'erreurs possible si l'item présenté à l'écran est correctement orthographié, c'est-à-dire, tel qu'il était écrit au cours de la phase de lecture. Les temps de réponse ainsi que le taux de réponses correctes sont enregistrés. Finalement, les participants réalisent deux épreuves de report partiel et de report global de séquences de six consonnes, tâches qui permettent de mesurer l'empan visuo-attentionnel (VA) de chaque sujet.

Les résultats obtenus pendant la phase test montrent une augmentation significative des performances en dictée et en décision orthographique, suggérant qu'il y a bien eu apprentissage orthographique incident pendant la phase de lecture. En effet, les scores reportés en dictée montrent une augmentation des performances avec le nombre d'occurrences. Par ailleurs, les résultats obtenus dans la tâche de décision orthographique indiquent que les pseudo-mots sont reconnus au-dessus du hasard lorsqu'ils ont été lus 3 ou 5 fois et que les performances – temps de réponse et taux de réponses correctes – sont d'autant plus élevées que le pseudo-mot a été lu plus souvent.

L'analyse du comportement oculomoteur généré pendant la phase de lecture indique que les temps de traitement – durée du regard et durée de fixation – et le nombre de fixations sont plus élevés pour les pseudo-mots que pour les mots et que ces mesures décroissent en fonction du nombre d'occurrences. Parallèlement, les résultats obtenus montrent une plus forte diminution du nombre de fixations et de la durée du regard pour les pseudo-mots que pour les mots, suggérant que les variations du comportement oculomoteur sont associées à l'apprentissage orthographique et non simplement à un effet de répétition. Par ailleurs, l'analyse des données oculométriques révèle une diminution massive des temps de traitement (durée du regard, durée de single fixation et durée de seconde fixation) entre la première et la deuxième occurrence pour la durée du regard et la durée de deuxième fixation et entre la deuxième et la troisième occurrence pour la durée des fixations uniques. Ces résultats nous permettent d'affiner les connaissances sur le décours temporel de l'apprentissage orthographique, suggérant que l'apprentissage orthographique survient dès les tous premiers stades du processus d'apprentissage.

Finalement, une analyse exploratoire de l'effet de l'empan VA sur l'apprentissage orthographique – et par conséquent, sur le comportement oculomoteur – est réalisée. Pour cela, les participants ont été scindés en deux groupes d'effectif égal en fonction de leur score aux tâches de mesure d'empan VA. Les résultats révèlent un effet significatif des capacités VA sur les temps de réponse reportés pour la tâche de décision orthographique. Plus l'empan VA est faible, plus les temps de réponse sont élevés. Par ailleurs, l'analyse des mouvements oculaires montre des

temps de traitement (durée du regard et durée de deuxième fixation) d'autant plus élevés que les capacités VA sont faibles. Enfin, l'évolution du nombre de bonnes réponses dans les tâches de mesure de l'apprentissage orthographique ne montrent pas de différence entre les deux groupes. Ces résultats, bien qu'exploratoires, suggèrent qu'un empan VA élevé induit une diminution du temps de traitement d'un nouveau mot sans modifier la dynamique d'apprentissage.

L'ensemble des résultats montre un meilleur apprentissage de l'orthographe lexicale des nouveaux mots avec le nombre d'occurrences, entraînant une diminution des temps de traitement et du nombre de fixations au cours de l'apprentissage. L'observation d'une diminution plus marquée en début d'apprentissage pourrait être liée à des influences lexicales top-down induites par la création d'une représentation orthographique en mémoire lexicale dès la première rencontre avec le nouveau mot. Enfin, les résultats exploratoires obtenus suggèrent que les capacités VA pourraient jouer un rôle dans le processus d'apprentissage orthographique de nouveaux mots. Plus précisément, il semble que des capacités plus élevées d'empan VA permettent un traitement plus efficace de l'orthographe des mots (temps de fixation plus courts) sans interagir avec le processus de mémorisation. Des études ultérieures devront être menées afin d'étayer cette hypothèse auprès de populations présentant une plus grande variabilité de leurs capacités d'empan VA.

Abstract

While conceptual models of orthographic learning postulate that visual word analysis is a necessary step when processing a new written word, only a few studies have examined eye movements during repeated reading of new words and their implicit memorization. The present study used eye-tracking to examine the exposure-by-exposure time course of implicit orthographic learning while reading. The eye movements of 42 adult skilled readers were recorded while they had to read 30 isolated pseudo-words to which they were exposed one, three or five times. Known words were further presented in similar conditions to disentangle changes in eye movements that reflect orthographic learning from those that reflect a simple repetition effect. The quality of orthographic memorization of these new words was later assessed by a spelling-to-dictation task and an original orthographic decision task. To study the impact of visual attention span (VA span) on orthographic learning, participants were split into two groups with higher or lower VA span. Increased performance with exposures on the offline measures showed that efficient orthographic learning did occur. The orthographic decision paradigm was highly sensitive to the learning process. The three eye-tracking measures of gaze duration, single fixation and second-of-two fixation duration showed early learning effects. More gradual effects were further found on the number of fixations. Higher VA span participants showed faster orthographic decision and reduced processing times during novel word reading. Major changes in the oculomotor measures sensitive to orthographic learning occur from the first to the second exposure, highlighting the relevance of an exposure-by-exposure in-depth analysis.

1 Introduction

To become a fluent reader and a good speller, children have to memorize the orthographic form of thousands of words. Orthographic learning mainly occurs with exposure to printed words via reading experience. This learning is mostly incidental, occurring while the reader does not pay particular attention to the written word and does not go through any particular strategy while reading. Despite the key role of orthographic learning in both reading – for the visual recognition of printed words – and writing – for word-specific spelling, we still have little insight into how learning affects new printed word processing and what are the cognitive mechanisms at play. In this paper, we explore the incidental learning of new words using eye-tracking, allowing studying the learning process as reading unfolds, in an online manner. Furthermore, while most previous studies have emphasized the impact of phonological decoding on orthographic learning, we here examine whether the visual attention capacity, as measured through tasks of visual attention span (VA span), has a role on the incidental learning of novel words.

A diversity of methods has been used to explore incidental orthographic learning. The most popular is the self-teaching paradigm proposed by David Share in his seminal experiments (Share, 1999, 2004). In this paradigm, participants are administered a learning phase followed by a test phase. During the learning phase, they are asked to read aloud meaningful stories in which target pseudo-words – introduced as names for fictitious places, animals, fruits, etc. – are embedded in short texts. Each story introduces a single target pseudo-word that can be repeated a few times per text. Orthographic learning is assessed during the following test phase that is proposed immediately or some days/months after text reading. Three tasks were initially designed to gauge the acquisition of the target pseudo-word orthographic form: an orthographic choice task, that required identifying the target pseudo-word alongside a homophonic foil (e.g.,

YAIT-YATE), a naming task, in which the targets (e.g., YAIT) and their foils were randomly displayed and naming reaction times recorded, and a spelling task.

Results of Share's seminal studies showed that orthographic learning occurred after only a few encounters with the novel word. Evidence of orthographic learning has been reported after a single encounter (Nation et al., 2007; Share, 2004) and one to four encounters appear to be sufficient for relatively durable memorization of the target specific orthography (Bowey & Muller, 2005; Nation et al., 2007; Share, 1999, 2004; Tucker et al., 2016). Further, studies revealed that orthographic learning could occur out of context. Effective learning was not only reported for novel words embedded in meaningful texts but also in scrambled passages (Cunningham, 2006) or when presented in isolation (Bosse et al., 2015; Nation et al., 2007; Share, 1999). Most previous studies explored incidental orthographic learning in children (from the first grade to later grades). Written word learning was also explored in adults but mainly using intensive repetitions with explicit instruction and/or artificial languages (Taylor, Plunkett, & Nation, 2011).

One issue with the orthographic learning paradigm is that it does not straightforwardly provide evidence that learning was effective following exposition to the novel words. The validity and reliability of the tasks used to assess orthographic learning have been debated (Castles & Nation, 2008). Among the tasks initially used by Share (1995, 1999), the orthographic choice task is one of the most popular, but simultaneous presentation of alternative orthographies in this paradigm may induce strategy-based processing without tapping the word recognition process. Moreover, performance tends to be high and thus not sensitive enough to orthographic learning. In contrast, performance on the spelling task was typically reported as drastically low in children (Bosse et al., 2015; Cunningham, 2006) and the task not sensitive enough to reflect orthographic learning (Castles & Nation, 2008). Some alternative solutions to these more standard tasks have been proposed, like measures of length effects (Martens & De Jong, 2008) or semantic knowledge (H. Joseph et al., 2014), but these measures were also found to lack sensitivity.

In the present study, an original "orthographic decision" task was administered as a potential new alternative. In this task, targets and homophonic foils were displayed one at a time and participants had to respond either YES if the written form corresponded to the spelling of the learned novel word, or NO otherwise. Presentation of a single written form at a time in this orthographic decision task prevents the strategy-based responses reported in the orthographic choice task and more naturally taps the word recognition process. Moreover, response accuracy and reaction times were both recorded to increase task sensitivity. Orthographic learning was further evaluated through a spelling task where participants were asked to spell the target novel words. The use of the spelling task with children was debated, but higher performance and better sensitivity was here expected for adults. This task which requires fully specified orthographic information about the target is particularly relevant to assess target-specific orthographic memorization. While the memorization tasks provide off-line measures of orthographic learning, eye movements appear as providing insights on the online learning process.

Most research on eye movements in reading have focused on real word processing in text reading (Rayner, 1998, 2009). Some studies have investigated the effect of a single encounter with pseudo-words on eye movements (Chaffin, Morris, & Seely, 2001; Lowell & Morris, 2014; Wochna & Juhasz, 2013). While there was no pseudo-word effect on fixation location (Lowell & Morris, 2014), more refixations were reported for pseudo-word processing and extra processing-time was needed for pseudo-words as compared to familiar words. These findings are consistent with reports of word frequency effect on eye movements (Rau, Moll, Snowling, & Landerl,

2015; Schilling et al., 1998; Williams & Morris, 2004), suggesting that the pseudo-words were processed as words would be.

Only a few studies have investigated how eye movement measures are affected by repeated exposure to the same novel word in conditions of incidental learning. H. Joseph et al. (2014) explored eye movements while English-speaking adult participants read pseudo-words embedded in sentences. The stimuli were bisyllabic 6-letter pseudo-words that were presented in short stories (i.e., 15 exposures to each of the 16 pseudo-word) during the 5-day sessions of the learning phase. Eye movement tracking revealed an effect of repeated exposure on pseudo-word processing time, shorter first fixation duration and shorter single fixation duration being reported over repeated sessions. Pagan and Nation (2019) measured eye movements as adult participants silently read meaningful sentences containing rare English words that were presented four times during the learning phase. Their results qualitatively show an effect of number of encounters on eye movements, namely shorter processing time across exposures, as previously reported by H. Joseph et al. (2014). Eye movement variations with repeated exposures were also incidentally observed by Gerbier, Bailly, and Bosse (2015, 2018) but using a reading-while-listening paradigm in 6th grade French pupils. In this paradigm, participants processed pseudo-words embedded in meaningful texts while listening to the spoken version of the text through headphones. Twelve 2-syllable pseudo-words were presented four times each in different stories. Online eye movement monitoring revealed a decrease of the first fixation duration and a tendency to shift the first fixation location to the right across repetitions. While the quality of orthographic learning was only assessed through tasks of pseudo-word semantic processing in the two former studies, performance on tasks of both semantic processing and orthographic choice attested that comprehension was good and orthographic learning while reading was efficient in the latter studies.

Overall, eye movement monitoring while reading novel words appears as particularly useful for the real-time study of orthographic learning (Nation & Castles, 2017). However, previous studies mainly focused on the effect of the learning context on eye movements, focusing on the order-of-acquisition (H. Joseph et al., 2014), contextual diversity, spacing and retrieval practice (Pagan & Nation, 2019) or the diversity of semantic contexts (H. Joseph & Nation, 2018). In several studies (Gerbier et al., 2015, 2018; H. Joseph & Nation, 2018; H. Joseph et al., 2014), the number of exposures to targets was varied but none of them provided an in-depth analysis of eye-movement changes from exposure to exposure. Moreover, none of them allowed to distinguish the real effect of an orthographic learning process from a simple item repetition effect. In the present study, we monitored eye movements while participants read both novel words and known words that were presented in isolation and to which they were exposed from one to five times. For the first time, this paradigm allows an exposure-by-exposure tracking of eye-movement changes during novel word orthographic learning, thus providing direct insight on the online learning process. For the first time, comparative analysis of the time course of eye movements on the target novel words and on known real words allowed distinguishing the changes that can be attributed to orthographic learning from those that just reflect an item repetition effect.

The self-teaching hypothesis (Share, 1995, 1999) asserts that phonological decoding is the central mechanism by which orthographic knowledge is acquired. When exposed to a novel word, readers have to decode the word, using their general letter-sound mapping knowledge to generate the novel word's spoken form. The theory posits that every time the novel word is successfully decoded, the reader has the opportunity to memorize its orthographic form. Some computational models that have implemented the self-teaching mechanism (Pritchard et al., 2018; Ziegler et al., 2014) illustrate the critical role of phonological decoding in the development

of orthographic knowledge. This central role is supported by experimental findings showing that accurate decoding is a powerful predictor of incidental orthographic learning (Ricketts, Bishop, Pimperton, & Nation, 2011) and that orthographic learning is lower in conditions of concurrent articulation (de Jong, Bitter, van Setten, & Marinus, 2009; Kyte & Johnson, 2009; Share, 1999). However, despite the unquestionable role of phonological decoding in orthographic knowledge acquisition, there is also evidence that decoding does not guarantee orthographic learning when successful and does not systematically prevent orthography acquisition when unsuccessful (Castles & Nation, 2006, 2008; Nation et al., 2007; Tucker et al., 2016).

Beyond decoding skills, other factors like prior orthographic knowledge and print exposure do predict the quality of orthography learning (Cunningham, 2006; Stanovich & West, 1989), suggesting the involvement of additional orthographic (Cunningham, Perry, & Stanovich, 2001) or visual-orthographic (Share, 2008) processing skills during self-teaching. The impact of visual-orthographic processing on orthographic learning was directly addressed by Bosse et al. (2015) in a self-teaching paradigm. In their study, either the whole letter-string of printed bi-syllable pseudo-words was simultaneously displayed, or the first and second syllables were presented successively one at a time during the learning phase. In both conditions, children were asked to utter the spoken form of the entire pseudo-word after presentation. For the pseudo-words that were accurately decoded, results revealed better orthographic learning when the entire pseudo-word letter-string was simultaneously available for visual processing during the learning phase. Beyond decoding skills, this suggests that orthographic learning further depends on the learner's capacity to attend simultaneously to the whole sequence of letters of the novel word while reading.

A number of previous studies have also provided indirect evidence that orthographic acquisition relates to the capacity for multiletter simultaneous processing, that is, the VA span (Bosse et al., 2007). Individuals with higher VA span show more efficient word recognition skills, thus faster reading (Antzaka et al., 2017; Bosse & Valdois, 2009; Lobier, Dubois, & Valdois, 2013; Valdois, Roulin, & Bosse, submitted), more accurate irregular word reading (Bosse & Valdois, 2009) and smaller length effects (van den Boer, de Jong, & van Meeteren, 2013); for computational modeling, see Ginestet et al. (2019). Exploration of their eye movements during text reading revealed lesser fixations and larger saccades, suggesting that more letters were simultaneously processed per fixation (Prado, Dubois, & Valdois, 2007). Overall, these findings suggest that higher VA span would allow the identification of a higher number of letters simultaneously while reading, which would boost novel word processing and facilitate orthographic memorization. The present study for the first time explores the potential impact of VA span on orthographic learning.

Our main purpose was to provide new insights on the process of implicit orthographic learning. Eye movement monitoring was used during the learning phase to track orthographic learning online through the analysis of cross-exposure effects on eye movements. The paradigm we used – repeated exposition to novel words presented without context– offered the opportunity to track the exposure-by-exposure evolution of eye movements for an in-depth analysis of the online learning process. Furthermore, real words were presented mixed to the target pseudo-words during the reading phase and in similar conditions of eye movement recording for a baseline measure to which pseudo-word processing was compared. Any between-exposure variation of eye movements for words was interpreted as mainly reflecting repetition effects. By comparison, any differential effect for the target pseudo-words would be interpreted as evidence of orthographic learning. The off-line tasks were administered after the learning phase to ensure that any variation of eye movements as a function of the number of exposures might be interpreted as reflecting on-line orthographic learning.

As previously reported in self-teaching paradigms, we expected better performance on the off-line measures of orthographic learning with increased exposure to the target pseudo-words. Better learning would translate to fewer fixations and shorter processing time across exposures during reading. In line with behavioral evidence that orthographic learning occurs from the first exposure (Share, 2004), significant eye movement changes were expected to occur very early, mainly between the first and second exposures. The words of mid-to-high frequency used as baseline were expected to have stable orthographic memories, thus limiting online memorization effects. Accordingly, interactions between the number of exposures (1, 3 or 5) and the item type (words or pseudo-words) are expected on number of fixations and processing time, as a marker of novel word orthographic acquisition clearly different from a simple item repetition effect.

Another goal of the present study was to explore the cognitive skills that affect implicit orthographic learning. We here focused on the potential influence of multiletter simultaneous processing skills on target pseudo-word processing and orthographic learning, through the administration of off-line tasks of VA span. We predicted eye movements and orthographic learning to vary depending on the participants' VA span skills. Participants with higher VA span would be more prone to identify a higher number of letters simultaneously while reading, which would affect both eye movement measures and orthographic learning. As undergraduate students with expert reading skills, participants showed little inter-individual variation in their VA span. They were thus split into two groups with higher or lower VA span for comparison of their eye movements while reading. We predicted that participants with a higher VA span would show higher orthographic memorization skills, thus higher performance in spelling and in orthographic decision. They might further exhibit faster learning over time than lower VA span participants, which might result in faster decrease in number of fixations and processing time across exposures.

2 Method

2.1 Participants

Forty-six undergraduate students ($n_{Female} = 36$) participated in the experiment. All participants were native French-speakers with normal or corrected-to-normal vision and no known learning or reading disorders. They received a 20 € compensation for their participation. Informed written consent was obtained from each participant and the study was approved by the ethics committee for research activities involving humans (CERNI number: 2018-Avis-02-06-01) of the Grenoble-Alpes University. It was thus performed in accordance with the ethical standards of the declaration of Helsinki.

2.2 Protocol

Each participant was first engaged in an incidental learning phase followed by a test phase. During the learning phase, they were asked to read aloud real control words and target pseudo-words that were mixed and displayed 1, 3 or 5 times each on the screen while their eye movements were recorded. To ensure incidental learning, the participants were not informed prior to the learning phase that the purpose of the study was to test their orthographic knowledge of the target pseudo-words. Two tasks of spelling-to-dictation and orthographic decision were used in the test phase to estimate their pseudo-word orthographic learning. The two tasks were

systematically administered in the same order, spelling-to-dictation first. Global and partial letter-report tasks were further administered to measure the participants' VA span.

2.3 The learning phase

2.3.1 Material

Pseudo-words A list of pseudo-words was designed to be used for implicit learning. Thirty legal bi-syllabic 8-letter pseudo-words were generated by reference to the French lexical database LEXIQUE (New, Pallier, Ferrand, & Matos, 2001). They were built up from existing trigrams (mean $f_{trigram}$ = 1,655; SD $f_{trigram}$ = 1,176; range: 327–4,144) to have no orthographic neighbors (i.e., none differed from a real word by a single letter). None was homophone to a real word. Each pseudo-word contained at least two ambiguous graphemes, that is, graphemes corresponding to phonemes that could be written in at least two different ways. As a result, the pseudo-words accurate spellings could not be derived from their phonological form. For example, the target pseudo-word *GOUCIONT* corresponding to the phonological form / gusjɔ̃ / would be written *GOUSSION* when following the most frequent phoneme-to-grapheme mappings in French. Moreover, at least another pseudo-word of the list included the same phoneme but corresponding to another ambiguous grapheme that was not the most often associated to this phoneme in French. For example, the phoneme / ɔ̃ /, spelled *ONT* in *GOUCIONT*, was spelled *ON* in *FAINGION*. As a result, to be correctly spelled, the pseudoword-specific orthographic form had to be memorized. Furthermore, to prevent any systematic visual exploration strategy during the learning phase, the ambiguous graphemes' positions were varied between items, so that they could be located at the beginning, middle left, middle right or at the end of the pseudo-word letter-string.

To investigate the effect of the number of exposures (1, 3 or 5), three different subsets of pseudo-words (noted A, B and C; see Test Lists in 4.1) were created. The three lists were matched in orthographic features – position and number of ambiguous graphemes and mean trigram frequency (respectively, 1627.4, 1668.6, 1668.5). They were further matched in "orthographic difficulty", controlling for the number of possible spellings that complied with the pseudo-word pronunciation.

Control words The thirty control words were 8-letter long with no orthographic neighbors. All were of medium frequency (per million, mean f_W = 35.57; SD f_W = 18.86). As for pseudo-words, the word list was divided into three sublists (noted A, B and C; see Test Lists in 4.2) balanced with respect to their mean frequencies (respectively, 34.44, 36.03, 35.93).

2.3.2 Apparatus

All stimuli were displayed on a screen (DELL, Round Rock, Texas, United States) using a 1680×1050 px resolution and a 60 Hz refresh rate. Stimuli were presented on a black background in white lowercase Courier New font (size 32 px) and experiments were designed with the open-source experiment builder Opensesame (Mathôt, Schreij, & Theeuwes, 2012; v.3.1.2).

During the reading phase, the participants' eye movements were monitored using a RED250 eye tracker (SMI® company, Teltow, Germany) at a viewing distance of 67 cm. The system was interfaced with a laptop (Latitude E6530, DELL, Round Rock, Texas, United States) and gaze position recording was performed by the iViewX software (SMI® company, Teltow, Germany). Each character covered a horizontal visual angle of 0.47°, and each 8-letter stimulus covered a

3.76° visual angle. To minimize head movements, participants kept their forehead pressed on a frontal support.

2.3.3 Procedure

A first listening phase without visual display was administered to each participant. The pronunciation of the target pseudo-words was heard through headphones. The "canonical" pronunciation provided through headphones followed the most frequent grapheme-to-phoneme correspondences. Participants were instructed to listen carefully to the pronunciation of the new words (pseudo-words) as they would have to read them aloud in the second phase. The rationale for providing "canonical" pronunciations was to avoid having different participants refer to different pronunciations.

After the initial listening phase, the eye-tracker was calibrated, using a five-point calibration procedure. Then, the learning phase started. Each trial began with the presentation of a fixation cross. Participants were instructed to keep their eyes on the cross; doing so long enough (randomized time ranged from 400 ms to 600 ms) would trigger the display of the stimulus on the screen.

A stimulus was composed of a control word or a pseudo-word and a digit from 1 to 9 (see Figure 4.1). The two items (word or pseudo-word and digit) were simultaneously presented on the horizontal line, the digit always right of the written item. To avoid any phenomenon of spatial habituation or anticipation, the position of the fixation cross was varied randomly from 204 px to 804 px horizontally left to the center of the screen. The distance between the fixation cross and the letter-string and between the letter-string and the digit was fixed, and set to 9.41 cm (8.03° at a viewing distance of 67 cm).

Participants were asked to read aloud as naturally as possible the written word or pseudo-word, then the digit. Then, they pressed the space bar of the keyboard to trigger the next trial. No feedback was provided during the task. Drift correction and re-calibration of the eye-tracker could be performed at any time by the experimenter if necessary, with drift correction systematically performed every 10 trials.

Ten items, pseudo-words and real words, were displayed once each; ten items were displayed three times each and ten five times; for a total of 180 trials. The trials were presented with a different, random order for each participant. The number of exposures was counterbalanced across subjects, so that, for example, Participant 1 would read the pseudo-words and control words from Lists A, B and C respectively once, three times and five times, whereas Participant 2 would read Lists A, B and C respectively five times, once and three times, etc. The reading task was preceded by 4 practice trials (see Practice Lists in 4.1 and 4.2) during which feedback was provided.

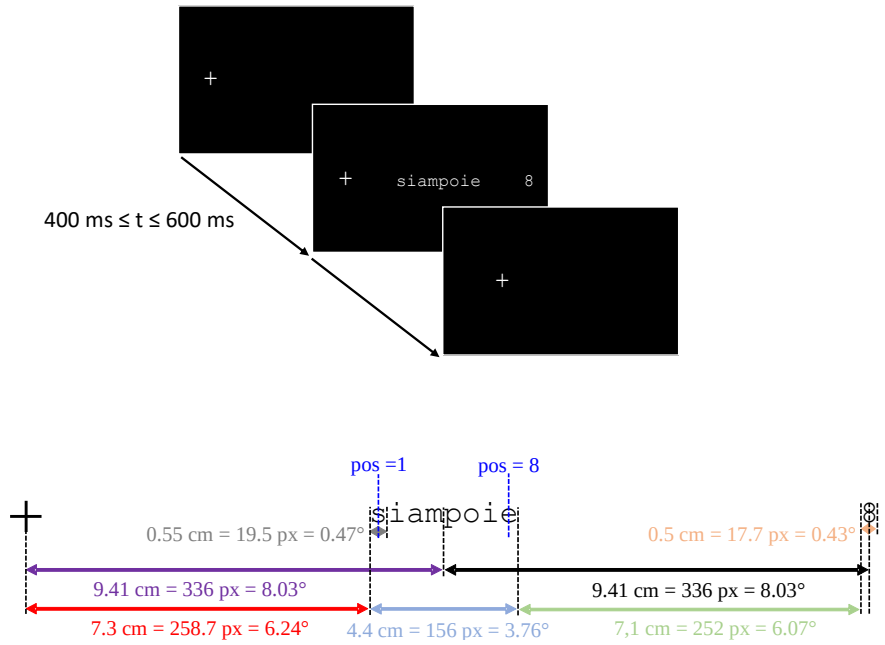

Figure 4.1 – Experimental design. Top: illustration of the time course of a trial during the learning phase. Bottom: distances (in cm, pixels and degrees) between items and reference frame for counting letter positions (illustrated by pos=1 for the first and pos=8 for the last letter.

2.4 Test phase

2.4.1 Unexpected pseudo-word spelling-to-dictation test

Immediately after the learning phase, participants performed a spelling-to-dictation task to measure how well they have memorized the experimental pseudo-words. The oral pronunciation of each pseudo-word was successively presented through headphones, once each, without repetition, and in a random, different order for each participant. Participants had to type each pseudo-word immediately after its pronunciation on the computer keyboard. They were instructed that the pseudo-words were those to which they had been previously exposed to and were explicitly asked to spell them as they were spelled during the learning phase. Overall, 30 pseudo-words (see Test Lists in 4.1) were written by each participant. No time limit was imposed and participants triggered the next trial at their convenience. No feedback was provided to the participants. Each spelling strictly identical to the target pseudo-word orthography was scored as correct.

2.4.2 Orthographic decision test

Target pseudo-words' learning was further assessed through an orthographic decision task. In this task, each target pseudo-word was paired with a pseudo-homophone.

Stimuli Thirty pseudo-homophones (see Test List in 4.3) were generated following the same criteria as for the pseudo-words. The pseudo-homophones were matched to the pseudo-words

in trigram frequency (mean $f_{trigram}$ = 1,391; SD $f_{trigram}$ = 1,040; range: 307–3,829). The written form of each pseudo-homophone included at least one ambiguous grapheme that differed from that of the related target pseudo-word (e.g., *SAITTEAU-CEITTEAU*, *PHACRAIS-PHACRAIT*, *SIEMPOIT-SIAMPOIE*).

Protocol Each item (target pseudo-word or pseudo-homophone) was successively displayed on the computer screen one at a time. The participants were asked to decide as quickly and as accurately as possible whether the displayed item was spelled as the target pseudo-word or not. Each trial began with a fixation cross displayed at the center of the computer screen during 500 ms. The fixation cross was immediately replaced by a forward mask composed of eight hash marks (########) for 500 ms. The target was then displayed, centered on fixation, until the participant's response.

Presentation order was randomized and different for each participant. No feedback was provided to the participants. The two possible response buttons were the left button (for a "Yes" response) and the right button (for a "No" response) of a serial response (SR) box. A total of 60 targets were presented: the 30 target pseudo-words read during the learning task (see Test Lists in 4.1) and the 30 pseudo-homophones (see Test List in 4.3). The task was preceded by 4 practice trials (see Practice List in 4.1 and 4.3) during which feedback was provided. Accuracy and reaction times were recorded.

2.5 VA span tasks

The participants were administered global and partial report tasks to estimate their VA span. The tasks followed the protocol described by Antzaka et al. (2017) which we only summarize in the following. At each trial, a 6-consonant string was briefly displayed centered on the fixation point. The consonant string was built-up from 10 consonants (B, P, T, F, L, M, D, S, R, H) with no repeated letter. The consonant string was presented in black uppercase Arial font on a white background. Inter-letter spacing was increased to avoid crowding (0.57° inter-consonant space). Twenty-four strings were successively presented in the global report task, 72 in the partial report task. In both tasks, trials began with a fixation dot displayed at the center of the screen during 1 s, immediately followed by a blank screen during 50 ms. The 6-consonant string was then displayed for 200 ms.

In the global report task, participants were told to verbally report as many letters as possible immediately after string presentation, regardless of the letter position in the string. In the partial report task, a single vertical bar appeared for 50 ms (1.1° below one of the letter positions) at the offset of the consonant string, indicating the position of the letter to be reported. The experimenter typed the participants' response without providing any feedback. The experimenter then proceeded to the next trial by pressing the Enter key. Both tasks were preceded by 10 practice trials during which feedback was provided. The strings were displayed in a random order that differed for each participant.

Accuracy was recorded for the two tasks of VA span as the number of target letters accurately reported, which induces a maximum score of 144 (for 24 trials × 6 letters per trials) and 72 (one letter per trials) for the global and partial report task respectively. In order to give the same importance to each task, the total score of VA span (TS_{VAS}, expressed as a percentage) was calculated, for each participant, using the following relation:

$$TS_{VAS} = \frac{(Global_{score} + 2 \times Partial_{score}) \times 100}{2 \times 144} \tag{4.1}$$

3 Results

We first consider results of the orthographic memorization tasks to evaluate the incidental learning of the pseudo-words' orthographic form. Next we analyze eye movements during the reading phase, focusing on how they are affected by the number of exposures to target pseudo-words. Last, we compare the eye movement patterns of the participants with lower or higher visual attention (VA) span to explore whether and how VA span affected orthographic learning while reading.

