

CEDoc : Lettres, Sciences Humaines, Arts et Science de l'Éducation

Formation Doctorale : Sciences du langage, Cognition et Didactiques des langues

Spécialité : Langue et Littérature Françaises

L'intégration spatiale de l'information orthographique lors de la lecture de mots en langue arabe : Étude chez l'adulte et l'enfant d'école primaire

Thèse pour l'Obtention du Doctorat

Nom et Prénom de l'Étudiant : ZEGHLI Hicham
CNE : 0321830048

Sous le Direction des Professeurs
M. Mostafa BOUANANI
et
M. Bernard LÉTÉ

Année universitaire 2021/2022



THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE LYON

En cotutelle internationale avec

L'UNIVERSITÉ SIDI MOHAMED BEN ABDELLAH

Opéré par

L'UNIVERSITÉ LUMIÈRE-LYON 2

Discipline : Sciences cognitives

École doctorale : Neurosciences et Cognition (ED 476 NSCO)

Unité de recherche : Laboratoire d'Étude des Mécanismes Cognitifs (UR EMC)

Présentée et soutenue par

Hicham ZEGHLI

Le 17/12/2021

L'intégration spatiale de l'information orthographique lors de la lecture de mots en langue arabe : Étude chez l'adulte et l'enfant d'école primaire

Directeurs de thèse : Pr. Bernard LÉTÉ & Pr. Mostafa BOUANANI

Membres du jury

- | | |
|---|--------------------|
| - Pr. Souad EL AYACHI, Université Sidi Mohamed Ben Abdellah de Fès | Examinatrice |
| - Pr. Séverine CASALIS, Université Lille | Rapporteuse |
| - Pr. Yousri MARZOUKI, Université du Qatar | Rapporteur |
| - Pr. Stéphanie DUCROT, Université Aix-Marseille | Examinatrice |
| - Pr. Mostafa BOUANANI, Université Sidi Mohamed Ben Abdellah de Fès | Directeur de thèse |
| - Pr. Bernard LÉTÉ, Université Lumière-Lyon 2 | Directeur de thèse |

À mes parents...

"Rien d'aventuré, rien d'atteint."

Remerciements

Je tiens à exprimer ma profonde reconnaissance à mes co-directeurs de thèse,

M. Mostafa BOUANANI, pour m'avoir soutenu, conseillé et motivé tout au long de ce parcours. Je le remercie pour m'avoir fait confiance, et mis sur le chemin de cette belle aventure, vers le voyage et le savoir.

M. Bernard LÉTÉ, pour sa disponibilité et son suivi permanent, sa grande patience et sa précieuse confiance ; pour ses remarques et précisions qui ont alimenté mes réflexions et grâce auxquelles j'ai appris tellement de choses. Je le remercie pour sa convivialité, son bon accueil au sein du laboratoire, son accompagnement tout au long de mes séjours et son soin de m'offrir les bonnes conditions de travail.

Mes grands remerciements vont également

À Mme Stéphanie DUCROT et M. Jacques JAYEZ d'avoir accepté d'être présents en tant que jury à mon conseil de suivi de thèse.

Aux professeurs M. Benaïssa ZARHBOUCH et M. Imsail ALAOUI pour leur importante collaboration qui m'était d'une grande aide.

À tous les membres du jury qui ont accepté d'évaluer ce travail, et qui me font l'honneur d'être présents à la soutenance.

Un merci tout particulier est adressé

À Christophe, pour m'avoir accordé son temps et avoir partagé avec moi ses expériences dans le cadre de ma thèse ; pour les échanges enrichissants et sa gentillesse.

À Brice et Emma pour leur grande sympathie ; pour avoir toujours pensé à moi à mon absence.

À Stéphanie pour m'avoir accompagné au début.

À Eddy pour m'avoir accueilli dans ses cours.

Je tiens finalement à remercier tous les élèves et étudiants ayant participé à ce travail, ainsi que les directeurs des écoles m'ayant accueilli dans leurs établissements.

Résumé

Ce travail de thèse porte sur l'intégration spatiale de l'information orthographique lors de la lecture de mots en arabe. À l'aide du paradigme *Flanking Letter Lexical Decision* de Dare et Shillcock (2013), trois questions principales ont été abordées : 1) la question de la nature du codage orthographique dans le système d'écriture de l'arabe ; 2) la question de la mise en place progressive d'une expertise en lecture chez les apprentis-lecteurs arabophones, traduite par un traitement parallèle des lettres ; 3) la question de l'impact de la direction de la lecture sur les traitements visuo-attentionnels en arabe. L'Expérience 1 conforte le modèle en bigrammes ouverts de Grainger et Van Heuven (2003) en confirmant le rôle des bigrammes dans le traitement orthographique en arabe. L'Expérience 2 révèle une mise en place tardive du traitement parallèle des lettres chez l'apprenti-lecteur arabophone comparativement à l'apprenti-lecteur français. Les Expériences 3 et 4 ne démontrent aucun biais attentionnel lié à la direction de la lecture en arabe. Ces données recueillies en arabe ont été confrontées à celles obtenues en langues à alphabet latin avec le même paradigme. L'objectif étant de mieux comprendre les mécanismes cognitifs universels sous-tendant la lecture, et les mécanismes spécifiques sous-tendus par les caractéristiques du système orthographique arabe.

Mots-clés

Système d'écriture arabe – Reconnaissance de mots écrits – Traitement orthographique – Apprentissage de la lecture – Flankers – Traitement parallèle des lettres.

Abstract

This thesis work focuses on the spatial integration of orthographic information during word reading in Arabic. Using Dare and Shillcock's (2013) *Flanking Letter Lexical Decision* paradigm, three main questions were addressed: 1) the question of the nature of orthographic encoding in the Arabic orthographic system; 2) the question of the gradual development of reading expertise in Arabic beginning readers, reflected in parallel letter processing; and 3) the question of the impact of reading direction on visual-attentional processing in Arabic. Experiment 1 supports Grainger and Van Heuven's (2003) open bigram model by confirming the role of bigrams in orthographic processing in Arabic. Experiment 2 reveals a late onset of parallel letter processing in Arabic beginning readers compared to French beginning readers. Experiments 3 and 4 show no attentional bias related to the direction of reading in Arabic. These data collected in Arabic were compared to those obtained in Latin alphabet languages with the same paradigm. The aim was to better understand the universal cognitive mechanisms underlying reading, and the specific mechanisms constrained by the characteristics of the Arabic orthographic system.

Keywords

Arabic script – Visual word recognition – Orthographic processing – Learning to read – Flanking letters – Parallel processing of letters.

Table des matières

Présentation générale	1
Introduction	4
1. Chapitre 1 : La reconnaissance des mots écrits en écriture latine	7
1.1. Les systèmes orthographiques	7
1.2. Le codage des lettres et de leurs positions	8
1.2.1. Un codage des lettres en position absolue ou en position relative ?	9
1.2.2. Le modèle à bigrammes ouverts de Grainger et Van Heuven (2003)	10
1.2.3. Un traitement parallèle ou séquentiel des lettres ?	12
1.3. Les contraintes oculomotrices et attentionnelles en lecture	13
1.3.1. Les contraintes oculomotrices	13
1.3.1.1. Le champ visuel	13
1.3.1.2. Les mouvements oculaires : saccades et fixations	14
1.3.2. Les contraintes attentionnelles	15
1.3.2.1. L'empan perceptif (perceptual span)	15
1.3.2.1.1. Flexibilité de la taille de l'empan perceptif	16
1.3.2.1.2. Développement de la taille de l'empan perceptif chez l'apprenti-lecteur	17
1.3.2.2. Phénomène de masquage latéral (crowding)	18
1.3.2.3. La position optimale de fixation (Optimal Viewing Position)	20
1.4. Traitement de l'information parafovéale chez l'adulte	23
1.4.1. Paradigme d'étude de mouvements oculaires (Eye-tracking)	23
1.4.2. Paradigme d'amorçage masqué	25
1.4.3. Paradigme FLLD (Flanking letter lexical decision)	25
1.5. Traitement de l'information parafovéale chez l'apprenti-lecteur	27
2. Chapitre 2 : La reconnaissance des mots écrits en arabe	30
2.1. La langue arabe	30
2.1.1. Aperçu historique	30
2.1.2. L'alphabet arabe	31
2.1.3. Caractéristiques morphologiques	34
2.1.4. Synthèse	35
2.2. La reconnaissance des mots en écriture arabe	36
2.3. Le traitement orthographique du mot écrit en arabe	39
2.3.1. La reconnaissance des lettres	39
2.3.2. Le codage de la position des lettres en arabe	41
2.4. Les traitements visuo-attentionnels et la position optimale de fixation (<i>Optimal Viewing Position</i>) en écriture arabe	44
2.4.1. Traitements visuo-attentionnels	44

2.4.2.	Position optimale de fixation	44
3.	Chapitre 3 : Reconnaissance des mots écrits et apprentissage de la lecture en langue arabe..	47
3.1.	Organisation de l'enseignement primaire au Maroc	47
3.2.	Le problème de la diglossie pour l'apprenti-lecteur arabe	48
3.3.	Le traitement des signes diacritiques.....	49
3.4.	La perceptibilité des lettres en écriture arabe	51
4.	Chapitre 4 : Synthèse théorique.....	54
5.	Chapitre 5 : Expérience 1. Étude de l'effet des flankers chez l'adulte	58
5.1.	Objectif	58
5.2.	Méthode.....	58
5.2.1.	Participants.....	58
5.2.2.	Matériel	58
5.2.3.	Procédure	60
5.3.	Résultats	61
5.3.1.	Analyse sur les mots	62
5.3.2.	Analyse sur les pseudo-mots.....	63
5.4.	Synthèse de l'expérience 1.....	64
6.	Chapitre 6 : Expérience 2. Étude de l'effet des flankers chez l'apprenti-lecteur.....	66
6.1.	Objectif	66
6.2.	Méthode.....	66
6.2.1.	Participants.....	66
6.2.2.	Matériel	67
6.2.3.	Procédure	67
6.3.	Résultats	67
6.3.1.	Analyse sur les mots	67
6.3.2.	Analyse sur les pseudo-mots.....	71
6.4.	Synthèse de l'expérience 2.....	74
7.	Chapitre 7 : Étude de l'effet de la direction de la lecture sur le traitement des mots chez l'adulte	76
7.1.	Objectif	76
7.2.	Expérience 3 : Direction de la lecture – Présentation unilatérale de bigrammes.....	76
7.2.1.	Méthode.....	76
7.2.1.1.	Participants.....	76
7.2.1.2.	Matériel	76
7.2.1.3.	Procédure	77
7.2.2.	Résultats	77
7.2.2.1.	Analyse sur les mots	78
7.2.2.2.	Analyse sur les pseudo-mots.....	79

7.3.	Expérience 4 : Direction de la lecture – Présentation bilatérale de mots.....	80
7.3.1.	Méthode.....	80
7.3.1.1.	Participants.....	80
7.3.1.2.	Matériel.....	80
7.3.1.3.	Procédure.....	81
7.3.2.	Résultats.....	81
7.3.2.1.	Analyse sur les mots.....	82
7.3.2.2.	Analyse sur les pseudo-mots.....	83
7.4.	Synthèse des expériences 3 et 4.....	84
8.	Chapitre 8 : Discussion générale et perspectives de recherche.....	86
8.1.	Rappel des objectifs et synthèse des résultats.....	86
8.2.	Perspectives de recherche.....	90
	Conclusion.....	94
	Références.....	95
	Annexe.....	105

Présentation générale

La lecture est le résultat d'un apprentissage long reposant sur la mise en place de mécanismes complexes. Depuis plusieurs décennies, les chercheurs essaient de comprendre les processus mis en jeu au cours de la lecture, ainsi que le développement de ces processus au cours de l'apprentissage de la lecture. La reconnaissance des mots constitue une étape essentielle dans le traitement du langage écrit. Elle implique le codage correct de deux types d'information : l'identité des lettres et leur position dans le mot. Par exemple, on peut différencier deux voisins orthographiques tels que "lire" et "dire" sur la base de l'identité des lettres, alors que l'information positionnelle est également importante pour faire la distinction entre deux anagrammes tels que "orange" et "organe".