3.1 Statistical models

The data were analyzed by means of generalized linear mixed effects (glme) modeling using the `lmer` function for reaction times (RT) data and the `glmer` function for score data, from the `lme4` package of the R computing environment (R Core Team, 2018; RStudio version 1.0.143). Participants and items were introduced as random factors. In order to analyze normally distributed data, all RTs and temporal data – first-fixation duration, second-fixation duration and gaze duration – were log-transformed and a Poisson regression modeling was performed on the number of fixations before statistical analyses.

All models related to orthographic memorization tasks included the number of exposures as fixed factor. In all models related to eye-tracking data, number of exposures, item type and their interaction were specified as fixed factors. In order to distinguish the number of exposure effect from the trial order effect, the order of presentation of items (continuous variable from 1 to 180) was also included in the models as covariable. Finally, since we explored the effect of VA span on all measures only using pseudo-words data, number of exposures, VA span group and their interaction were specified as fixed factors.

Post-hoc comparisons were performed using Tuckey contrasts in orthographic memorization tasks – spelling-to-dictation and orthographic decision. For all eye-tracking measures, post-hoc analyses modeling local interaction between the first and second exposure and between the second and third exposure were performed using similar linear mixed models than previously described.

In order to report statistical results from models that converged successfully (Bates, Kliegl, Vasishth, & Baayen, 2015; Matuschek, Kliegl, Vasishth, Baayen, & Bates, 2017), the order of presentation of items was not included in the model for the analysis of the number of fixations and random factor "item" was not included in the model for the analysis of local interactions as well as in all models related to accuracy in the dictation task.

Finally, data from all the tasks performed by 42 participants were analyzed, due to the exclusion of four participants following clean-up of eye movement data (see subsection 3.3.1).

3.2 Performance in Orthographic memorization

Results from both the target pseudo-word spelling-to-dictation and orthographic decision tasks are summarized in Table 4.1. Results show a significant main effect of the number of exposures on performance in the spelling-to-dictation task ($z = 3.93$, $p < .001$). The pseudo-words were more accurately spelled when they had been more often encountered during the reading phase. Post-hoc comparisons showed that participants made more accurate spellings after five exposures to the same pseudo-word than after a single exposure ($z = 3.91$, $p < .001$). Spelling performance did not significantly differ from one to three exposures ($z = 2.11$, $p = .087$) or from three to five exposures ($z = 1.87$, $p = .146$). Spelling results suggest that incidental orthographic learning was effective during the reading phase.

Table 4.1 – Performance on the target pseudo-words (mean and standard deviation) as a function of the number of exposures for the spelling-to-dictation and the orthographic decision tasks.

Exposures	PW spelling-to-dictation	Orthographic decision			
		Targets PW		Homophones	
	% correct	RTs	% correct YES response	RTs	% correct NO response
1	11.4 (9.0)	1231 (470)	54.4 (15.6)	1284 (492)	58.0 (17.1)
3	16.4 (15.0)	1114 (419)	72.5 (16.3)	1284 (500)	48.4 (17.5)
5	21.4 (16.2)	1007 (254)	78.7 (13.3)	1229 (395)	50.6 (16.5)

With respect to the orthographic decision task, trials with log-duration more than 2.5 standard-deviations from the mean were excluded from analyses (1.35 % of the trials). Both the percentage of correct responses and RTs obtained in the orthographic decision task are presented in Table 4.1. In the absence of orthographic learning, choices in orthographic decision were expected to be merely random (around 50 %). A one-sample t-test showed that the proportion of correct choices (combining all exposure levels) was significantly above chance for the target pseudo-words ($t = 11.24$, $p < .001$) but not for their homophones ($t = 1.32$, $p = .195$). The above-chance-level performance on the targets was found for both the three-exposure ($t = 8.92$, $p < .001$) and the five-exposure ($t = 14.04$, $p < .001$) conditions.

Further analyses were performed on response accuracy and RTs but only for the target pseudo-words, due to random performance on the homophones. An effect of the number of exposures was expected on the two measures if orthographic learning occurred during the reading phase. Indeed, target pseudo-words were recognized more accurately ($z = 7.97$, $p < .001$) and faster ($t = -7.44$, $p < .001$) across exposures. Post-hoc comparisons show that the number-of-exposures effect on both response accuracy and RTs was significant from one to three exposures (accuracy: $z = 5.69$, $p < .001$; RTs: $z = -3.94$, $p < .001$) and from three to five exposures (accuracy: $z = 2.40$, $p = .044$; RTs: $z = -3.71$, $p < .001$). Above-chance-level performance on the target pseudo-words and more efficient pseudo-word recognition (accuracy and RTs) across exposures both support that orthographic learning did occur during the reading phase. Evidence for significant learning effects after only three exposures suggest that three or perhaps two exposures are sufficient for orthographic learning.

3.3 Analyses of eye movements during the learning phase

3.3.1 Data cleaning

Although we recorded eye-tracking data for the two eyes, only data from the right eye were analyzed. Parsing raw data from the eye-tracker (i.e., event detection) was done using the BeGaze software (SMI® company, Teltow, Germany). The High-Speed algorithm was used, keeping default parameter values: minimum saccadic duration was 22 ms, peak speed threshold of a saccade was 40°/s and minimum fixation duration was 50 ms. A custom-designed software was further used to select the relevant fixations for subsequent analyses. Visual inspection of the vertical coordinates of each participant led to remove those fixations located too far from the horizontal line. Using this method, we removed 1.71 % of the total number of fixations (835 fixations out of 48,865 fixations). To explore eye movements during online orthographic learning, we focused the analyses on the oculomotor events occurring during the first-pass of letter-string processing. The oculomotor events were taken into account for further analyses if

occurring within an area of interest defined as follows: fixations further than 30 px (0.72°) to the left of the left boundary of the first letter, or further than 30 px (0.72°) to the right of the right boundary of the last letter, were excluded from analyses.

For all analyses, trials that showed more than one blink on the target pseudo-word were excluded. We further excluded trials with more than 8 saccades or 9 fixations that were considered as extreme outliers according to a quartile-based criterion (more saccades than $Q_3 + 3 \times (Q_3 - Q_1)$). Trials with total log-duration further than 2.5 standard-deviations from the mean were further excluded (range: 79–2,402 ms). Finally, trials containing at least one fixation log-duration greater than 2.5 standard-deviation from the mean (i.e., more than 1,055 ms) were excluded from the analyses. The whole data from four participants were discarded due to the low percentage of trials remaining after data cleaning ($< 65\%$). Overall, 90.66 % of the trials (6854 out of 7560 trials, see Table 4.2 for details) from 42 participants were kept for further analyses. Learning effects were first explored on the whole population before focusing on potential differential effects depending on the participants' VA span abilities.

Figure 4.2 summarizes results for the different eye movement measures of interest across exposures: number of fixations, gaze duration (the sum of all first pass fixations made on the item), and fixation duration. Fixation durations were analyzed separately for trials characterized by a single fixation and for trials with multiple fixations. Most previous studies (e.g., Pagan & Nation, 2019) that distinguished single fixation trials from multiple fixation trials focused on the analysis of the first-of-multiple fixation while ignoring the number and duration of subsequent fixations. We here decided to restrict the analysis to a more homogeneous corpus of trials, considering trials with exactly one or two fixations. As shown on Table 4.2, two-fixation trials predominated in the corpus. This allowed an in depth analysis of both the first and second fixation duration.

For all eye movement measures, results are provided for the target pseudo-words and the control words as a function of the number of exposures. A differential exposure effect for the target pseudo-words and control words was expected as a marker of new word orthographic learning. As performance on memorization tasks provided evidence for effective orthographic learning occurring across the first three exposures, we were in particular attentive to differences in oculomotor measures between the first and second or third exposure.

Number of fixations As shown on Figure 4.2(A), the number of fixations recorded during the reading phase varied depending on item type (main effect: $z = 12.44$, $p < .001$) and number of exposures (main effect: $z = -3.82$, $p < .001$) . More importantly, the item-type by number-of-exposure interaction was significant ($z = -2.21$, $p = .027$), showing a steeper decrease in number of fixations across exposures for pseudo-words than for words. Post-hoc analyses show no local interactions, suggesting a gradual decline over the whole five exposures.

Gaze Duration Gaze Duration for target pseudo-words and control words, as a function of the number of exposures, is plotted in Figure 4.2(B). The two main effects of item-type ($t = 15.30$, $p < .001$) and number-of-exposures ($t = -8.25$, $p < .001$) were significant, as well as the interaction between these two factors ($t = -4.53, p < .001$) showing steeper gaze duration decrease with increased exposure for pseudo-words than for words. Post-hoc analyses show that the interaction was mainly driven by sharper gaze duration decline between the first and second exposures ($t = -2.00$, $p = .046$; non significant interaction between the second and third exposure: $t = -1.56$, $p = .119$).

Table 4.2 – Number and percentage of trials per item type and exposure (in columns: W for words, PW for pseudo-words, and exp1 to exp5 for the five exposures) and number of fixations (in rows). The total indicates the number and the percentage of trials kept after removing outliers.

	Size	**PW exp1**	**PW exp2**	**PW exp3**	**PW exp4**	**PW exp5**	**W exp1**	**W exp2**	**W exp3**	**W exp4**	**W exp5**
1 fixation	N	152	149	145	76	91	382	313	310	163	160
	%	12.1	17.7	17.3	18.1	21.7	30.3	37.3	36.9	38.8	38.1
2 fixations	N	358	288	318	173	173	515	346	351	179	177
	%	28.4	34.3	37.9	41.2	41.2	40.9	41.2	41.8	42.6	42.1
3 fixations	N	276	176	175	84	74	186	83	80	35	38
	%	21.9	21.0	20.8	20.0	17.6	14.8	9.9	9.5	8.3	9.0
4 fixations	N	168	86	71	35	25	48	22	22	8	9
	%	13.3	10.2	8.5	8.3	6.0	3.8	2.6	2.6	1.9	2.1
$\geq$ 5 fixations	N	138	50	50	22	14	29	16	8	4	3
	%	11.0	6.0	6.0	5.2	3.3	2.3	1.9	1.0	1.0	0.7
Total	N	1092	749	759	390	377	1160	780	771	389	387
	%	86.7	89.2	90.4	92.9	89.8	92.1	92.9	91.8	92.6	92.1

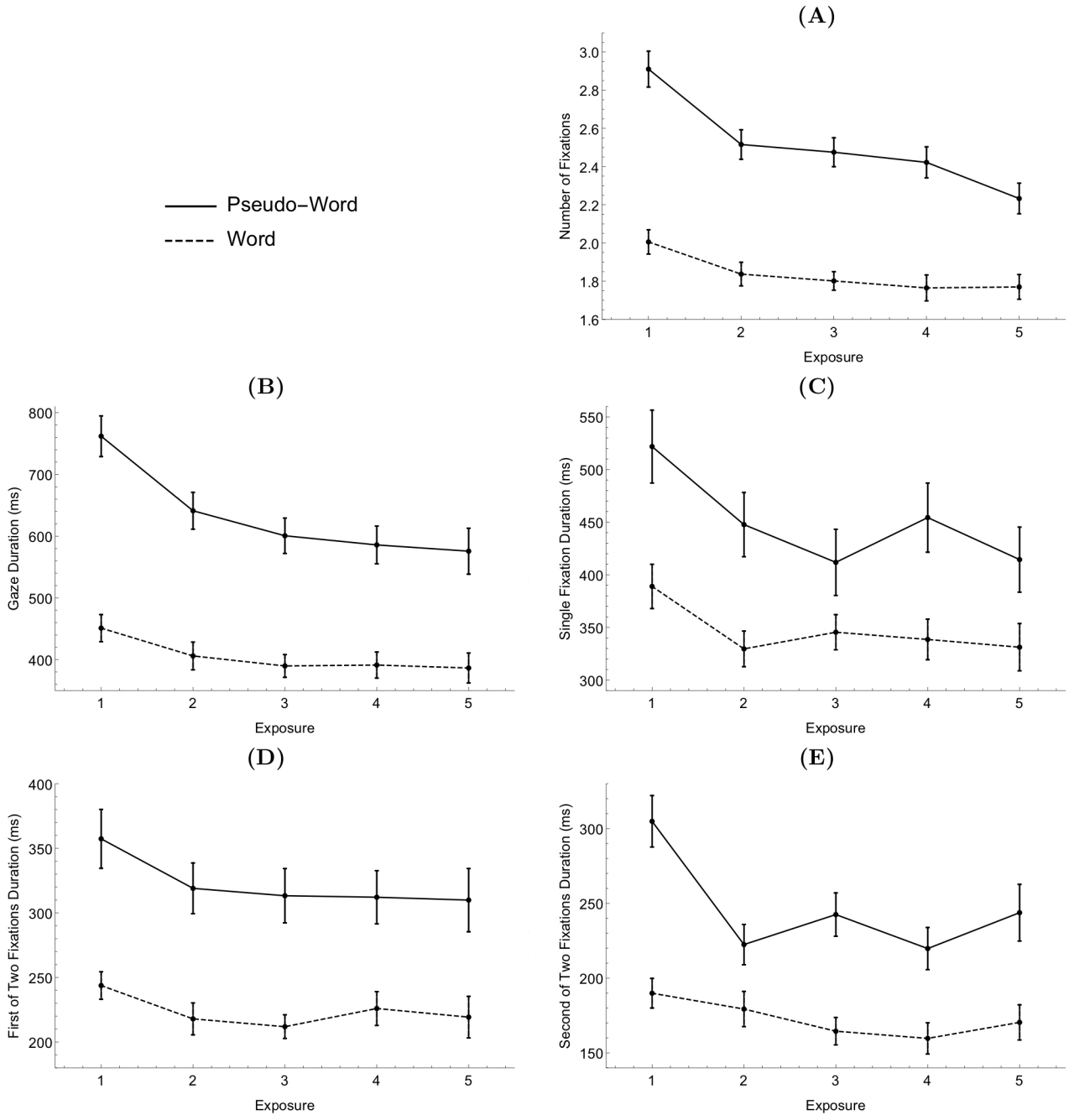

Figure 4.2 – Eye movement measures recorded during the reading task as a function of the number of exposures. **(A)** Number of fixations, **(B)** Gaze duration, **(C)** Single fixation duration, **(D)** First-of-two fixation duration and **(E)** Second-of-two fixation duration. All times are in ms and vertical bars are for standard errors.

Fixation Duration in Single Fixation Trials The analysis of single fixation duration according to item-type and number-of-exposures was performed on the 1,941 trials (25.7 % of data) having a single fixation (see Table 4.2 for details). Results are presented in Figure 4.2(C). The analysis showed longer single fixation duration for pseudo-words than for words ($t = 7.23$, $p < .001$) and a significant main duration decrease across exposures ($t = -4.28$, $p < .001$). The item-type by number-of-exposure interaction was not significant ($t = -0.91$, $p = 0.365$), suggesting that the duration of single fixations similarly decreased for words and pseudo-words across exposures. Local interaction analyses reveal a similar decrease of fixation duration for words and pseudo-words between the first and second exposure ($t = 1.82$, $p = .069$) but a sharper decrease on target pseudo-words between the second and third exposure ($t = -3.37, p < .001$).

Fixation Duration in Two-Fixation Trials Two-fixation trials represent 2,878 trials or 38.1 % of data (see Table 4.2 for details). Figure 4.2(D) and Figure 4.2(E) illustrate the number-of-exposure effect on target pseudo-words and control words for the first and second fixation of the two-fixation trials respectively. Results show that the first-of-two fixation duration lasted longer for pseudo-words than for words ($t = 9.35$, $p < .001$) and duration decreased across exposures ($t = -3.35$, $p < .001$). The item-type by number-of-exposure interaction was not significant ($t = -0.90, p = .369$). Post-hoc analyses show that none of the local interactions between the first and second ($t = -0.83, p = .409$) or second and third ($t = 0.02, p = .985$) exposure was significant.

Analysis of the second-of-two fixation duration showed a main effect of item-type ($t = 10.50$, $p < .001$) and number-of-exposures ($t = -3.50$, $p < .001$), but no significant interaction between item-type and number-of-exposures ($t = -1.72$, $p = .086$). However, there was a significant local interaction between the first and second exposure ($t = -3.38$, $p < .001$) but not between the second and third exposure ($t = 1.74$, $p = .082$), showing sharper decline of the second-of-two fixation duration between the first two exposures to target pseudo-words.

3.4 Exploratory analysis of potential VA Span Effect

We further evaluated the potential influence of the participants' VA span abilities on orthographic learning (spelling-to-dictation and orthographic decision) and eye movement measures for target pseudo-words. To address the potential influence of VA span, participants were split into two groups with higher or lower VA span, using the median value (see Equation 4.1 for TS_{VAS} calculation; $TS_{VAS}(med) = 79$ % ; $TS_{VAS}(mean) = 78.2$ % ; $TS_{VAS}(SD) = 7.9$ %). The two groups were characterized by a mean VA span performance of 71.8 % (SD = 4.6 %) and 84.6 % (SD = 4.8 %) respectively.

Figure 4.3 illustrates performance on the orthographic decision task and on eye movement measures for which significant VA span Group effects were reported, namely RTs from the orthographic decision task, gaze duration and second-of-two fixation duration from the learning phase. No significant Group effect or Group by number-of-exposures interaction was found for any of the other memorization or eye movement measures.

3.4.1 Memorization tasks

We first checked whether orthographic learning was more effective depending on the participants' VA span. Results from the spelling-to-dictation task revealed no significant Group effect ($z = 1.73$, $p = .085$) on accurate spellings and no interaction between the Group

and the number-of-exposures ($z = -1.42$, $p = .155$). Similarly, no significant Group effect ($z = -0.28$, $p = .779$) and no interaction ($z = 0.57$, $p = .569$) were found in orthographic decision accuracy. However, The Group effect was significant on RTs, see Figure 4.3(A). Participants with a higher VA span recognized target spellings faster than participants with a lower VA span ($z = -3.11$, $p = .003$). The group effect on RTs did not interact with the number-of-exposures ($z = 1.30$, $p = .193$).

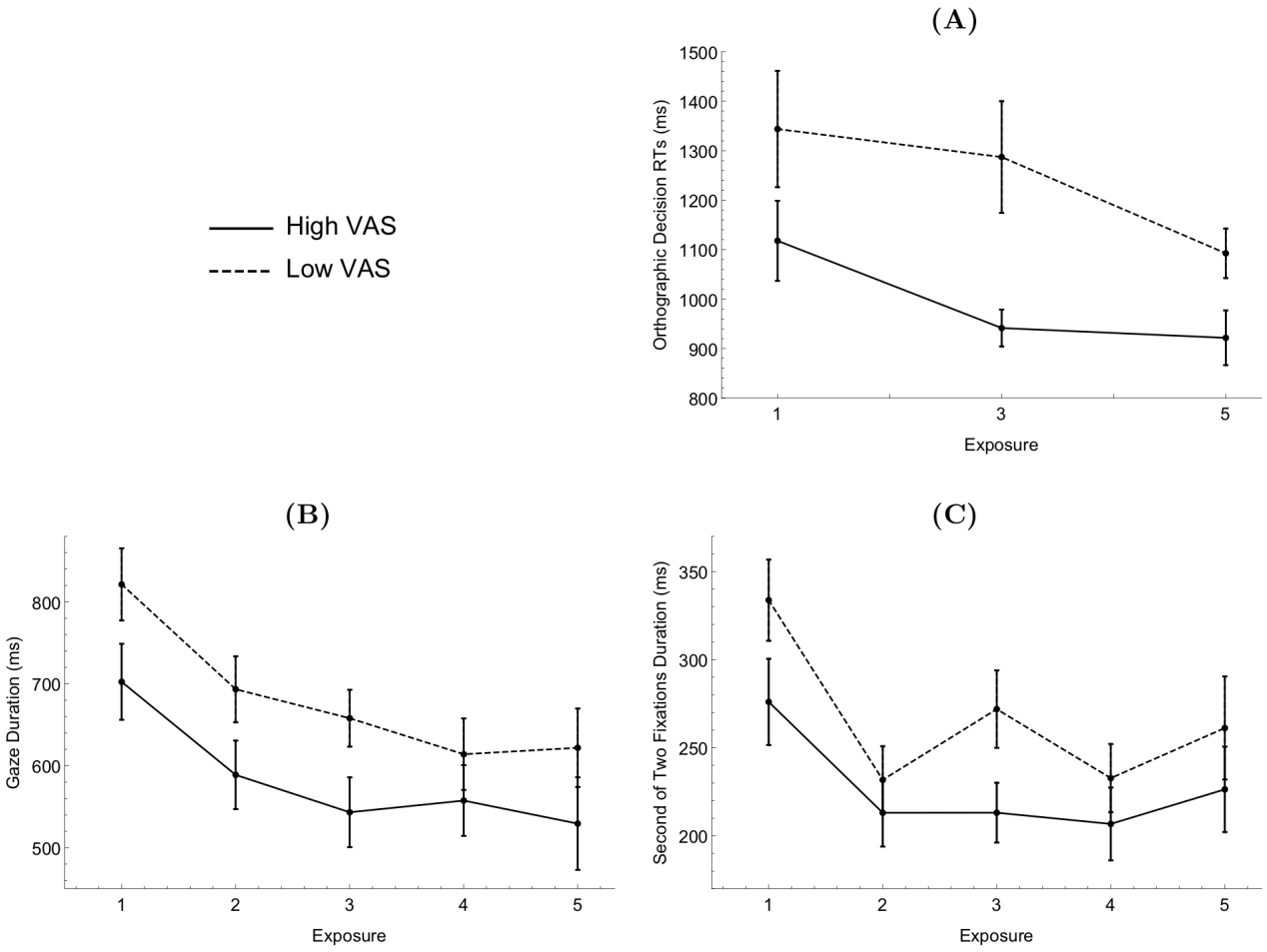

Figure 4.3 – Results represented as a function of the number of exposures and VA span group. From top left to bottom right: **(A)** RTs in ms from the orthographic decision task, **(B)** Gaze duration in ms from the learning phase and **(C)** Second-of-two fixation duration in ms from the learning phase. Vertical bars are for standard errors.

3.4.2 Eye movement measures

Figure 4.3(B) represents gaze duration for the two VA span groups as a function of the number of exposures. Gaze duration was significantly longer for participants with lower VA span ($t = -2.02$, $p = .050$) but the two VA span groups showed very similar decrease in gaze duration across exposures ($t = -0.17$, $p = .867$).

No main effect of VA span group on fixation duration was found either for single fixation trials ($t = -0.64$, $p = .527$) or for the first-of-two fixation trials ($t = -1.09$, $p = .281$). There

was no Group by number-of-exposures interaction either (respectively, $t = -0.79$, $p = .432$ and $t = -0.63$, $p = .530$). However, as illustrated on Figure 4.3(C), the second-of-two fixation duration was significantly longer for the group with lower VA span ($t = -2.37$, $p = .021$). The group effect was similar across exposures ($t = 1.08$, $p = .279$).

4 Discussion

In this study, we investigated incidental learning of new words using eye-tracking. Through a paradigm inspired from the self-teaching paradigm initiated by Share (1999), French-speaking skilled readers were exposed to novel words, presented in isolation for one, three or five exposures. A conventional spelling-to-dictation task and an original orthographic decision task were used to measure orthographic learning after the reading phase. Further, VA span tasks were administered to investigate the potential impact of visual attention capacity on the processing of new words while reading and as a result, on the orthographic learning process.

Some of the current behavioral and eye-tracking findings are in line with previously reported data, showing that the paradigm we used was appropriate to provide insights on the speed and quality of orthographic learning. Such evidence is threefold. First, results from the offline orthographic memorization tasks suggest that incidental orthographic learning began very early during processing, namely across the three first exposures to the novel word. Fast orthographic learning is convergent with findings in previous research (Nation et al., 2007; Share, 2004). However, low accuracy performance in spelling-to-dictation suggests that memorization of the perfect whole spelling of a new 8-letter word is a longer process which is not completed after five exposures to the new written form. Spelling was consistently reported as a difficult task, less prone to highlight orthographic learning, in particularly in children (Bosse et al., 2015; Cunningham, 2006; Share, 1999). Second, in line with the previous eye-tracking studies that compared word and pseudo-word processing (Lowell & Morris, 2014; Rayner, 2009), we show that the number of fixations is higher for pseudo-words than for words. The item-type effect is quite consistent and extends to all the eye movement measures. These findings suggest that lack of semantic and orthographic representations for pseudo-words prevents lexical access which disrupts first-pass processing and leads to longer processing times. Third, we show that the higher the number of exposures to the target pseudo-word, the lower the processing time. There is strong agreement that visual processing of a new word is influenced by the number of encounters with this word (see H. Joseph & Nation, 2018; H. Joseph et al., 2014 and, for qualitatively consistent observation, see, Pagan & Nation, 2019). However, variations in the oculomotor measures carried out on real words also suggest some sensitivity to the number of exposures. Given the range of frequency of our control words, it seems difficult to assimilate these variations to any orthographic learning. This effect may be more likely due to repeated reading of the same set of well-known words in a short time, inducing familiarization to the set of stimuli through repetition. If that were the case, such a process could also account for some of the variations of the oculomotor measures for the pseudo-words. The use of control words is thus critical to disentangle repetition effects from learning effects. Comparison of exposure effects on words and pseudo-words in the current study revealed higher decrease in processing time – as evidenced by the number of fixations and gaze duration– over exposures for pseudo-words than for words. This finding clearly shows that a repetition effect cannot account alone for the oculomotor behavior observed on pseudo-words. Evidence for differential processing time during pseudo-word processing attests that decreased processing time with increasing exposures does reflect orthographic learning.

In the following, we focus on the novelty of the current findings, underlying the relevance of the orthographic decision task and exposure-by-exposure track of eye-movements during the learning phase to reveal orthographic learning. We last discuss how VA span influences the learning process.

A first main contribution of the current study is to provide novel ways of measuring orthographic learning. Based on previous work of Wang et al. (2011), Tamura et al. (2017) and Wang, Nickels, Nation, and Castles (2013), we designed a new orthographic decision task as an alternative to the standard but more debated task of orthographic choice (Castles & Nation, 2008; Tucker et al., 2016). Instead of using a multiple choice paradigm as in the standard task, targets and homophonic foils were presented one-at-a-time to prevent strategic influences, like responses based on phonological and/or orthographic comparison. It was assumed that the processing of isolated items, targets or foils, in orthographic decision would more directly trigger the recognition system and would increase sensitivity to orthographic learning. Current evidence suggests that the novel task likely reflects orthographic learning. First, better and faster target pseudo-word recognition was reported across exposures. Second, reported effects on both accuracy and RTs between the first and third exposure suggest good task sensitivity to the early steps of orthographic learning. Last, significant changes between the third and fifth exposures suggests sensitivity all along the learning process. The orthographic decision task thus appears as a promising offline alternative to track the evolution of implicit learning during reading.

Orthographic learning was further assessed in our adult participants through a spelling-to-dictation task. Despite relatively low performance on this task, target pseudo-words' accurate spelling improved with the number of exposures. Contrary to previous evidence from children (Bosse et al., 2015; Cunningham, 2006), this suggests that the task may be useful to evaluate orthographic learning in adults. However, significant effects on spelling performance were only reported between the first and fifth exposure, suggesting limited sensitivity to the learning process, even in adults. This lack of sensitivity contrasts with previous evidence of learning effects from the third exposure in the orthographic decision task. This finding provides further support to the orthographic decision task, that appears as more sensitive to orthographic learning than the spelling-to-dictation task. Whether the novel orthographic decision task challenges the widely used orthographic choice task warrants further investigation, through direct comparison of performance in the two tasks. Future studies would further explore whether the orthographic decision paradigm is appropriate to measure orthographic learning in children.

A second main contribution of our work is to provide insights on the time course of orthographic learning through the exposure-by-exposure analysis of eye movements during the learning phase. Although changes in eye movements across different phases of repeated print exposure to novel words was examined in previous research (see H. Joseph & Nation, 2018; H. Joseph et al., 2014 and, for qualitatively consistent observation, see Pagan & Nation, 2019), our study is the first to provide an exposure-by-exposure, in-depth analysis of this online process. The current findings suggest such analysis is particularly relevant and informative. While the offline measures of orthographic memorization suggest relative quick learning of the novel words (after only a few exposures), the exposure-by-exposure exploration of the learning process likely highlights very early evidence of orthographic learning. Such exploration further allows investigating whether and how the different eye movement measures respond to additional exposures during implicit learning. The analysis of eye movements in our adult participants revealed early changes with additional exposure on the two measures of gaze duration and single fixation duration. A sharper duration decline was observed between the first and second exposure to the novel word for the former and between the second and third exposure for the

latter. Thus, specific changes in time processing for pseudo-words that can be attributed to orthographic learning occur during the very first stages of the learning process. These findings support prior evidence of rapid orthographic learning (Cunningham, 2006; Share, 2004). Robust learning starts after just one exposure not only in transparent languages as Hebrew (Share, 2004) but also in non transparent languages as English (Cunningham, 2006) or French (current data). Rapid and automatic orthographic learning thus appears as a property of the learning system that is independent of language transparency. Current findings further confirm that efficient early orthographic learning can occur in the absence of semantic context (Landi, Perfetti, Bolger, Dunlap, & Foorman, 2006; Nation et al., 2007).

Another originality of the current work was to capitalize on the predominance of two fixation trials for an analysis of both the first and second fixation in conditions where novel words were fixated twice and only twice. Most previous studies in the field analyze the duration of the first of multiple fixations without providing insights on later processing (Bertram, 2011), see however Mousikou and Schroeder (2019) for an attempt in an eye tracking study on morphological processing. An analysis of the first-of-multiple fixation duration was actually carried out in the present study; results were not reported here as they were very similar to those reported for the first-of-two fixation. Interestingly, the analysis showed no specific effect of orthographic learning on the first-of-two fixation duration but a significant effect on the duration of the second-of-two fixation, suggesting that analyses restricted to the first-of-multiple fixation may lack sensitivity to the learning process. Lack of item-specific exposure effect on processing time for the first-of-two fixation suggests that the gaze duration decline previously reported between the first and second exposure would primarily reflect decline in the second-of-two fixation duration.

Overall, the exposure-by-exposure analysis of eye movements reveals an early effect of orthographic learning on gaze duration, single fixation duration and the second-of-two fixation duration. Fixation number is a further eye movement measure that is sensitive to orthographic learning but the number of fixations gradually declines across exposures without any specific effect emerging from the first encounters with the novel word. Decrease in processing time from the first to the second exposures suggests that duration eye-movement measures are mainly affected by the creation of an orthographic trace of the target novel word in memory. Longer processing time at the first exposure is likely to reflect time needed for bottom-up information extraction on letter identity and creation of a specific orthographic representation in memory. As a result, the major changes observed in processing time between the first and second exposure may primarily reflect improved letter identification at the second exposure, due to top-down influence from the newly acquired orthographic representation. There is no evidence that further exposures to the same novel word have strong impact on processing time. However, increase in novel word familiarity over repeated exposures may facilitate word recognition thus resulting in a gradual decrease in number of fixations over time. Thus, the exposure-by-exposure analysis of eye movements during the learning phase appears as particularly useful to better understand the critical links between eye-movement patterns and the orthographic learning process. Moreover, contrary to previous research that minimized the impact of visual processing on learning (Share, 1999; Ziegler et al., 2014), current findings clearly argue for a major role of visual processing on new orthographic knowledge acquisition.

A last novelty of the current study was to explore whether differences in VA span do affect orthographic learning. For this purpose, our set of participants was split into two groups with higher or lower VA span skills. Comparison of the two groups' performance on the offline measure of orthographic decision revealed that adult readers with a higher VA span recognized novel word spellings faster than their lower VA span peers. The higher VA span group further showed shorter gaze duration and shorter second-of-two fixation duration, thus showing a VA

span effect on two of the eye-movement duration measures previously identified as sensitive to orthographic learning. However, the number-of-exposure effect was similar in the two groups, thus independent of the participants' VA span.

Although the role of VA span in orthographic learning requires further investigations, the current findings suggest that the amount of visual attention available for multiletter processing increases the likelihood that a novel word orthographic representation will be established in memory. The rational for such assumption is that the more attention is allocated to the novel word, the faster is letter identity processing, which boosts relevant letter information encoding and facilitates orthographic learning. This aligns with previous evidence that the amount of attention available for novel word processing is critical to orthographic knowledge acquisition (Landi et al., 2006) and that readers with higher VA span read faster (Antzaka et al., 2017; Bosse & Valdois, 2009; Lobier et al., 2013) and show more efficient lexical reading skills (van den Boer et al., 2013). Better spelling performance in higher VA span readers further supports a contribution of visual attention to efficient orthographic knowledge acquisition (van den Boer, Elsje, & de Jong, 2015). While the assumption of a causal link between VA span skills and reading acquisition has been debated (Goswami, 2015; Lobier & Valdois, 2015), recent report that pre-readers VA span is a significant and independent predictor of future reading skills (Valdois et al., submitted) and that reading improves following specific VA span training (Zoubrinetzky, Collet, Nguyen Morel, Valdois, & Serniclaes, 2019) do support a causal relationship. The overall findings have strong implications for models of learning to read, suggesting that visual attention is critical for the acquisition of word-specific orthographic knowledge, which allows beginning readers to shift from slow phonological decoding to fast word recognition and reading by sight.

Acknowledgments

This work has been supported by a French Ministry of Research MESR grant for EG.

Conflict of Interest Statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix: List of items

1. Pseudo-words used in the learning phase, the spelling-to-dictation of pseudo-words and the orthographic decision: List A: broufand, chaiquau, deinrint, faingion, gouciont, nauplois, ploitart, speirain, quinsard, tramoint; List B: ceitteau, chanquet, coirtint, drottont, flommais, glounein, priquoin, quarlant, siampoie, trimpond; List C: bussiond, cherrein, ciercard, claffand, fentroit, phacrait, prinnant, tauppart, scrodain, trancare

2. Control words used in the learning phase: List A: uniforme, portrait, enceinte, mouchoir, surprise, complice, scandale, chanteur, immeuble, avantage; List B: physique, revanche, horrible, boutique, sensible, fauteuil, chocolat, mensonge, solution, voyageur; List C: prochain, grandeur, nocturne, lointain, religion, empereur, division, quartier, province, jugement

3. Homophones used in the orthographic decision: broufant, cheicaud, dainrein , phingion, goussion, nopploie, ploittar, spairint, quainsar, trammoin, saitteau, chenquai, quoirtin, drautond, flaummet, glounnin, pricoint, quarland, siempoit, trimppon, busciont, chairain, siercart, claphant, phantroi, phacrais, prinnand, thauppar, scraudin, trenquar

Chapitre 5

BRAID-Learn : un modèle computationnel d'apprentissage de l'orthographe lexicale

Note

Ce Chapitre reprend un manuscrit en préparation.

Résumé

Si de nombreux modèles computationnels modélisent la reconnaissance de mots, la lecture de mots isolés ou encore les mouvements oculaires au cours de la lecture de textes, peu de modèles proposent une implémentation du processus d'apprentissage orthographique. Récemment, deux modèles simulant l'acquisition de nouvelles traces orthographiques ont été développés (Pritchard et al., 2018; Ziegler et al., 2014). Ces modèles, développés à partir des modèles double-voie, implémentent l'hypothèse d'auto-apprentissage de Share (1995, 1999, 2004). Ils postulent un rôle majeur des traitements phonologiques dans l'apprentissage orthographique, supposant un rôle secondaire des traitements visuels. Dans ces modèles, le décodage phonologique réussi d'un nouveau mot garantit la mémorisation « parfaite » (c'est-à-dire, sans erreur) de sa représentation orthographique. Finalement, la sous-spécification des traitements visuels implémentés ne leur permet pas de rendre compte, notamment, des difficultés d'apprentissage associées à un déficit visuo-attentionnel ni même de la variation des mouvements oculaires observés au cours de la lecture répétée de nouveaux mots.

Notre objectif principal est de proposer un nouveau modèle computationnel d'apprentissage orthographique, le modèle *BRAID-Learn*, permettant de rendre compte à la fois de l'apprentissage de l'orthographe lexicale et des mouvements oculaires en lecture de mots et de nouveaux mots. Dans cet article, nous présentons la structure du modèle *BRAID-Learn* et les hypothèses implémentées. Grâce à une description mathématique de ses mécanismes, nous décrivons le processus d'apprentissage orthographique. Puis, afin d'évaluer la pertinence des hypothèses implémentées, nous mesurons la capacité du modèle à simuler les variations du comportement oculomoteur observé comportementalement au cours de l'apprentissage implicite de nouveaux mots.

Pour modéliser l'apprentissage orthographique, nous faisons l'hypothèse que pour mémoriser l'orthographe d'un nouveau mot, l'information perceptive sur les lettres qui composent

le nouveau mot doit être optimale. Développé à partir du modèle de reconnaissance de mots BRAID qui inclut un gradient d'acuité visuelle, un mécanisme d'interférences latérales entre lettres voisines (*crowding*) et un mécanisme d'attention visuelle, le modèle *BRAID-Learn* postule un rôle majeur des traitements visuels lors du traitement de mots écrits isolés. Le processus d'apprentissage repose sur trois mécanismes. Le mécanisme d'exploration visuo-attentionnelle autorise les déplacements visuo-attentionnels dont les caractéristiques – position du point de focalisation attentionnelle et dispersion de la distribution attentionnelle – permettent d'optimiser l'accumulation d'information perceptive sur les lettres. Le mécanisme de modulation de l'influence lexicale top-down module la quantité d'informations lexicales en fonction de la familiarité du stimulus. Finalement, le mécanisme d'apprentissage de traces orthographiques modélise le processus de mémorisation des représentations orthographiques par mise à jour des connaissances lexicales.

La pertinence de ces mécanismes est évaluée à travers deux séries de simulations. En se basant sur l'étude de Ginestet et al. (submitted), le modèle est exposé cinq fois à 30 nouveaux mots et à 30 mots connus. Le nombre de déplacements attentionnels (ou fixations) et la durée du traitement (ou durée du regard) simulés sont comparés aux résultats comportementaux.

Les résultats des simulations montrent que la totalité des mots connus sont reconnus par le modèle et que le modèle apprend correctement 25 des 30 nouveaux mots. Les cinq mots non appris ont en fait été assimilés à des mots connus avec lesquels ils partagent des caractéristiques orthographiques, conduisant à des erreurs de lexicalisation. Les patterns oculomoteurs simulés sont en accord avec ceux observés comportementalement par Ginestet et al. (submitted) montrant 1) un nombre de fixations et une durée du regard supérieurs pour les nouveaux mots que pour les mots connus, 2) une diminution du nombre de fixations et de la durée du regard au cours des présentations plus marquée pour les nouveaux mots et, 3) une diminution plus forte de ces mesures entre la première et la seconde présentation.

Cependant, une comparaison quantitative des résultats montre que le modèle *BRAID-Learn* ne capture pas avec précision l'ensemble des caractéristiques oculomotrices. D'une part, le nombre de fixations, et par conséquent, la durée du regard générés lors du traitement des nouveaux mots au cours de la première occurrence sont supérieurs à ceux observés comportementalement. D'autre part, ces mesures sont supérieures aux données comportementales pour les mots quel que soit le nombre d'occurrences.

L'ensemble des résultats montre que le modèle *BRAID-Learn* reproduit fidèlement les différents patterns de variation du comportement oculomoteur observés comportementalement lors de la lecture répétée de nouveaux mots et de mots connus. Ceci est d'autant plus intéressant que cet apprentissage se fait sur la base de mécanismes non spécifiques au type d'items. Finalement, l'absence d'apprentissage observé pour cinq nouveaux mots ainsi que la surestimation du nombre de fixations et de la durée du regard générées par le modèle suggèrent des pistes d'amélioration du modèle *BRAID-Learn*.

Abstract

To describe orthographic learning, the main theoretical paradigm is the Self-Teaching theory, which posits that novel words are learned when they are decoded, that is to say, when their phonological form is successfully obtained. However, such a focus on phonological decoding is at odds with recent observations, that highlight a relation between visuo-attentional and oculomotor behaviors and orthographic learning. In this paper, we develop the *BRAID-Learn* model, an extension of the BRAID model that enables orthographic learning. The BRAID model is a probabilistic and hierarchical model with a detailed account of visual processing (including acuity and crowding effects, and a visuo-attentional model), but no phonological representations of words. The *BRAID-Learn* model rests on three main mechanisms: first, visual exploration of stimulus is controlled so as to optimize information gain about the letters of the stimulus; second, top-down lexical influence is modulated as a function of stimulus familiarity; third, after exploration, learning integrates the perceived letter with the previous orthographic representation of the stimulus. We conduct simulations of the *BRAID-Learn* model on a set of stimuli taken from a previous experiment investigating orthographic learning in expert readers. We show that the model is able to account qualitatively for the observed patterns, with novel words requiring more fixations initially than words, followed by a rapid decrease, as their orthographic representation is learned. Our results suggest that oculomotor and visuo-attentional exploration is an intrinsic portion of orthographic learning, understudied so far.

1 Introduction

Phonological decoding – the use of spelling-sound mapping knowledge to translate letter strings into phonemes – is a first major step of reading acquisition allowing beginning readers to decode the new words they encounter while reading. The laborious phonological serial decoding of beginning and awkward readers contrasts with the fluent and immediate recognition of individual words that characterizes expert reading. Moving from slow phonological decoding to fluent reading depends on the child orthographic learning skills (Castles et al., 2018). However, the mechanisms by which orthographic learning occurs and how they can be modelled remains under-specified.

The self-teaching theory provided evidence for one of the mechanisms at play (Share, 1995, 1999). The theory postulates that each successful decoding of a novel word provides an opportunity to learn the novel word orthographic form. Accordingly, phonological decoding is viewed as the primary cognitive mechanism involved in orthographic learning. Explicit learning of spelling-sound correspondences allows children to decode the novel word which bootstraps orthographic knowledge acquisition. A few computational models have implemented the self-teaching mechanism (Pritchard et al., 2018; Ziegler et al., 2014). By application of grapheme-phoneme mappings, the phonemes corresponding to the printed word are activated, which in turn yields activation of the corresponding phonological word in long-term memory. Then, a new orthographic representation is created and the association of the new orthographic word representation with the phonological word can be learned. Ziegler et al. (2014) showed how word-specific orthographic knowledge might be successfully acquired while starting with very limited knowledge on spelling-sound correspondences. Pritchard et al. (2018) showed how contextual and semantic information contributes to single word identification to facilitate irregular-word learning. However, both a force and a limit of these implementations of how children self-learn novel orthographic words is the emphasis on phonological decoding while

avoiding explicit modeling of the visual mechanisms involved in novel word letter-string processing. In both models, a complete and immediate identification of the letters that compose the novel word is postulated, as if information on word-letter identity was fully available one-shot while reading. The main contribution of the current study will be to implement the visual and visual attention mechanisms involved in orthographic processing to show how orthographic information on the novel word is implicitly acquired while reading.

Support to the current approach is twofold. First, behavioral evidence from self-teaching studies, developmental dyslexia research and animal studies suggests that word orthographic learning is not entirely parasitic on phonological decoding. Second, experimental studies using eye tracking in conditions of implicit orthographic learning clearly show that orthographic information on novel words is not immediately available but accumulates gradually across successive encounters with the novel word.

Although the self-teaching theory ascribes a central role to phonological decoding in orthographic knowledge acquisition (Cunningham, 2006; de Jong et al., 2009; Kyte & Johnson, 2009; Nation et al., 2007; Share, 1999), there is evidence – and Share himself acknowledges – that orthographic learning is not fully explained by decoding ability (Castles & Nation, 2006, 2008). In particular, factors that relate to visual word processing, like "orthographic processing" and "print exposure" have been identified as contributing to the development of orthographic knowledge, beyond phonological skills (Cunningham et al., 2001), see (Castles & Nation, 2006) for a review.

Further evidence against phonological processing as the sole basis of orthographic learning comes from developmental dyslexia. On the one hand, prototypical cases of phonological dyslexia have been reported who demonstrate fully developed word-specific orthographic knowledge despite major phonological deficit (Howard, 1996). On the other hand, there are cases of surface dyslexia who show a major deficit of irregular-word reading and spelling despite normal phonological skills (Inserm, 2007; Romani, Di Betta, Tsouknida, & Olson, 2008; Romani, Ward, & Olson, 1999; Valdois et al., 2003). This suggests that very poor decoding skills do not necessarily prevent orthographic learning and that having good phonological skills does not guarantee normal development of lexical orthographic knowledge. Of particular interest for the present purpose, search for the cognitive deficits associated with developmental surface dyslexia revealed that selective orthographic deficit was associated with a deficit of distinct visual elements simultaneous processing, dubbed a visual-attention (VA) span deficit (Bosse, 2005; Dubois et al., 2010; Valdois et al., 2003). Further evidence that VA span more specifically relates to reading subskills that reflect word-specific orthographic knowledge – like irregular word reading (Bosse & Valdois, 2009), reading speed (Lobier et al., 2013) or the length effect in word reading (van den Boer et al., 2013) – supports a potential contribution of VA span to word-specific orthographic knowledge acquisition. More direct evidence comes from studies showing a link between VA span and spelling acquisition (van den Boer et al., 2015) and from studies showing that VA span modulates novel word orthographic learning (Bosse et al., 2015; Chaves et al., 2012; Ginestet et al., submitted). Without minimizing the role of phonological skills in orthographic acquisition, these findings suggest that visual factors independently contribute to the development of word-specific orthographic knowledge. Data from animal studies further suggest that the contribution of visual processing skills to orthographic knowledge acquisition may have been underestimated since animals can acquire impressive orthographic knowledge in the absence of language and phonological skills (Grainger et al., 2012; Scarf et al., 2016). These findings highlight the urgency to better understand how visual processing and visual attention skills contribute to orthographic learning and self-teaching.

Otherwise, recent exploration of eye movements in conditions of novel word implicit learning

revealed that orthographic learning is visually a complex process. While implicit learning begins from the first encounter with the novel word (Bosse et al., 2015; Bowey & Muller, 2005; Cunningham, 2006; Nation & Castles, 2017; Share, 1999, 2004; Tucker et al., 2016), orthographic learning is not completed at the first exposure but requires multiple encounters. Strong variations in eye movements due to repeated exposure with the same novel word are reported across the first two or three exposures, but learning effects can be observed later on and five successive exposures can be insufficient for the novel word (or rare words) to be processed as a known word (Ginestet et al., submitted; H. Joseph & Nation, 2018; H. Joseph et al., 2014; Nation & Castles, 2017). Clearly, orthographic processing for implicit learning is a long process that contrasts with the simplified picture assumed by most computational models. Monitoring eye movements provided additional insights on the mechanisms at play. Gradual decrease in processing time (gaze duration and fixation duration) with successive encounters is the main indicator of orthographic learning. The reduction in processing time across exposures during online eye movement measures is associated with increased performance on offline measures of novel word spelling knowledge. This suggests that letter identification is boosted from exposure to exposure through top-down influence due to gradual reinforcement of the novel word orthographic representation (see Ginestet et al., submitted; H. Joseph & Nation, 2018; H. Joseph et al., 2014 and, for qualitatively consistent observation, see, Pagan & Nation, 2019). Available data thus suggests that orthographic learning is a gradual, not an all-or-nothing, process that relies on close interactions between bottom-up processing for the extraction of letter information from the novel printed word and top-down lexical influences, including the influence of the under-acquisition novel word orthographic representation.

Therefore, our aim in the current study is to propose a new computational model of online orthographic learning while reading. We develop *BRAID-Learn*, an original extension of an existing word recognition model, the BRAID model (for Bayesian model of word Recognition with Attention, Interference and Dynamics; Phénix, 2018), which includes mechanisms of visual acuity and lateral interference together with a visual attention component. The *BRAID-Learn* model describes how visual and visual attention mechanisms contribute to novel word orthographic processing and word learning. We then use the *BRAID-Learn* model to simulate orthographic acquisition of new words, and compare our simulation with a previously-acquired set of experimental observations (Ginestet et al., submitted).

2 The BRAID-Learn model

In this section, we define and present the *BRAID-Learn* model, as an extension of the BRAID model. Since both models share a similar structure, we first provide a brief description of the BRAID model and, second, we present the mechanisms added into BRAID to model orthographic learning.

2.1 The BRAID model

A full description of the BRAID model is provided elsewhere (Ginestet et al., 2019; Phénix, 2018), and beyond the scope of this paper. Instead, we briefly describe some salient features of the BRAID model that are relevant to understanding the proposed extension, *BRAID-Learn*.

In a nutshell, BRAID is a probabilistic, hierarchical model of visual, attentional and lexical knowledge and it allows simulating tasks such as letter recognition, word recognition and lexical decision. We have shown previously (Ginestet et al., 2019; Phénix, 2018; Phénix et

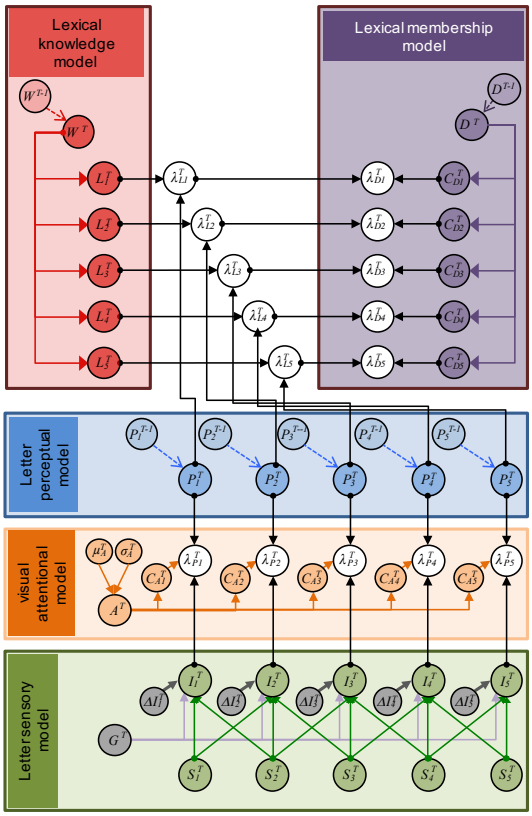

Figure 5.1 – Graphical representation of the structure of the BRAID model. Each of the five colored blocks represents a submodel; each node of the graph represents a variable of the model; and each arrow represents a probability distribution of the model. The graphical schema presented here corresponds to a time-slice, at time instant T (note the superscripts $T-1$ and T in some nodes) of the BRAID model configured for a 5-letter stimulus (note the subscripts from 1 to 5, in variables such as variables S_1^T to S_5^T). See text for details.

al., 2018) that the BRAID model was able to account for classical effects of the field, including the frequency effect, the letter superiority effect and the length effect in lexical decision. The BRAID model can be seen as instantiating previous models in a unified mathematical, probabilistic framework, allowing extending them; for instance, the BRAID model features an original visuo-attentional layer, allowing studying how visuo-attentional properties affect letter and word perception.

Technically, BRAID is defined by a joint probability distribution, linking sensory, perceptual and lexical probabilistic variables. This joint probability distribution is defined thanks to conditional independence hypotheses, which allow delineating 5 submodels and their connections; this forms the structure of the model (see Figure 5.1). We now describe some features of each 5 submodels of the BRAID model, and how it is then used, thanks to Bayesian inference, to simulate letter recognition, word recognition and lexical decision.

2.1.1 The five submodels of the BRAID model

The "Letter Sensory" submodel This submodel concerns low-level visual processing of letter stimuli, S_1^1 to S_N^T, with subscripts 1 to N referring to spatial positions, and superscripts 1 to T referring to time instants (we will use $S_{1:N}^{1:T}$ as a shorthand for the whole set of these variables). From the stimulus, this submodel essentially infers "internal" representations of letter identity, in the form of discrete probability distributions, over variables $I_{1:N}^{1:T}$, with their

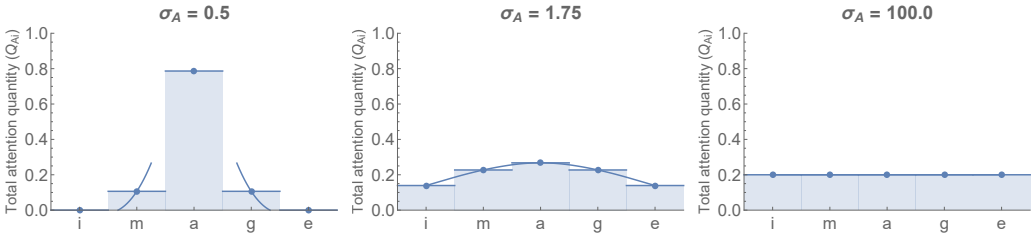

Figure 5.2 – Illustration of the attention distribution over the letter string of the stimulus word IMAGE for a position of attention $\mu_A = 3$, and for different values of attention dispersion σ_A: Left: $\sigma_A = 0.5$; Middle: $\sigma_A = 1.75$; Right: $\sigma_A = 100.0$.

domain the set of the 27 possible characters (26 letters plus a special character denoting an unknown or missing letter).

The core component of the letter sensory submodel is a confusion matrix, from stimuli to internal representations of letters, calibrated to match typical, expert reader performance in isolated letter recognition (Geyer, 1977). Several mechanisms modulate this for non-isolated letter recognition, modeling gaze position (with variable $G^{1:T}$), acuity gradient (that increases uncertainty as a function of excentricity from gaze), and interference from neighboring letters (yielding crowding effects).

The "Visual Attentional" submodel Using intermediate variables and probability distributions (technically, so called "coherence" (Bessière, Laugier, & Siegwart, 2008; Gilet et al., 2011) and "control" variables (Phénix, 2018)), the visual attentional submodel acts as a layer filtering the transfer of bottom-up information, i.e., from the "Letter Sensory" submodel to the "Letter Perceptual" submodel. This allows to modulate this information transfer differently for each position, modeling the visuo-attentional distribution. To do so, the probability distribution $P(A^t \mid \mu_A^t \ \sigma_A^t)$ at time t characterizes the spatial distribution of visual attention by a discretized and truncated Gaussian probability distribution. Its mean μ_A^t models the position of attentional focus (which we assume, in all the simulations presented here, to coincide with gaze position G^t), and its standard deviation σ_A^t models the attentional dispersion.

As Figure 5.2 shows, the smaller the value of σ_A, the more attention is focused on a small number of letters. For instance, with $\sigma_A = 0.5$, attention is focused, enhancing the perceptual accumulation of information about the 3rd letter, mostly (in our example, $\mu_A = 3$ and the stimulus is 5-letter long), to the detriment of external letters (e.g., the 1st and 5th are hardly processed). On the other hand, a large value of σ_A (for example, 100) simulates a uniform distribution of attention over the stimulus. In this case, the speed of perceptual information accumulation is equal for all letter positions.

The "Letter Perceptual" submodel The third submodel we describe is the letter perceptual submodel, in which evidence about letter identity is accumulated, over time, into probabilistic variables $P_{1:N}^{1:T}$. It can be seen as a series of Markov chains, one for each position n. Each such Markov chain, in essence, is a temporally evolving probability distribution, here over the discrete space of all 27 possible characters. This probabilistic model both has intrinsic dynamics, according to which information gradually decays towards a resting state with is the uniform distribution, and input information from "neighboring" submodels (i.e., those linked to it by probabilistic dependencies, see Figure 5.1). In the BRAID model, the letter perceptual submodels receives, first, perceptual information from the letter sensory submodel filtered

by the visual attentional submodel, in a bottom-up manner, and second, lexically predicted information from the lexical knowledge submodel, in a top-down manner.

The "Lexical Knowledge" submodel This submodel encodes, into the model, knowledge about a set of known words W, i.e., a lexicon. Over this space, a temporal model, again akin to a Markov chain, is defined. The initial state of this temporal model is the prior probability distribution $P(W^0)$, that encodes the frequency of words of W (as in the Bayesian Reader model Norris, 2006). The intrinsic dynamics of the distribution over W, as above for P, is also a gradual decay towards the initial state.

Finally, words w in W are associated with their corresponding letter sequence $L_{1:N}^{1:T}$ by a probabilistic model, such that, in each position, the correct letter at that position for this word has a high probability value (0.974), and all other alternatives have small probability values (0.001), to allow for possible spelling errors in the stimulus.

The "Lexical Membership" submodel The 5th and final submodel might be interpreted as a "word familiarity detector". It features a third and final Markov chain, over variable D, which is Boolean, with the "True" value representing that a word belongs to the lexicon. The initial, prior distribution $P(D^0)$ is uniform, representing a 50/50 chance that the input stimulus is a known word (it is a viable assumption in many experimental setups, although probably not realistic in ecological situations). Variables $D^{1:T}$ are related to Boolean variables $C_{D_{1:N}}^{1:T}$, in a probabilistic model that represents knowledge about whether a sequence of letters corresponds to a known word, or not: for a known word, all variables $C_{D_{1:N}}^{1:T}$ are assumed to be "True"; on the contrary, for a sequence of letters that is not a known word, at least one of the variables $C_{D_{1:N}}^{1:T}$ is assumed to be "False". These patterns of values serve as templates, to be compared with values of the coherence variables between the perceptual and lexical submodels, so that "observing" the flow of information between these two submodels allows to infer whether the input stimulus is a known word or not.

2.1.2 Probabilistic questions to simulate cognitive tasks

The BRAID model expresses, using probability distributions, knowledge related to letter identity, how known words are related to their corresponding letter sequences, and how to describe whether a sequence of letters belong to the known lexicon. This knowledge is then used in several cognitive processes, which we simulate by computing probabilistic distributions of interest using Bayesian inference. We call this "asking a probabilistic question" to the model.

For instance, the first cognitive task we consider is letter recognition. It is modeled by the following probabilistic question:

$$Q_{P_n^T} = P(P_n^T \mid [S_{1:N}^{1:T} = s]\ [G^{1:T} = g]\ \mu_A^{1:T}\ \sigma_A^{1:T}\ [\lambda_{P_{1:N}}^{1:T} = 1]\ [\lambda_{L_{1:N}}^{1:T} = 1])\ , \tag{5.1}$$

which can be read as: What is the probability distribution over the perceived letter at position n, for any time step t, given the stimulus letter sequence s, gaze position g the current attentional distribution (μ_A, σ_A), and given that information is allowed to propagate from the stimulus to the lexical submodel ($[\lambda_{P_{1:N}}^{1:T} = 1]$, $[\lambda_{L_{1:N}}^{1:T} = 1]$)?

We do not provide here the mathematical expression that Bayesian inference yields as an answer to this question (Phénix, 2018). However, the resulting computation can be interpreted as in classical, three-layer models with lexical, top-down influence: the sensory letter submodel extracts information about letter identity from the sequence stimulus; part of this perceptual information, depending on the attentional distribution, is propagated and accumulated into the

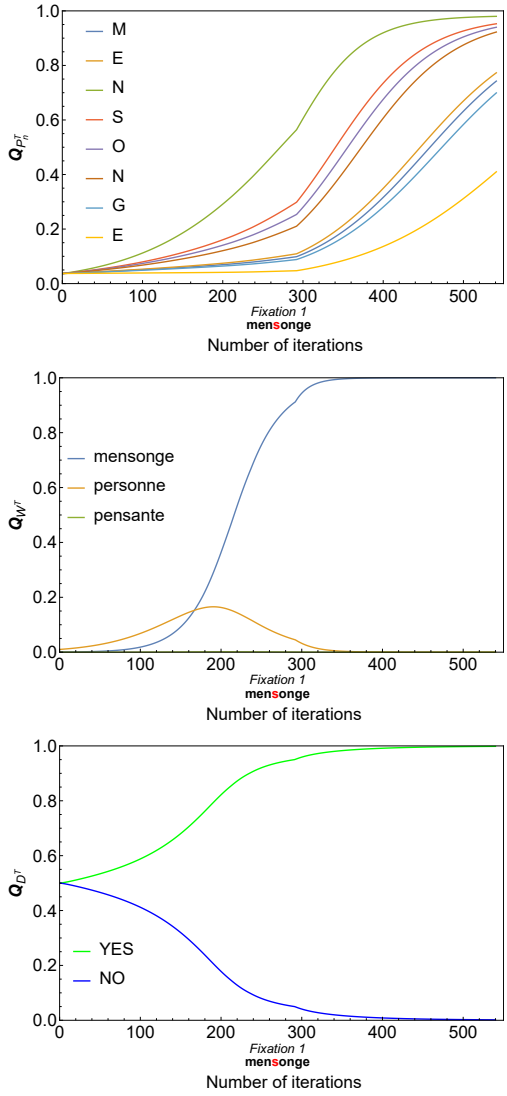

Figure 5.3 – Evolution of inference for $Q_{P_n^T}$ (top; Eq. (5.1)), Q_{W^T} (middle; Eq. (5.2)) and Q_{D^T} (bottom; Eq. (5.3)) as a function of simulated time (x-axis) for the 8-letter stimulus *MENSONGE*, with $g = \mu_A = 4$ (eye and attention are positioned over letter S, indicated in red under each plot) and $\sigma_A = 1.75$ (default value for attentional dispersion). For letter recognition (top plot), only the probability value of the correct letter at each position is shown. For word recognition (middle plot), only the probability values of the three most probable words are shown. For lexical decision, for each time-step, the whole probability distribution over the Boolean lexical membership variable D^T is shown.

dynamic models of the perceptual layer submodel. These propagate to the lexical submodel, gradually changing the probability distribution over words which, in a feedback manner, informs the perceptual layer submodel.