Les recherches sur le codage de l'information orthographique se sont essentiellement focalisées sur l'écriture latine des langues indo-européennes. Les caractéristiques de chaque langue à l'origine de la diversité des systèmes d'écriture imposent pourtant des contraintes dans la reconnaissance des mots qui sont rarement prises en compte par les chercheurs. L'arabe standard moderne, une langue alphabétique de la famille des langues sémitiques, est reconnu en tant que langue officielle par 25 états, ce qui le place en troisième position après l'anglais et le français. À cette importance quantitative, s'ajoute une importance qualitative qui provient de ses propriétés structurelles distinctives, de sa directionnalité inversée par rapport à l'écriture latine et de la spécificité de son système orthographique. Cela lui confère depuis quelques années un intérêt croissant auprès des chercheurs pour mieux comprendre les mécanismes cognitifs impliqués dans le traitement du mot écrit et l'apprentissage de la lecture, et ainsi mieux appréhender les aspects universaux des mécanismes cognitifs impliqués et les aspects spécifiques sous-tendus par le système d'écriture arabe.

Ce travail de thèse s'inscrit dans cette perspective et vise à étudier plus spécifiquement les mécanismes d'intégration spatiale de l'information orthographique pendant la lecture de mots en écriture arabe à l'aide du paradigme *Flanking Letter Lexical Decision* (ci-après FLLD) développé par Dare et Shillcock (2013). Dare et Shillcock (2013) ont demandé à leurs participants de faire des décisions lexicales sur des mots et des non-mots cibles entourés de deux lettres à gauche et à droite des stimuli cibles. Ces flankers pouvaient être reliés orthographiquement au mot cible (par exemple, RO ROCK CK) ou non reliés (par exemple, PA

ROCK TH). Les participants pouvaient ignorer les bigrammes du fait qu'ils n'étaient pas pertinents pour la tâche. Dare et Shillcock ont constaté que les décisions lexicales étaient facilitées par des bigrammes reliés orthographiquement, non seulement lorsqu'ils respectaient leur ordre dans la cible (comme dans l'exemple ci-dessus), mais également lorsque l'ordre était inversé (par exemple, CK ROCK RO). Le paradigme a été utilisé depuis par Grainger et collaborateurs (e.g., Cauchi, Lété, & Grainger, 2020 ; Grainger, Mathôt, & Vitu, 2014 ; Snell, Bertrand, & Grainger, 2018 ; Snell, Cauchi, Grainger, & Lété, 2020) et les résultats, interprétés comme reflétant l'intégration spatiale des informations orthographiques entre la cible et les flankers, ont permis d'améliorer les modèles de traitement du mot écrit. Le paradigme est d'utilisation simple, y compris auprès d'enfants. Il permet d'étudier la nature des informations pouvant être intégrées à gauche et à droite d'un point de fixation (i.e., orthographique, phonologique, sémantique, morphologique, ...) en manipulant les caractéristiques linguistiques des flankers.

À l'aide de ce paradigme FLLD, il s'agira en particulier de déterminer si les effets observés en écriture latine peuvent refléter des mécanismes cognitifs universels sous-tendant la lecture ou s'il s'agit de mécanismes spécifiques sous-tendus par les caractéristiques du système orthographique. Nous chercherons également à savoir à quel moment de l'apprentissage de la lecture ces mécanismes se mettent en place et s'ils dépendent des habitudes de lecture, et donc de la direction d'allocation des traitements visuo-attentionnels.

Ce travail de thèse s'est organisé autour de trois questions principales.

1. La question de la nature du codage orthographique dans le système d'écriture de l'arabe : il s'agit de tester la robustesse et l'universalité du modèle à bigrammes ouverts de Grainger et Van Heuven (2003) en cherchant à savoir s'il peut être également un cadre explicatif des processus de codage orthographique impliqués dans la reconnaissance de mots écrits en arabe.
2. La question de l'apprentissage de la lecture en arabe : il s'agit d'étudier la trajectoire développementale des processus d'intégration spatiale des informations orthographiques lors de l'apprentissage de la lecture, en cherchant à cerner la transition entre un traitement séquentiel des informations orthographiques chez les lecteurs débutants, vers un traitement en parallèle des informations orthographiques tout au long de l'apprentissage de la lecture en cycle primaire.

3. La question de l'impact de la direction de la lecture en arabe sur les mécanismes d'intégration de l'information orthographique : il s'agit d'étudier l'effet de la direction de la lecture sur les traitements visuo-attentionnels, en cherchant à déterminer si les effets révélés dans un système orthographique qui se lit et s'écrit de gauche à droite, s'observent également dans un système qui se lit et s'écrit de droite à gauche.

Le présent document est structuré en trois grandes parties.

Dans la première partie de synthèse théorique, le Chapitre 1 présente l'état actuel des connaissances sur la reconnaissance des mots écrits en écriture latine. Le Chapitre 2 est consacré au système d'écriture arabe. Le Chapitre 3 fait le point de façon synthétique sur l'apprentissage de la lecture en arabe. Enfin, le Chapitre 4 présente une synthèse du cadre théorique permettant de positionner notre partie expérimentale et de justifier nos choix méthodologiques.

La deuxième partie présente les études comportementales réalisées. L'expérience 1 présentée au Chapitre 5 explore l'effet des flankers lors du traitement du mot écrit chez l'adulte. L'expérience 2 présentée au Chapitre 6 explore le même effet chez l'apprenti-lecteur. Les expériences 3 et 4 présentées au Chapitre 7 concernent pour leur part l'effet de la direction de la lecture sur le traitement du mot écrit chez l'adulte.

La troisième partie est dédiée à la discussion générale et aux perspectives de recherches (chapitre 8).

Introduction

Le traitement orthographique du mot écrit repose essentiellement sur les mécanismes de codage de l'identité des lettres et de leur position dans le mot. Deux hypothèses ont été proposées sur le codage positionnel : un codage en position absolue où chaque lettre est codée à sa bonne position dans le mot et un codage en position relative où chaque lettre peut être codée à une position différente mais pas trop éloignée de sa position réelle permettant de lire un mot dont deux lettres sont *pemrutées*. Les recherches en écriture latine ont donné des résultats plutôt en évidence pour un codage en position relative. Afin d'expliquer ces résultats, le modèle de Grainger et van Heuven (2003) introduit la notion de codage en bigrammes ouverts dont les effets ont été confortés avec le nouveau paradigme FLLD de Dare et Shillcock (2013).

Le traitement orthographique du mot écrit en *Arabe Standard Moderne* reste encore très peu étudié comparativement au traitement morphologique qui concentre la quasi-totalité des recherches du fait que l'une des caractéristiques principales de l'arabe est sa structure basée sur des racines consonantiques. La plupart des preuves de l'importance de ces unités morphémiques dans la reconnaissance du mot écrit en arabe proviennent des résultats obtenus en amorçage morphologique. D'autres preuves plus récentes en amorçage orthographique contestent pourtant l'importance de la racine. La dimension morphologique a été prise en compte également dans les quelques recherches sur le codage positionnel des lettres en arabe. Les résultats sont pourtant loin d'être convergents. En tâche de décision lexicale, le maintien de l'ordre des consonnes de la racine semble important pour obtenir des effets de transposition de lettres, impliquant un codage absolu. Cependant, en tâche *Same - Different*, des effets de transposition de lettres ont été obtenus même avec l'altération de l'ordre des consonnes de la racine dans le mot, démontrant plutôt un codage relatif de la position des lettres.

La plupart des recherches sur le traitement orthographique se sont penchées sur le lecteur expert. L'étude des trajectoires développementales du traitement orthographique du mot écrit reste peu investie, y compris dans les recherches en écriture latine. Néanmoins, les quelques recherches qui s'y rapportent ont révélé que les effets observés chez l'adulte dans l'intégration de l'information orthographique pendant la lecture se mettent en place assez

précocement (e.g., Snell et al., 2020). En écriture arabe, les études sur le traitement orthographique dans une perspective développementale sont quasiment inexistantes. Les travaux qui ont été menés ont porté principalement sur les aspects liés à la situation de diglossie pour l'apprenti-lecteur et sur la spécificité du graphisme de l'écriture arabe.

La question de l'impact de la direction de la lecture en écriture arabe est restée longtemps peu abordée. On observe cependant aujourd'hui un développement important des recherches dans ce domaine notamment grâce au recueil de mouvements oculaires (pour une synthèse récente voir AlJassmi, Hermena & Paterson, 2021). Pour les écritures latines se lisant de gauche à droite, une abondante littérature montre qu'il existe une distribution asymétrique de l'attention à droite du mot, dans le sens de la lecture. Cette asymétrie vers la droite semble s'inverser vers la gauche en écriture arabe mais, comme nous le verrons au Chapitre 2, les données restent encore trop parcellaires et nécessitent d'être confortées par d'autres travaux.

Le cadre théorique est structuré en trois chapitres.

Le Chapitre 1 apporte quelques éclaircissements théoriques sur les mécanismes du codage orthographique en écriture latine, de l'analyse des traits visuels à la reconnaissance du mot, en passant par la question du codage, relatif vs absolu, de la position des lettres dans le mot. Des arguments sont présentés en faveur d'un codage relatif de la position des lettres, étayés par le modèle à bigrammes ouverts de Grainger et Van Heuven (2003). Des arguments sont apportés également par différents paradigmes (mouvements oculaires ; amorçage masqué ; FLLD) en faveur de l'orientation des traitements visuo-attentionnels vers la droite (dans le sens de la lecture des langues latines) et du rôle du prétraitement parafovéal dans le traitement fovéal lors de la lecture experte et chez l'apprenti-lecteur, permettant ainsi un traitement parallèle des informations orthographiques.

Le Chapitre 2 est consacré à la langue arabe. Après avoir détaillé les caractéristiques de l'arabe standard moderne, une synthèse des travaux portant essentiellement sur le codage de l'identité et de la position des lettres dans le mot est présentée. Les traitements visuo-attentionnels sont également abordés en rapport avec la directionnalité de la lecture. On y verra que la singularité orthographique et morphologique de l'arabe, construite sur la base de racines, fournit des preuves mitigées de l'importance des consonnes de la racine et du codage absolu de leur position dans la reconnaissance de mots. Quelques autres preuves font

apparaître une inversion de l'orientation des ressources attentionnelles à gauche vers le sens de la lecture, contrairement à l'écriture latine.

Le Chapitre 3 recense les contraintes que rencontre l'apprenti-lecteur arabophone au cours de l'apprentissage de la lecture, liées à la situation de diglossie et à la complexité du système orthographique arabe.

1. Chapitre 1 : La reconnaissance des mots écrits en écriture latine

1.1. Les systèmes orthographiques

Dans les langues alphabétiques, un nombre limité de lettres permet de transcrire les mots parlés de la langue. Ce nombre peut varier d'une langue à l'autre (*e.g.*, 26 dans les langues européennes, 28 en arabe) à cause, notamment, du lien qu'entretiennent les lettres avec les sons qui leur correspondent. Un apprenti-lecteur va devoir apprendre à discriminer, reconnaître et identifier les lettres qui composent les mots de sa langue lorsqu'elles sont présentées de façon isolée. Leur reconnaissance, lorsqu'elles se présentent dans la séquence du mot, représente une difficulté supplémentaire et met en jeu des mécanismes spécifiques de reconnaissance visuelle qui composent la première étape de traitement pour la reconnaissance du mot. Par exemple, dans la langue française, un seul graphème peut représenter plusieurs phonèmes : le C dans le mot *cadeau* est prononcé [K] mais dans le mot *facile*, il est prononcé [S]. À l'inverse, d'autres langues adoptent des systèmes de relation graphème-phonème plus réguliers. En arabe, les représentations et la forme des lettres sont variables selon leur position dans le mot (isolée, initiale, médiane ou finale). Cependant, les relations entre lettres et sons sont relativement systématiques (Awadh, 2016).

La complexité des relations entre graphèmes et phonèmes définit le degré de transparence des langues. Les langues transparentes sont définies par des relations relativement systématiques et régulières alors que les langues opaques sont caractérisées par des systèmes de relations complexes.