Figure 5.3(top) illustrates the temporal accumulation of perceptual information about letters composing the 8-letter long French word *MENSONGE* (*LIE*), in a simulation where the eye position g and visual attention focus position μ_A are assumed to be on the fourth letter (S) for the whole simulation (and with attention dispersion at its default value, $\sigma_A = 1.75$). We see that perceptual information gradually accumulates towards the correct recognition of all letters, but that it does so slower as distance to the position of the eye and of the attentional focus increases.

The second task, word recognition, is modeled in a similar manner, by considering the probabilistic question:

$$Q_{W^T} = P(W^T \mid [S_{1:N}^{1:T} = s]\ [G^{1:T} = g]\ \mu_A^{1:T}\ \sigma_A^{1:T}\ [\lambda_{P\,1:N}^{1:T} = 1]\ [\lambda_{L\,1:N}^{1:T} = 1])\ . \tag{5.2}$$

Contrary to letter recognition, in word recognition the "target space", i.e. the probabilistic distribution of interest, is over the word space W. The result of inference, in this case, is similar to the inference for letter perception, with the same flow of information, from the stimulus, up to the lexical submodel, with a feedback to the letter perception submodel.

Coming back to the example of processing the stimulus *MENSONGE*, simulation of word recognition leads to the progressive activation of the corresponding word of the lexical space (W = *MENSONGE*) and its lexical competitors, such as W = *PERSONNE* (*PERSON*) and W = *PENSANTE* (*THINKING*), as shown Figure 5.3(middle).

The third and final cognitive task is lexical decision, that is to say, recognizing whether the input letter sequence matches that of a known word. The probabilistic question is:

$$Q_{D^T} = P(D^T \mid [S_{1:N}^{1:T} = s]\ [G^{1:T} = g]\ \mu_A^{1:T}\ \sigma_A^{1:T}\ [\lambda_{D\,1:N}^{1:T} = 1]\ [\lambda_{L\,1:N}^{1:T} = 1])\ . \tag{5.3}$$

As previously, a stimulus is given, gaze position and attention distributions are set, and information is allowed to propagate into the model. However, here, we do not assume that there is a match between the stimulus and a known word; instead, by involving the lexical membership submodel ($\lambda_{D\,1:N}^{1:T} = 1$), the probability distribution over variables $\lambda_{L\,1:N}^{1:T}$ is evaluated, in essence, performing error detection in the stimulus with respect to all possible known words. Here, information flows through the whole BRAID architecture: as previously, from the stimulus to the lexical submodel and back down to the perceptual letter submodel, with the added involvment of the lexical membership model as an observer.

We reprise once more our example where the stimulus *MENSONGE* is processed. Simulating lexical decision according to Eq. (5.3) (see Figure 5.3(bottom)) shows that, in this example, the probability that the input is a word increases faster than the probability of the corresponding word in the lexical space (compare the bottom and middle plots of Figure 5.3). In other words, assuming an identical decision threshold for both processes would yield shorter response times for lexical decision than for word recognition: the model would recognize the input as a word, before knowing exactly what word it is. Such an observation is consistent with human observations (Phénix, 2018).

2.2 The *BRAID-Learn* model

The *BRAID-Learn* model is an extension of the BRAID model, that incorporates three new mechanisms allowing learning orthographic representations of visually presented new words. Its

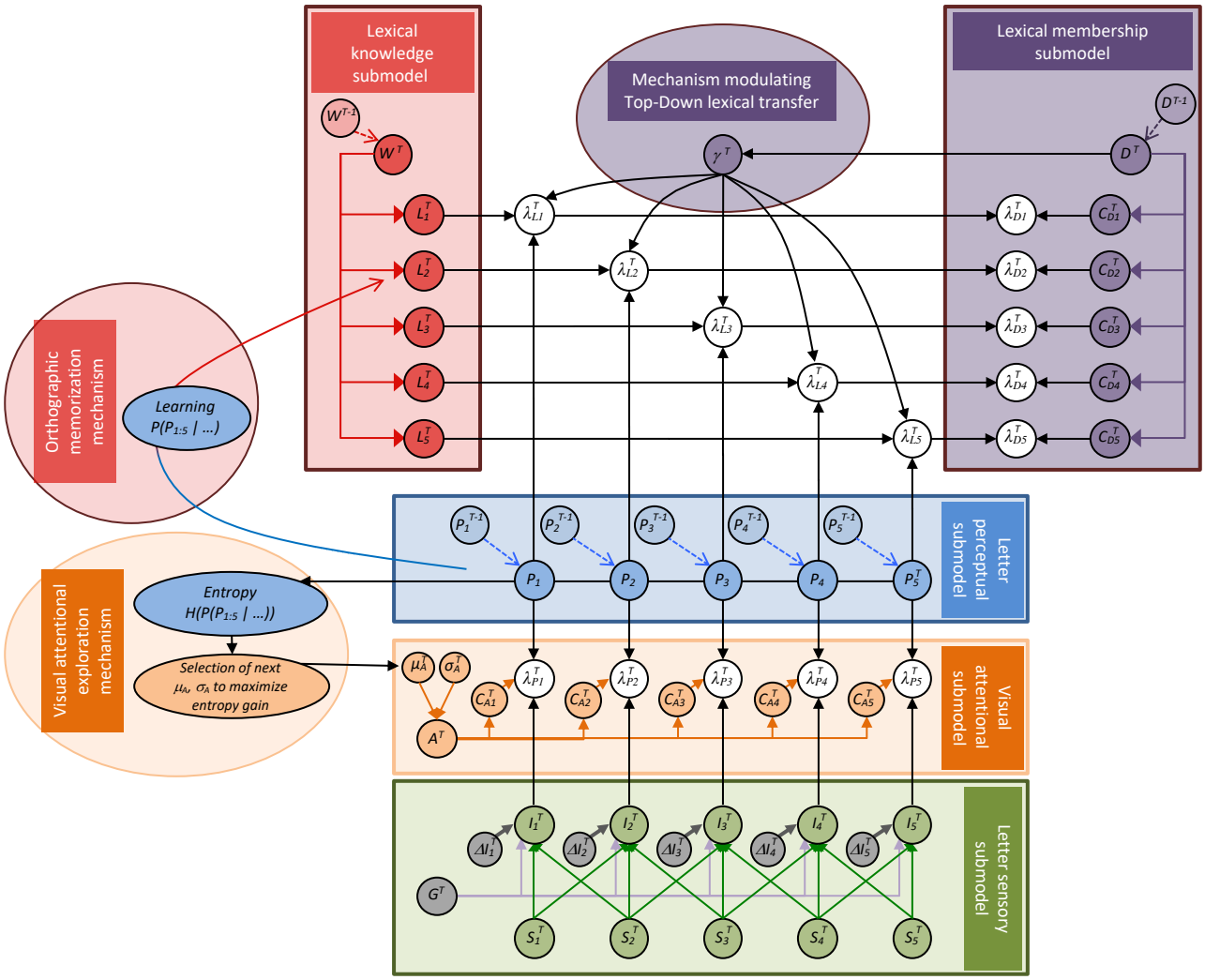

Figure 5.4 – Graphical representation of the *BRAID-Learn* model. The five colored blocks are the same submodels as in the BRAID model, with the same graphical convention (see Figure 5.1). To this architecture, the *BRAID-Learn* model adds three mechanisms, represented as colored ovals, that transfer and transform (colored and black arrows going through the ovals) information contained in portions of the BRAID model.

main assumption is that the model's aim is to accumulate efficient information about letters of the stimulus, so that, when faced with a novel word, this information can then be learned as an orthographic trace paired with a newly allocated point of the lexicon W. Therefore, the three main mechanisms of the *BRAID-Learn* model concern how it accumulates information about letters, how novelty detection influences stimulus processing, and, finally, how the resulting perceived traces are used to learn a new orthographic trace. Figure 5.4 shows a graphical representation of the *BRAID-Learn* model (to compare with the BRAID model, see Figure 5.1).

2.2.1 Efficient accumulation of perceptual evidence about letters

To model the accumulation of perceptual evidence about the letters in a stimulus sequence, we consider the letter recognition task of Eq. (5.1). It is defined as a function of the current attentional distribution parameters, i.e. its position $\mu_A^{1:T}$ and its dispersion $\sigma_A^{1:T}$. Of course, fixing a unique attentional distribution and gaze position throughout stimulus processing can yield inefficient processing. For long words (e.g., 8-letter long), concentrating attention leaves

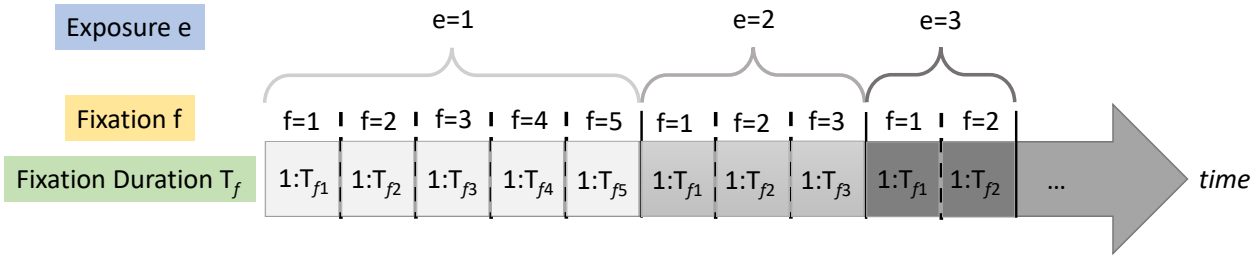

Figure 5.5 – Schematic representation of the time course of the sequencing of several fixations and exposures. Each time a word is encountered (each exposure e), it is fixated once or more times (fixations f), and each such fixation consists in a (variable) duration T_f where the eye and attention positions are fixed. This results in a varying overall gaze duration for each exposure (sum of all T_fs).

almost no perceptual processing available for some letters, and spreading attention maximally (i.e., distributing attention uniformly) yields massive, unrealistic length effects (Ginestet et al., 2019). Furthermore, and as described previously, it is well-known that eye-movements are observed in natural settings, for instance for long words and during new word processing (Lowell & Morris, 2014).

Therefore, the first and main mechanism of *BRAID-Learn* is a visuo-attentional control mechanism, that is to say, the model controls and changes its attentional distribution and gaze position over time, so as to accumulate perceptual evidence efficiently. To describe the sequencing of several fixations, we refine our temporal notation. A simulation from time-steps 0 to T is broken down as a series of exposures to a stimulus letter sequence, e from 1 to E, each exposure consisting of a variable number of fixations f from 1 to F and each fixation of variable length, from 1 to T_f time-steps (see Figure 5.5).

During one exposure, at the end of each fixation, the model selects the attentional distribution parameters that would provide the most efficient accumulation of perceptual evidence to be yet gathered. The classical mathematical measure of the information content of a discrete probability distribution $P(X)$ is its entropy, noted $H(P(X))$ and defined by:

$$H(P(X)) = -\sum_X \left(P(X)\log P(X)\right) \ . \tag{5.4}$$

The lower the entropy of a probability distribution, the greater its information content: for a given variable X, it is maximal for the uniform distribution over X, which encodes maximal uncertainty, and 0 for Dirac distributions, which encode maximal certainty. Therefore, decreasing entropy amounts to gaining information.

The *BRAID-Learn* model, therefore, also aims at optimizing information gain by maximizing entropy decrease. Mathematically, before fixation $f+1$, we enumerate a range of possible values for upcoming attention position $\mu_A^{T,f+1,e}$ and dispersion $\sigma_A^{T,f+1,e}$; for each such possible future attention distribution, and assuming that the input stimulus will not change during next fixation, we simulate letter recognition in each position n with

$$\begin{aligned} &P_{\text{next}}(n, \mu_A^{f+1,e}, \sigma_A^{f+1,e}) \\ &\quad = \ P(P_n^{T,f+1,e} \mid [S_n^{T,f+1,e} = s]\ [G^{T,f+1,e} = g^{f+1,e}]\ \mu_A^{T,f+1,e}\ \sigma_A^{T,f+1,e}) \ . \end{aligned} \tag{5.5}$$

Recall that we assume that gaze position and attention position coincide, so that $g^{T,f+1,e} = \mu_A^{T,f+1,e}$. We can then compute the entropy gain between the predicted and current distribution over letters, and average it across positions, for all possible attention distribution parameters; we note this $\overline{\Delta H}(\mu_A^{T,f+1,e}, \sigma_A^{T,f+1,e})$.

To model the physical "motor cost" of performing the visuo-attentional displacement to each enumerated future fixations, we use a straightforward measure, considering only the magnitude of the supposed displacements of gaze and attention: $MC(\mu_A^{T,f+1,e}) = |\mu_A^{T,f+1,e} - \mu_A^{T,f,e}|$. We use this measure to penalize large displacements of gaze and attentional positions, so that the overall gain measure GT that the model maximizes is a weighted combination of information gain penalized by motor cost:

$$\overline{GT}(\mu_A^{T,f+1,e}, \sigma_A^{T,f+1,e}) = (1-\alpha)\overline{\Delta H}(\mu_A^{T,f+1,e}, \sigma_A^{T,f+1,e}) - \alpha MC(\mu_A^{T,f+1,e}) \; . \tag{5.6}$$

Finally, the model selects, for its next fixation, attentional parameters and gaze position that maximize measure $\overline{GT}$.

Having described how, at any point in time, the next fixation parameters are selected, we define the initial parameters and termination criterion. Whatever the stimulus, whether it is a word or not, and since the model, at initialization, has no knowledge of the stimulus type, the parameters for the first fixation are identical. Therefore, whatever the exposure e, We assume that gaze and attention "land" at position $\mu_A^{T,1,e} = 3$. This initial position is the rounded value closest to the one from our previous experimental observations (3.01 across all item types, i.e., for words and non-words, and across all repetition exposures) in which expert readers had to read 8-letter words and non-words (Ginestet et al., submitted). For the initial dispersion of visual attention, we apply the usual default value in the BRAID model: $\sigma_A^{T,1,e} = 1.75$.

We define two termination criteria: the first defines how long each fixation is going to last, and the second is used to decide that no further fixations are going to be performed. Concerning fixation duration, and to introduce variability in simulations, we assume that the model aims at having as short fixations as possible (later on, during data analyses, back-to-back fixations on the same spatial position are aggregated and counted as a single fixation on this position; aiming for short fixations is not a theoretical claim, instead it just yields temporal granularity in our simulations). Initial simulations have shown that, in the first few iterations of the predictive evaluation of entropy gain, the "winning parameters" were numerically close, until a clear set of value emerged and, most of the time, stayed ahead until the maximal window of predictive computation (set at $T = 290$ iterations, based again on experimental observations (Ginestet et al., submitted)). We thus detect the time-step T_f for which the predicted winning parameter values have been stable for 20 (arbitrarily) previous time-steps. Finally, we set the minimal duration T_f to be at 50 iterations. The upcoming fixation is then performed with these winning parameters for that duration.

The second termination criterion prevents a further fixation when its expected information gain is below a threshold. Since fixations are of varying duration T_f, this is scaled as a function of T_f. We have empirically calibrated our stop criterion to correspond to 1 nat of information gain $\overline{\Delta H}$ for a fixation of 250 iterations (our simulations use the natural base e for entropy calculation, which is therefore measured in nats instead of bits). Therefore, our termination threshold is $T_f/250$; whenever a fixation is selected and associated to an information gain below this threshold, it is not performed by the model, and the current exposure e is considered terminated.

2.2.2 Modulation of lexical influence during word learning

The second main ingredient of the *BRAID-Learn* model is a mechanism to modulate the amount of top-down lexical information during word processing, as a function of word familiarity. A simplified description of the desired mechanism is as follows: if the input letter sequence is a known word, then strong top-down lexical information can be fed back to the letter perceptual

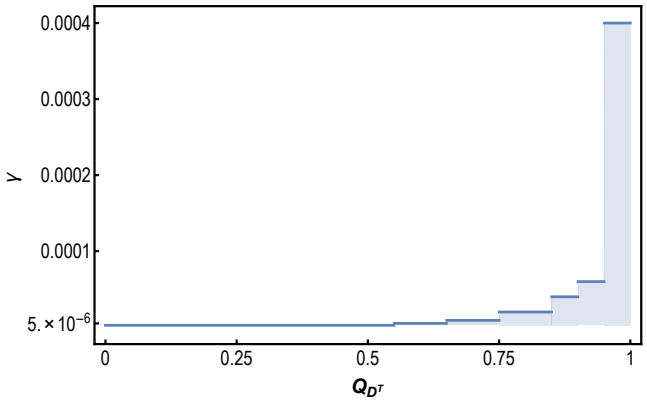

Figure 5.6 – Graph of the function of parameter γ (y-axis), that pilots the amount of top-down transfer of lexical information, as a function of the probability that $D^T = \text{true}$ (x-axis), computed by Eq. (5.3), that represents probability of lexical membership and word familiarity.

sub-model, to speed up their identification, in turn speeding up word recognition. On the other hand, if the input letter sequence is not a known word, then top-down lexical information should be diminished to try to avoid generalization toward the closest word in the lexicon, as it would yield illusory letter percepts, resulting in failure to veridically process the letters of the input novel word.

For the sake of brevity, we do not describe here the probabilistic model that allows modulating the top-down influence from the lexical knowledge sub-model to the letter perceptual model in *BRAID-Learn*. It involves building an asymmetric layer of coherence variables between these sub-models, and piloting, via control variables, the amount of information propagating top-down; this mechanism is mathematically similar to how we control, in the visual attentional sub-model, the amount of information propagating bottom-up from the letter sensory submodel to the letter perceptual sub-model. We note γ the parameter introduced by this mechanism; the higher γ, the more there is top-down lexical information transfer.

Finally, we modulate γ as a function of how likely it is that the input letter sequence corresponds to a known word. In the model, this information is already represented, by the probability distribution, at each instant, over the lexical membership variable D^T. Piloting γ as a function of D^T can be interpreted as using the "lexical decision" variable space to modulate lexical influences over letter perception. Note that this does not mean that lexical decision is performed per se, as no decision threshold is involved, and the task does not consist in deciding whether the input is a word or not; instead, we assume that lexical membership is assessed in an on-going manner, even during letter and word recognition, and modulates the information flow of these tasks. Here, the probability distribution over D^T can be interpreted as an online evaluation of lexical membership and of word familiarity.

To define the mathematical relationship between D^T and γ, our main assumption is that top-down lexical influence increases for familiar words. In mathematical terms, γ is a monotonously increasing function of the probability that $D^T = \text{true}$. Furthermore, empirical exploration shows that γ needs to have small values; the lexical knowledge model contains a lot of information (it is of low entropy, as it consists of almost-Dirac distributions) and injecting it too fast into the letter perceptual letters results in trumping sensory evidence by lexical predictions. For instance, when $\gamma = 1$, and whatever the input letter sequence, the probability distributions over letters at the perceptual layer converge towards the letters of the most frequent word of the lexicon in a few iterations. We chose to implement the relation giving γ as a function of the probability that $D^T = \text{true}$ (as evaluated by Eq. (5.3)) by a piece-wise, monotonously

increasing constant function, shown Figure 5.6.

We note that the chosen function includes a sudden increase for γ when the probability that D^T = true passes .95. When γ increases in such a manner, this increases the top-down lexical influence, so that the probability distribution over letters P_n^t suddenly receives more lexical evidence. In our simulations, this results in noticeable increases in the slopes of curves representing the evolution of probabilities for letters, words and lexical membership (e.g., see Figure 5.3 at iteration $t = 291$; note that it appears to coincide with the 290 iterations chosen as the duration for the first fixation, but it actually does not).

2.2.3 Memorization and update of orthographic traces

Finally, a third mechanism allows the update of lexical knowledge which is the last step in the learning process of *BRAID-Learn*. It takes effect once an exposure is considered terminated, that is, once one of the termination criteria of visuo-attentional exploration is fulfilled. The lexical knowledge sub-model is updated to learn the perceived letters, integrating them into the already available probabilistic model available for that word, if it was already known, or using them to create a new lexical trace, if the input sequence was detected as a new word by the lexical decision question (Eq. (5.3)).

Mathematically, at the end of exposure e, and for each position n, the complete probability distribution about the perceived letter, $P(P_n^{T,f,e} \mid [S_n^{T,f,e} = s]\ [G^{T,f,e} = g]\ \mu_A\ \sigma_A)$, is combined with the previous probability distribution about the letter at that position, $P(L_n^e \mid [W^e = w])$, in the lexical sub-model, for the recognized word w:

$$P(L_n^{e+1} \mid [W^{e+1} = w]) = \Big[P(L_n^e \mid [W^e = w]).(e-1) + P(P_n^{T,f,e} \mid [S_n^{T,f,e} = s]\ [G^{T,f,e} = g]\ \mu_A\ \sigma_A)\Big]/e \tag{5.7}$$

The model also increments by 1 (arbitrarily) the estimated frequency count of word w, in the prior probability distribution of the lexical sub-model.

If the input letter sequence was recognized as a new word by lexical decision, then a new entry w_{new} is allocated in word space W, and the initial letter trace for that word is simply the probability distributions over its perceived letters after this first exposure.

3 Simulation of the orthographic learning : the effect of the repeated reading of novel words on eye movements

We now present simulation results from the *BRAID-Learn* model. After a short illustration of its behavior on a example, allowing us to present our evaluation tools, we evaluate the ability of the *BRAID-Learn* model to account for human orthographic learning by comparing our model simulations of visuo-attentional trajectories during novel word exploration with a set of previously acquired behavioral data Ginestet et al. (submitted).

3.1 Simulating orthographic learning: illustrative example

We now show simulation results. We applied the *BRAID-Learn* model first on a word already part of the known lexicon, the word *MENSONGE*, and second on a novel word to learn, *SCRODAIN*. In both cases, we analyze the simulation results both in terms of the output behavior, that is to say, the eye-movements generated, and further, by showing how internal probability distributions evolved dynamically during the course of the simulation.

3.1.1 Applying *BRAID-Learn* to a known word

We use the same stimulus word as the one we used previously, to illustrate the three tasks of letter recognition, word recognition and lexical decision (see Figure 5.3). However, here, instead of processing the stimulus with a fixed, central eye and attention positions, we let the *BRAID-Learn* model select visuo-attentional parameters to optimize the accumulation of perceptual evidence over letters.

The simulation yields two fixations for processing the word *MENSONGE*. The first one is dictated by default parameters of the *BRAID-Learn* model: whatever the word type, the first fixation for an 8-letter long stimulus is at position $g = \mu_A = 3$ (over the "N" of *MENSONGE*), with attentional dispersion $\sigma_A = 1.75$, and lasts 290 iterations. The second fixation, selected by optimizing the predicted perceptual information gain, is at position 7 (over the "G" of *MENSONGE*), with attentional dispersion $\sigma_A = 2.0$, and lasts 250 iterations.

At the end of the second fixation, the termination criterion is met and the model proceeds to orthographic learning. In this case, the stimulus is a word, and correctly corresponds to the one recognized by the model, so that the lexical representation for word $W = MENSONGE$ is updated from the acquired perceptual representation over letters (note that the *BRAID-Learn* model performs this update irrespective of item type). By default, lexical representations for known words are defined in BRAID by "almost-Dirac" probability distribution, that is to say, the probability of the correct letter at each position is close to one (0.974) with the residual probability mass equally distributed on other letters. After the second fixation, the perceptual probability distribution over letters is not as strongly peaked as this initial lexical distribution; for instance, in position $n = 2$, the final probability is: $Q_{P_2^{540}} = P(P_2^{540} \mid [S_{1:N}^{1:T} = MENSONGE]\ [G^{1:290} = 3]\ [G^{291:540} = 7]\ [\mu_A^{1:290} = 3]\ [\mu_A^{291:540} = 7]\ [\sigma_A^{1:290} = 1.75]\ [\sigma_A^{291:540} = 2.0]\ [\lambda_{P\,1:N}^{1:T} = 1]\ [\lambda_{L\,1:N}^{1:T} = 1]) = 0.763$. Applying the lexical update of Eq. (5.8) yields a slight decrease of the stored lexical probability value predicting letter "E" at position 2.

Subsequent exposures to the word *MENSONGE* yield exactly the same performed fixations; lexical updates continue slightly and slowly decreasing the probability value for letters at their positions.

The time-course evolution of probability distributions over letters, over words, and of lexical decision during the simulation are shown Figure 5.7. We observe that the *BRAID-Learn* model has almost no effect on the dynamical evolution of word recognition (compare Figure 5.3 and Figure 5.7 middle plots) and lexical membership (compare Figure 5.3 and Figure 5.7 bottom plots), except for a slight increase in slope of probability curves at the beginning of the second fixation (iterations 290 to 310). This indicates that the selected fixation is slightly advantageous for word identification and lexical decision, as it slightly speeds up convergence toward high probability values.

For letter recognition, in contrast, the effect of the *BRAID-Learn* model is more drastic (compare Figure 5.3 and Figure 5.7 top plots). The first fixation mostly allowed identification of the letter directly under the fixation position (the "N" at position 3). In contrast, the second fixation, at position 7, almost boosts all remaining letters. Indeed, letters "N" and "G" at positions 6 and 7 are rapidly identified. Finally, the remaining letters, even far from fixation, also see their probabilities ramp up and converge to high values, thanks to the lexical influence, at this stage in full effect, thanks to the very high probability value for the word *MENSONGE* in lexical space.

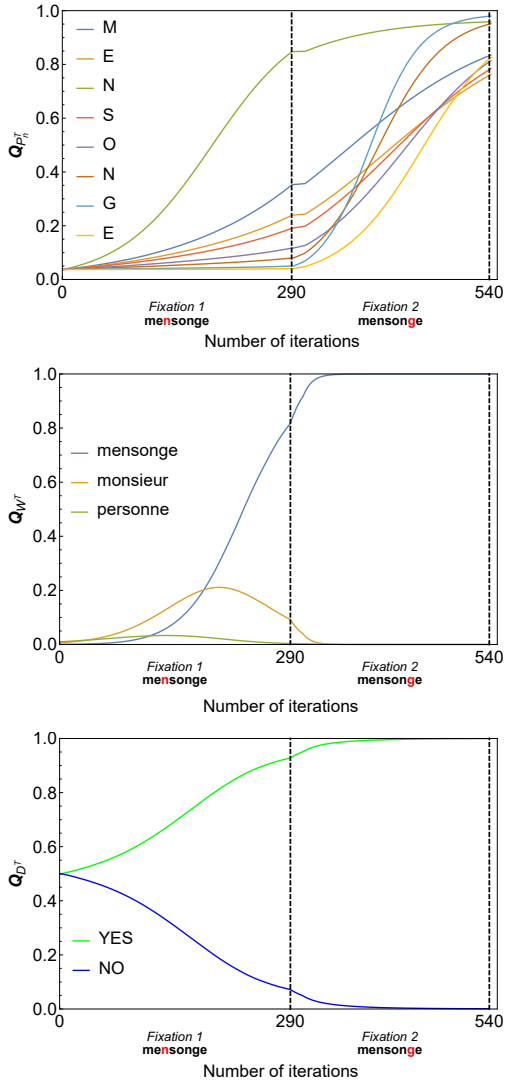

Figure 5.7 – Evolution of inference for $Q_{P_n^T}$ (top; Eq. (5.1)), Q_{W^T} (middle; Eq. (5.2)) and Q_{D^T} (bottom; Eq. (5.3)) as a function of simulated time (x-axis) for the 8-letter stimulus *MENSONGE*, with fixations computed by the *BRAID-Learn* model. Graphical representation is identical to the one of Figure 5.3, with an added vertical, dashed line for delimiting different fixations.

3.1.2 Applying *BRAID-Learn* to a novel word

We now apply the *BRAID-Learn* model to an 8-letter non-word stimulus, the letter sequence *SCRODAIN*. For the first exposure, the simulation yields 5 different fixations before the termination criterion is met; for the second exposure, 3 fixations are needed; for the third and subsequent exposures, 2 fixations are needed, in positions 3 then 7, exactly as in the previous example *MENSONGE*, in which the stimulus was a known word. Details about fixations for the first three exposures to stimulus *SCRODAIN* are shown Figure 5.8.

Figure 5.9 shows how Total Gain evolves as a function of exposures and fixations. We observe a stabilization of expected Total Gain after the third exposure, as the system converges towards a regime where stimulus *SCRODAIN*, having been already encountered three times, is associated with a lexical representations precise enough so that the stimulus is treated as a known word.

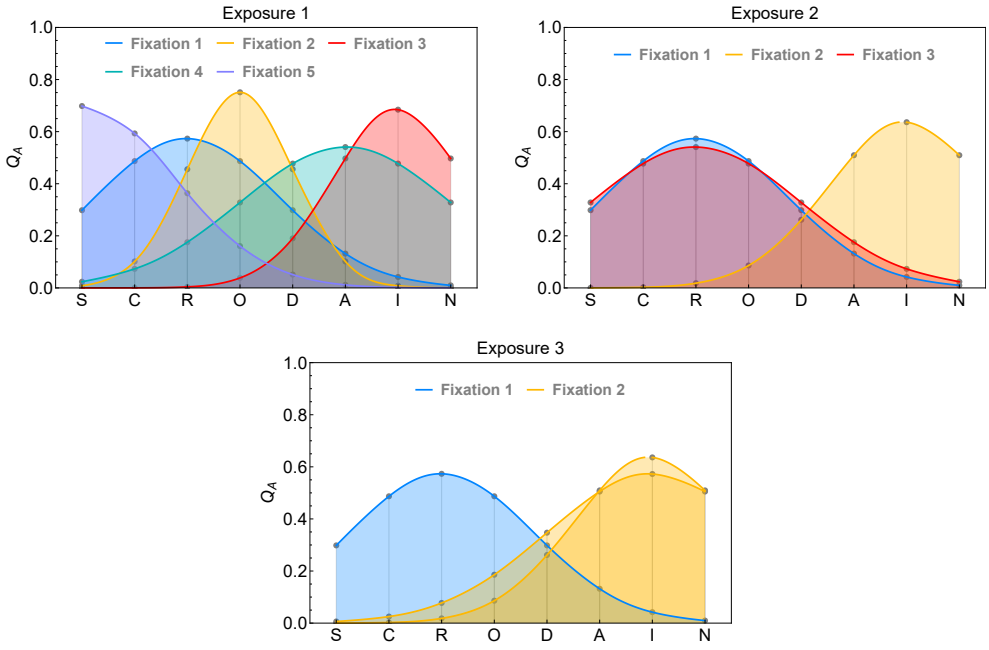

Figure 5.8 – Evolution of visuo-attentional parameters selected by the *BRAID-Learn* model during the first three exposures to the 8-letter stimulus non-word *SCRODAIN*. Each plot represents the probability value Q_A attributed to each position by the attentional model, following a Gaussian probability distribution. Each Gaussian mean value provides the selected position for attention focus $\mu_A^{T,f+1,e}$ and gaze position $g^{T,f+1,e}$: the first exposure (top left) yields 5 fixations (positions 3, 4, 7, 6 then 1); the second exposure (top right) yields 3 fixations (position 3 then position 7 then back to position 3); the third exposure yields 3 fixations (position 3 then 7 and 7 again), with the last two aggregated in our analyses, as they coincide in position. Each variance provides the selected value for the dispersion of the visual attention distribution, $\sigma_A^{T,f+1,e}$.

However, the first exposure appears to be different, with a Total Gain inferior to that of subsequent exposures. Recall that the Total Gain measure mostly captures the expected information gain during stimulus processing. During the first exposure to stimulus *SCRODAIN*, it is correctly identified as being a non-word, which, via the modulation of γ, suppresses the top-down transfer of information from the lexical submodel to the perceptual letter submodel. Consequently, during first exposure, the only source of information about letter originates from sensory processing, contrary to subsequent exposures, where it originates both from sensory processing and lexical predictions. Information gain during first exposure is therefore slower, overall, than for further exposures; the termination criterion based on information accumulation speed is thus attained for higher values of remaining information. This explains how the first exposure has a smaller Total Gain value to reach before termination, compared to further exposures.

Figure 5.10 shows the time-course evolution of probability distributions over letters, over words and of lexical decision during the first exposure to stimulus *SCRODAIN*. We observe that, when processing terminates, the stimulus is correctly recognized as a new word (the probability that D^T is false is high), and all its letters are correctly identified (each probability distribution over letters, at each position, has a high value on the correct letter identity).

Since the stimulus is recognized as a new word, the probability distribution over words is switching between hypotheses, with no clear convergence to a single, winning hypothesis. This

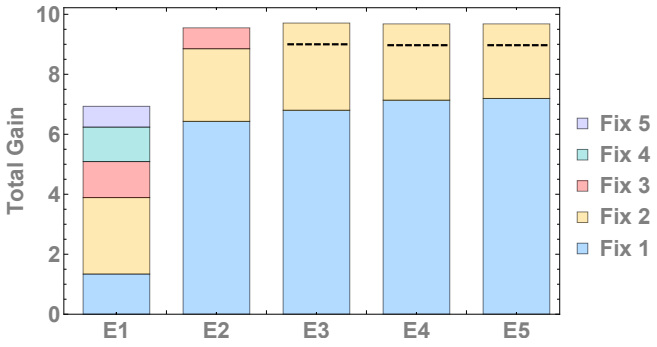

Figure 5.9 – Total Gain as a function of exposure number (noted E1 to E5) and fixation number (noted Fix 1 to Fix 5), during the processing of the 8-letter stimulus non-word *SCRODAIN*. The dashed horizontal line represents subsequent fixations that occur on the same spatial position, and are thus aggregated in following analyses.

is the expected behavior, since, during first exposure to a novel word, by definition, the word space W does not contain a point corresponding to the stimulus. Instead, the most likely hypotheses in word space are close neighbors to the stimulus, with the best one depending on processing stage, and more specifically, depending on current letter perception and fixation position. For example, consider iteration 401: few letters are well identified and gaze and attention are centered on the 4th position (the O of *SCRODAIN*). At this point, the most probable word is *PARODIAI*, which shares with *SCRODAIN* the R, O and D, which are the best perceived letters.

3.2 Simulation of the orthographic learning based on Ginestet et al. (submitted) study

We now conduct simulations to compare the overall behavior of the *BRAID-Learn* model with experimental observations gathered previously (Ginestet et al., submitted). We first recall the main results of this behavioral experiment, then compare with simulation results.

3.2.1 Behavioral experiment

In their study, Ginestet et al. (submitted) examined the effect of repeated reading of novel words on eye movements. During the reading phase, adult French-speakers were exposed to thirty bisyllabic 8-letter novel words (among which was our previous example stimulus, *SCRODAIN*) and thirty bi or trisyllabic 8-letter known words (control words, containing our previous example word, *MENSONGE*). Novel words were constructed from existing trigrams (mean $f_{trigram}$ = 1,655; SD $f_{trigram}$ = 1,176; range: 327–4,144). Thus, they are legal with respect to the French language. None was homophone to a real word and none have orthographic neighbors. Each pseudo-word contains at least two ambiguous graphemes that can be written in at least two different ways. The thirty 8-letter words were used as control. All words have no orthographic neighbors and were of medium frequency (per million, mean $f_W = 35.57$; SD $f_W = 18.86$). All items used (novel and known words) are provided in 4.3.

Words and novel words were randomly presented one, three or five times in isolation. Orthographic learning was further assessed through two post-test tasks: a spelling-to-dictation task and an orthographic decision task.

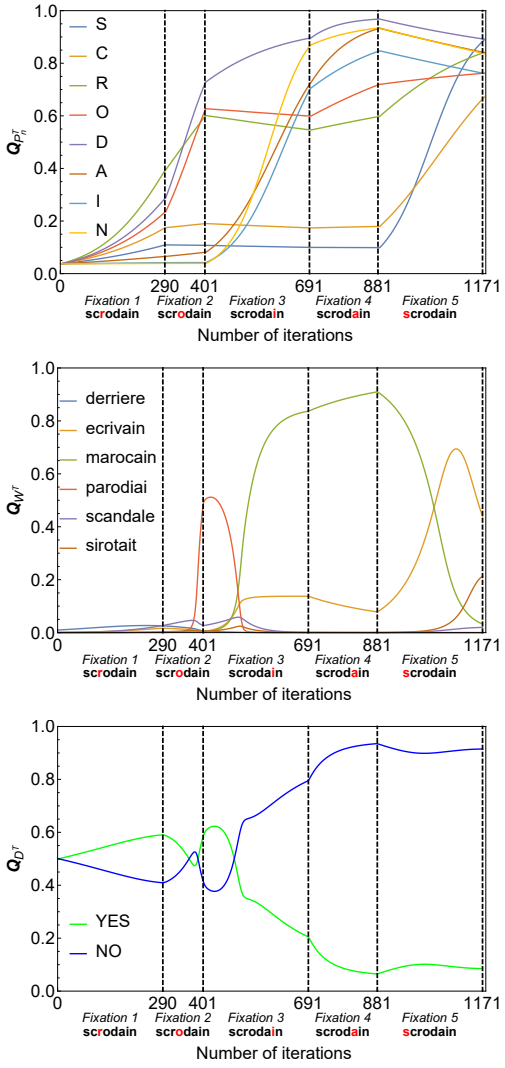

Figure 5.10 – Evolution of inference for $Q_{P_n^T}$ (top; Eq. (5.1)), Q_{W^T} (middle; Eq. (5.2)) and Q_{D^T} (bottom; Eq. (5.3)) as a function of simulated time (x-axis) for the first exposure to the 8-letter stimulus non-word *SCRODAIN*, with fixations computed by the *BRAID-Learn* model. Graphical representation is identical to the one of Figure 5.7.

Results from the dictation and orthographic decision tasks showed, first, that orthographic learning occurred and, second, that there was a significant item-type (word vs novel word) by number-of-exposure (1 to 5) interaction, for both gaze duration and number of fixations. This suggested that eye movement changes across exposures could be attributed to orthographic learning. Main effects of both item-type and number-of-exposure were reported for all eye-tracking measures (number of fixations, gaze duration, single fixation duration, first-of-two fixation duration and second-of-two fixation duration), showing higher processing times for novel words than for known words. Furthermore, a continuous decrease of processing times was qualitatively observed across exposures, and reached statistical significance for some exposure-to-exposure comparisons.

Post-hoc analyses further suggested that orthographic learning occurred during the very early stages of the learning process. Indeed, a sharper decrease was reported for pseudo-words than for word 1) between the first and second exposure for gaze duration, first-of-two fixation duration and second-of-two fixation duration and 2) between the second and third exposure for

single fixation duration. No local interaction was reported for the number of fixations neither between the first and the second exposure nor between the second and the third exposure. These observations suggest that the orthographic trace of a novel word is memorized from its first encounter and that each next encounter allows a reinforcement of the memorized orthographic representation.

3.2.2 *BRAID-Learn* simulations

Material and method The same words and novel words as those that were used in the behavioral experiment were presented five times to the *BRAID-Learn* model. As participants of the behavioral experiments were adult French-speakers, the model was configured with lexical knowledge from a French lexicon (Ferrand et al., 2010), that is to say, its known words and frequency distribution were identified from that database. Other parameters of the model were set as described in Section 2.

Simulated results Overall, the model correctly recognized all of the 30 presented words. Concerning novel words, 25 out of the 30 were correctly learned after the first exposure and correctly recognized (and re-learned) during subsequent exposures. For the remaining 5 novel words however, lexicalization errors occur, that is to say, the model incorrectly identified the stimulus as being a previously known word (final probability of lexical familiarity above .9). In that case, no new point was allocated in the word space W to learn the novel word stimulus. Instead, the most probable word, in all cases a close neighbor of the stimulus in W (e.g., *FRANCAIS* (french) for *TRANCARE*), was chosen as the most likely hypothesis, yielding incorrect merging of the current perceptual trace with the lexical representation of the recognized word.

The simulated behavior is highly systematic for the studied 8-letter words, which are always processed using two fixations, at position 3 (set by calibration) then position 7 (chosen by the entropy gain maximization mechanism). Let us note that this is not a general property of the *BRAID-Learn* model, as it does not contain a mechanism defining this stereotypical behavior. Instead, this behavior is the result of the entropy gain maximization, applied to our 30 8-letter words. Empirical observations indicate that it is likely to generalize to words, when their form conforms to statistical regularities; in contrast, words containing, for instance, low frequency trigrams yield more fixations.

For novel words, simulations are more variable. Some novel words, during first exposure, require 5 fixations (as in the above example of *SCRODAIN*, see Section 3.1.2). This varies between 3 fixations for novel word *PHACRAIT* to 6 fixations for novel word *PRIQUOIN*. Overall though, the number of required fixations decreases for all words along exposures. Most of the novel words are processed, at the fifth exposure, with 2 fixations at positions 3 then 7, exactly as words are processed.

Since some parameters of the *BRAID-Learn* model, such as the duration of the first fixation (290 iterations) were calibrated from the experimental observations of Ginestet et al. (submitted), that we aim to reproduce, we of course exclude these dimensions from our analyses. Instead, we only consider how the number of fixations and overall gaze duration evolves as a function of exposures, and qualitatively assess their observed and simulated patterns. (Of course, we note that overall gaze duration indeed contains, as a component, the duration of the first fixation, that is equalized between the model and behavioral experiment.) Finally, we also exclude from analyses the 5 novel words that the model failed to learn. Results are presented Figure 5.11.

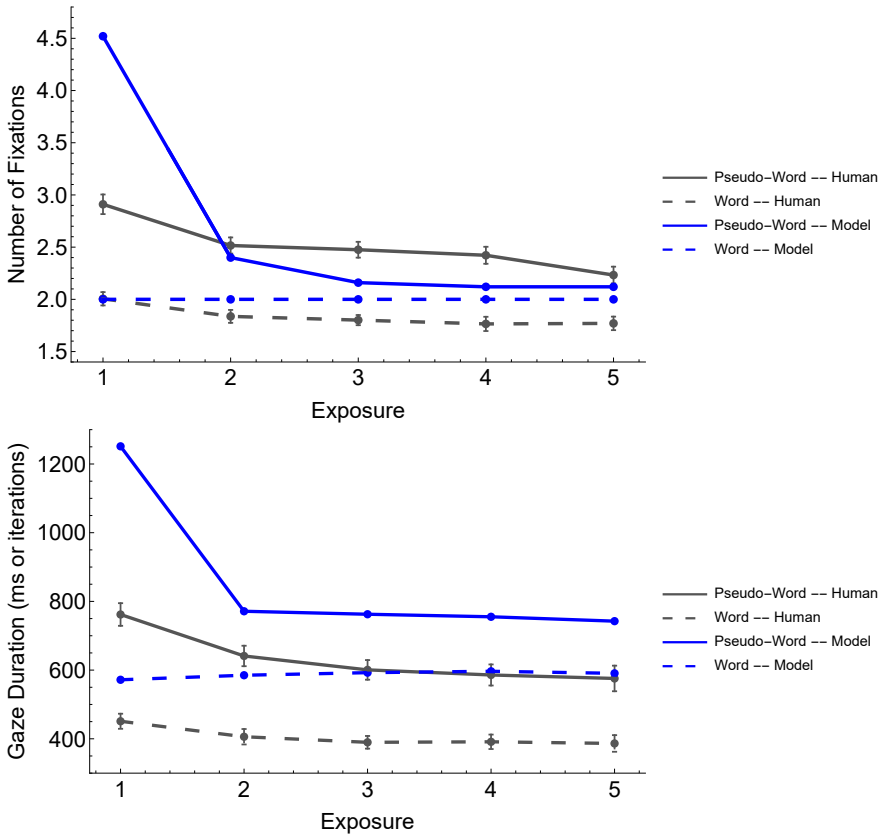

Figure 5.11 – Number of fixations and gaze duration, as a function of exposures, from the experiment of Ginestet et al. (submitted) (black lines) and simulated by the *BRAID-Learn* model (blue lines), for word stimuli (dashed lines) and pseudo-word stimuli (solid lines).

We now provide statistical analyses of the simulated results. To analyze the numbers of fixations, we apply generalized linear models (`glm`); to analyze gaze duration measures, we apply linear models (`lm`). In both cases, we use the `lme4` package of the R environment (R Core Team, 2018; RStudio version 1.0.143). To make simulated measures normally distributed before statistical analyses, a Poisson regression was applied to numbers of fixations and a log-transformation was applied to gaze duration measures; these steps are identical to data manipulation steps of the behavioral observations.

All statistical models include the number of exposures, item type and their interactions as fixed factors. Two post-hoc analyses were conducted. First, using similar statistical models than those described above, local comparisons between the first and second exposures, and between the second and third exposures were assessed. Second, contrasts of item type were evaluated on all measures (with exposure number as a fixed factor).

We first describe statistical analyses concerning the number of fixations. Results show a main effect of item type on the number of fixations ($z = 4.87; p < .001$) but no main effect of exposure ($z = 0.0; p = 1.0$; this is a degenerate value for the statistical model, resulting from lack of variability: whatever the item, whatever the exposure number, words are systematically processed using two fixations). The number of fixations decreases faster for pseudo-words than for words ($z = -3.40; p < .001$). Post-hoc comparisons show that this decrease is significant between the first and second exposures ($z = -2.61; p = .009$), but not for subsequent exposures (e.g., between the second and third, $z = -0.403; p = 0.687$). Finally, results also show an overall decrease of the number of fixations along exposures for pseudo-words ($z = -4.88; p < .001$).

We now analyze gaze duration measures. Results show a main effect of item type on

gaze duration ($z = 19.20; p < .001$) but no main effect of exposure ($z = 0.95; p = .345$). Gaze duration decreases more rapidly for pseudo-words than for words ($z = -9.30; p < .001$). Post-hoc comparisons show that this interaction occurs mainly between the first and second exposures ($z = -11.59; p < .001$), with no statistically significant interaction between the second and third exposures ($z = -0.32; p = .748$). Finally, there is a main effect of exposure on gaze duration for pseudo-words ($z = -8.37; p < .001$), and a weaker, main effect of exposure on gaze duration for words, in which gaze duration actually increases for words as a function of exposures ($z = 2.119; p = .0358$).

4 Discussion

4.1 Summary

In this study, we have developed and presented an original model of orthographic learning, called *BRAID-Learn*. It is an extension of the BRAID model, that includes three mechanisms: the first, main mechanism controls visuo-attentional displacements so as to efficiently accumulate perceptual evidence about the letters of the current stimulus; the second mechanism modulates the strength of the top-down influence of lexical knowledge on the dynamically built representation of perceived letters; the third and final mechanism concerns either the creation of a new lexical representation, for learning a word encountered for the first time, or the update of lexical representations of previously encountered words.

To compare with experimental observations gathered previously (Ginestet et al., submitted), we have applied the *BRAID-Learn* model to process a set of 8-letter French words and French-like novel words. We have shown that the *BRAID-Learn* model successfully performs recognition of the known words, and orthographic learning of most of the novel words. Simulations predict different visuo-attentional and oculomotor behaviors for words and novel words. For word processing, there are almost no variation in oculomotor measures, and their exploration is stereotypical in all aspects, that is to say, concerning the number, location and duration of fixations. For novel word processing in contrast, the first exposure yields long processing duration, with many fixations, exploring almost all letters of the stimulus. At the end of this first exposure, if the word is correctly identified as being novel, a lexical representation matched with the new word is created, that is sufficient to affect and modulate the processing behavior for subsequent exposures. Consequently, the number of fixations and gaze duration sharply decrease then stabilize, converging towards processing the new word as a known word.

The overall behavior of the *BRAID-Learn* model was further compared, quantitatively, to experimental observations (Ginestet et al., submitted). Statistical analyses yield favorable comparison, with several effects significant both in experimental and in simulated behaviors: as observed in human participants, the *BRAID-Learn* model predicts different oculomotor behaviors for words and novel words, with novel words requiring more fixations and longer gaze duration for the first exposure, and the discrepancy rapidly decreasing for subsequent exposures. However, matching between simulations and observations is not perfect in every regard, as several measures are quantitatively different (e.g., for the first exposure, the number of fixations is higher in the model than in experimental observations).

4.2 Discussion of methodological and theoretical ingredients used

4.2.1 Visuo-attentional control to optimize perceptual evidence accumulation

The main theoretical claim of the *BRAID-Learn* model we have defined is to assume that visuo-attentional and eye positions are controlled, so as to optimize information gain, as a means to make orthographic learning possible and efficient. We now break down this claim into three components.

Consider the first portion of this claim, that is to say, the simple idea that visuo-attentional characteristics and eye movements are controlled, at all. This is of course classical in models of text reading, where simulating realistic eye movements is a central issue. In contrast, eye positions and visuo-attentional distributions are seldom described in models of isolated word recognition, or of lexical decision. A few exceptions are the MORSEL (Multiple Object Recognition and attention SELection; Mozer & Behrmann, 1990), MTM (Multi Trace Model; Ans et al., 1998) and BRAID (Phénix, 2018) models, that describe attentional distribution over the stimulus, the SERIOL model (Sequential Encoding Regulated by Inputs to Oscillations within Letter Units; C. Whitney, 2001), that takes into account eye position and visual acuity effect, and, finally, a pixel-based model (Bernard, del Prado Martin, Montagnini, & Castet, 2008) for finding the optimal gaze position despite deficits in the visual field (e.g., scotoma). It remains that the main models of visual word recognition and reading under-specify the visual processing involved, and thus do not represent how the eye and visuo-attentional characteristics affect word and letter perception. Of course, since the *BRAID-Learn* model we have presented is an extension of BRAID, it is well-suited to modeling such effects.

Consider now the second portion of our theoretical claim, that concerns how visuo-attentional characteristics and eye movements are controlled. We have assumed that the controller would aim at optimizing the speed of perceptual evidence accumulation. Such an assumption is common in a wide variety of domains, including computational modeling of oculomotor behavior during natural scene visual perception (Lee & Stella, 2000; Raj, Geisler, Frazor, & Bovik, 2005), visual search modeling (Colas, Flacher, Tanner, Bessière, & Girard, 2009; Friston, Adams, Perrinet, & Breakspear, 2012; Najemnik & Geisler, 2005; Navalpakkam, Koch, Rangel, & Perona, 2010), modeling visual exploration of objects (shape-matching, Renninger, Verghese, & Coughlan, 2007) and, closer to our focus, reading and visual word recognition modeling (Bernard et al., 2008; Legge, Hooven, Klitz, Mansfield, & Tjan, 2002; Legge, Klitz, & Tjan, 1997; Salvucci, 2001).

Further, we have assumed that information gain would be measure using entropy and its evolution over time. Again, using entropy gain maximization to drive sensory exploration and behavior has a long history in various domains. Reviewing all numerous applications of this method is way beyond the scope of this paper; however, we provide here some illustrations of its pervasiveness. For instance, it is the centerpiece in active sensing (e.g., Ferreira, Castelo-Branco, & Dias, 2012) and active navigation techniques (e.g., Fox, Burgard, & Thrun, 1998) in robotics. It is also used in online data analysis and optimal experimental design (Harquel et al., 2017; Kontsevich & Tyler, 1999; Lesmes, Jeon, Lu, & Dosher, 2006). Centering once more to our precise focus, it has also been applied to word recognition and text reading, in particular in models of eye movement control in reading (Bernard et al., 2008; Legge et al., 2002, 1997; Salvucci, 2001). For instance, in the Mr. Chips model, Legge et al. (2002, 1997) assumed that saccade length is selected to minimize uncertainty about the fixated word and that refixations occur until entropy is zero or below some threshold, to represent that the fixated word is perfectly identified. Interestingly for our approach, Bernard et al. (2008) simulated differences in reading strategy for normal readers and central scotoma patients during word recognition.

For this, they assumed that word recognition occurs through an optimal reading strategy, in which fixation locations are selected by readers in order to maximize information gain about letters. We employ an identical concept in the *BRAID-Learn* model.

Finally, consider our main theoretical claim in its entirety, by adding the third component, whereby optimizing gaze and visuo-attentional displacements would enable efficient orthographic learning of novel words. Simulations have indeed backed up that claim, as experimental results have provided learning of novel words, with qualitatively realistic time-course and dynamics. More precisely, in a few exposures, the predicted visuo-attentional and gaze measures align between words and newly learned words, so that new words are acquired rapidly, after only single-digit exposures. In the model, lexical representations also get similar quickly between words and newly learned words. At this stage of our study, the fine details are yet to be accounted for, of course. Nevertheless, we believe that our results validate our main theoretical assumption, in that that observed eye and visuo-attentional displacements can be explained by a simple mathematical principle representing information gain optimization. This was the current study's main purpose.

4.2.2 Modulation of top-down lexical influence by lexical familiarity

The second main theoretical ingredient of the *BRAID-Learn* model concerns the modulation, by lexical familiarity, of the top-down influence of lexical knowledge on letter perception. In the model, we have implemented a mechanism such that lexical influence is stronger when the stimulus appears to be a known word. During orthographic learning of a new word, when exposures accumulate, the lexical representation of the word gets internalized and its entropy decreases (i.e., the lexical probability distributions converge towards quasi-Dirac probability distributions), so that lexical influence speeds up letter recognition and processing times (gaze duration and number of fixations), overall, decrease.

In our simulations, this mechanism produced a sharp decrease of number of fixations and gaze duration as early as the second exposure, in line with previous experimental observations Ginestet et al. (submitted) and previous conclusions Lowell and Morris (2014). In this last study, authors report a stronger length effect on first-pass fixation duration for novel words than for words. According to the authors, this suggests that the additional fixations and refixations observed during the first exposure to a novel word would reflect the memorization of a new orthographic representation. In turn, this new orthographic representation, acquired during the first exposure, would speed up processing as early as during the second exposure, by their top-down influence. In *BRAID-Learn*, we have implemented exactly such a mechanism, with exactly such an effect.

Another component of lexical representations, which, in the BRAID and *BRAID-Learn* models, takes the form of a prior probability distribution over words, captures word frequency. Indeed, it has been observed that word frequency is also a factor that modulates eye movements, also quite possibly through top-down lexical influence. Several previous studies have focused on this effect, by comparing visual processing of novel words, low and/or high frequency words (Chaffin et al., 2001; Kliegl, Grabner, Rolfs, & Engbert, 2004; Lowell & Morris, 2014; Rau et al., 2015; Reingold, Reichle, Glaholt, & Sheridan, 2012; Schilling et al., 1998; Williams & Morris, 2004; Wochna & Juhasz, 2013). All of these studies observed that the higher the word frequency, the lower the processing time and number of fixations, suggesting that visual processing is sensitive to the number of encounters (i.e., word frequency). However, no frequency effect was found on the landing position of the eye (Lowell & Morris, 2014). Further, in their study, Schilling et al. (1998) compared the effect of word frequency in naming, lexical decision and reading time (single fixation, first fixation and overall gaze duration measures). Their results

indicated a stronger word frequency effect on reaction times for lexical decision than for naming and reading. However, Schilling et al. (1998) reported significant correlation between reaction times in lexical decision and eye fixation data and concluded that "these tasks incorporate a common lexical access process that is affected similarly by whatever produces differences in the word frequency effect across subjects".

The BRAID model, in line with these studies, already assumed a top-down lexical influence of all components of orthographic knowledge (frequency prior and predicted letter probability distributions) on letter perception. This was, for instance, critical to account for classical word superiority effects (e.g., Grainger & Jacobs, 1994) and its modulation from context. The top-down influence was dictated by the rules of Bayesian inference applied to the BRAID model, as would be the case in any cyclic Bayesian network (Lauritzen & Spiegelhalter, 1988; Murphy, 2002; Pearl, 1988). In BRAID, such cycles arise because it contains hierarchically linked Markov chains, over the letter space, the word space and the lexical membership space. This is also in line with the interactive activation account (McClelland & Rumelhart, 1981), including its probabilistic interpretation (McClelland, 2013).

Somewhat surprisingly, we note that a debate persists, to this day, about the relevance of such feedback loops during sensory processing (see, e.g., Magnuson, Mirman, Luthra, Strauss, and Harris (2018) contra Norris, McQueen, and Cutler (2018)). It is certain that, in "flat" architectures, such as in optimal-Bayesian models, that only consider a prior distribution term and a unique likelihood term, no feedback loop is warranted by the rules of Bayesian inference. However, in more complex architectures, such as in three-layer models, including the BRAID model, the propagation of probabilistic information during inference is more complex, and sometimes involves feedback loops. It has therefore been our modeling assumption to consider such feedback loops in our models; in our case, lexical top-down influence on letter perception. We have observed, in our simulations, that this top-down influence could either have helpful effects (in the *BRAID-Learn* case, by speeding up letter identification of known words) or deleterious effects (lexicalization errors, preventing the *BRAID-Learn* model to successfully recognize regular novel words as being, indeed, new).

On this classical issue, we have thus adopted the classical stance to include top-down lexical influences. However, and, in a more original fashion, we have proposed in the *BRAID-Learn* model that this influence would be modulated as a function of how likely the stimulus is to belong to the set of known words. This allowed tuning the model so that, from the same mathematical principle of perceptual evidence gain maximization, it would yield different oculomotor behaviors for words and novel words (i.e., there is no modeling ingredient specific to item type, other than the lexical influence modulation function). We acknowledge that, in the current preliminary study and simulations, the model appears to be slightly off, in terms of its sensibility for detecting novel words as such. Indeed, with only 25 words learned out of 30 presented novel words, the model appears eager to lexicalize, which could maybe be adequate to describe rapid reading where novel words are not expected, but does not describe realistically observations from experiments in which isolated words are presented. Fine tuning the modulation function remains to be done; nevertheless, we find the general principle, whereby portions of the model control the information flow in other portions of the model, to be a promising and original one.

4.2.3 Eye movements to study orthographic learning

The third point we address in this discussion is mainly methodological. Recently proposed as a useful means to observe the mechanisms at play during orthographic learning (Nation & Castles, 2017), studying eye movements performed during implicit learning of novel words remains an original technique, seldom employed to this day. One of the first studies implicitly linking

eye movements variations to orthographic learning was performed by H. Joseph et al. (2014), who focused on the order-of-acquisition effect on orthographic learning. While monitoring eye movements during reading, adult English-speakers were exposed to sixteen bisyllabic 6-letter novel words presented fifteen times, in meaningful texts, during a five-day exposure phase. Results showed a significant decrease of all processing times (i.e., gaze duration, single fixation duration, first fixation duration and total reading time) over days.

Similar observations have been reported in two studies during the repeated reading of rare words considered as novel words. H. Joseph and Nation (2018) examined the effect of semantic context on orthographic learning. Six words that were "not well known to the children" (4^{th} and 5^{th} grades English-speakers) were embedded in sentences and presented fourteen times: Results showed an exposure effect on total reading time, lower gaze duration and total reading time during post-test compared to pre-test. More recently, Pagan and Nation (2019) focused on contextual diversity, spacing and retrieval practice effects on orthographic learning. Adult English-speakers were exposed to forty-two rare words embedded in six meaningful sentences. Their results qualitatively show shorter processing times across exposures, indicating that orthographic learning occurred.

This technique, and the general observation that eye movement variations are due in part to the progressive orthographic memorization, are also features of the experiment that we based our simulations on in this article Ginestet et al. (submitted). These experimental observations clearly suggest two different causes of eye movement variations as a function of exposures. First, simple repetition affects words (and likely, novel words as well), slightly reducing the number of fixations and overall gaze duration, and second, orthographic learning of novel words massively reduces these measures after the first exposure.

The *BRAID-Learn* model, since it is focused on modeling orthographic learning, has been able to successfully account for this second cause of eye movement variations, but has not addressed the first cause. Indeed, a simple repetition effect of words would be easily modeled by a mechanism that temporarily increases the likelihood that a recently presented word is going to be presented next. For instance, temporary manipulation of word frequencies, or a probability distribution representing a context as a function of recent words, would implement such a mechanism. The *BRAID-Learn* model, as it was defined in this study, contains no such mechanism, and thus was unable to reproduce the repetition effect for words observed in participants.

4.3 Current limits and future road-map

We have already noted previously several instances where simulations from the *BRAID-Learn* model did not quite match the experimental observations of Ginestet et al. (submitted), suggesting that the model could be improved by further calibration. More precisely, we have noted that the model failed to learn 5 out of the 30 presented novel words and that it did not account for repetition effects; that the number of fixations and overall gaze duration appears higher in the model than in experimental observations during the first exposure; that the model of orthographic learning arbitrarily increases frequency by 1 for each exposure to a novel word (making newly learned words extremely rare compared to the rest of the set of known words, despite their actual high frequency in the context of the experimental setup); that the model predicts a slight increase in oculomotor measures for words, opposite to experimental observations (as the mathematics for orthographic learning actually degrade the quasi-Dirac probability distribution of the lexical representation of words); the termination criteria were somewhat arbitrarily designed and calibrated; the landing position and duration of the first fixation were calibrated

from experimental data and not predicted by the model; the relative motor cost in the total gain measure was arbitrarily set and the model does not include a physically realistic model of eye movements.