La notion de transparence réfère à la relation entre unités orthographiques et unités phonologiques ; elle fait le plus souvent référence à la correspondance graphème-phonème (Share, 2008 ; Zielger & Goswami, 2005). Dans une orthographe transparente, les correspondances graphème-phonème sont beaucoup plus consistantes que dans une langue opaque. Dans les langues alphabétiques, on peut établir le degré de transparence de chaque langue selon son degré d'écart par rapport à une correspondance systématique et univoque entre lettres et phonèmes. Cela dépend du degré de prédiction de la prononciation du mot en fonction de son orthographe (Lieberman, Fitch, Halwes & Erickson, 1980). Dans une orthographe totalement transparente, chaque lettre ou groupe de lettres formant un

graphème correspond à un seul et unique phonème et, vice versa, chaque phonème ne peut se transcrire que d'une seule façon.

Seymour, Aro et Erskine (2003) ont comparé le degré de transparence en anglais et dans 12 autres langues européennes. Selon leur classification, le finlandais, l'espagnol, l'italien et plusieurs autres langues (grec, norvégien, allemand, islandais) se situent du côté des langues les plus transparentes du continuum alors que le français et l'anglais se situent parmi les langues les plus opaques. Le degré de transparence orthographique pourrait influencer sur l'efficacité des traitements : les orthographes transparentes activent des informations plus phonologiques, alors que les orthographes opaques activent des informations essentiellement orthographiques (Katz & Frost, 1992).

Dans toutes les études citées jusqu'ici, les classifications ne concernent que les langues alphabétiques et n'incluent pas les langues sémitiques comme l'arabe qui répondent à des principes d'organisation fondamentalement différents (Awadh, 2016) que nous présenterons plus en détail au Chapitre 2 consacré à l'arabe.

1.2. Le codage des lettres et de leurs positions

La plupart des auteurs (e.g., McClelland & Rumelhart, 1981 ; Paap et al., 1982 ; Grainger & Jacobs, 1996 ; Bowers, 2000) s'accordent sur le fait que la lettre représente l'unité de base de la perception des mots écrits et que son identité abstraite (indépendamment de la typographie, de la taille, de la casse) est codée lors des premiers stades de la reconnaissance d'un mot (Evetts & Humphreys, 1981). Outre l'information sur l'identité, l'information concernant la position de la lettre dans le mot doit être également codée. C'est grâce à ce système de codage que nous sommes capables de distinguer des mots sensiblement différents composés avec les mêmes lettres comme "sac" et "cas". De très nombreuses recherches ont essayé de comprendre comment l'identité et la position des lettres sont traitées dans la reconnaissance d'un mot écrit.

Deux hypothèses ont été proposées sur le codage de la position des lettres : un codage en position absolue où chaque lettre est codée à sa bonne position dans le mot et un codage en position relative où chaque lettre peut être codée à une position différente (e.g. pour 1234 : 1324, 1423, ...). Dans ce dernier cas, l'écart d'une lettre par rapport à sa position

d'origine (1 position, 2 positions, ...) va déterminer la possibilité et la facilité d'une reconnaissance correcte.

1.2.1. Un codage des lettres en position absolue ou en position relative ?

Une première hypothèse de McClelland et Rumelhart (1981), dans leur modèle *Interactive Activation (IA)*, proposait que l'identité abstraite de la lettre ainsi que sa position dans la séquence étaient codées en même temps. Selon le modèle, chaque lettre serait traitée par un canal indépendant qui extrait et analyse les traits visuels de la lettre dans une position spécifique, ce qui signifie qu'il existerait un canal de traitement pour la lettre "N" du mot en 5 lettres "NUAGE" en première position, et un canal différent pour la lettre "N" du mot "ANGLE" en deuxième position, et ainsi de suite. Ce type de codage apparaît peu économique puisqu'il implique la duplication multiple de l'ensemble des lettres de l'alphabet afin de pouvoir coder toutes les positions de suites de lettres.

Une deuxième hypothèse stipule un codage de la position des lettres non pas en position absolue mais plutôt en position relative (Grainger & Segui, 1990 ; Humphreys, Evett, & Quinlan, 1990 ; Peressotti & Grainger, 1995, 1999). Cette hypothèse a été émise à la suite de certains résultats obtenus avec le paradigme d'amorçage¹ orthographique. En manipulant la position des lettres partagées par l'amorce et la cible, certains auteurs ont testé l'influence de la position absolue par rapport à la position relative des lettres dans la séquence. Ces effets sont mesurés relativement à une condition contrôle où l'amorce et la cible ne partagent aucune lettre.

Humphreys, Evett et Quinlan (1990) furent les premiers à montrer que les effets d'amorçage variaient en fonction à la fois du nombre et de la position des lettres partagées entre l'amorce (non-mot) et la cible (mot) dans une tâche d'identification perceptive. Dans

¹ Le paradigme d'amorçage est une technique expérimentale utilisée pour préciser les processus sous-tendant la reconnaissance visuelle des mots. Le principe est le suivant : un premier stimulus, l'amorce, est présenté juste avant un second stimulus, la cible, que le participant doit traiter (Forster & Davis, 1984). La relation entre l'amorce et la cible est manipulée. Par exemple, l'amorce et la cible peuvent être reliées orthographiquement (e.g., aile-AIDE) ou phonologiquement (e.g., roze-ROSE). On compare l'influence de l'amorce reliée à l'influence d'une amorce non reliée (e.g., four-CIEL). La manipulation du temps de présentation de l'amorce permet d'étudier les processus conscients de nature délibérée et stratégique (l'amorce est visible, le temps de présentation de l'amorce est supérieur à 66 ms) ou non conscients très rapides et automatiques (l'amorçage est dit masqué, l'amorce est subliminale, le temps de présentation de l'amorce est de l'ordre de 17 à 66 ms (Ferrand & Grainger, 1992, 1993). La tâche du participant peut être une décision lexicale, une dénomination, une identification perceptive ou une catégorisation sémantique.

ces conditions, les amorces et les cibles n'ont pas la même longueur de sorte que la position absolue des lettres change alors que la position relative, préservant l'ordre des lettres, est maintenue. Par exemple, ils ont observé pour la cible "black" (12345) de 5 lettres que l'amorce "blck" (1245) de 4 lettres, qui préserve l'ordre relatif des lettres, accélérât son identification, comparé à des conditions ne préservant pas la position relative des lettres comme "bcak" (1435). Des résultats comparables ont été obtenus par Peressotti et Grainger (1999) en décision lexicale avec amorçage masqué : des amorces conservant la position relative des lettres comme "blcn" (1346) facilitaient le traitement des mots cibles comme "balcon", comparé à des amorces non reliées. De même, l'insertion de lettres différentes ou de caractères (B-LC-N) pour maintenir la position absolue des lettres ne facilitait pas plus le traitement comparé à la condition où seule la position relative est préservée. Grainger et Van Heuven (2003) rapportent également des résultats en faveur d'un codage de la position relative des lettres sur des cibles comme "silence" de 7 lettres lorsque l'amorce "slene" (13457) préservait la 1^{ère} et la dernière lettre, et les 3 lettres du milieu (dans le bon ordre), comparé à une condition totalement non reliée.

Ces résultats suggèrent que les lettres sont codées en fonction de leurs positions relatives préservant l'ordre des lettres plutôt qu'en fonction de leurs positions absolues. Ils sont confortés par d'autres études plus récentes qui ont démontré que, lorsqu'une amorce partageait toutes les lettres du mot cible, l'effet d'amorçage était robuste malgré des changements dans l'ordre des lettres, phénomène appelé *l'effet de la transposition des lettres* "the transposed-letter effect" (Acha & Perea, 2008 ; Andrews, 1996 ; Perea & Lupker, 2003a, 2003b, 2004 ; Schoonbaert & Grainger, 2004). Par exemple, il a été observé, lors d'une tâche de décision lexicale, que le mot "judge" était identifié plus vite s'il était précédé du pseudo-mot "jugde" que du pseudo-mot "junpe", de même que "casino" est plus vite identifié lorsqu'il est précédé de "caniso" que "caviro". On perçoit ainsi souvent des pseudo-mots (construits à partir d'un mot qui a subi une transposition de deux lettres) comme étant les mots réels à la base de leur construction (e.g., "RELOVUTION" est perçu comme "REVOLUTION") (Perea & Fraga, 2006 ; Perea, Rosa, & Gomez, 2005).

1.2.2. Le modèle à bigrammes ouverts de Grainger et Van Heuven (2003)

Afin d'expliquer ces effets d'amorçage orthographique et de transposition, Grainger et Van Heuven (2003) ont proposé un modèle impliquant un codage contigu et non-contigu de

combinaison de paires de lettres, appelées "bigrammes ouverts". Ainsi, le mot "SILENCE" est codé par l'ensemble suivant de bigrammes ouverts non ordonnés par pas de trois, avec une distance maximale de deux lettres intermédiaires entre chaque bigramme : SI, SL, SE, IL, IE, IN, LE, LN, LC, EN, EC, EE, NC, NE, CE. C'est l'activation combinée de cet ensemble non ordonné de bigrammes ouverts qui fournit de façon non-ambiguë l'ordre correct des lettres composant le mot. Ce modèle comporte trois niveaux de représentations (cf. Figure 1) : le niveau des lettres, la carte des positions relatives des lettres, et les représentations orthographiques globales (lexique orthographique).

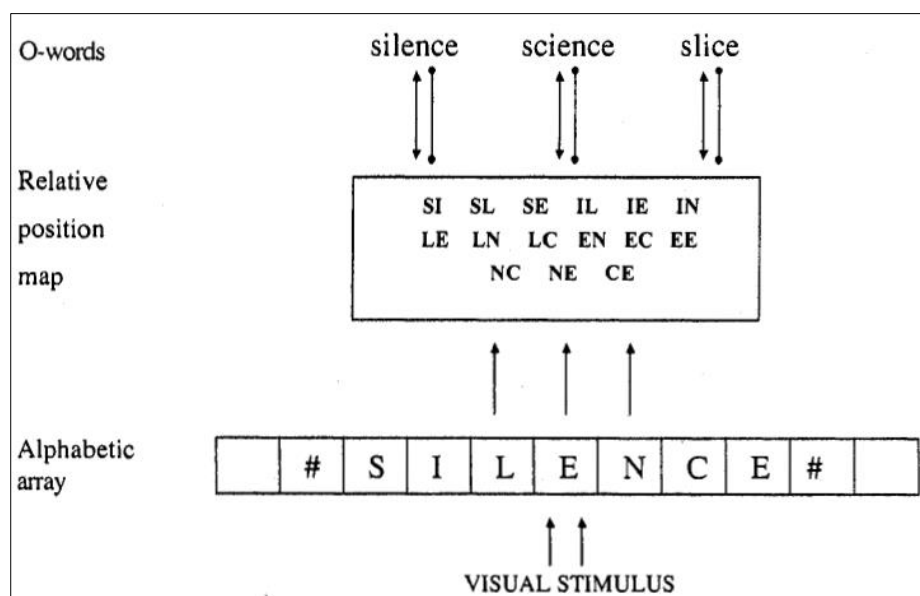


Figure 1. Architecture du modèle à bigrammes ouverts de Grainger et Van Heuven (2003 ; figure p.14).

Le premier niveau est un ensemble de détecteurs de caractères alphabétiques (et d'espaces) qui traite simultanément toutes les lettres d'un mot. Chaque détecteur fournit de l'information sur l'identité d'une lettre à un emplacement rétinien donné. Cette information est ensuite envoyée au niveau de la carte des positions relatives des lettres (ensemble de représentations utilisant des bigrammes ouverts) qui combine l'information des différentes régions rétiniennes de façon à donner une information sur la position relative des lettres dans le mot. Les bigrammes ouverts activés à ce niveau envoient de l'activation aux représentations orthographiques lexicales (*i.e.*, tous les mots dont l'orthographe est connue) compatibles avec ces unités. Ces unités sont connectées par des liens bidirectionnels excitateurs et inhibiteurs aux bigrammes ouverts. Par exemple, le bigramme ouvert OR activerait les représentations des mots MORT, TONER, etc. Les connexions inhibitrices servent à bloquer l'activation de mots qui ne sont pas compatibles avec l'entrée visuelle.