Even though it makes the current simulations of the *BRAID-Learn* model only match qualitatively the available observations, it suggests that the model could account for a wider variety of reading situations than the narrowly defined experimental situation of Ginestet et al. (submitted). Recall that this experiment only involved expert readers learning meaningless novel words, without semantic or syntactic context, and with no explicit constraints on their reading speed or accuracy. It is likely that every element of the previous assertions can be varied experimentally, affecting the resulting observed behaviors.

For instance, we have noted the model's sensitivity to lexicalization errors, and how it could be tuned by adjusting the top-down influence and its modulation function. It could be the case that different reading styles could be captured by different calibrations of this portion of the model. Indeed, consider the continuum between fast, error-prone reading on the one hand to slow, methodical reading on the other hand. Increasing the amount of top-down lexical influence reduces processing times, at the cost of increasing probability of lexicalization errors; this captures a portion of this well-known speed and accuracy trade-off.

Another example concerns the difference between expert and learning readers. It is very likely that expert readers and children still learning how to read would exhibit different oculomotor behaviors, even on the exact same material and in the exact same experimental situation. Indeed, previous studies (Blythe, Liversedge, Joseph, White, & Rayner, 2009; H. S. Joseph, Liversedge, Blythe, White, & Rayner, 2009) have already shown higher number of fixations and processing times in children than in adults. Our simulations have shown that the *BRAID-Learn* model could predict a high number of fixations for the first exposure to a novel word. This was unrealistic for describing the expert reader behavior, making the model probably behave more like a learning reader in that situation. "Fixing" the model would thus be different whether the target was to describe child-like or expert-like behavior.

However, to properly describe learning acquisition by children, tuning the existing *BRAID-Learn* model is not going to be sufficient. Indeed, in expert readers, we could, as a first approximation, hope that semantic representations (as they are absent since we studied isolated word recognition) and phonological representations (as they are mostly redundant with the orthographic code) could be ignored. This is certainly not the case when modeling reading acquisition, as semantic and phonological knowledge are learned "in advance" of orthographic knowledge: in the typical situation, a word's sound form and meaning are known before its written form is acquired. Therefore, representations, as they are learned, and because of their different learning dynamics, are likely to interfere in non-trivial, and non-negligible manner. This is the focus of on-going work, in which we aim to marry a variant of the BRAID model that includes phonological representations (the *BRAID-Phon* model) with the current *BRAID-Learn* model, to account for reading acquisition.

Overall, the fact that the *BRAID-Learn* model fails to precisely describe the available set of experimental observations suggests that it could capture a wider set of situations than the one of this specific data set, without resorting to ad-hoc mechanisms. Furthermore, it suggests a number of possible experimental manipulations to explore, providing promising avenues for future experimental and modeling studies.

Appendix: List of items

Lists of Items used in simulations

Pseudo-words : broufand, chaiquau, deinrint, faingion, gouciont, nauplois, ploitart, speirain, quinsard, tramoint, ceitteau, chanquet, coirtint, drottont, flommais, glounein, priquoin, quarlant, siampoie, trimpond, bussiond, cherrein, ciercard, claffand, fentroit, phacrait, prinnant, tauppart, scrodain, trancare.
Control words : uniforme, portrait, enceinte, mouchoir, surprise, complice, scandale, chanteur, immeuble, avantage, physique, revanche, horrible, boutique, sensible, fauteuil, chocolat, mensonge, solution, voyageur, prochain, grandeur, nocturne, lointain, religion, empereur, division, quartier, province, jugement.

Chapitre 6

Discussion

1 Principales contributions

L'objectif principal de ce travail de thèse était de développer un nouveau modèle computationnel de traitement de mots et de leur apprentissage orthographique nommé *BRAID-Learn*. Après une revue de littérature permettant de préciser nos hypothèses de recherche, nous avons présenté trois articles dont les différents objectifs contribuent au développement du modèle *BRAID-Learn*. Dans cette section, nous rappelons les deux principales contributions de ce travail de thèse ainsi qu'un court résumé des contributions annexes.

1.1 *BRAID-Learn* : un modèle d'apprentissage orthographique

Les modèles computationnels de reconnaissance de mots, de lecture à haute voix et de mouvements oculaires en lecture de texte se sont développés indépendamment les uns des autres, parfois en s'ignorant. En conséquence, la grande majorité des modèles actuels sont spécifiques, simulant une partie des effets reportés dans la littérature. Par exemple, les modèles d'apprentissage orthographique récemment développés par Ziegler et al. (2014) et Pritchard et al. (2018) ne modélisent pas les mouvements oculaires observés pendant la lecture de nouveaux mots. A l'inverse, le seul objectif des modèles de contrôle des mouvements oculaires en lecture de texte (Engbert et al., 2005; Reichle et al., 2009) est de simuler les effets classiques (fréquence, prédictibilité, longueur) sur le comportement oculomoteur. Ils ne simulent donc pas les variations des mouvements oculaires observés lors de la lecture répétée de nouveaux mots, donc en condition d'apprentissage implicite.

Le modèle *BRAID-Learn* que nous proposons implémente des mécanismes permettant de rendre compte à la fois de l'apprentissage de l'orthographe lexicale et des mouvements oculaires en lecture de mots et de pseudo-mots. Développé à partir du modèle de reconnaissance de mots BRAID qui inclut un gradient d'acuité visuelle, des interférences latérales entre lettres voisines et un sous-modèle d'attention visuelle, le modèle *BRAID-Learn* met l'emphase sur les traitements visuels opérés lors du traitement de mots écrits isolés. A travers ce modèle, nous faisons l'hypothèse qu'un apprentissage orthographique réussi résulte de la mémorisation de la représentation orthographique d'un nouveau mot induite par l'accumulation « efficace » d'information perceptive sur les lettres qui le composent. Le processus d'apprentissage modélisé dans le modèle *BRAID-Learn* repose sur trois mécanismes principaux : un mécanisme d'exploration attentionnelle, un mécanisme de modulation lexicale (effet top-down) et un mécanisme de mémorisation.

Les deux premiers sont impliqués, au cours du processus d'apprentissage, dans l'accumula-

tion d'information perceptive sur les lettres. Le premier mécanisme guide l'exploration visuo-attentionnelle en déterminant et contrôlant les déplacements oculaires (position du regard) et visuo-attentionnels (position du point de focalisation attentionnelle et dispersion de la distribution attentionnelle). Pour cela, un principe d'optimisation du gain d'entropie est appliqué. En d'autres termes, à chaque rencontre avec le nouveau mot, chaque fixation permet de maximiser le gain d'information perceptive sur les lettres. Dans le modèle *BRAID-Learn*, la construction de traces perceptives résulte d'informations sensorielles Bottom-Up, filtrées par le sous-modèle d'attention visuelle et d'informations lexicales top-down, représentant les connaissances lexicales orthographiques sur les mots. Le second mécanisme fait l'hypothèse d'une modulation de l'influence lexicale top-down – sur l'accumulation d'information perceptive sur les lettres – par la familiarité du stimulus. Autrement dit, plus le mot est détecté comme familier par le sous-modèle dynamique d'appartenance lexicale, plus l'accumulation d'information perceptive sera influencée par les connaissances lexicales. Il est à noter que les influences lexicales ne sont jamais totalement supprimées, permettant au modèle de tenir compte des régularités statistiques orthographiques lors du traitement d'un stimulus, quel qu'il soit.

Lorsque le gain d'entropie (c'est-à-dire, d'information perceptive) est jugé insuffisant, les déplacements visuo-attentionnels sont stoppés. Le processus d'apprentissage touche à sa fin ; le troisième et dernier mécanisme d'apprentissage de traces orthographiques procède alors à la mémorisation de la représentation orthographique du nouveau mot. Cela se traduit soit, par la création d'un nouveau point dans l'espace des mots du sous-modèle des connaissances lexicales lors de la première rencontre avec le nouveau mot, soit, par la mise à jour des connaissances lexicales sur le nouveau mot lors de rencontres ultérieures.

Deux séries de simulations ont été effectuées afin d'évaluer la capacité du modèle *BRAID-Learn* à simuler l'effet de la lecture répétée sur le comportement oculomoteur observé comportementalement (Chapitre 4 ; Ginestet et al., submitted). Pour cela, 30 nouveaux mots (pseudo-mots) bisyllabiques de 8 lettres, composés d'au moins deux graphèmes ambigus et orthographiquement légaux en langue française, ont été présentés cinq fois au modèle. Le modèle a, de plus, été exposé à 30 mots connus (mots contrôles) de même longueur et de fréquence moyenne ($f = 35.57$ par million). Deux mesures ont été analysées : le nombre de « fixations » ou nombre de déplacements attentionnels et la durée du regard ou durée totale des fixations visuo-attentionnelles.

Les résultats des simulations montrent que le modèle *BRAID-Learn* reproduit, quantitativement, l'apprentissage orthographique de nouveaux mots en reproduisant les tendances observées comportementalement au cours de la lecture répétée de nouveaux mots. A l'image des données comportementales, le modèle génère des patterns différents en fonction du type d'item, simulant un plus grand nombre de fixations et une durée de fixation plus élevée pour les nouveaux mots que pour les mots connus.

Les simulations portant sur les nouveaux mots montrent une forte diminution de l'ensemble des mesures au cours des présentations. Lors de la première rencontre avec un nouveau mot, le modèle génère de nombreux déplacements attentionnels – et par conséquent, un temps de traitement plus élevé –, permettant l'accumulation d'information perceptive sur l'ensemble du mot. La diminution massive du nombre de déplacements attentionnels et du temps de traitement qui est observée entre la première et la seconde occurrence suggère que la représentation orthographique mémorisée lors de la première occurrence permet une accélération de l'accumulation d'information perceptive sur les lettres grâce aux influences lexicales (effet top-down). Les patterns simulés sont en accord avec les données comportementales, suggérant un apprentissage orthographique rapide, qui se manifeste dès la première présentation. Plus généralement, la représentation orthographique mémorisée est renforcée à chaque nouvelle exposition au mot.

En retour, chaque nouvelle rencontre avec le nouveau mot bénéficie des connaissances lexicales accumulées au cours des présentations antérieures.

Enfin, les résultats simulés, obtenus sur les mots connus, montrent que deux fixations sont systématiquement requises, une en début de mot et une en fin de mot. Ces observations sont en accord avec les données simulées dans le Chapitre 3, montrant que deux fixations permettent un traitement plus réaliste des mots de 7 lettres ou plus (c'est-à-dire, qui rendent mieux compte des effets de longueur observés). La simulation d'un nombre de fixations constant au cours des présentations successives est également en accord avec les données comportementales présentées dans le Chapitre 4 qui montrent l'absence d'effet du nombre d'occurrences sur le nombre de fixations.

Néanmoins, une comparaison quantitative des données simulées et comportementales montre que les résultats des simulations ne sont pas totalement compatibles avec les données comportementales. Le nombre de déplacements attentionnels et la durée totale de traitement sont notamment supérieurs aux données expérimentales pour les mots, quel que soit le nombre d'expositions. Ces indicateurs sont également supérieurs pour les nouveaux mots à la première exposition. Nous revenons sur ce point de discussion et plus généralement sur les limites actuelles du modèle, dans la Section 2 de ce chapitre.

Bien que davantage de comparaisons entre données comportementales et simulées soient nécessaires pour conclure sur la pertinence des hypothèses postulées dans le modèle proposé, les résultats obtenus suggèrent que les mécanismes implémentés dans le modèle *BRAID-Learn* permettent, à travers un unique modèle, de simuler la reconnaissance de mots, l'apprentissage orthographique et les mouvements oculaires lors du traitement de mots isolés, connus ou nouveaux.

1.2 Le rôle de l'attention visuelle

Dans le modèle d'apprentissage que nous proposons dans ce travail de thèse, nous faisons l'hypothèse que les caractéristiques visuelles et visuo-attentionnelles – représentant les déplacements attentionnels et donc, oculaires – s'ajustent afin d'optimiser l'accumulation d'information perceptive sur les lettres. Comme les résultats présentés dans le Chapitre 5 le suggèrent, le mécanisme d'attention visuelle implémenté dans le modèle *BRAID-Learn* joue un rôle essentiel dans le processus d'apprentissage orthographique.

Les simulations présentées dans le Chapitre 5 avaient pour objectif principal d'évaluer la capacité du modèle *BRAID-Learn* à simuler le comportement oculomoteur observé expérimentalement chez l'adulte normo-lecteur, au cours de la lecture répétée de nouveaux mots.

Dans notre mécanisme d'optimisation du gain d'entropie, nous évaluons un ensemble de valeurs possibles pour la distribution attentionnelle future. La position du point de focalisation attentionnelle peut être localisée sur chacune des lettres qui composent le nouveau mot ($\mu_A \in [1; 8]$) et la dispersion σ_A de la distribution attentionnelle peut varier entre 1.0 et 2.0, ce qui garantit, au minimum, une répartition de la quantité d'attention sur environ 4 lettres et, au maximum, sur l'ensemble du mot (si l'on considère une fixation attentionnelle centrée sur le mot, c'est-à-dire, $\mu_A = 4.5$). Le domaine de définition de σ_A est compatible avec les résultats des simulations antérieures (voir, Phénix, 2018 mais aussi Chapitre 3 pour une calibration spécifique des paramètres associés au sous-modèle d'appartenance lexicale) censées modéliser les compétences attentionnelles d'un adulte normo-lecteur. Les résultats des simulations montrent que seul le décodage d'un nouveau mot peut sélectionner la solution d'une distribution attentionnelle très piquée (c'est-à-dire, $\sigma_A = 1.0$). Ces solutions « piquées », tendent à disparaître avec les expositions : pour les nouveaux mots, à partir de la quatrième occurrence, elle n'est plus

jamais sélectionnée. Pour les mots également, les solutions à grande dispersion sont favorisées quelque soit l'exposition (σ_A entre 1.5 et 2.0).

Les résultats des simulations et les patterns visuo-attentionnels générés par le modèle nous amènent à faire une prédiction en situation de distribution de l'attention réduite, soit spatialement en limitant sa dispersion, soit en diminuant la quantité attentionnelle totale à allouer. Une telle réduction ralentirait l'accumulation d'information perceptive sur l'ensemble du mot, générant un plus grand nombre de fixations et ralentissant, voire, pour les cas extrêmes, empêchant l'apprentissage orthographique de nouveaux mots.

Cette hypothèse est étayée aussi bien par deux résultats comportementaux (Chapitre 4) et simulés (Chapitre 3), décrits dans ce manuscrit. Premièrement, l'analyse exploratoire de l'effet de l'empan VA sur la mémorisation de nouvelles traces lexicales orthographiques et sur les mouvements oculaires présentée dans le Chapitre 4, montre que de plus faibles capacités visuo-attentionnelles induisent des temps de traitement (durée du regard et durée de seconde fixation sur deux) plus élevés. Cependant, le ralentissement du processus de reconnaissance observé ne semble altérer ni le processus d'apprentissage ni sa dynamique. Toutefois, ces données ayant été obtenues auprès d'adultes normo-lecteurs présentant des capacités visuo-attentionnelles jugées dans la norme, il paraît intéressant d'approfondir la question du rôle de l'attention visuelle dans l'apprentissage orthographique. Nous présentons, dans la Section 3 de ce chapitre, une étude comportementale (prévue cette année) et la simulation des données obtenues permettant de répondre à cet objectif.

Deuxièmement, dans le Chapitre 3, nous avons étudié l'effet de longueur sur les mots en décision lexicale en fonction de la quantité d'attention visuelle allouée sur les lettres. Comportementalement, cet effet montre une augmentation des temps de réponse avec la longueur des mots. Nous simulons la tâche de décision lexicale sur 1 200 mots de longueur comprise entre 4 et 11 lettres à travers trois séries de simulations dans lesquelles les caractéristiques visuo-attentionnelles varient.

Dans une première série de simulations (notée 2.A dans le Chapitre 3), nous admettons une distribution uniforme des ressources visuo-attentionnelles et une quantité totale d'attention visuelle fixe, quelle que soit la longueur des mots. Autrement dit, la quantité d'attention visuelle allouée à chacune des lettres composant le mot non seulement, est identique quelle que soit la position de la lettre dans le mot (relativement à la position du point de focalisation attentionnnelle) mais elle est d'autant plus petite que le mot est long. Dans une seconde série de simulations (notée 2.B dans le Chapitre 3), l'attention visuelle est distribuée uniformément sur les lettres du mot et la quantité totale d'attention visuelle allouée à un mot est fonction de sa longueur. En d'autres termes, la quantité d'attention visuelle allouée à chaque lettre est identique pour toutes les lettres qui composent le mot et quelle que soit la longueur du mot. Finalement, dans une troisième série de simulations (notée 2.C dans le Chapitre 3), nous considérons une distribution normale, mais piquée, de l'attention visuelle. Ainsi, l'attention visuelle n'est allouée qu'à une seule lettre du mot située sous le point de focalisation attentionnelle. Pour ce dernier cas uniquement, le modèle simule des déplacements visuels et visuo-attentionnelles sur chacune des lettres qui composent le mot jusqu'à ce que le modèle décide si le stimulus est un mot connu ou non. Ce procédé caractérise une lecture sérielle, c'est-à-dire, lettre-à-lettre du mot présenté.

Les résultats obtenus montrent qu'un déficit visuo-attentionnel, simulé par une répartition uniforme des ressources attentionnelles indépendante de la longueur du mot (simulation 2.A) prédit un effet de longueur exagéré relativement à celui observé comportementalement. Cette même tendance est observée, malgré la réalisation de multiples fixations, lorsque l'attention visuelle ne permet le traitement que d'une seule lettre à la fois (simulation 2.C). A l'inverse, une

répartition uniforme des ressources attentionnelles en fonction de la longueur du mot (simulation 2.B) prédit un effet de longueur inversé. Autrement dit, plus le mot est long, plus le temps de réponse est court. Une modulation de l'effet de longueur est donc simulée uniquement par modification de la quantité d'attention visuelle allouée à chacune des lettres composant le mot.

En conclusion, ces résultats suggèrent un rôle critique de l'attention visuelle dans le traitement d'un mot. L'augmentation des temps de traitement observée dans le cas de ressources attentionnelles limitées nous amène à penser que le modèle *BRAID-Learn* devrait prédire, dans ce même cas, un apprentissage d'autant plus ralenti que le mot est long, conduisant probablement à des difficultés voire une absence d'apprentissage dans le cas de mots longs.

1.3 Autres contributions

Bien qu'ayant pour objectif principal de contribuer au développement du modèle d'apprentissage orthographique *BRAID-Learn*, les travaux décrits aux Chapitres 3 et 4 ont permis de répondre à des objectifs secondaires à ce travail de thèse. Dans cette section, nous présentons succinctement les différentes contributions associées à chacun de ces chapitres.

1.3.1 Continuum entre lecture lettre-à-lettre et lecture globale grâce au modèle visuo-attentionnel

Premièrement, les résultats obtenus dans le Chapitre 3 à travers la simulation de l'effet de longueur observé sur les mots en décision lexicale remettent en question l'hypothèse formulée par les modèles computationnels de lecture de type double-voie (Coltheart et al., 2001; Perry et al., 2007). Dans ces modèles, une augmentation des temps de traitement des mots en fonction de leur longueur résulte d'une lecture sérielle du mot, par la voie sous-lexicale, c'est-à-dire, par décodage phonologique. Cette hypothèse suggère que la tâche de décision lexicale nécessite l'activation de la représentation phonologique des mots. Les modèles computationnels de lecture qui postulent un traitement parallèle des lettres (Ans et al., 1998; Seidenberg & McClelland, 1989) quant à eux, ont échoué dans la simulation de l'effet de longueur observé en décision lexicale.

Dans les simulations de la tâche de décision lexicale du Chapitre 3, nous montrons que le modèle BRAID simule avec succès l'effet de longueur lorsque le modèle réalise deux fixations visuo-attentionnelles pour les mots longs (à partir de 7 lettres), à savoir, une en début de mot et une en fin de mot. Par contre, un effet de longueur exagéré est observé lorsqu'une seule fixation centrale est autorisée en lecture globale. Cela suggère qu'un effet de longueur ne traduit pas nécessairement un traitement sériel du mot, contrairement à ce que suggèrent les modèles classiques. Nos résultats sont par contre en accord avec la discussion récente des effets de longueur et de leur interprétation proposée par C. Whitney (2018). Donc, le modèle BRAID simule avec succès l'effet de longueur observé en décision lexicale bien qu'il n'implémente pas de traitement phonologique. Nos résultats 1) proposent une alternative à l'hypothèse d'un décodage phonologique ralenti pour expliquer l'augmentation des temps de réponse observés comportementalement en décision lexicale, 2) suggèrent que l'attention visuelle pourrait jouer un rôle essentiel dans le processus de reconnaissance de mot et 3) suggèrent qu'une seconde fixation permet de résoudre un compromis entre durée et efficacité du traitement visuel, induit par les phénomènes limitants du traitement visuel simultané de lettres tels que l'acuité visuelle, une réduction de l'attention visuelle et le *crowding*.

Cela ouvre la perspective d'utiliser les caractéristiques visuo-attentionnelles pour passer d'un traitement sériel à un traitement global ou d'un traitement global à un traitement sériel. Ces traitements impliquent des mécanismes spécifiques correspondant à deux branches

de traitements distincts dans les modèles double-voie. Ainsi, au lieu d'une architecture rigide impliquant la sélection d'un mode de lecture parmi les deux disponibles, la variation des paramètres visuo-attentionnels permettrait de passer graduellement d'une lecture globale, lorsque l'attention visuelle tente de couvrir l'ensemble du stimulus, ce qui fonctionne efficacement pour les mots courts ou fréquents, à une lecture sérielle dans le cas des mots longs. Un composant attentionnel unique permettrait également de rendre compte du traitement lettre-à-lettre du stimulus observé en condition pathologique ou dans les étapes précoces de l'apprentissage. Les cas intermédiaires, accessibles dans notre conception mathématique de la distribution attentionnelle, ont été étudiés largement dans ce travail, et ont notamment permis la description de l'évolution graduelle des comportements oculomoteurs et visuo-attentionnels dans l'apprentissage orthographique des nouveaux mots.

1.3.2 Méthodologie pour étudier l'apprentissage orthographique

Si les données oculométriques recueillies grâce à l'étude comportementale présentée dans le Chapitre 4 servent à la calibration et l'évaluation du modèle d'apprentissage *BRAID-Learn*, cette étude apporte de nouveaux éléments théoriques et méthodologiques sur l'étude des processus cognitifs impliqués dans l'apprentissage orthographique. Nous détaillons deux contributions – autres que celles déjà évoquées.

Premièrement, la méthodologie adoptée dans le Chapitre 4 a permis de nous assurer qu'il y avait bien apprentissage orthographique et que les patterns oculomoteurs observés sur les mots nouveaux étaient bien le reflet de cet apprentissage. Les quelques études qui ont examiné l'influence de la lecture répétée de nouveaux mots sur le comportement oculomoteur (voir H. Joseph & Nation, 2018; H. Joseph et al., 2014, et pour une observation consistante qualitativement, Pagan & Nation, 2019) ne permettent pas de conclure avec un haut niveau de certitude que les variations du comportement oculomoteur reflètent bien le mécanisme d'apprentissage orthographique. En effet, soit les études n'incluent pas l'analyse de données sur des mots contrôles (H. Joseph & Nation, 2018; H. Joseph et al., 2014; Pagan & Nation, 2019), soit, l'étude ne contrôle pas, *a posteriori*, la qualité de la mémorisation orthographique (H. Joseph et al., 2014; Pagan & Nation, 2019).

Dans notre étude expérimentale, le comportement oculomoteur observé pendant la lecture répétée de nouveaux mots est comparé à celui observé pendant la lecture de mots connus ce qui permet de dissocier un effet de répétition (induit par la lecture répétée d'un nombre restreint de stimuli) d'un effet d'apprentissage. Les résultats montrent une diminution des mesures oculaires, telles que le nombre de fixations et les durées de traitement, plus importante au cours des présentations pour les nouveaux mots que pour les mots connus. Les résultats obtenus lors des post-tests mesurant la qualité de la mémorisation orthographique montrent que plus les nouveaux mots sont lus, plus les performances en dictée et en décision orthographique sont élevées, suggérant qu'il y a bien eu apprentissage orthographique incident pendant la phase de lecture. Associé aux données oculométriques, cela suggère que les variations du comportement oculomoteur sont bien induites par l'apprentissage orthographique. Ainsi, à notre connaissance, l'étude que nous proposons est la première à associer des données mesurées sur les variations du comportement oculomoteur et l'apprentissage orthographique.

Une deuxième contribution de ce chapitre est de proposer une tâche originale permettant d'évaluer l'apprentissage orthographique en phase test. La majorité des études comportementales portant sur l'apprentissage orthographique adoptent un protocole standard proposé initialement par Share (1999). Ce protocole comporte deux phases : la phase d'apprentissage incident par lecture répétée de nouveaux mots (isolés ou intégrés dans un texte) et la phase de mesure de la qualité de l'apprentissage orthographique des nouveaux mots lus pendant la phase

précédente. Classiquement, trois tâches (post-tests) sont utilisées pour évaluer l'apprentissage orthographique effectif : une dictée des nouveaux mots, une tâche de dénomination et une tâche de choix orthographique (voir 1.1 pour une description complète de ces tâches).

La fiabilité de cette dernière tâche a récemment été remise en question (Castles & Nation, 2008; Tucker et al., 2016; Wang et al., 2011). Selon Wang et al. (2011), d'une part, la présence simultanée de l'item cible, de son homophone et de deux distracteurs, pourrait favoriser le développement de stratégies de réponse dans la prise de décision, et d'autre part, la probabilité de choisir la réponse correcte passe de 25 à 50 % si le participant a une bonne représentation phonologique du nouveau mot lu précédemment.

Pour pallier ces critiques, nous avons proposé, dans l'étude du Chapitre 4, d'évaluer l'apprentissage orthographique par une tâche « originale » de décision orthographique. Développée à partir des travaux de Wang et al. (2011) (voir Tamura et al., 2017; Wang et al., 2013 pour d'autres applications), cette dernière constitue une alternative à la tâche de choix orthographique. Dans cette tâche, les nouveaux mots cibles présentés pendant la phase de lecture et un homophone de même longueur sont présentés aléatoirement, un à un, à l'écran. Le participant doit décider la plus vite et le plus précisément possible si l'item présenté à l'écran est orthographié tel qu'il l'était lors de la phase d'apprentissage.

Les résultats montrent une augmentation significative des performances des participants (temps de réponse plus court et taux de bonnes réponses plus élevé) avec le nombre de fois où le nouveau mot a été présenté en phase d'apprentissage incident. De plus, les résultats suggèrent une grande sensibilité de la tâche de décision orthographique à l'apprentissage orthographique, montrant une augmentation significative des performances lorsque les nouveaux mots ont été lus trois fois relativement à ceux lus une fois et lorsque les nouveaux mots ont été lus cinq fois relativement à ceux lus trois fois. Pour conclure, ces observations suggèrent que la tâche de décision orthographique proposée permet d'évaluer avec précision l'évolution de l'apprentissage orthographique incident, validant la fiabilité et l'utilité de cette tâche en tant qu'alternative à la tâche de choix orthographique.

2 Limites actuelles du modèle *BRAID-Learn*

La comparaison de données simulées aux données comportementales permet d'évaluer la pertinence des hypothèses implémentées dans un modèle computationnel. C'est dans cette optique que nous proposons deux séries de simulations dans le Chapitre 5 dans lesquelles le modèle d'apprentissage orthographique *BRAID-Learn* simule la lecture répétée (cinq fois) de 30 nouveaux mots et 30 mots connus. Si les résultats obtenus montrent que le modèle reproduit qualitativement les effets de la lecture répétée de nouveaux mots (c'est-à-dire, l'apprentissage incident de nouveaux mots) sur le comportement oculomoteur, il n'en demeure pas moins que les résultats simulés mettent en évidence des différences quantitatives entre les données simulées et les données comportementales.

Dans cette section, nous ne discuterons pas de l'ensemble des limites actuelles du modèle *BRAID-Learn* dont une liste – probablement non exhaustive – est déjà proposée dans la discussion du Chapitre 5. A la place, nous choisissons d'en analyser quelques-unes. Nous les groupons en distinguant trois origines possibles : une calibration des paramètres non optimale, des hypothèses implémentées non suffisantes et des mécanismes cognitifs absents.

2.1 Calibration des paramètres du modèle

La grande majorité des paramètres du modèle *BRAID-Learn* sont communs à ceux du modèle BRAID, calibrés sur la base de données neurophysiologiques ou simulées (Ginestet et al., 2019; Phénix, 2018). Cela a permis une simulation adéquate des données comportementales relevant du domaine de la reconnaissance des mots ou de la décision lexicale (voir, par exemple, Chapitre 3). Cependant, le relatif manque de données comportementales dans le domaine de l'apprentissage orthographique limite la possibilité de calibrer correctement les paramètres du modèle. Ainsi, nous avons fixé la valeur d'une majorité des paramètres du modèle *BRAID-Learn* soit *a priori*, soit sur la base de simulations explorées empiriquement. Nous identifions deux résultats dans nos simulations dont l'écart relatif aux données comportementales pourrait être amélioré par une calibration plus adéquate des paramètres du modèle *BRAID-Learn*.

Premièrement, nous avons observé que le modèle n'apprenait correctement que 25 nouveaux mots sur les 30 traités. L'absence d'apprentissage observée pour ces 5 nouveaux mots est systématiquement due à une erreur de lexicalisation. Autrement dit, le modèle associe la représentation orthographique du nouveau mot à celle d'un mot connu (par exemple, le modèle reconnaît le mot « FRANCAIS » à la place de « TRANCARE »). Il s'agit de mots orthographiquement proches des nouveaux mots, qui ne diffèrent que de deux ou trois lettres, situées, dans la grande majorité des cas, aux extrémités du mot.

Ainsi, malgré la réalisation de plusieurs fixations lors de la première rencontre avec le nouveau mot, l'accumulation d'information perceptive sur les lettres traitées en premier augmente la probabilité que le stimulus soit un mot connu, entraînant une augmentation du poids des influences lexicales. Finalement, le modèle reconnaît un mot déjà présent dans le lexique orthographique (exemple : « FRANCAIS ») et n'apprend pas le nouveau mot (exemple : « TRANCARE »). A la place, il procède à la mise à jour des connaissances lexicales du mot reconnu le plus probable. Ces observations suggèrent que la fonction pilotant la force des influences lexicales en fonction de l'évaluation de la familiarité lexicale du stimulus (γ en fonction de la probabilité de D^T, voir Figure 5.6) est peut-être mal ajustée. En effet, pour les valeurs intermédiaires, et donc utilisées initialement dans le traitement du stimulus, de probabilité d'appartenance lexicale, l'influence top-down résultante est vraisemblablement un peu trop élevée. L'abaisser permettrait d'éviter les erreurs de lexicalisation et donc, les cas où l'apprentissage n'est pas réalisé.

Deuxièmement, le nombre de fixations et par conséquent, la durée du regard simulés lors de la première occurrence d'un nouveau mot sont supérieurs aux valeurs comportementales. Cette surestimation du temps de traitement et du nombre de déplacements visuo-attentionnels ne s'observe que lors de la première rencontre avec le mot, suggérant que les critères d'arrêt fixés du processus simulent un apprentissage trop performant dès la première occurrence. Dans le modèle *BRAID-Learn*, les déplacements visuo-attentionnels sont stoppés si le gain total d'information perceptive obtenu à l'issu d'une fixation est inférieur à un seuil fixé à priori. Plus le seuil est faible, plus la quantité d'informations à accumuler est élevée. Ainsi, si à l'issue d'une fixation le seuil n'est pas atteint, le modèle réalise une nouvelle fixation, répétant les déplacements visuo-attentionnels jusqu'à ce que le gain d'information dépasse le seuil. Nous supposons que le seuil est trop faible, simulant un apprentissage trop méthodique dans son exploration du stimulus, et donc résultant en des traces perceptives trop bonnes à l'issue de la première présentation. Au lieu de modéliser l'apprentissage implicite, la valeur actuelle pourrait représenter un apprentissage explicite, dans lequel le lecteur a pour consigne de mémoriser l'orthographe d'un nouveau mot.

2.2 Hypothèses implémentées dans le modèle

Dans le modèle *BRAID-Learn*, nous faisons l'hypothèse qu'un apprentissage orthographique réussi résulte de la construction efficace de traces perceptives. Deux mécanismes permettent de répondre à cet objectif : le mécanisme d'exploration visuo-attentionnelle et le mécanisme de modulation de l'influence lexicale top-down. En accord avec notre hypothèse, le mécanisme d'exploration visuo-attentionnelle choisit les caractéristiques visuelles (position du regard) et attentionnelles (position du point de focalisation attentionnelle et la dispersion de la distribution attentionnelle) permettant de maximiser le gain d'information perceptive au cours d'une fixation.

Si ce mécanisme semble suffisant pour générer de la variabilité dans les durées des fixations à partir de la deuxième, des simulations préliminaires (non décrites dans ce manuscrit) suggèrent que l'implémentation et la calibration actuelles ne permettent pas au modèle de rendre compte de l'effet de fréquence du mot sur la durée de première fixation. Selon cet effet, la durée de première fixation est d'autant plus faible que la fréquence du mot est plus élevée (pour un exemple, voir Lowell & Morris, 2014). Pour cette raison, nous avons fixé la durée de la première fixation à 290 itérations quel que soit le type d'item et quelle que soit la fréquence des mots. Cette valeur correspond à la moyenne des temps de fixation obtenue dans l'étude du Chapitre 4 lors de la première occurrence, tout type d'items confondus.

Nous écartons l'hypothèse selon laquelle cet effet pourrait être obtenu par une augmentation du poids des influences lexicales au cours du processus. En effet, une telle modification entraînerait une augmentation des erreurs de lexicalisation, apparaissant donc contradictoire avec nos conclusions précédentes. A la place, nous supposons que cette observation peut signifier qu'un critère d'arrêt d'une fixation (définissant la durée d'une fixation) basé uniquement sur le gain d'information perceptive n'est pas suffisant.

Nous avons développé le modèle *BRAID-Learn* dans l'optique de simuler l'apprentissage orthographique ; dans ce contexte, nous avons fait l'hypothèse d'un principe d'accumulation d'information, et l'avons appliqué à l'espace des distributions de probabilités sur l'identité des lettres perçues. Nous avons appliqué cet unique principe à la simulation du traitement des nouveaux mots et, par principe de parcimonie, l'avons également appliqué pour le traitement des mots connus. Cependant, il parait envisageable que l'espace cible d'optimisation du gain d'entropie puisse varier en fonction de la tâche. On peut imaginer par exemple qu'en lecture normale de texte, le système chercherait prioritairement à accumuler de l'information dans l'espace sémantique, alors qu'il le ferait dans l'espace W des mots connus (Bernard et al., 2008) en reconnaissance de mots isolés, et dans l'espace D d'appartenance lexicale en décision lexicale. Enfin, en reconnaissance des lettres comme pour l'apprentissage orthographique, cette accumulation porterait sur l'espace P des lettres perçues.

Notre expérience comportementale du Chapitre 4 mêle des stimuli qui sont des mots et des pseudo-mots. Ainsi, lors de la première fixation, lorsque la nature de l'item présenté n'est pas encore connue, il n'est pas évident que l'objectif soit directement l'identification des lettres. De nombreux mécanismes sont envisageables. Par exemple, le système pourrait initialement optimiser sur l'espace W, par défaut, puis basculer sur l'espace P s'il apparaît que le stimulus n'est pas un mot connu. Une alternative serait d'optimiser initialement sur l'espace D, pour ensuite savoir s'il faut basculer sur l'espace W ou P en fonction de la nature du stimulus ; etc.

Un tel mécanisme pourrait apporter de la variabilité dans les durées de première fixation, et une différenciation de ces durées en fonction de la nature du stimulus. Cela pourrait conduire à la diminution de la durée de fixation lorsque le stimulus est rapidement reconnu ou détecté comme familier, comparativement aux stimuli non-familiers.

2.3 Mécanismes cognitifs absents

Le modèle *BRAID-Learn*, présenté comme un modèle « hybride », à mi-chemin entre les modèles d'apprentissage orthographique et les modèles de contrôle de mouvements oculaires, a pour objectif principal de modéliser l'apprentissage orthographique de nouveaux mots en simulant le comportement oculomoteur observé au cours de la lecture répétée de ces nouveaux mots.

A l'image des modèles de contrôle des mouvements oculaires en lecture de texte (Engbert et al., 2005; Reichle et al., 2009), le modèle *BRAID-Learn* met l'emphase sur les processus cognitifs impliqués dans le traitement visuel, supposant un rôle majeur de l'attention visuelle dans la réalisation des mouvements oculaires (et donc, dans l'apprentissage orthographique).

Pour répondre à cet objectif, les mécanismes implémentés dans le modèle *BRAID-Learn* sont dédiés à la modélisation des mouvements oculaires intra-mot, à savoir, les refixations et les durées de fixations observées lors de la lecture de mot isolé. Les résultats simulés obtenus montrent que le modèle prédit correctement les patterns oculomoteurs observés lors de la lecture répétée de nouveaux mots et de mots connus. Cependant, il serait présomptueux de suggérer que le modèle *BRAID-Learn* tient compte de l'ensemble des caractéristiques oculomotrices associées à la lecture de mots. En effet et à titre d'exemple, les déplacements oculomoteurs simulés dans le modèle *BRAID-Learn* sont supposés instantanés. En d'autres termes, le temps de la réalisation d'une saccade est nul. Il n'est donc pas pris en compte dans le temps de traitement d'un stimulus au cours d'une présentation. Ajouter un temps de déplacement saccadique augmenterait le temps de traitement d'un stimulus sans pour autant changer les patterns observés.

De plus, nous avons systématiquement appliqué l'hypothèse simplificatrice selon laquelle la position du regard g et du point de focalisation attentionnelle μ_A coïncident à chaque instant, y compris lors de la réalisation d'une saccade intra-mot. Cette implémentation n'est pas en accord avec les données comportementales qui montrent un déplacement du point de focalisation attentionnelle précoce, c'est-à-dire, réalisé avant le déplacement du regard (Rayner, 2009). De ce point de vue, le modèle de contrôle de mouvements oculaires en lecture de texte EZ-Reader est plus abouti, découplant le déplacement de l'attention visuelle et du regard et estimant à 25 ms la durée de réalisation d'une saccade oculaire. Notons que, dans *BRAID-Learn*, rien ne force à faire coïncider la position du regard et de la focalisation attentionnelle ; un mécanisme décrivant la dynamique relative de ces deux positions pourrait être ajouté à notre modèle.

Finalement, à l'inverse des modèles développés récemment (Pritchard et al., 2018; Ziegler et al., 2014) qui supposent un rôle majeur du décodage phonologique lors de l'apprentissage orthographique d'un nouveau mot, le modèle *BRAID-Learn* n'implémente pas le traitement phonologique des mots. Les résultats obtenus montrent que, malgré l'absence de représentations phonologiques pré-existantes (c'est-à-dire, avant apprentissage), le modèle apprend correctement plus de 80 % des nouveaux mots qui lui sont présentés. Cette observation suggère de remettre en question l'hypothèse d'auto-apprentissage implémentée dans les modèles de Ziegler et al. (2014) et Pritchard et al. (2018) qui postule que l'apprentissage orthographique ne survient que grâce au décodage phonologique réussi du nouveau mot. Pour autant, nous ne remettons pas en question l'implication des traitements phonologiques dans l'apprentissage orthographique. Par exemple, alors que nous pouvons simuler l'apprentissage orthographique dans le cas d'un déficit des traitements visuo-attentionnels (observé dans la dyslexie de surface), l'absence d'un mécanisme phonologique implémenté dans le modèle *BRAID-Learn* ne nous permet pas de simuler cet apprentissage dans le cas d'un déficit phonologique (observé dans la dyslexie phonologique).

Pour conclure, les mécanismes implémentés dans le modèle *BRAID-Learn* et les résultats obtenus ouvrent de nouvelles perspectives dans l'étude des processus cognitifs impliqués dans l'apprentissage orthographique, en montrant une implication des traitements visuo-attentionnels

dans le processus d'apprentissage. Cependant, certains mécanismes impliqués dans l'apprentissage orthographique demeurent absents du modèle, ne permettant pas de simuler les effets associés à ces mécanismes. Nous revenons sur le développement de ces mécanismes dans la Section 3.

3 Perspectives

Dans cette section, nous présentons des simulations et études comportementales que nous envisageons de développer prochainement. Puis, nous présentons les travaux actuellement en cours dans l'équipe dans le cadre du développement du modèle *BRAID-Learn*. Finalement, nous ouvrons des perspectives sur l'étude des conséquences des trajectoires lexicales sur l'apprentissage orthographique et phonologique, faisant le lien avec les méthodes d'apprentissage de la lecture.

3.1 Simulations et étude comportementale envisagées à court terme

La comparaison de données simulées à celles obtenues comportementalement permet soit d'évaluer la pertinence des hypothèses implémentées dans un modèle computationnel, soit de calibrer les paramètres du modèle. Un des intérêts majeurs de l'utilisation des méthodes computationnelles réside dans leur pouvoir prédictif. Autrement dit, il est tout à fait envisageable d'utiliser le modèle *BRAID-Learn* pour prédire des effets jamais observés comportementalement, suggérant la mise en place de nouvelles études comportementales pour évaluer la pertinence des hypothèses du modèle.

3.1.1 Prédire des effets grâce à la modélisation computationnelle

Les deux séries de simulations que nous proposons ici permettraient de prédire, chez le lecteur expert, comment les caractéristiques lexicales et orthographiques (régularité orthographique et voisinage orthographique) de nouveaux mots à apprendre auraient un effet sur le comportement oculomoteur au cours de la lecture répétée de ces stimuli et donc au cours de leur apprentissage orthographique.

Effet de la régularité orthographique Considérons par exemple l'effet de la régularité orthographique de nouveaux mots – c'est-à-dire, la fréquence des lettres, des bigrammes, des trigrammes, etc. de ce nouveau mot, permettant de calculer à quel point il « ressemble » orthographiquement aux autres mots existant de la langue – sur le comportement oculomoteur au cours de leur apprentissage. Si à notre connaissance aucune étude n'a étudié cet effet, les données reportées par Chaffin et al. (2001) suggèrent que la régularité orthographique des nouveaux mots influence les mouvements oculaires. En effet, les auteurs observent une durée totale de traitement (somme des durées de fixation en tenant compte des régressions faites sur le mot) plus longue pour les nouveaux mots que pour les mots connus peu fréquents, mais une durée de première fixation et une durée du regard similaires pour les deux types d'items. Ils ont supposé que l'absence d'effet de fréquence – pourtant classiquement observé – sur ces deux mesures pourrait être due aux caractéristiques orthographiques des nouveaux mots. En effet, les nouveaux mots étaient orthographiquement légaux, construits en respectant les régularités graphotactiques de l'anglais. Par ailleurs, les nouveaux mots utilisés ont été sélectionnés par des participants (autres que ceux ayant passé la tâche) en fonction de leur familiarité subjective.

Selon Chaffin et al. (2001), cette observation est en accord avec les hypothèses implémentées dans le modèle de contrôle des mouvements oculaires en lecture de texte EZ-Reader (Reichle et

al., 2009). Ce modèle postule que les mouvements oculaires sont déclenchés par l'accès lexical. Ainsi, le modèle initie une saccade oculaire lorsque le seuil d'activation lexicale est franchi. Par conséquent, plus un nouveau mot respecte les régularités graphotactiques de sa langue, plus il sera traité comme un mot familier, expliquant l'absence d'effet de fréquence sur le comportement oculomoteur entre les mots connus peu fréquents et les nouveaux mots observé par Chaffin et al. (2001).

Le mécanisme de modulation des influences lexicales implémenté suggère que le modèle *BRAID-Learn* pourrait rendre compte des observations de Chaffin et al. (2001). En effet, le poids des influences lexicales est fonction de la familiarité du stimulus. Plus leur poids est élevé, plus l'accumulation d'information perceptive est influencée par les connaissances lexicales. Pour autant, le poids des influences lexicales n'est jamais nul, même lorsque le modèle détecte, avec une forte probabilité, le stimulus comme non familier. Ainsi, un nouveau mot dont l'orthographe respecte les régularités statistiques de la langue devrait être détecté comme suffisamment familier par le modèle pour que les influences lexicales accélèrent le traitement visuel de ce nouveau mot.

Ainsi, lors de la lecture répétée de nouveaux mots, le modèle *BRAID-Learn* devrait prédire un apprentissage facilité par la régularité orthographique. D'une part, nous devrions observer un temps de traitement et un nombre de fixations inférieurs, à la première occurrence, pour les nouveaux mots réguliers relativement aux nouveaux mots ne respectant pas les régularités orthographiques de la langue. D'autre part, nous devrions observer une diminution plus rapide du temps de traitement et du nombre de fixations pour les nouveaux mots orthographiquement réguliers.

Effet du voisinage orthographique Le voisinage orthographique d'un mot désigne classiquement le nombre de mots du lexique qui ne diffèrent que d'une seule lettre et qui ont la même longueur (Coltheart, Davelaar, Jonasson, & Besner, 1977). Plus récemment, Yarkoni, Balota, and Yap (2008) proposent une définition plus flexible, étendant le voisinage orthographique à l'ensemble des mots générés par transposition de deux lettres, insertion, substitution ou suppression d'une lettre.

Actuellement, il n'existe pas de consensus sur l'effet du voisinage orthographique en reconnaissance de mots ou en décision lexicale (Phénix et al., 2018). Selon les études, et selon les conditions expérimentales, on observe un effet facilitateur (c'est-à-dire, une diminution des temps de réponse) ou inhibiteur (c'est-à-dire, une augmentation des temps de réponse). A travers plusieurs séries de simulations, le modèle BRAID a pu être utilisé pour rendre compte de la réversibilité de l'effet de la fréquence du voisinage en décision lexicale.

A notre connaissance, aucune étude ne s'est intéressée à l'effet du voisinage sur l'apprentissage orthographique. Récemment, les résultats reportés par Tamura et al. (2017) montrent que l'apprentissage de nouveaux mots, voisins orthographiques de mots connus peu fréquents, engendrent un effet facilitateur en décision lexicale avec amorçage lorsque les nouveaux mots ont été lus quatre fois et que cet effet s'inverse lorsqu'ils ont été lus douze fois. Selon les auteurs, l'apparition d'un effet inhibiteur est due à la compétition lexicale entre les mots nouvellement appris et leurs voisins orthographiques. Ils concluent que le processus d'engagement lexical, défini comme « l'émergence de processus interactifs entre les mots nouvellement appris et les mots existants », est un processus lent, nécessitant un plus grand nombre d'expositions aux nouveaux mots que la « simple » mémorisation de la représentation orthographique.

Compte-tenu de la complexité du modèle *BRAID-Learn*, il nous paraît difficile de prédire le comportement du modèle lors de l'apprentissage de nouveaux mots voisins orthographiquement de mots connus. Il nous semble donc intéressant d'explorer cette question par la réalisation de

simulations. De telles simulations permettraient également d'explorer la quantification de la mesure même de voisinage. En effet, les mesures classiquement utilisées représentent le voisinage par un nombre unique qui est typiquement élevé lorsqu'un mot a beaucoup de voisins. Cependant, la fréquence des voisins, leur distance relative au mot considéré et leur nombre interagissent certainement de manière complexe. Le modèle *BRAID-Learn* pourrait être utilisé pour tenter de dissocier ces dimensions en fonction de leurs effets comportementaux sur l'apprentissage orthographique, et affiner la ou les mesures mathématiques du voisinage.

3.1.2 Etudier l'effet de l'empan visuo-attentionnel sur l'apprentissage orthographique

Les résultats exploratoires décrits dans le Chapitre 4 montrent une augmentation de la durée du regard sur les nouveaux mots pour les participants présentant un empan visuo-attentionnel plus faible. Pour autant, le processus de mémorisation orthographique n'est pas altéré.

Afin d'approfondir notre exploration du rôle de l'attention visuelle dans l'apprentissage orthographique, nous proposons, dans une expérience en cours, d'étudier l'effet de l'empan visuo-attentionnel sur les mouvements oculaires et sur la mémorisation orthographique au cours de la lecture répétée de nouveaux mots. Pour cela, en utilisant le même protocole et le même matériel expérimental que celui décrit dans le Chapitre 4, de jeunes adultes francophones présentant des troubles de la lecture associés à un déficit visuo-attentionnel, seront exposés 1, 3 ou 5 fois aux nouveaux mots et aux mots connus présentés isolément.

Utiliser un protocole identique nous permettra de comparer les données pathologiques recueillies à celles obtenues chez le normo-lecteur dans le Chapitre 4. Globalement, nous faisons l'hypothèse qu'un lecteur dyslexique devrait obtenir de moins bonnes performances dans les tâches de mesure de l'apprentissage orthographique effectif relativement à celles observées chez le normo-lecteur. Par ailleurs, nous nous attendons à observer des temps de traitement et un nombre de fixations plus élevés ainsi qu'une diminution moins marquée de ces mesures au cours des présentations chez le lecteur dyslexique relativement au normo-lecteur. Ces résultats suggéreraient un rôle majeur dans le processus d'apprentissage orthographique de nouveaux mots.

Finalement, les nouvelles données recueillies pourront être simulées par le modèle *BRAID-Learn* offrant une nouvelle opportunité d'évaluer la pertinence des mécanismes implémentés et plus particulièrement, d'évaluer le rôle de l'attention visuelle dans le processus d'apprentissage. Plus précisément, nous étudierons si la diminution de la dispersion de la distribution attentionnelle est suffisante pour rendre compte des performances des sujets dyslexiques, ou s'il faut l'accompagner d'une diminution de la quantité d'attention disponible.

3.2 *BRAID-Acq*, un mariage de *BRAID-Phon* et *BRAID-Learn*

Nous l'avons déjà noté, le modèle *BRAID-Learn* que nous avons développé dans ce travail apparaît comme un modèle plausible de l'apprentissage orthographique incident chez le lecteur expert, puisque c'est à des données comportementales de cette nature que nous l'avons principalement comparé. Une question déjà discutée plus haut concerne la généralité de ses mécanismes et hypothèses ; nous avons argumenté qu'il pourrait servir de base pour décrire un modèle de l'apprentissage de la lecture chez l'enfant.

Cependant, pour cela, les représentations phonologiques et sémantiques paraissent inévitables. Dans un travail en cours dans l'équipe (doctorat d'Ali Saghiran), le modèle *BRAID-Phon*, une extension du modèle BRAID qui contient des représentations phonologiques, est développé. Les simulations de ce modèle concernent, évidemment, la lecture à haute voix de

mots écrits (par exemple, tâche de dénomination), et tentent de reproduire des effets connus, dans la littérature, relatifs à cette tâche. Le modèle a également pour ambition de simuler le décodage oral de nouveaux mots.

Cependant, l'apprentissage n'est pas encore traité par ce modèle ; un mariage des modèles *BRAID-Learn* et *BRAID-Phon* est envisagé. L'objectif serait le développement d'un modèle d'apprentissage incluant à la fois un sous-modèle de reconnaissance de mots pleinement spécifié dans ses dimensions visuo-attentionnelles, un sous-modèle phonologique et un mécanisme d'acquisition d'informations orthographiques nouvelles. Notons *BRAID-Acq* ce modèle.

Dans *BRAID-Acq*, nous pourrions alors simuler l'apprentissage orthographique dans une condition où un mot est déjà connu sous sa forme phonologique mais non écrite, ce qui parait la situation écologique la plus fréquente, mais aussi, grâce à la symétrie des règles probabilistes de calcul, la situation inverse où un mot est connu orthographiquement avant d'être entendu. Cette situation parait moins répandue, mais pourrait décrire la situation de prononciation des noms propres, voire même des noms communs dans certaines langues irrégulières (par exemple, en anglais). Cette situation, qui, sur le plan théorique, pourrait être le miroir de la situation écologique, a, à notre connaissance, rarement été étudiée expérimentalement.

Cette piste ouvre la voie à des études simulées « longitudinales », c'est-à-dire dans lesquelles on ferait apprendre au modèle des séries conséquentes de nouveaux mots. L'objectif serait de simuler les « pseudo » étapes de l'apprentissage de la lecture telles qu'elles sont décrites dans les études développementales. Nous pourrions étudier divers effets comportementaux (effets de longueur, de voisinage orthographique, de similarité phonologique, etc.) et leur évolution selon la base de connaissance apprise et son développement. Le modèle serait alors notamment évalué quant à sa capacité à rendre compte des trajectoires d'apprentissage décrites dans la littérature, qu'elles soient normales ou déficitaires, et de l'évolution différenciée des profils de lecture selon la transparence des langues.

3.3 Application : vers des méthodes d'enseignement de l'apprentissage de la lecture

L'étude des effets des trajectoires d'apprentissage ouvre une perspective applicative que nous souhaitons décrire ici. Par « trajectoire d'apprentissage », nous voulons dire la séquence des items présentés à un apprenant. Cela couvre de nombreuses dimensions : quels sont les mots à apprendre initialement, et ceux à apprendre plus tard ; quels sont les régimes de répétition des mots ; vaut-il mieux répartir les mots de sorte à les disperser maximalement dans l'espace des possibles orthographiques (ou phonologiques), de sorte à explorer l'espace des combinaisons existantes dans une langue, ou bien vaut-il mieux les concentrer, et « explorer de proche en proche » ?

Etudier expérimentalement ces questions parait évidemment extrêmement coûteux et, en pratique, hors de portée. En revanche, nous pourrions utiliser notre modèle comme substitut expérimental. Sa base d'apprentissage pourrait être notamment manipulée afin d'identifier les facteurs qui favorisent le passage d'une lecture lente et laborieuse à la lecture fluide de l'expert.

Cela permettrait, en retour, de définir une base d'apprentissage optimisée dont l'efficacité pourrait être testée en classe et qui, une fois validée, pourrait être intégrée à de nouvelles méthodes d'apprentissage de la lecture. Cela ouvre des perspectives pour potentiellement améliorer les méthodes d'apprentissage de la lecture pour les enfants, mais aussi pour définir des méthodes appropriées pour pallier des déficits développementaux. Par exemple, des déficits phonologiques et des déficits visuo-attentionnels pourraient être compensés par des trajectoires particulières d'apprentissage de la lecture, spécifiquement définies pour « contre-balancer les

déficits ». L'idée de baser la définition d'une partie d'une méthode d'enseignement de la lecture sur l'étude et la simulation d'un modèle mathématique de l'apprentissage orthographique et phonologique nous semble originale, et nous l'espérons, prometteuse.

Bibliography

Acha, J., & Perea, M. (2008). The effects of length and transposed-letter similarity in lexical decision: Evidence with beginning, intermediate, and adult readers. *British Journal of Psychology*, *99*(2), 245–264.

Andreu, S., & Steinmetz, C. (2016). Les performances en orthographe des élèves en fin d'école primaire (1987-2007-2015). *Note d'information*, *28*.

Ans, B., Carbonnel, S., & Valdois, S. (1998). A connectionist multiple-trace memory model for polysyllabic word reading. *Psychological Review*, *105*(4), 678–723.

Antzaka, A., Lallier, M., Meyer, S., Diard, J., Carreiras, M., & Valdois, S. (2017). Enhancing reading performance through action video games: The role of visual attention span. *Scientific reports*, *7*(1), 14563.

Arguin, M., & Bub, D. (2005). Parallel processing blocked by letter similarity in letter by letter dyslexia: A replication. *Cognitive Neuropsychology*, *22*(5), 589–602.

Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*(2), 283.

Barton, J. S., Hanif, H. M., Björnström, L. E., & Hills, C. (2014). The word length effect in reading: A review. *Cognitive Neuropsychology*, *31*(5-6), 378-412.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv preprint arXiv:1506.04967*.

Bernard, J.-B., & Castet, E. (2019). The optimal use of non-optimal letter information in foveal and parafoveal word recognition. *Vision research*, *155*, 44–61.

Bernard, J.-B., del Prado Martin, F. M., Montagnini, A., & Castet, E. (2008). A model of optimal oculomotor strategies in reading for normal and damaged visual fields. In *Deuxième conférence française de neurosciences computationnelles, "neurocomp08"*.

Bertram, R. (2011). Eye movements and morphological processing in reading. *The Mental Lexicon*, *6*(1), 83–109.

Besner, D., Risko, E. F., Stolz, J. A., White, D., Reynolds, M., O'Malley, S., & Robidoux, S. (2016). Varieties of attention: Their roles in visual word identification. *Current Directions in Psychological Science*, *25*(3), 162–168.

Bessière, P., Laugier, C., & Siegwart, R. (Eds.). (2008). *Probabilistic reasoning and decision making in sensory-motor systems* (Vol. 46). Berlin: Springer.

Bessière, P., Mazer, E., Ahuactzin, J. M., & Mekhnacha, K. (2013). *Bayesian programming*. Boca Raton, Florida: CRC Press.

Blythe, H. I., Liversedge, S. P., Joseph, H. S., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision research*, *49*(12), 1583–1591.

Bosse, M.-L. (2005). De la relation entre acquisition de l'orthographe lexicale et traitement visuo-attentionnel chez l'enfant. *Rééducation orthophonique*, *222*, 9–30.

Bosse, M.-L., Chaves, N., Largy, P., & Valdois, S. (2015). Orthographic learning during reading:

The role of whole-word visual processing. *Journal of Research in Reading*, *38*(2), 141–158.

Bosse, M.-L., Tainturier, M. J., & Valdois, S. (2007). Developmental dyslexia: The visual attention span deficit hypothesis. *Cognition*, *104*(2), 198–230.

Bosse, M.-L., & Valdois, S. (2009). Influence of the visual attention span on child reading performance: A cross-sectional study. *Journal of Research in Reading*, *32*(2), 230–253.

Bowey, J., & Muller, D. (2005). Phonological recoding and rapid orthographic learning in third-graders' silent reading: A critical test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *92*(2), 203–219.

Campbell, R., & Butterworth, B. (1985). Phonological dyslexia and dysgraphia in a highly literate subject: A developmental case with associated deficits of phonemic processing and awareness. *The Quarterly Journal of Experimental Psychology*, *37*(3), 435–475.

Carrasco, M. (2011). Visual attention: The past 25 years. *Vision Research*, *51*, 1484–1525.

Castles, A., Coltheart, M., Savage, G., Bates, A., & Reid, L. (1996). Morphological processing and visual word recognition evidence from acquired dyslexia. *Cognitive Neuropsychology*, *13*(7), 1041–1058.

Castles, A., & Nation, K. (2006). How does orthographic learning happen? In S. Andrews (Ed.), *From inkmarks to ideas: Current issues in lexical processing* (pp. 151–179). Hove and New-York: Psychology Press.

Castles, A., & Nation, K. (2008). Learning to be a good orthographic reader. *Journal of Research in Reading*, *31*(1), 1–7.

Castles, A., Rastle, K., & Nation, K. (2018). Ending the reading wars: Reading acquisition from novice to expert. *Psychological Science in the Public Interest*, *19*, 5–51.

Chaffin, R., Morris, R., & Seely, R. E. (2001). Learning new word meanings from context: A study of eye movements. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*(1), 225–235.

Chaves, N., Totereau, C., & Bosse, M.-L. (2012). Acquérir l'orthographe lexicale: quand savoir lire ne suffit pas. *ANAE – Approche Neuropsychologique des Apprentissages chez L'Enfant*, *118*, 271–279.

Colas, F., Flacher, F., Tanner, T., Bessière, P., & Girard, B. (2009). Bayesian models of eye movement selection with retinotopic maps. *Biological Cybernetics*, *100*(3), 203–214.

Colé, P., & Valdois, S. (2007). L'acquisition de la lecture et ses troubles. *Psychologie de développement cognitif de l'enfant*, 187–221.

Coltheart, M., Davelaar, E., Jonasson, J. T., & Besner, D. (1977). Access to the internal lexicon. In S. Dornic (Ed.), *Attention and perfomance VI* (pp. 535–555). Hillsdale, NJ: Erlbaum.

Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. C. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204–256.

Cunningham, A. E. (2006). Accounting for children's orthographic learning while reading text: Do children self-teach? *Journal of Experimental Child Psychology*, *95*, 56–77.

Cunningham, A. E., Perry, K. E., & Stanovich, K. E. (2001). Converging evidence for the concept of orthographic processing. *Reading and Writing: An interdisciplinary Journal*, *14*, 549–568.

Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*(3), 713.

Dehaene, S., Cohen, L., Sigman, M., & Vinckier, F. (2005). The neural code for written words: a proposal. *Trends in Cognitive Sciences*, *9*(7), 335–341.

de Jong, P. F., Bitter, D. J., van Setten, M., & Marinus, E. (2009). Does phonological recoding occur during silent reading and is it necessary for orthographic learning? *Journal of Experimental Child Psychology*, *104*(3), 267–282.

Dubois, M., Kyllingsbæk, S., Prado, C., Musca, S. C., Peiffer, E., Lassus-Sangosse, D., & Valdois, S. (2010). Fractionating the multi-character processing deficit in developmental dyslexia: Evidence from two case studies. *Cortex*, *46*, 717–738.