Dans ce modèle, le codage de la position des lettres repose sur l'identité de la première et la dernière lettre dans le mot, et un minimum d'information concernant la position relative des lettres internes, ce qui permet d'expliquer par exemple l'amorçage orthographique obtenu avec des amorces conservant seulement la position relative des lettres de la cible. Ainsi, l'amorce "slece" pour la cible "silence" active les bigrammes ouverts SL, SE, LE, LC, EC et EE, qui sont présents dans le mot cible, sauf le bigramme SC qui ne l'est pas.

L'activation de la représentation orthographique par un nombre élevé de bigrammes ouverts explique également l'effet de facilitation obtenu en amorçage avec lettres transposées. Avec le même principe, l'amorce "ocaen" (codée par les bigrammes ouverts **OC-OA-OE-CA-CE-CN-AE-AN-EN**) obtenue par transposition de deux lettres de la cible "OCEAN" (codée par les bigrammes ouverts **OC-OE-OA-CE-CA-CN-EA-EN-AN**) facilite plus la reconnaissance de la cible vu le nombre de bigrammes qu'ils partagent (en gras), comparée à l'amorce à deux lettres substituées "ocuon" (codée par les bigrammes **OC-OU-OO-CU-CO-CN-UO-UN-ON**) qui ne partage que deux bigrammes avec la cible.

D'autres modèles comme SERIOL de Whitney (2001) ou SOLAR de Davis et Bowers (2006) peuvent également expliquer ces effets d'amorçage, là où le modèle IAM de McClelland et Rumelhart (1981) avait échoué (pour une synthèse, voir Grainger, 2008).

1.2.3. Un traitement parallèle ou séquentiel des lettres ?

Le traitement est dit parallèle lorsque toutes les unités d'une séquence sont traitées simultanément, c'est-à-dire que le traitement des différentes unités s'initie et se termine en même temps. En revanche, on parle de traitement sériel lorsque les unités sont traitées les unes à la suite des autres ; une unité ne pouvant commencer à être traitée que lorsque le traitement de la précédente est terminée. Dans le cadre du traitement orthographique du mot écrit, cette question fait grand débat entre les chercheurs. Certains modèles, comme le modèle SERIOL de Whitney (2001) postulent un encodage sériel de gauche à droite des lettres (pour les langues se lisant dans ce sens), alors que d'autres, comme le modèle à bigrammes ouverts de Grainger et van Heuven (2003) vu plus haut, postulent un encodage en parallèle de l'ensemble des lettres qui composent la séquence.

Adelman et Marquis (2010) et Marzouki et Grainger (2014) ont montré que le traitement des lettres dans la séquence d'un mot ou d'un pseudo-mot commence simultanément sur toutes les lettres dès les premières vingt millisecondes du traitement. Ils

ont montré néanmoins qu'un traitement parallèle ne signifie pas pour autant une identification uniforme sur l'ensemble des lettres à traiter, certaines lettres pouvant être mieux identifiées que d'autres. En outre, certains auteurs suggèrent que le traitement parallèle est utilisé seulement pour la lecture de mots familiers et proposent l'utilisation d'un traitement séquentiel pour la lecture de pseudo-mots (Coltheart & Rastle, 1994 ; Coltheart et al., 2001) ou de mots peu fréquents (Ferrand & New, 2003 ; Valdois et al., 2006). D'autres proposent que le système de lecture a recours au traitement en parallèle avec les pseudo-mots également (Harm & Seidenberg, 2001, 2004 ; McClelland & Rumelhart, 1981 ; Plaut et al., 1996 ; Seidenberg & McClelland, 1989). Enfin, certains proposent l'utilisation du traitement séquentiel avec tous types de stimuli verbaux (Whitney, 2001, 2008b).

Dans cette perspective, un traitement parallèle ne devrait pas être affecté par des modifications du nombre d'éléments à traiter. En revanche, lors d'un traitement sériel, l'augmentation du nombre d'éléments à traiter pourraient induire un coût lors du déplacement attentionnel et donc affecter le traitement en le ralentissant ou diminuant les performances produisant un "effet de longueur" sur les performances. Il a été mis en évidence que la présence d'un tel effet en lecture traduit une stratégie de lecture sérielle alors que l'absence d'un tel effet traduit une stratégie de lecture parallèle (Cohen, Dehaene, Vinckier, Jobert, & Montavont, 2008 ; Juphard et al., 2004). De nombreuses études ont révélé une absence d'effet longueur pour les mots (Coltheart et al., 2001 ; Valdois et al., 2006 ; Cohen et al. 2003 ; Weekes, 1997), mais selon Whitney (2001) cette absence d'effet pourrait être justifiée par un traitement sériel très rapide ("*cover shifting*", voir aussi Wolfe, 2003).

1.3. Les contraintes oculomotrices et attentionnelles en lecture

1.3.1. Les contraintes oculomotrices

1.3.1.1. Le champ visuel

Le champ visuel est la portion de l'espace vue par un œil immobile dirigé droit devant. Il totalise près de 180°. C'est autour du centre de fixation du regard que la vision est la plus nette. Elle se dégrade progressivement en s'éloignant de ce point avec une diminution drastique d'environ 50% par degré d'écartement. Ce champ visuel se compose de 3 zones de vision (cf. Figure 2) :

- La zone de *vision fovéale* : c'est la vision la plus nette, de plus forte acuité. Elle couvre 2° du champ visuel, soit 1° de part et d'autre du point de fixation.
- La zone de *vision parafovéale* : elle couvre en moyenne 10° du champ visuel, c'est-à-dire qu'elle s'étend jusqu'à 5° de chaque côté du point de fixation.
- La zone de *vision périphérique* : elle débute environ à partir de 5° du centre de la fixation.

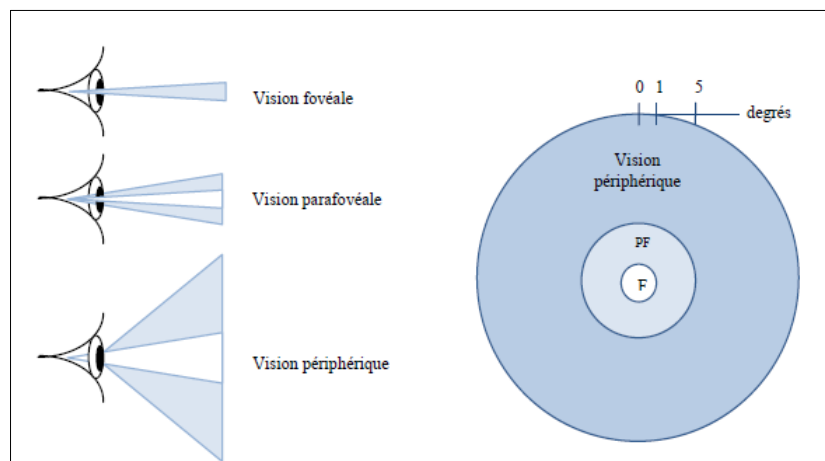


Figure 2. Le champ visuel en vision fovéale, parafovéale et périphérique.

En raison de la limitation de notre acuité visuelle, l'information ne peut être perçue avec une forte précision que sur la partie centrale du champ visuel : c'est la fovéa, zone qui correspond à un angle de 2° autour du point de fixation, soit à 6 ou 7 caractères au total, pour un texte se situant à 30 cm de l'œil de lecteur. L'acuité totale, qui inclut la zone parafovéale, s'étend à 5° autour du point de fixation, soit 14-15 caractères (Rayner, 1986). Ainsi, nous déplaçons constamment nos yeux lors de l'activité de lecture (saccades et fixations) dans le but d'amener très rapidement le prochain mot en zone fovéale.

1.3.1.2. *Les mouvements oculaires : saccades et fixations*

Pendant la lecture, des périodes de relative stabilité appelées fixations alternent avec des périodes de mouvements rapides appelés saccades. Lorsque nous portons attention sur un texte, nos yeux se déplacent 4 à 5 fois par seconde. Chaque déplacement donne lieu à une fixation à un endroit différent.

La durée des fixations est très variable (Rayner, 1998). En lecture, les fixations durent en moyenne 250 ms (entre 50 et 500 ms typiquement) et sont fortement influencées par le contenu linguistique. La durée des saccades est beaucoup plus brève et insensible au contenu

linguistique et aux traitements cognitifs. Ce sont des mouvements balistiques. Pratiquement aucune information visuelle n'est extraite pendant les saccades (phénomène de suppression saccadique). L'information visuelle est extraite seulement pendant les fixations (Wolverton et Zola, 1983). Une saccade de 2°, typique en lecture, dure environ 30 ms. La plupart des saccades (90%) sont dirigées vers la droite, dans le sens de la lecture (pour les langues qui se lisent de gauche à droite). Environ 10 % de saccades sont dirigées vers la gauche (Rayner & Pollatsek, 1989). Ces saccades, appelées saccades régressives ou de régressions, permettent au lecteur de relire des passages de texte qu'il a déjà lu.

1.3.2. Les contraintes attentionnelles

1.3.2.1. L'empan perceptif (*perceptual span*)

L'empan perceptif se définit comme l'espace autour du point de fixation à l'intérieur duquel l'information visuelle utile peut être extraite (Rayner & Liversedge, 2004). Ses limites horizontales sont généralement mesurées au moyen de la technique de la fenêtre mobile (*moving window*) (cf. Figure 3) lors d'une situation naturelle de lecture. Cette technique (McConkie & Rayner, 1975 ; Rayner & Bertera, 1979 ; Rayner, 1998 pour une revue) consiste à masquer (par des X) ou à modifier les lettres situées en périphérie, au-delà d'une certaine limite à gauche et à droite du point de fixation pendant qu'un sujet lit un texte affiché sur un écran. Pour lire, le sujet ne peut donc utiliser que l'information préservée située à l'intérieur de la fenêtre, au centre de son champ visuel. L'enregistrement simultané des mouvements oculaires permet de déplacer la fenêtre en fonction de la position de fixation du sujet : dès que le sujet effectue une saccade et fixe une nouvelle position sur la ligne de texte, l'information précédemment fixée est masquée et les lettres dans la fenêtre autour du nouveau point fixé sont affichées normalement, alors que les lettres en périphérie sont masquées.

La taille de l'empan perceptif correspond à l'étendue maximale de la fenêtre à partir de laquelle la vitesse de lecture et/ou les caractéristiques des mouvements oculaires (*i. e.* l'amplitude des saccades et les durées de fixation) ne sont plus modifiées par rapport à une condition contrôle, sans fenêtre. En effet, si la vitesse de lecture est réduite ou si les caractéristiques des saccades/fixations sont modifiées, une partie de l'information visuelle utile est masquée et la fenêtre est plus petite que l'empan perceptif. Lorsque la fenêtre

comprend jusqu'à 3-4 caractères à gauche du point de fixation et 14-15 caractères à sa droite, aucune perturbation de la vitesse de lecture n'est observée. Cependant, si la fenêtre est plus étroite, on observe un ralentissement du rythme de lecture (Denbuurman, Roerema, & Gerrissen, 1981 ; McConkie & Rayner, 1975, 1976 ; Rayner, Well, & Pollatsek, 1980 ; Underwood & McConkie, 1985).

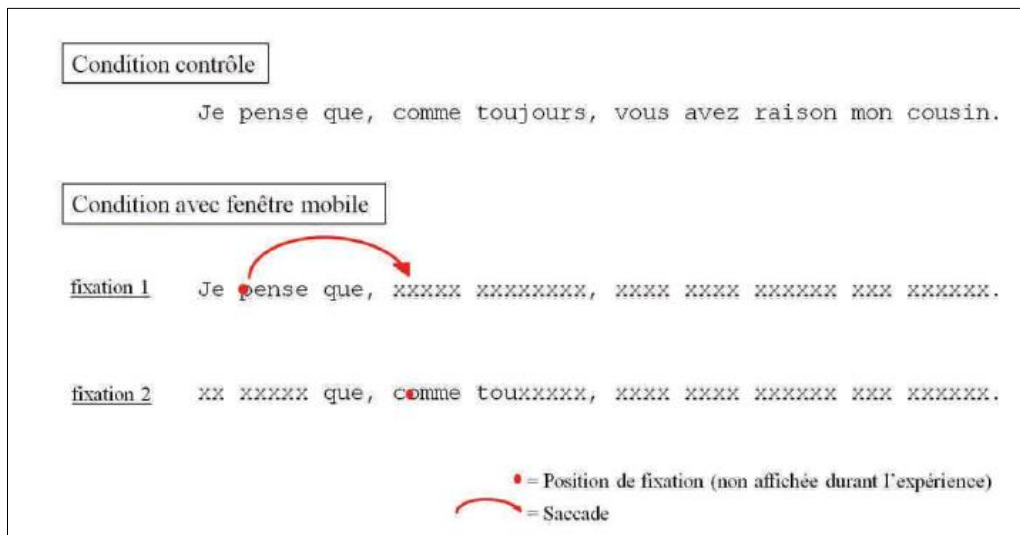


Figure 3. Paradigme de la fenêtre mobile. Exemple de phrase affichée sur l'écran ; lorsque le sujet effectue une saccade du mot "pense" vers le mot "comme". Ici la modification du texte, qui commence au-delà de sept espaces caractères à gauche et à droite du point de fixation, consiste à remplacer les lettres par des "x", sans masquer les espaces inter-mot. Figure extraite de Yao-N'Dré (2013), p.54.