Duncan, J., Bundesen, C., Olson, A., Humphreys, G., Chavda, S., & Chibuya, H. (1999). Systematic analysis of deficits in visual attention. *Journal of Experimental Psychology: General*, *128*(4), 450–478.

Engbert, R., Longtin, A., & Kliegl, R. (2002). A dynamical model of saccade generation in reading based on spatially distributed lexical processing. *Vision research*, *42*(5), 621–636.

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: a dynamical model of saccade generation during reading. *Psychological review*, *112*(4), 777.

Facoetti, A., & Molteni, M. (2001). The gradient of visual attention in developmental dyslexia. *Neuropsychologia*, *39*(4), 352–357.

Ferrand, L. (2000). Reading aloud polysyllabic words and nonwords: The syllabic length effect reexamined. *Psychonomic Bulletin & Review*, *7*(1), 142–148.

Ferrand, L., Brysbaert, M., Keuleers, E., New, B., Bonin, P., Méot, A., ... Pallier, C. (2011). Comparing word processing times in naming, lexical decision, and progressive demasking: Evidence from Chronolex. *Frontiers in psychology*, *2*, 306.

Ferrand, L., Méot, A., Spinelli, E., New, B., Pallier, C., Bonin, P., ... Grainger, J. (2017). MEGALEX: A megastudy of visual and auditory word recognition. *Behavior Research Methods*, 1–23.

Ferrand, L., & New, B. (2003). Syllabic length effects in visual word recognition and naming. *Acta Psychologica*, *113*(2), 167–183.

Ferrand, L., New, B., Brysbaert, M., Keuleers, E., Bonin, P., Méot, A., ... Pallier, C. (2010). The French Lexicon Project: Lexical decision data for 38,840 French words and 38,840 pseudowords. *Behavior Research Methods*, *42*(2), 488–496.

Ferreira, J. a. F., Castelo-Branco, M., & Dias, J. (2012). A hierarchical Bayesian framework for multimodal active perception. *Adaptive Behavior*, *20*(3), 172–190.

Fiset, D., Gosselin, F., Blais, C., & Arguin, M. (2006). Inducing letter-by-letter dyslexia in normal readers. *Journal of Cognitive Neuroscience*, *18*(9), 1466–1476.

Fox, D., Burgard, W., & Thrun, S. (1998). Active markov localization for mobile robots. *Robotics and Autonomous Systems*, *25*(3–4), 195–207.

Frederiksen, J. R., & Kroll, J. F. (1976). Spelling and sound: Approaches to the internal lexicon. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(3), 361.

Friston, K., Adams, R., Perrinet, L., & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in psychology*, *3*, 151.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K. E. Patterson, J. C. Marshall, & M. Coltheart (Eds.), *Surface dyslexia* (pp. 301–330). New York, NY, USA: Routledge: Taylor and Francis Group.

Gaillard, R., Naccache, L., Pinel, P., Clémenceau, S., Volle, E., Hasboun, D., ... others (2006). Direct intracranial, fMRI, and lesion evidence for the causal role of left inferotemporal cortex in reading. *Neuron*, *50*(2), 191–204.

Gerbier, E., Bailly, G., & Bosse, M.-L. (2015). Using Karaoke to enhance reading while listening: Impact on word memorization and eye movements. In *Speech and language technology for education (slate)* (pp. 59–64).

Gerbier, E., Bailly, G., & Bosse, M.-L. (2018). Audio-visual synchronization in reading while listening to texts: Effects on visual behavior and verbal learning. *Computer Speech and Language*, *47*, 79–92.

Geyer, L. (1977). Recognition and confusion of the lowercase alphabet. *Perception & Psychophysics*, *22*(5), 487–490.

Gilet, E., Diard, J., & Bessière, P. (2011). Bayesian action-perception computational model: Interaction of production and recognition of cursive letters. *PLoS ONE*, *6*(6), e20387.

Ginestet, E., Phénix, T., Diard, J., & Valdois, S. (2019). Modeling the length effect for words in lexical decision: The role of visual attention. *Vision research*, *159*, 10–20.

Ginestet, E., Valdois, S., Diard, J., & Bosse, M.-L. (in press). Comprendre l'apprentissage de l'orthographe lexicale et ses difficultés pour mieux l'enseigner : apports et limites des dernières modélisations computationnelles. *ANAE – Approche Neuropsychologique des Apprentissages chez L'Enfant*.

Ginestet, E., Valdois, S., Diard, J., & Bosse, M. L. (submitted). Incidental learning of novel words in adults: Effects of exposure and visual attention on eye movements. —.

Gomez, P., Ratcliff, R., & Perea, M. (2008). The Overlap model: a model of letter position coding. *Psychological Review*, *115*(3), 577–601.

Goswami, U. (2015). Sensory theories of developmental dyslexia: three challenges for research. *Nature Reviews Neuroscience*, *16*, 43–54.

Grainger, J., Dufau, S., Montant, M., Ziegler, J. C., & Fagot, J. (2012). Orthographic processing in Baboons (*Papio Papio*). *Science*, *336*(6078), 245–248.

Grainger, J., Dufau, S., & Ziegler, J. C. (2016). A vision of reading. *Trends in Cognitive Sciences*, *20*(3), 171–179.

Grainger, J., & Jacobs, A. M. (1994). A dual read-out model of word context effects in letter perception: Further investigations of the word superiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1158.

Harquel, S., Diard, J., Raffin, E., Passera, B., Dall'Igna, G., Marendaz, C., ... Chauvin, A. (2017). Automatized set-up procedure for transcranial magnetic stimulation protocols. *NeuroImage*, *153*, 307–318.

Howard, D. (1996). Developmental phonological dyslexia: Real word reading can be completely normal. *Cognitive Neuropsychology*, *13*(6), 887–934.

Hudson, P. T., & Bergman, M. W. (1985). Lexical knowledge in word recognition: Word length and word frequency in naming and lexical decision tasks. *Journal of Memory and Language*, *24*(1), 46–58.

Hutzler, F., Ziegler, J. C., Perry, C., Wimmer, H., & Zorzi, M. (2004). Do current connectionist learning models account for reading development in different languages? *Cognition*, *91*(3), 273–296.

Inserm. (2007). Dyslexie, dysorthographie, dyscalculie. Bilan des données scientifiques. *Paris: Les éditions Inserm*, *XV*, 842p - Expertise collective.

Jorm, A. F., & Share, D. L. (1983). An invited article: Phonological recoding and reading acquisition. *Applied psycholinguistics*, *4*(2), 103–147.

Joseph, H., & Nation, K. (2018). Examining incidental word learning during reading in children: The role of context. *Journal of Experimental Child Psychology*, *166*, 190–211.

Joseph, H., Wonnacott, E., Forbes, P., & Nation, K. (2014). Becoming a written word: Eye movements reveal order of acquisition effects following incidental exposure to new words during silent reading. *Cognition*, *133*(1), 238–248.

Joseph, H. S., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*,

49(16), 2078–2086.

Juhasz, B. J., & Rayner, K. (2003). Investigating the effects of a set of intercorrelated variables on eye fixation durations in reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(6), 1312–1318.

Juphard, A., Carbonnel, S., & Valdois, S. (2004). Length effect in reading and lexical decision: Evidence from skilled readers and a developmental dyslexic participant. *Brain and Cognition*, *55*(2), 332–340.

Kliegl, R., Grabner, E., Rolfs, M., & Engbert, R. (2004). Length, frequency, and predictability effects of words on eye movements in reading. *European journal of cognitive psychology*, *16*(1-2), 262–284.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Research*, *39*, 2729–2737.

Kyte, C. S., & Johnson, C. J. (2009). The role of phonological recoding in orthographic learning. *Journal of Experimental Child Psychology*, *93*(2), 166–185.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. *Cognitive psychology*, *6*(2), 293–323.

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): still no identification without attention. *Psychological review*, *111*(4), 880.

Lacroux, C. M. (2015). La prise en compte des fautes d'orthographe dans les dossiers de candidature par les recruteurs: une étude empirique par la méthode des protocoles verbaux. *@GRH*(1), 73–97.

Landi, N., Perfetti, C. A., Bolger, D. J., Dunlap, S., & Foorman, B. R. (2006). The role of discourse context in developing word form representations: A paradoxical relation between reading and learning. *Journal of Experimental Child Psychology*, *94*(2), 114–133.

Lauritzen, S., & Spiegelhalter, D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems (with discussion). *Journal of the Royal Statistical Society series B*, *50*, 157–224.

Lee, T. S., & Stella, X. Y. (2000). An information-theoretic framework for understanding saccadic eye movements. In *Advances in neural information processing systems* (pp. 834–840).

Legge, G. E., Hooven, T. A., Klitz, T. S., Mansfield, J. S., & Tjan, B. S. (2002). Mr. Chips 2002: New insights from an ideal-observer model of reading. *Vision research*, *42*(18), 2219–2234.

Legge, G. E., Klitz, T. S., & Tjan, B. S. (1997). Mr. Chips: an ideal-observer model of reading. *Psychological review*, *104*(3), 524.

Lesmes, L. A., Jeon, S.-T., Lu, Z.-L., & Dosher, B. A. (2006). Bayesian adaptive estimation of threshold versus contrast external noise functions: The quick *TvC* method. *Vision Research*, *46*, 3160–3176.

Lien, M.-C., Ruthruff, E., Kouchi, S., & Lachter, J. (2010). Event frequent and expected words are not identified without spatial attention. *Attention, Perception and Psychophysics*, *72*(4), 973–988.

Lobier, M., Dubois, M., & Valdois, S. (2013). The role of visual processing speed in reading speed development. *PLoS ONE*, *8*(4), e58097.

Lobier, M., & Valdois, S. (2015). Visual attention deficits in developmental dyslexia cannot be ascribed solely to poor reading experience. *Nature Reviews Neuroscience*, *16*(4), 225.

Lowell, R., & Morris, R. K. (2014). Word length effects on novel words: Evidence from eye movements. *Attention, Perception, & Psychophysics*, *76*(1), 179–189.

Magnuson, J. S., Mirman, D., Luthra, S., Strauss, T., & Harris, H. D. (2018). Interaction in

spoken word recognition models: Feedback helps. *Frontiers in Psychology*, *9*, 369.

Mancheva, L., Reichle, E. D., Lemaire, B., Valdois, S., Ecalle, J., & Guérin-Dugué, A. (2015). An analysis of reading skill development using E-Z Reader. *Journal of Cognitive Psychology*, *27*(5), 657–676.

Martens, V. E., & De Jong, P. F. (2008). Effects of repeated reading on the length effect in word and pseudoword reading. *Journal of Research in Reading*, *31*(1), 40–54.

Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, *44*(2), 314–324.

Matuschek, H., Kliegl, R., Vasishth, S., Baayen, H., & Bates, D. (2017). Balancing type I error and power in linear mixed models. *Journal of Memory and Language*, *94*, 305–315.

McCann, R. S., Folk, C. L., & Johnston, J. C. (1992). The role of spatial attention in visual word processing. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1015–1029.

McClelland, J. L. (2013). Integrating probabilistic models of perception and interactive neural networks: a historical and tutorial review. *Frontiers in Psychology*, *4*, 503.

McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. an account of basic findings. *Psychological Review*, *88*(5), 375–407.

Morrison, R. E. (1984). Manipulation of stimulus onset delay in reading: evidence for parallel programming of saccades. *Journal of Experimental psychology: Human Perception and performance*, *10*(5), 667.

Mousikou, P., & Schroeder, S. (2019). Morphological processing in single-word and sentence reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(5), 881–903.

Mozer, M. C., & Behrmann, M. (1990). On the interaction of selective attention and lexical knowledge: A connectionist account of neglect dyslexia. *Journal of Cognitive Neuroscience*, *2*(2), 96–123.

Murphy, K. (2002). *Dynamic Bayesian networks: Representation, inference and learning* (Ph. D. thesis). University of California, Berkeley, Berkeley, CA.

Najemnik, J., & Geisler, W. S. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387.

Nation, K., Angell, P., & Castles, A. (2007). Orthographic learning via self-teaching in children learning to read English: Effects of exposure, durability, and context. *Journal of Experimental Child Psychology*, *96*(1), 71–84.

Nation, K., & Castles, A. (2017). Putting the learning into orthographic learning. In K. Cain, D. L. Compton, & R. K. Parrila (Eds.), *Theories of reading development* (pp. 147–168). Amsterdam / Philadelphia, PA: John Benjamins Publishing Company.

Navalpakkam, V., Koch, C., Rangel, A., & Perona, P. (2010). Optimal reward harvesting in complex perceptual environments. *Proceedings of the National Academy of Sciences*, *107*(11), 5232–5237.

New, B., Ferrand, L., Pallier, C., & Brysbaert, M. (2006). Reexamining the word length effect in visual word recognition: New evidence from the English Lexicon Project. *Psychonomic Bulletin & Review*, *13*(1), 45–52.

New, B., Pallier, C., Ferrand, L., & Matos, R. (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUE™. *L'année psychologique*, *101*(3), 447–462.

Norris, D. (2006). The Bayesian reader: Explaining word recognition as an optimal bayesian decision process. *Psychological Review*, *113*(2), 327–357.

Norris, D. (2013). Models of visual word recognition. *Trends in cognitive sciences*, *17*(10),

517–524.

Norris, D., McQueen, J., & Cutler, A. (2018). Commentary on "interaction in spoken word recognition models". *Frontiers in Psychology*, *9*, 1568.

O'Regan, J. K., & Jacobs, A. M. (1992). Optimal viewing position effect in word recognition: A challenge to current theory. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(1), 185–197.

Pagan, A., & Nation, K. (2019). Learning words via reading: Contextual diversity, spacing and retrieval effects in adults. *Cognitive Science*, *43*.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Francisco, CA, USA: Morgan Kaufmann.

Perry, C., & Ziegler, J. C. (2002). Cross-language computational investigation of the length effect in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(4), 990–1001.

Perry, C., Ziegler, J. C., & Zorzi, M. (2007). Nested incremental modeling in the development of computational theories: the CDP+ model of reading aloud. *Psychological Review*, *114*(2), 273–315.

Perry, C., Ziegler, J. C., & Zorzi, M. (2010). Beyond single syllables: Large-scale modeling of reading aloud with the Connectionist Dual Process (CDP++) model. *Cognitive Psychology*, *61*(2), 106–51.

Perry, C., Ziegler, J. C., & Zorzi, M. (2013). A computational and empirical investigation of graphemes in reading. *Cognitive Science*, *1*, 29.

Perry, C., Zorzi, M., & Ziegler, J. C. (2019). Understanding dyslexia through personalized large-scale computational models. *Psychological science*, *30*(3), 386–395.

Peterson, R. L., Pennington, B. F., & Olson, R. K. (2013). Subtypes of developmental dyslexia: Testing the predictions of the dual-route and connectionist frameworks. *Cognition*, *126*(1), 20–38.

Phénix, T. (2018). *Modélisation bayésienne algorithmique de la reconnaissance visuelle de mots et de l'attention visuelle* (Unpublished doctoral dissertation). Univ. Grenoble Alpes.

Phénix, T., Diard, J., & Valdois, S. (2016). Les modèles computationnels de lecture. In M. Sato & S. Pinto (Eds.), *Traité de neurolinguistique* (pp. 167–182). Louvain-la-Neuve, Belgium: De Boeck supérieur.

Phénix, T., Valdois, S., & Diard, J. (2018). Reconciling opposite neighborhood frequency effects in lexical decision: Evidence from a novel probabilistic model of visual word recognition. In T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings of the 40th annual conference of the cognitive science society* (pp. 2238–2243). Austin, TX: Cognitive Science Society.

Phénix, T., Valdois, S., & Diard, J. (soumis). The role of attention in visual word recognition : A Bayesian programming approach. —.

Plaut, D. C. (1999). A connectionist approach to word reading and acquired dyslexia: Extension to sequential processing. *Cognitive Science*, *23*(4), 543–568.

Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: computational principles in quasi-regular domains. *Psychological Review*, *103*(1), 56–115.

Pollatsek, A., Juhasz, B. J., Reichle, E. D., Machacek, D., & Rayner, K. (2008). Immediate and delayed effects of word frequency and word length on eye movements in reading: A reversed delayed effect of word length. *Journal of Experimental Psychology: Human Perception and Performance*, *34*(3), 726–750.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring

the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*(1), 1–56.

Prado, C., Dubois, M., & Valdois, S. (2007). The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span. *Vision Research*, *47*, 2521–2530.

Pritchard, S. C., Coltheart, M., Marinus, E., & Castles, A. (2018). A computational model of the self-teaching hypothesis based on the dual-route cascaded model of reading. *Cognitive Science*, 1–49.

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Raj, R., Geisler, W. S., Frazor, R. A., & Bovik, A. C. (2005). Contrast statistics for foveated visual systems: Fixation selection by minimizing contrast entropy. *JOSA A*, *22*(10), 2039–2049.

Rau, A. K., Moll, K., Snowling, M. J., & Landerl, K. (2015). Effects of orthographic consistency on eye movement behavior: German and English children and adults process the same words differently. *Journal of Experimental Child Psychology*, *130*, 92–105.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372–422.

Rayner, K. (2009). The 35th Sir Frederick Bartlett lecture: Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.

Rayner, K., & Johnson, R. L. (2005). Letter-by-letter acquired dyslexia is due to the serial encoding of letters. *Psychological Science*, *16*(7), 530–534.

Rayner, K., & Raney, G. E. (1996). Eye movement control in reading and visual search: Effects of word frequency. *Psychonomic Bulletin & Review*, *3*(2), 245–248.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological review*, *105*(1), 125.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and brain sciences*, *26*(4), 445–476.

Reichle, E. D., Warren, T., & McConnell, K. (2009). Using E-Z Reader to model the effects of higher level language processing on eye movements during reading. *Psychonomic Bulletin & Review*, *16*(1), 1–21.

Reingold, E. M., Reichle, E. D., Glaholt, M. G., & Sheridan, H. (2012). Direct lexical control of eye movements in reading: Evidence from a survival analysis of fixation durations. *Cognitive psychology*, *65*(2), 177–206.

Renninger, L. W., Verghese, P., & Coughlan, J. (2007). Where to look next? eye movements reduce local uncertainty. *Journal of vision*, *7*(3), 6–6.

Richardson, J. T. (1976). The effects of stimulus attributes upon latency of word recognition. *British Journal of Psychology*, *67*(3), 315–325.

Ricketts, J., Bishop, D. V., Pimperton, H., & Nation, K. (2011). The role of self-teaching in learning orthographic and semantic aspects of new words. *Scientific Studies of Reading*, *15*(1), 47–70.

Risko, E. F., Stolz, J. A., & Besner, D. (2010). Spatial attention modulates feature crosstalk in visual word processing. *Attention, Perception and Psychophysics*, *72*(4), 989–998.

Romani, C., Di Betta, A. M., Tsouknida, E., & Olson, A. (2008). Lexical and nonlexical processing in developmental dyslexia: A case for different resources and different impairments. *Cognitive neuropsychology*, *25*(6), 798–830.

Romani, C., Ward, J., & Olson, A. (1999). Developmental surface dysgraphia: What is the underlying cognitive impairment? *The Quarterly Journal of Experimental Psychology Section A*, *52*(1), 97–128.

Salvucci, D. D. (2001). An integrated model of eye movements and visual encoding. *Cognitive Systems Research*, *1*(4), 201–220.

Scarf, D., Boy, K., Reinert, A. U., Devine, J., Güntürkün, O., & Colombo, M. (2016). Orthographic processing in pigeons (columba livia). *Proceedings of the National Academy of Sciences*, *113*(40), 11272–11276.

Schilling, H. E., Rayner, K., & Chumbley, J. I. (1998). Comparing naming, lexical decision, and eye fixation times: Word frequency effects and individual differences. *Memory & Cognition*, *26*(6), 1270–1281.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568.

Seidenberg, M. S., & Plaut, D. C. (1998). Evaluating word-reading models at the item level: Matching the grain of theory and data. *Psychological Science*, *9*(3), 234–237.

Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, *55*(2), 151–218.

Share, D. L. (1999). Phonological recoding and orthographic learning: A direct test of the self-teaching hypothesis. *Journal of Experimental Child Psychology*, *72*(2), 95–129.

Share, D. L. (2004). Orthographic learning at a glance: On the time course and developmental onset of self-teaching. *Journal of Experimental Child Psychology*, *87*(4), 267–298.

Share, D. L. (2008). Orthographic learning, phonological recoding and self-teaching. *Advances in Child Development and Behavior*, 31–82.

Stanovich, K. E., & West, R. F. (1989). Exposure to print and orthographic processing. *Reading Research Quarterly*, 402–433.

Stein, J. (2014). Dyslexia: the role of vision and visual attention. *Current developmental disorders reports*, *1*(4), 267–280.

Stolz, J. A., & Stevanovski, B. (2004). Interactive activation in visual word recognition: Constraints imposed by the joint effect of spatial attention and semantics. *Journal of Experimental Psychology: Human Perception and Performance*, *30*, 1064–1076.

Tamura, N., Castles, A., & Nation, K. (2017). Orthographic learning, fast and slow: Lexical competition effects reveal the time course of word learning in developing readers. *Cognition*, *163*, 93–102.

Taylor, J. S. H., Plunkett, K., & Nation, K. (2011). The eye movements of dyslexic children during reading and visual search: Impact of the visual attention span. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *37*(1), 60–76.

Tucker, R., Castles, A., Laroche, A., & Deacon, H. (2016). The nature of orthographic learning in self-teaching: Testing the extent of transfer. *Journal of Experimental Child Psychology*, *145*, 79–94.

Valdois, S., Bidet-Ildei, C., Lassus-Sangosse, D., Reilhac, C., N'guyen-Morel, M.-A., Guinet, E., & Orliaguet, J.-P. (2011). A visual processing but no phonological disorder in a child with mixed dyslexia. *Cortex*, *47*(10), 1197–1218.

Valdois, S., Bosse, M.-L., Ans, B., Carbonnel, S., Zorman, M., David, D., & Pellat, J. (2003). Phonological and visual processing deficits can dissociate in developmental dyslexia: Evidence from two case studies. *Reading and Writing: An Interdisciplinary Journal*, *16*, 541–572.

Valdois, S., Carbonnel, S., Juphard, A., Baciu, M., Ans, B., Peyrin, C., & Segebarth, C. (2006). Polysyllabic pseudo-word processing in reading and lexical decision: converging evidence

from behavioral data, connectionist simulations and functional MRI. *Brain Research*, *1085*(1), 149–162.

Valdois, S., Roulin, J.-L., & Bosse, M. L. (submitted). Visual attention modulates reading acquisition. —.

van den Boer, M., de Jong, P. F., & van Meeteren, M. M. H. (2013). Modeling the length effect: Specifying the relation with visual and phonological correlates of reading. *Scientific Studies of Reading*, *17*(4), 243–256.

van den Boer, M., Elsje, v. B., & de Jong, P. F. (2015). The specific relation of visual attention span with reading and spelling in Dutch. *Learning and individual differences*, *39*, 141–149.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision research*, *41*(25-26), 3513–3533.

Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics*, *47*(6), 583–600.

Wang, H.-C., Castles, A., Nickels, L., & Nation, K. (2011). Context effects on orthographic learning of regular and irregular words. *Journal of Experimental Child Psychology*, *109*(1), 39–57.

Wang, H.-C., Nickels, L., Nation, K., & Castles, A. (2013). Predictors of orthographic learning of regular and irregular words. *Scientific Studies of Reading*, *17*(5), 369–384.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). Eye movements and the modulation of parafoveal processing by foveal processing difficulty: A reexamination. *Psychonomic Bulletin & Review*, *12*(5), 891–896.

Whitney, C. (2001). How the brain encodes the order of letters in a printed word: the SERIOL model and selective literature review. *Psychonomic Bulletin & Review*, *8*(2), 221–43.

Whitney, C. (2018). When serial letter processing implies a facilitative length effect. *Language, Cognition and Neuroscience*, *33*(5), 659–664.

Whitney, D., & Levi, D. M. (2011). Visual crowding: a fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences*, *15*(4), 160–168.

Williams, R., & Morris, R. (2004). Eye movements, word familiarity, and vocabulary acquisition. *European Journal of Cognitive Psychology*, *16*(1-2), 312–339.

Wochna, K. L., & Juhasz, B. J. (2013). Context length and reading novel words: An eye-movement investigation. *British Journal of Psychology*, *104*(3), 347–363.

Yarkoni, T., Balota, D., & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, *15*(5), 971–979.

Ziegler, J. C., Perry, C., & Zorzi, M. (2014). Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *369*(1634), 20120397.

Zoubrinetzky, R., Bielle, F., & Valdois, S. (2014). New insights on developmental dyslexia subtypes: Heterogeneity of mixed reading profiles. *PLoS ONE*, *9*(6), e99337.

Zoubrinetzky, R., Collet, G. M., Nguyen Morel, M. A., Valdois, S., & Serniclaes, W. (2019). Remediation of allophonic perception and visual attention span in developmental dyslexia: a joint assay. *Frontiers in Psychology*, *10*, 1502.

Titre — Modélisation bayésienne et étude expérimentale du rôle de l'attention visuelle dans l'acquisition des connaissances lexicales orthographiques

Résumé —

Dans cette thèse, nous étudions le rôle de l'attention visuelle lors de l'acquisition, par un lecteur expert, de nouvelles connaissances lexicales orthographiques. Notre contribution est double. D'une part, nous développons un nouveau modèle computationnel, probabiliste d'apprentissage orthographique. Notre modèle, nommé *BRAID-Learn*, est une extension de BRAID, un modèle probabiliste hiérarchique de la reconnaissance visuelle de mots et de décision lexicale. D'autre part, nous apportons des données expérimentales originales sur l'évolution des mouvements oculaires lors de l'apprentissage incident de formes orthographiques nouvelles et démontrons la capacité du modèle à rendre compte de ces observations. Notre contribution est décrite dans trois articles.

Dans le premier article, nous simulons l'effet de longueur tel qu'observé expérimentalement pour les mots en décision lexicale dans le French Lexicon Project. À travers 5 simulations, nous montrons que l'attention visuelle module l'effet de longueur et que réaliser plusieurs fixations attentionnelles lors du traitement des mots longs (à partir de 7 lettres) réduit l'effet de longueur au lieu de l'accentuer. Cette étude nous permet de calibrer les paramètres du sous-module de décision lexicale de notre modèle d'apprentissage.

Le second article porte sur l'étude expérimentale du comportement oculomoteur de lecteurs adultes experts lors de l'apprentissage incident de mots nouveaux. Nous montrons que le nombre de fixation et la durée de traitement évoluent en fonction du nombre d'exposition avec un nouveau mot, témoignant du renforcement progressif de sa représentation orthographique en mémoire. Des données exploratoires suggèrent qu'apprentissage orthographique et comportements oculomoteurs sont également modulés par les capacités visuo-attentionnelles des participants.

Dans le troisième et dernier article, nous présentons le modèle d'apprentissage *BRAID-Learn* et testons sa capacité à rendre compte des données oculomotrices précédemment décrites en condition d'apprentissage orthographique. Le modèle repose sur deux hypothèses originales. La première est que le système contrôle les paramètres visuo-attentionnels afin d'optimiser l'accumulation d'information perceptive sur l'identité des lettres du stimulus et, donc, de construire efficacement une nouvelle trace orthographique pendant l'apprentissage. La seconde hypothèse est que la familiarité lexicale, c'est-à-dire, la probabilité que le stimulus présenté soit un mot connu, module l'influence descendante des représentations lexicales sur la perception des lettres. Nous montrons que le modèle reproduit avec succès les observations, c'est-à-dire la diminution du nombre de fixations et du temps de traitement pour les mots nouveaux au fil des répétitions.

BRAID-Learn est le premier modèle d'apprentissage orthographique à établir un lien explicite entre acquisition orthographique et mouvements oculomoteurs en condition d'apprentissage incident. Une autre contribution importante de cette thèse est de montrer et préciser le rôle de l'attention visuelle dans l'apprentissage orthographique, suggérant que cette dimension pourrait être fortement impliquée dans le passage du mode de lecture analytique caractéristique de l'apprenant au mode de lecture global qui caractérise le lecteur expert.

Mots clés — Modèle computationnel, modélisation bayésienne, apprentissage orthographique, mouvements oculaires, attention visuelle, décision lexicale, influences top-down.

Title — Bayesian modeling and experimental study of the role of visual attention in the acquisition of orthographic lexical knowledge

Abstract —

In this thesis, we study the role of visual attention when an expert reader acquires new orthographic lexical knowledge. Our contribution is twofold. On the one hand, we develop an original computational, probabilistic model of orthographic learning. Our model, named *BRAID-Learn*, is an extension of BRAID, a hierarchical probabilistic model of visual word recognition and lexical decision. On the other hand, we gather original experimental data on the evolution of eye movements during incidental learning of new orthographic forms and demonstrate the ability of the model to account for these observations. Our contribution is described in three articles.

In the first article, we simulate the length effect as experimentally observed for words in lexical decision in the French Lexicon Project. Through 5 simulations, we show that visual attention modulates the length effect and that several attentional fixations during the processing of long words (7 letters or more) reduces the length effect, instead of accentuating it. This study allows us to calibrate the parameters of the lexical decision sub-model of our learning model.

The second article focuses on the experimental study of oculomotor behavior of expert adult readers during the incidental learning of new words. We show that the number of fixations and processing duration vary according to the number of exposures to a novel word, testifying to the progressive strengthening of its orthographic representation in memory. Exploratory data suggest that orthographic learning and oculomotor behaviors are also modulated by the visual-attentional abilities of the participants.

In the third and final article, we present the learning model *BRAID-Learn* and test its ability to account for previously described oculomotor data in orthographic learning conditions. The model is based on two original hypotheses. The first is that the system controls the visual-attentional parameters in order to optimize the accumulation of perceptual information on letters of the stimulus and, therefore, to efficiently build a new orthographic trace during learning. The second hypothesis is that lexical familiarity, that is, the probability that the stimulus presented is a known word, modulates the top-down influence of lexical representations on letter perception. We show that the model successfully reproduces the observations, namely the decrease of the number of fixations as well as processing duration for novel words across exposures.

BRAID-Learn is the first orthographic learning model to establish an explicit link between orthographic learning and eye movements observed during the incidental orthographic learning. Another contribution of this thesis is to show and clarify the role of visual attention in orthographic learning, suggesting that this dimension could be strongly involved in the transition from serial reading, that characterizes learning readers, to global reading, that characterizes expert readers.

Keywords — Computational modeling, Bayesian modeling, orthographic learning, eye movements, visual attention, lexical decision, top-down influences.