1.3.2.1.1. Flexibilité de la taille de l'empan perceptif

La taille de l'empan perceptif n'est pas fixe, la quantité totale d'informations extraites pendant une fixation peut varier en fonction des caractéristiques du système d'écriture. Il a été montré par exemple qu'à droite du point de fixation, les lecteurs japonais ne pouvaient pas traiter plus de 6 caractères (Ikeda & Saida, 1978) et les lecteurs chinois plus de 3 caractères (Inhoff & Liu, 1998). Les unités lexicales de ces systèmes d'écriture sont plus condensées que celles des langues alphabétiques. Comme le souligne Rayner (1997), pour formuler une même phrase, il faut utiliser davantage de caractères en langue anglaise qu'en langue japonaise ou chinoise.

La quantité d'information parafovéale prélevée peut dépendre de la difficulté du traitement de l'information fovéale (Häikiö et al., 2009 ; Henderson & Ferreira, 1990 ; Kennison & Clifton, 1995 ; Schroyens, Vitu, Brysbaert, & d'Ydewalle, 1999 ; White, Rayner & Liversedge, 2005 ; White, 2007). Henderson et Ferreira (1990) ont constaté que la quantité

d'information extraite en parafovéa est plus importante au cours de l'identification fovéale d'un mot de haute fréquence plutôt que de basse fréquence. Comme le soulignent Schotter, Angel et Rayner (2012), plus le traitement de l'information fovéale est difficile, plus il nécessite l'investissement de ressources attentionnelles, et moins celles-ci sont disponibles pour un traitement de l'information parafovéale.

L'expertise en lecture est un facteur qui détermine également la quantité d'information parafovéale pouvant être extraite. Par exemple, Rayner, Slattery et Bélanger (2010) ont rapporté que les lecteurs rapides présentent un empan perceptif plus étendu que les lecteurs moins rapides. Selon les auteurs, les lecteurs moins rapides ont davantage besoin de ressources attentionnelles que les autres lecteurs pour identifier et comprendre le mot qu'ils fixent, ce qui se fait probablement aux dépens du traitement de l'information parafovéale.

1.3.2.1.2. Développement de la taille de l'empan perceptif chez l'apprenti-lecteur

Comparées au nombre d'études menées chez le lecteur expert, les recherches sur les mouvements oculaires en lecture chez l'apprenti lecteur sont beaucoup moins nombreuses. La plupart des travaux s'intéressant aux trajectoires développementales des processus cognitifs impliqués dans la lecture confirment l'influence du niveau de lecture sur le développement et la modulation de l'empan perceptif.

L'empan perceptif des apprenti-lecteurs est plus restreint, du fait qu'ils accordent plus d'attention au traitement fovéal contrairement aux lecteurs experts (Häikiö, Bertram, Hyönä, & Niemi, 2009 ; Rayner, 1986). La première étude dans le domaine du développement de la taille de l'empan perceptif a été menée par Rayner (1986). Il a montré que l'empan perceptif chez les apprentis lecteurs est moins étendu que celui des lecteurs experts. La taille de cet empan s'étend à seulement environ 11 caractères à droite du point de fixation pour les lecteurs de grade 2 et 4, contre 14 caractères à droite du point de fixation pour les lecteurs de grade 6. Au cours du développement des aptitudes en lecture, on constate une progression de la taille de l'empan perceptif chez les apprentis lecteurs qui devient similaire à celle observée chez le lecteur expert.

Lors d'une étude longitudinale d'enfants du grade 1 au grade 4, Sperlich, Meixner et Laubrock (2016) (cité dans Leibnitz et al., 2017) ont mis en évidence un changement significatif

de la taille de l'empan perceptif uniquement entre les grades 2 et 3, suggérant que le développement de ce dernier pourrait être un processus qui s'étendrait par palier.

Dans le même sens, il a été montré que les enfants de grade 1 et 2 prétraient peu voire pas d'informations en parafovéa et que des difficultés de traitement lexical entraînaient une utilisation réduite des informations parafovéales (Reichle et al., 2013). De même, Marx, Hutzler, Schuster, et Hawelka (2016) ont montré que les compétences en lecture jouent un rôle déterminant dans le prétraitement de l'information parafovéale, les meilleurs décodeurs étant capables de prétraiter la plus grande quantité d'informations en parafovéa (cité dans Leibnitz et al., 2017).

Certaines données ont également lié le développement de l'empan perceptif au changement dans la vitesse de lecture (Sperlich, Schad, & Laubrock, 2015) ; plus la lecture est rapide, plus la taille de l'empan est importante. Carver (1990 ; cité dans Jeon, Hamid, Maurer & Lewis, 2010) rapporte que les apprentis lecteurs de 7 ans augmentent leur vitesse de lecture de 14 mots par minute en moyenne chaque année, jusqu'au grade 6. Néanmoins, les lecteurs experts ont une vitesse de lecture qui reste supérieure (Häikiö et al., 2009). Cette différence laisse penser que d'autres facteurs comme la rapidité de l'accès au lexique restent encore à développer chez ces mêmes apprentis lecteurs. Par ailleurs, après une année d'apprentissage de la lecture, l'apprenti lecteur présente un empan perceptif avec une asymétrie vers la droite (Rayner, 1986). La capacité à orienter davantage son attention dans le sens de la lecture de la langue s'observe donc précocement au cours du développement des aptitudes en lecture.

1.3.2.2. Phénomène de masquage latéral (*crowding*)

Contrairement à ce qu'on pensait, l'acuité visuelle se dégrade à l'intérieur même de la fovéa et cette dégradation se poursuit de façon continue jusqu'à une excentricité d'environ 14° (O'Regan, 1990, 1992). Par exemple, à la lecture du titre "MAFIA : L'TALIE JUGE SES PARRAINS" (titre de première page du quotidien *Libération* parue le 11 février 1986), on a tendance à ne pas remarquer qu'il manquait le "I" du mot "ITALIE". L'absence du "I" n'est pas perçue parce qu'elle subit ce que l'on appelle, "un masquage latéral" (Townsend, Taylor, & Brown, 1971). Ainsi, une lettre au centre d'une séquence peut être difficilement identifiable par le bruit des lettres environnantes, et à l'inverse, une lettre à l'extrémité d'une séquence sera plus facilement identifiée car une quantité moindre d'informations viendra la masquer.

En effet, les observations de Mason et Katz (1976) avaient montré que des présentations de symboles abstraits (e.g., % ; } ; \$) n'engendraient pas un avantage particulier pour les symboles situés aux extrémités comme on pouvait l'observer pour les lettres. Cette mise en évidence d'une fonction de position sérielle en forme de U inversé a été confirmée par Mason (1982) lors de présentations de stimuli symboliques. L'absence d'un avantage pour le traitement de symboles en bordure de vision dans ce type de présentation impliquait que ce phénomène était moins lié à des caractéristiques neurosensorielles du système perceptif qu'à des caractéristiques cognitives de ce système, et relèverait de ce fait de mécanismes spécifiques au traitement des mots.

D'autres travaux ont tenté également d'explorer cet effet afin de confirmer qu'il n'était pas spécifiquement lié à une tâche précise. Par exemple Tydgat et Grainger (2009) ont utilisé le paradigme de choix forcé à double alternative (*Two-Alternative Forced Choice, 2AFC*; Reicher, 1969) qui se distingue des tâches simples d'identification de stimuli. Dans cette tâche, chaque essai commençait par la présentation de deux barres verticales de fixation placées au-dessus et au-dessous du centre d'un masquage (#####). Après la disparition des barres et du masquage, le stimulus de cinq caractères apparaît immédiatement pour une durée de 100 ms. Le masquage réapparaît ensuite accompagné de deux caractères (un au-dessus et un au-dessous) dans une des cinq positions possibles (cf. Figure 4 à gauche). Les participants devaient décider lequel des deux caractères était présent dans la suite de caractères précédemment présentée.

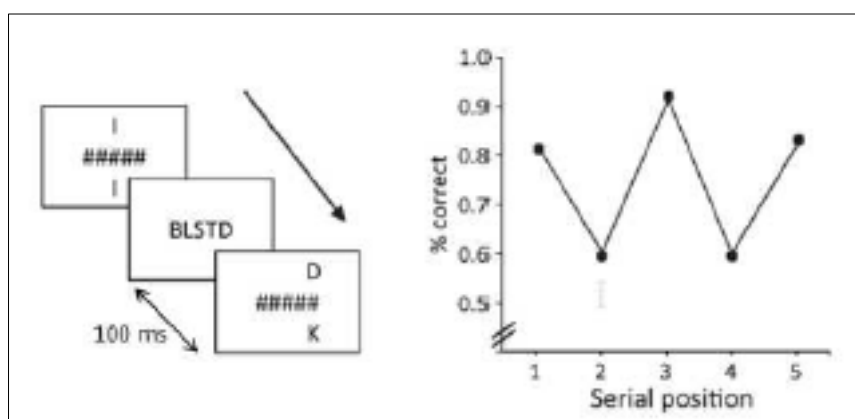


Figure 4. Tydgat et Grainger (2009). Représentation schématique du paradigme qui évalue la perception visuelle des lettres en fonction de sa position spatiale dans la séquence. Les résultats montrent une courbe en forme de W.

Au cours de plusieurs expériences, les auteurs ont utilisé ce paradigme en alternant des présentations de chaînes de 5 éléments composées de lettres, de chiffres et de symboles.

Les résultats de ces études ont révélé que les chaînes de stimuli symboliques engendraient des difficultés de report des éléments en bordure de vision alors que les chaînes de lettres et de chiffres n'en présentaient pas, ce qui concordait avec les observations de Mason et Katz (1976). Ainsi, des présentations de chaînes de 5 lettres engendraient de meilleurs taux d'identification pour la première lettre, la lettre centrale et la dernière lettre, alors que la seconde et la quatrième lettre étaient moins bien rapportées. Selon Grainger et Tydgate (2009), cette forme typique en W (cf. Figure 4 à droite) serait dû à l'influence conjointe de deux facteurs : l'acuité et le crowding. Le premier facteur s'expliquerait par une baisse de l'acuité visuelle en fonction de l'augmentation de la distance entre les lettres et le point de fixation oculaire. Le second facteur serait dû aux interférences liées au nombre de lettres présentes en bordure de vision. La présence du premier facteur engendrerait une perception pour les lettres situées en bordure de vision fortement diminuée. La présence du second facteur semblerait compenser la diminution de l'acuité pour les lettres situées en bordure avec un déploiement important de l'attention spatiale jusque dans des zones extrêmes de la chaîne. De la conjonction de ces deux facteurs résulterait la visibilité d'un mot (cf. Figure 5).

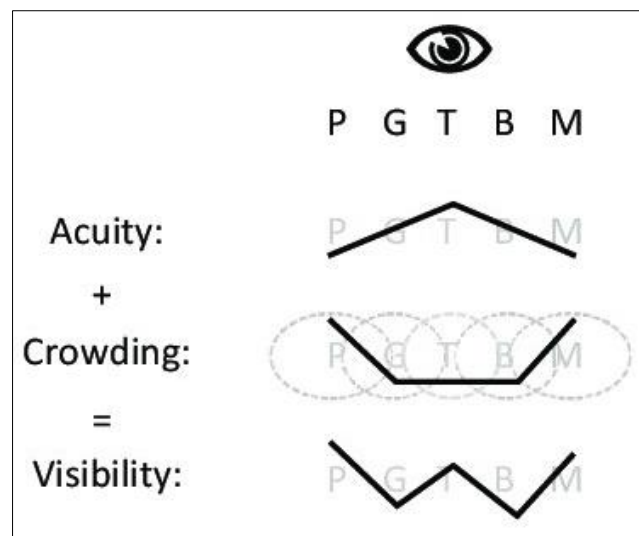


Figure 5. Grainger et al. (2016). Représentation de la visibilité du mot résultant de l'acuité et du crowding.

1.3.2.3. *La position optimale de fixation (Optimal Viewing Position)*

Certaines études sur les mouvements oculaires (e.g., O'Regan & Levy-Schoen, 1987 ; O'Regan et al., 1984) ont révélé que le temps total nécessaire pour traiter ou identifier un mot isolé est étroitement lié à la position de première fixation de l'œil dans le mot. Ainsi, le temps de traitement est minimum pour une position initiale de l'œil aux environs du centre-gauche

de ce mot (2^{ème} et 3^{ème} lettre pour un mot de cinq lettres), et la probabilité de ne faire qu'une seule fixation dans le mot pour l'identifier est plus importante pour cette même position initiale, ce qui valide l'hypothèse selon laquelle il existe une position privilégiée de l'œil dans le mot (OVP) quant à la rapidité et l'efficacité de traitement d'un mot isolé.

Cet effet a été obtenu également par Montant et al. (1998) grâce à un paradigme expérimental simple qui permet de contrôler la position du regard dans le mot (Nazir et al., 1991) : on demande au sujet de fixer un point au centre de l'écran (cf. Figure 6), puis le mot est présenté rapidement, de manière plus au moins décalée par rapport au point de fixation de sorte que l'œil du sujet peut être positionné sur chacune des différentes lettres du mot.

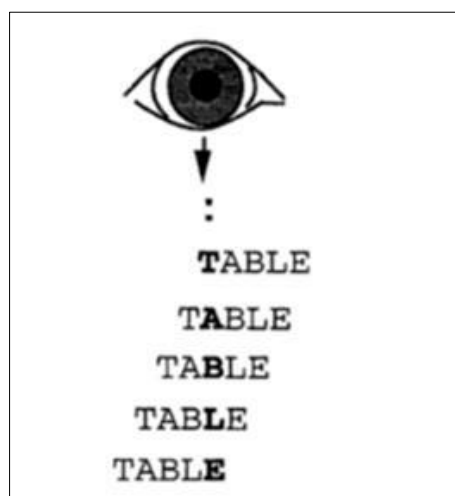


Figure 6 . Nazir et al. (1991). Technique de la position de fixation variable du regard.

En effet, les performances d'identification chutent au fur et à mesure que le regard s'éloigne de la position optimale (centre-gauche) et en particulier lorsque le regard fixe initialement le début ou la fin du mot, avec une asymétrie gauche-droite. Ainsi, la baisse des performances est beaucoup plus marquée quand le point fixé est à droite de l'OVP, en fin de mot, que lorsqu'il est à sa gauche, en début de mot. La courbe des réponses en fonction du point initiale de fixation est donc asymétrique (en forme *J* inversé) (cf. Figure 7).

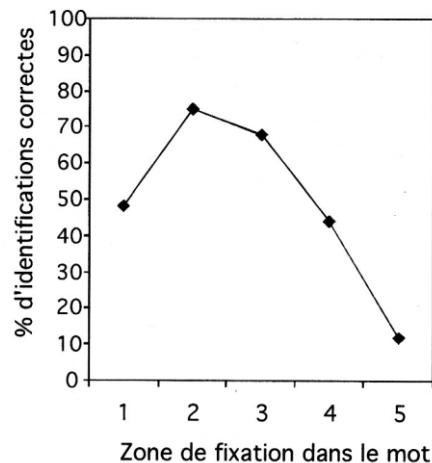


Figure 7. Montant et al. (1998). Courbe OVP obtenue dans une tâche d'identification perceptive pour un mot de 7 lettres chez des adultes

Cette asymétrie gauche de l'OVP peut être expliquée par plusieurs facteurs :

- **La direction de lecture.** Selon cette hypothèse, l'effet de position initiale de fixation est dû au fait que dans les langues lues de gauche à droite, les lecteurs ont tendance à déployer leur attention plus sur la droite du point de fixation (Pollatsek et al., 1981) puisque c'est là qu'ils vont extraire principalement la prochaine source d'information.
- **Les contraintes lexicales.** Elles sont liées à la structure informative du mot. Dans les langues comme le français et l'anglais, la plupart des mots peut être inférée à partir des premières lettres. Néanmoins, la structure morphologique peut influencer sensiblement cette position optimale. Par exemple, Holmes et O'Regan (1992) ont montré que lorsque la racine d'un mot se trouve au début (comme dans le mot suffixé "chanteur"), la position optimale est plus décalée vers la gauche du centre. En revanche, lorsque la racine est en fin de mot (comme dans le mot préfixé "reprendre"), la position optimale est décalée vers la droite du centre.
- **La spécialisation hémisphérique.** Cette hypothèse considère que l'aire de reconnaissance des mots écrits étant localisée dans l'hémisphère gauche, le champ visuel droit a un avantage pour la perception d'un support orthographique (Brysbaert, Vitu, & Schroyens, 1996). En effet, les informations perçues dans le champ visuel gauche transitent par le cortex visuel controlatéral, c'est-à-dire le droit, pour ensuite migrer vers l'hémisphère gauche. En revanche, les éléments perçus par le champ visuel droit n'ont pas à effectuer un transfert d'hémisphère pour être analysés.

1.4. Traitement de l'information parafovéale chez l'adulte

La question du traitement de l'information parafovéale chez l'adulte a été traitée à l'aide de différents paradigmes, y compris le nouveau paradigme FLLD. En effet, ceux-ci font apparaître des résultats convergents qui confortent leur validité.

1.4.1. Paradigme d'étude de mouvements oculaires (Eye-tracking)

La technique de *la fenêtre mobile* (McConkie & Rayner, 1975) avait montré que si on retirait des informations (par des systèmes de masques) de la zone parafovéale, la vitesse de lecture déclinait d'environ un tiers. Cette réduction de vitesse est due à l'accroissement des temps de fixation et à des saccades oculaires moins amples, donc à un plus grand nombre de fixations. Les informations sur les mots, fournies par la vision parafovéale, facilitent donc l'identification du mot lorsqu'il entre dans la zone fovéale (Inhoff & Rayner, 1986). L'apport de ces informations constitue le *bénéfice de prévisualisation* (*preview benefit*). L'étendue spatiale de ce bénéfice dépend de la difficulté à traiter le mot. Quand le mot fixé est difficile (long et/ou peu fréquent), les lecteurs ont un bénéfice de prévisualisation très faible. Toutefois, il est possible d'obtenir un bénéfice de prévisualisation sur deux mots à droite de celui fixé, s'ils sont très fréquents et courts (Rayner, 2009b).

Une deuxième technique appelée *technique des frontières* (*Boundary paradigm*, Rayner, 1975b) a permis d'examiner la nature de l'information extraite dans la zone parafovéale. La manipulation expérimentale consiste à remplacer un mot situé à une position critique de la ligne par une suite de lettres (amorce) tant que cette position critique n'est pas amenée en vision fovéale. Au moment où le regard croise une frontière (invisible) située quelques caractères avant la position critique, le stimulus correct est rétabli. On mesure le temps de fixation sur le mot situé à la position critique. Si une information est extraite de l'amorce en vision parafovéale, elle pourra conduire à un raccourcissement ou à un allongement de la fixation sur le mot critique.

Les premières études avaient utilisé une tâche de lecture de mots isolés (Rayner, 1978 ; Rayner, McConkie & Ehrlich, 1978 ; Rayner, McConkie, & Zola, 1980). Dans cette tâche, les participants devaient fixer une croix centrale. Au moment où l'item était présenté en vision parafovéale, il leur était demandé d'effectuer une saccade en direction de cet item. Durant la saccade, l'item en vision parafovéale était remplacé par un mot cible qui devait être prononcé le plus rapidement possible. Les résultats de ces études avaient montré que le temps

d'identification de la cible était plus court quand elle était identique à son aperçu (item en vision parafovéale) que quand elle ne l'était pas. Cet effet suggère que les informations de l'aperçu parafovéal facilitent l'identification du mot cible lors de sa fixation.

Pollatsek, Rayner et Balota (1986) avaient également montré que les durées de fixations du mot à la position critique (eg., "ribbon") étaient allongées lorsque l'amorce était visuellement dissimilaire ("flouon"). Il n'y avait cependant pas de différence sensible entre la condition où l'amorce était un non-mot visuellement similaire ("ribham") et la situation contrôle dans laquelle l'amorce était identique au mot cible.

Cet effet de similarité visuelle entre amorce et cible a amené certains chercheurs à vérifier s'il n'était pas dû à l'utilisation d'indices sur la forme générale du mot (contour, hampes...) en position parafovéale. Dans une étude plus ancienne, McConkie et Zola (1979) ont utilisé des textes présentés en TyPoGrApHiE aLtErNéE, et ils ont comparé une situation dans laquelle la typographie s'inversait au passage de la frontière (TyPoGrApHiE * tYpOgRaPhle) avec une situation où il n'y avait pas de changement. Les lecteurs ne détectaient pas le changement, et leur vitesse de lecture n'était pas affectée. Ces résultats démontrent que le bénéfice de l'amorce parafovéale ne peut pas être attribué à la forme générale des mots, mais bien à l'identification des lettres.

Certaines données plus récentes (Vitu, Brysbaert et Lancelin 2004) ont confirmé cet effet de bénéfice parafovéal sur des paires de mots orthographiquement reliés, qui ne se distinguent que par une seule lettre. Les résultats ont montré que les temps de fixation sur le mot cible (*i.e.*, en fovéa) étaient plus courts quand ce dernier était accompagné d'un autre (*i.e.*, en parafovéa) avec lequel il entretenait une similarité orthographique (e.g., pour et four), que quand ils n'entretenaient aucune similarité (e.g., pour et clan).

Johnson, Perea et Rayner (2007) ont exploré la présence de l'effet de transposition de lettres en parafovéa dans une situation normale de lecture silencieuse. Le mot présenté en parafovéa était soit identique à celui présenté en fovéa (e.g., clerk / clerk) ; soit deux de ses lettres internes ont été transposées (e.g., clerk / celrk) ; soit deux de ses lettres ont été substituées (e.g., clerk / cohkr). Les résultats ont montré que les temps de fixation étaient significativement plus courts dans les deux conditions *lettres identiques* et *lettres transposées*, comparées à la condition *lettres substituées*, avec aucune différence entre les deux premières, ce qui suggère que le traitement du mot en fovéa a été facilité par la ressemblance orthographique avec le mot présenté en parafovéa, en dépit de la transposition de ses lettres.

Les mêmes conditions (condition identique (e.g., coat / coat) ; condition lettres internes transposées (e.g., caot / coat) ; condition lettres internes substituées (e.g., ceit / coat) ont été répliquées dans une étude de Dare et Shillcock (2013), et les résultats ont été similaires.

1.4.2. Paradigme d'amorçage masqué

Une autre approche dans l'étude du traitement de l'information parafovéale utilise le paradigme d'amorçage. Cela consiste à présenter brièvement une amorce en vision parafovéale suivie d'une cible à identifier en vision centrale ou non. Le traitement de l'information parafovéale est inféré à partir de l'effet que l'amorce produit sur l'identification de la cible. Ce paradigme a été utilisé par Marzouki et Grainger (2008) avec une tâche de décision lexicale. Les amorces étaient présentées soit au centre, au point de fixation, soit à 2°, 4° ou 6° d'excentricité dans le champ visuel gauche ou droit. Les cibles étaient soit identiques, soit différentes aux amorces. Les cibles apparaissaient au même endroit que leur amorce dans une première expérience, et au point de fixation central dans une deuxième expérience. Dans les deux expériences, les cibles étaient mieux identifiées lorsqu'elles étaient identiques à l'amorce plutôt que différentes, qu'elles apparaissaient ou non au même endroit que leur amorce. Pour les auteurs, l'effet facilitateur suggère qu'indépendamment de sa position, l'amorce active une représentation orthographique pré-lexicale.

1.4.3. Paradigme FLLD (*Flanking letter lexical decision*)

Une preuve supplémentaire de l'effet de facilitation orthographique de la parafovéa sur la fovéa a été apportée récemment dans une étude de Dare et Shillcock (2013) utilisant un nouveau paradigme nommé *Flanking Letter Lexical Decision (FLLD)*. Ce paradigme est utilisé principalement pour étudier l'effet des informations parafovéales sur le traitement de mots en fovéa, le traitement parallèle de mots et la distribution spatiale de l'attention. Dare et Shillcock (2013) ont utilisé une tâche de décision lexicale sur une cible de quatre lettres non répétées située au centre de la vision (*i.e.* zone fovéale), entourée à gauche et à droite de bigrammes (*i.e.* zone parafovéale) présentés simultanément.

Trois conditions ont été manipulées dans leur étude : une condition Bigrammes adjacents dans laquelle les bigrammes suivaient l'ordre des lettres dans la cible (e.g., RO ROCK CK) ; une condition Bigrammes inversés dans laquelle l'ordre des bigrammes étaient inversés

(e.g., CK ROCK RO) et une condition Bigrammes différents (condition contrôle) dans laquelle les bigrammes n'avaient pas de relation orthographique avec la cible (e.g., DA ROCK SH) mais leur réunion constituait un autre mot (DASH). Les résultats ont montré que les deux premières conditions où les bigrammes entretenaient une relation orthographique avec la cible engendraient les temps de réaction les moins élevés, que l'ordre des bigrammes soit respecté ou non. Le même paradigme a été utilisé par Grainger, Mathôt et Vitu (2014) dans une étude en français qui a révélé des résultats similaires à ceux obtenus en anglais par Dare et Shillcock (2013).

Ces résultats renforcent l'idée que le bigramme, plutôt que la lettre, est une unité de traitement orthographique privilégiée et que le codage des lettres s'effectue en position relative plutôt qu'en position absolue. Notons que pour que l'effet d'inversion des bigrammes se manifeste, l'ordre des lettres dans le bigramme doit être préservé. Les résultats de ces deux études ont été interprétés sur les bases théoriques d'un nouveau modèle de traitement orthographique multi-mots² (Grainger et al., 2014) (cf. Figure 8). Le modèle postule un traitement parallèle des informations concernant le mot fixé, ainsi que celles à sa gauche et à sa droite, dans les limites imposées par l'attention spatiale. Il introduit également un codage en bigrammes à lettre unique (bigrammes formés par la combinaison d'un espace inter-mots (noté #) et d'une lettre).

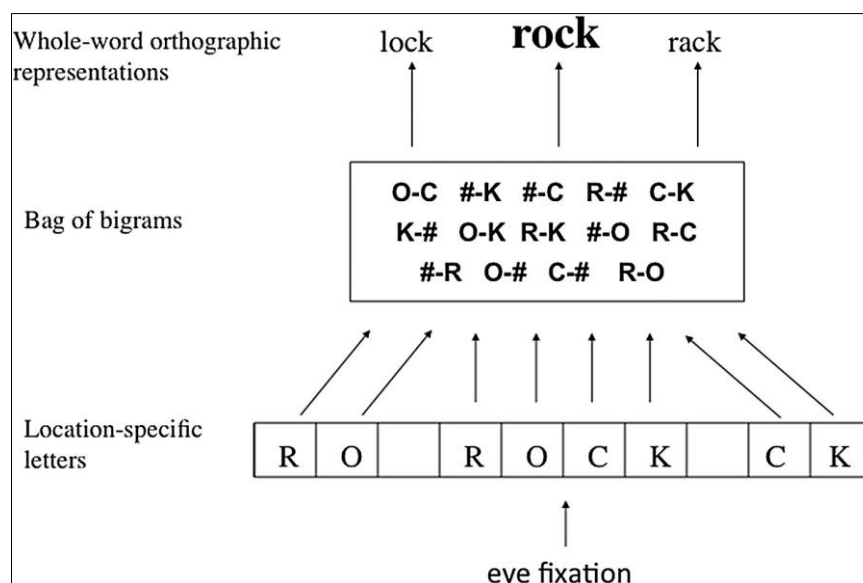


Figure 8. Architecture du modèle de traitement orthographique multi-mots de Grainger et al. (2014 ; figure p.2).

² Extension du modèle à bigrammes ouverts de Grainger et Van Heuven (2003).

Les détecteurs de l'emplacement des lettres (*Location-specific letters*) opèrent en parallèle sur un ensemble de mots, signalant qu'une lettre donnée (ou un espace inter-mots) est présente dans un emplacement donné, relatif au point de fixation (*Eye fixation*). Ces informations activent un ensemble de paires ordonnées de combinaison de caractères contigus et non contigus (lettres et espaces inter-mots) stocké sous forme d'un ensemble non ordonné de bigrammes ouverts (*Bag of bigrams*) formés seulement à l'intérieur du mot, et non pas entre les mots. Ainsi, à la lecture de la séquence "gray mouse", les bigrammes "g -r" et "g-y" sont activés, mais pas le bigrammes "y-m". Les bigrammes activent à leur tour les représentations orthographiques du mot entier (*whole-word orthographic representations*).

Le paradigme *FLLD* a permis également de tester la distribution de l'attention en lecture de mots. Dans une étude, Snell et Grainger (2018) ont manipulé sept conditions: une condition Cible isolée sans flankers (e.g., rock) ; quatre conditions bilatérales³ : condition Flankers répétés (e.g., rock rock rock), condition Flankers différents (e.g., step rock step), condition Mixte flanker répété à gauche (e.g., rock rock step), condition Mixte flanker répété à droite (e.g., step rock rock) ; et deux conditions unilatérales⁴ : condition Flanker unique répété à gauche (e.g., rock rock), condition Flanker unique répété à droite (e.g., rock rock). Les résultats ont montré un biais attentionnel à droite analogue à celui obtenu dans une situation naturelle de lecture avec recueil de mouvements oculaires (McConkie & Rayner, 1976 ; Rayner, 1998) : les temps de réaction étaient significativement moins élevés pour la condition Mixte flanker répété à droite par rapport à la condition Mixte flanker répété à gauche, et pour la condition Flanker unique répété à droite par rapport à la condition Flanker unique répété à gauche.

1.5. Traitement de l'information parafovéale chez l'apprenti-lecteur

Bien que l'information parafovéale ne soit pas directement fixée, son rôle dans la lecture a maintes fois été mis en évidence chez le lecteur expert (Rayner, 1998 pour une revue). Durant ces dernières années, ce champ de recherche n'a pas été largement investi chez l'apprenti lecteur. Toutefois, quelques recherches soutiennent directement ou indirectement l'utilité de l'information parafovéale chez l'apprenti lecteur (Rayner, 1986 ; Häikiö, Bertram, Hyönä, & Niemi, 2009 ; Häikiö, Bertram & Hyönä, 2010).

³ Conditions avec *flankers* à gauche et à droite de la cible.

⁴ Conditions avec *flankers* soit à gauche soit à droite de la cible.

Certaines données récentes en français (Khelifi, Sparrow, & Casalis, 2012, 2015) concernant le développement de l'attention en lecture attestent que la distribution de l'attention chez les apprentis lecteurs ressemble à celle observée chez le lecteur expert. Dans deux groupes d'expériences, les auteurs ont pu mettre en évidence une sensibilité des apprentis-lecteurs aux informations parafovéales, et un bénéfice de ces informations lors du traitement de mots en fovéa. L'objectif du premier groupe d'expériences était de montrer si l'apprenti lecteur, de grades 3 et 5, prélève l'information latérale en situation de lecture de mots isolés en vision centrale. Dans une tâche d'identification perceptive, les sujets étaient amenés à identifier des mots en vision centrale, accompagnés ou non de distracteurs présentés soit à gauche, soit à droite de la cible à identifier. L'effet délétère du distracteur latéral sur le traitement fovéal était utilisé comme indice révélateur de la sensibilité à l'information latérale. Les données des deux groupes de participants ont confirmé cette sensibilité qui s'est observée au travers d'une diminution significative du taux d'identification de mots cibles en présence de distracteurs. Ce qui pourrait suggérer que tout ou une partie de l'information latérale a été extraite et traitée.

Un deuxième groupe d'expériences a été mené chez des lecteurs de deux niveaux (grades 3 et 5) pour examiner le traitement de l'information parafovéale chez l'apprenti lecteur en lecture de mots isolés, et sa contribution dans les processus d'accès au lexique. Une tâche de décision lexicale avec un amorçage en vision parafovéale a été utilisée. Avant de recevoir un mot cible en vision fovéale, un aperçu parafovéal était délivré aux participants durant une brève période dans le champ visuel gauche ou droit. Le traitement de l'information parafovéale était évalué à partir de l'effet (facilitateur ou perturbateur) de ce traitement sur la reconnaissance ultérieure d'un mot cible présenté en région fovéale. L'aperçu parafovéal était manipulé de manière à présenter soit toutes les lettres, soit une partie des lettres ou soit encore aucune lettre de la cible à identifier. Les données recueillies ont montré que les apprentis lecteurs sont capables d'extraire de l'information parafovéale : le mot cible était identifié plus rapidement lorsque les participants recevaient un aperçu approprié (identique ou partiellement identique à la cible) plutôt qu'inapproprié (différent de la cible). Ce résultat suggère que le bénéfice de l'aperçu parafovéal peut être mis en évidence assez précocement au cours de l'apprentissage de la lecture, dès la troisième année du cycle.

Le même constat a été fait à l'issue d'une étude toujours en français de Snell, Cauchi, Grainger et Lété (2020) dans laquelle des élèves de primaire du CP au CM2 (grades 1 à 6) ont

été testés avec le paradigme *FLLD*. Chaque essai commençait par la présentation d'une croix de fixation au centre de l'écran pendant une durée de 1000 ms. La cible et les bigrammes l'entourant (flankers) étaient présentés pendant une durée relative au niveau d'apprentissage (Grade 1, 300 ms ; Grade 2 et plus, 250 ms). Les participants avaient une durée maximale de 3500 ms pour répondre. Les résultats des trois conditions manipulées (Bigrammes adjacents (e.g., RO ROCK CK) ; Bigrammes inversés (e.g., CK ROCK RO) ; Bigrammes différents (e.g., DA ROCK SH) ont révélé chez les participants de tous les niveaux une différence non significative entre les conditions Bigrammes adjacents et inversés. Les temps de réaction de ces deux conditions étaient significativement plus courts par rapport à la condition Bigrammes différents. L'effet de l'intégration spatiale de l'information orthographique par lequel la cible était reconnue plus rapidement lorsqu'elle était entourée de bigrammes adjacents que lorsqu'elle était entourée de bigrammes différents suggère que chez l'apprenti-lecteur, comme chez le lecteur expérimenté, la distribution spatiale de l'attention est déployée au-delà du mot fixé même à un stade précoce de l'apprentissage (dès le CP), permettant ainsi un traitement parallèle des informations fovéales et parafovéales en une seule fixation.

2. Chapitre 2 : La reconnaissance des mots écrits en arabe

2.1. La langue arabe

2.1.1. Aperçu historique

La langue arabe, dans sa forme écrite, est l'une des plus anciennes langues vivantes dans le monde actuel. Avant l'avènement de l'islam, l'arabe était en contact avec les langues appartenant à la famille sémitique comme le phénicien, l'hébreu, l'araméen, etc. Au sein de cet ensemble, elle appartient au sous-groupe du sémitique méridional. Avec l'extension de l'empire arabo-islamique à partir du VII^{ème} siècle, l'arabe est devenu l'une des principales langues véhiculaires⁵ du monde. Elle a connu, au cours de sa longue histoire, d'importantes périodes de création littéraire. C'est grâce à l'immense apport coranique et à toute la production religieuse qu'il a engendrée que l'arabe dit "classique" a succédé à l'arabe ancien. Cette mutation est attestée en particulier par le développement de l'écriture et par la codification de la langue et l'apparition des premiers dictionnaires et traités de grammaire.

L'évolution de l'arabe classique pendant un millénaire a fini par engendrer "l'arabe standard moderne", langue de la presse écrite et audio-visuelle, des débats politiques, des textes scientifiques et des textes littéraires. Parlée dans la plupart des pays arabes, l'arabe standard moderne n'est en revanche presque jamais la langue des échanges quotidiens. Les diverses formes d'arabe dialectal diffèrent les unes des autres et de l'arabe moderne sur le plan de la prononciation, du vocabulaire et de la grammaire.

La langue arabe peut donc être considérée comme renvoyant à un terme générique regroupant plusieurs variétés :

- L'arabe classique : la langue du Coran, parlée au VII^{ème} siècle.
- L'arabe standard moderne : une forme un peu différenciée de l'arabe classique qui constitue la langue écrite. Il s'agit de la variété retenue comme langue officielle dans tous les pays arabes et comme langue commune entre eux. Elle est également la langue employée dans la plupart des écrits administratifs, médiatiques, scientifiques, techniques, littéraires ainsi que dans la majorité des journaux télévisés.

⁵ Une langue véhiculaire est une langue qui permet les échanges entre des groupes parlant des langues différentes. C'est le contraire d'une langue vernaculaire parlée à l'intérieur d'un seul groupe.

- L'arabe dialectal : langue parlée au quotidien dans les pays arabes. D'un pays à l'autre, voire d'une région à l'autre dans un même pays, les dialectes arabes diffèrent. Il emprunte de nombreux vocables aux langues étrangères. Toutefois, de nombreux mots sont assez similaires entre les deux registres du fait de leurs bases linguistiques communes. Néanmoins, de nombreuses différences formelles subsistent. Le système phonologique de l'arabe dialectal est largement altéré par rapport à celui de l'arabe standard. Ces altérations portent en particulier sur les voyelles (changement de timbre, suppression ou ajout, transformation de leur longueur).

L'arabe dialectal et l'arabe standard peuvent être considérés comme deux langues différentes, la maîtrise des deux registres s'apparentant de fait à une sorte de bilinguisme (Baccouche, 2003 ; Saiegh-Haddad, 2004, 2005).

2.1.2. L'alphabet arabe

L'arabe est une langue sémitique qui s'écrit de droite à gauche. Son alphabet est composé de 28 lettres (cf. Tableau 1). Chaque lettre est représentée par différentes formes (allographes) selon qu'elle est isolée, au début, au milieu ou à la fin du mot.

L'arabe possède une écriture semi-cursive : 22 lettres s'écrivent reliées aux lettres qui les précèdent et qui les suivent dans le mot, et 6 lettres (ا و ز ر د ذ) ne se relient pas aux lettres qui les suivent (e.g., شروق [ʃruːq] = levée de soleil). Il en résulte que certains mots pourraient s'écrire en plusieurs segments, ou même en lettres isolées (e.g., ورق [wrq] = papier). Une grande similarité visuelle est observée entre les lettres. Elles peuvent s'organiser en neuf groupes en fonction de leur similitude visuelle (Boudelaa et al., 2019) (cf. Tableau 2). Celles des groupes 1 à 8 partagent la même forme de base, avec pour seules différences le nombre et la position des points qui les accompagnent. Par exemple dans le groupe 1, la forme de base des trois lettres est identique, mais la lettre ب [b] possède un seul point en dessous, alors que les deux lettres ت [t] et ث [θ] possèdent respectivement deux et trois points au-dessus. Les dix lettres du groupe 9 ont une forme de base unique qui n'est partagée ni entre elles, ni avec les lettres des autres groupes.

Chapitre 2 - La reconnaissance des mots écrits en écriture arabe

Tableau 1: Alphabet arabe et variations allographiques de ses lettres en fonction de leur position dans le mot.

Lettre	Nom	Fin	Milieu	Début	API
ا	alif	ـا	ـا	ا	a:
ب	ba	ـب	ـب	ب	b
ت	ta	ـت	ـت	ت	t
ث	ta (tha)	ـث	ـث	ث	θ
ج	ġim (jim)	ـج	ـج	ج	ʒ
ح	Ḥa	ـح	ـح	ح	ħ
خ	ḥa (kha)	ـخ	ـخ	خ	x
د	dal	ـد	ـد	د	d
ذ	ḍal (dhal)	ـذ	ـذ	ذ	ð
ر	ra	ـر	ـر	ر	r
ز	zay	ـز	ـز	ز	z
س	sin	ـس	ـس	س	s
ش	šin (shin)	ـش	ـش	ش	ʃ
ص	Ṣad	ـص	ـص	ص	s ^ʕ
ض	Ḍad	ـض	ـض	ض	d ^ʕ
ط	Ṭa	ـط	ـط	ط	t ^ʕ
ظ	Ẓa	ـظ	ـظ	ظ	z ^ʕ , ð ^ʕ
ع	ʿayn	ـع	ـع	ع	ʕ
غ	ġayn (ghayn)	ـغ	ـغ	غ	ɣ
ف	fa	ـف	ـف	ف	f
ق	qaf	ـق	ـق	ق	q
ك	kaf	ـك	ـك	ك	k
ل	lam	ـل	ـل	ل	l
م	mim	ـم	ـم	م	m
ن	nun	ـن	ـن	ن	n
ه	ha	ـه	ـه	ه	h
و	waw	ـو	ـو	و	w, u:
ي	ya	ـي	ـي	ي	j, i:
ء	hamza	أ و إ ي			ʔ

Tableau 2 : Groupes des lettres en arabe selon leur similitude visuelle.




Groupes	Exemple	Groupes	Exemple
1	ب ت ث /b, t, θ/	5	ش س /s, ʃ/
2	ج ح خ /j, ħ, x/	6	ض ص /s ^ʕ , d ^ʕ /
3	ذ د /d, ð/	7	ظ ط /t ^ʕ , ð ^ʕ /
4	ز ر /r, z/	8	غ ع /ɣ, ʕ/
9	أ ي و ه ن م ل ك ق ف /f, q, k, l, m, n, h, w, j, ʔ/		

Une seule lettre de l'alphabet arabe, (ا) (alif), entre dans la catégorie des voyelles longues. Vingt-cinq lettres représentent des consonnes. Parmi celles-ci, 2 lettres (و , ي) (waw,

ya) sont des semi-consonnes qui peuvent représenter également des voyelles longues ([i:]; [u:]) selon le contexte orthographique.

Les voyelles courtes ne font pas partie de l'alphabet mais sont représentées par un système facultatif de signes diacritiques qui s'ajoutent au mot. Ces signes sont présents pour des nécessités didactiques chez les enfants dans les premières étapes de l'apprentissage de la lecture. Elles sont au nombre de trois (cf. Tableau 3).

Tableau 3. Codage des trois voyelles courtes en arabe.

Signe diacritiques codant la voyelle	Nom de la voyelle	Prononciation (API)
	<i>Fatḥa</i> (Trait placé au-dessus de la lettre)	[a]
	<i>Damma</i> (Nœud placé au-dessus de la lettre)	[u]
	<i>Kasra</i> (Trait placé au-dessous de la lettre)	[i]

En plus des signes diacritiques représentant les voyelles courtes, il existe d'autres signes complémentaires :

- *Sukoon* [◌◌◌◌◌◌] : ce signe indique l'absence de voyelles.
- *Shadda* [◌◌◌◌◌◌◌◌] : ce signe marque le dédoublement de la lettre.
- *Madda* [◌◌◌◌◌◌◌◌◌◌◌] : ce signe est écrit au-dessus de la lettre *alif* pour indiquer son allongement.

Ces particularités alphabétiques donnent lieu à deux formes d'écriture : une forme voyellée (avec signes diacritiques) et une forme non voyellée (sans signes diacritiques) (cf. Tableau 4). Lorsque ces signes diacritiques sont présents pour des nécessités didactiques, l'orthographe arabe offre des relations transparentes entre graphème et phonème. Cependant, l'écriture standard destinée au lecteur expert ne note que le squelette du mot formé par les consonnes et les voyelles longues, ce qui donne lieu à un nombre de mots de formes identiques mais de sens différents (cf. Tableau 4 pour un exemple), et rend de ce fait les relations graphème/phonème moins transparentes (Abu-Rabia, Share & Mansour, 2003 ; Elbeheri & Everatt, 2007). Le lecteur doit alors se baser sur le contexte pour comprendre correctement le sens du texte.

Tableau 4 : Exemple d'une forme voyellée et d'une forme non voyellée de la racine كتب [ktb].

Forme	Prononciations possibles (API)	Sens
Voyellée (كَتَبَ)	[kataba]	il a écrit
Non voyellée (كتب)	[kataba]	il a écrit
	[kutiba]	il a été écrit
	[kutub]	livres

2.1.3. Caractéristiques morphologiques

L'arabe est une langue à morphologie non concaténative⁶. Son système morphologique se caractérise par une structure bidimensionnelle composée d'une racine et d'un schème⁷. La racine est généralement composée de trois consonnes, les racines à deux ou quatre consonnes sont peu nombreuses. Cette séquence trilitère, imprononçable, code une signification générale et abstraite. Par exemple, la suite كتب [ktb] signifie "ce qui se rapporte à l'écrit".

La formation des mots du lexique arabe nécessite obligatoirement l'ajout d'un schème à la racine. Le schème contient des informations sémantiques et syntaxiques particulières. Son application fournit ainsi une forme lexicale et grammaticale aux mots en complétant le sens général de la racine. Il s'agit d'un pattern dérivationnel (verbal ou nominal) qui permet de dériver les différents éléments de la langue (verbes, substantifs et adjectifs, adverbes...). Il peut être composé de voyelles, ou d'une combinaison de voyelles et de consonnes (préfixes, infixes et suffixes). C'est donc le schème qui véhicule les signes diacritiques absents en écriture non vocalisée. À l'opposé du français où la racine peut se modifier partiellement ou totalement d'un mot à un autre (e.g., revenir/retour, nager/natation, aller/je vais, j'irai...), l'arabe garde toujours sa racine de base dans la dérivation (Bouanani, 2015).

Pour désigner ces schèmes, les grammairiens arabes emploient comme modèle la racine de trois consonnes فعل (qui signifie l'idée d'agir, de faire) qui sert de patron à toutes les formes dérivées d'une racine. Les trois consonnes du modèle représentent les consonnes de

⁶ La morphologie non concaténative, également appelée morphologie discontinue, est une forme de formation et d'inflexion de mot dans laquelle la racine est modifiée sans impliquer l'enchaînement séquentiel de morphèmes.

⁷ Le schème est l'ensemble de consonnes et/ou de voyelles qui s'ajoutent aux consonnes de la racine afin de former des mots.

n'importe quelle racine trilitère. Ainsi, dans l'exemple 3 (cf. Tableau 5), le nom كَاتِب ([ka:tib] = écrivain) est formé par le croisement de la racine كَتَب [ktb] et du schème nominal فاعِل [Ca:CaC] .

Tableau 5: Exemple de formes orthographiques construites à partir de la racine كَتَب [ktb].

Mot	(API)	Sens	Schème appliqué ⁸	Valeur du schème	Type du schème
1) كَتَبَ	[kataba]	Il a écrit	فَعَلَ [CaCaCa]	L'accompli	Schémes
2) كَاتَبَ	[ka:taba]	Il a écrit à quelqu'un	فَاعَلَ [Ca:CaCa]	Valeur transitive	verbaux
3) كَاتِب	[ka:tib]	Écrivain	فَاعِل [Ca:CaC]	Nom d'agent	Schémes
4) مَكْتَب	[maktab]	Bureau	مَفْعَل [maCCaC]	Lieu de l'action	nominaux

Il existe de nombreux schèmes en arabe. Les schèmes verbaux sont peu nombreux. Les grammairiens comptent ainsi quatorze formes verbales dérivées mais le nombre de schèmes les plus fréquents est plus restreint (moins d'une dizaine). Chaque forme exprime des modalités verbales différentes telles que l'intensité (e.g., كَسَّر [kassara] = fracasser), le factitif (e.g., كَتَّبَ [kattaba] = faire écrire à quelqu'un), la réciprocité (e.g., كَاتَبَ [ka:taba] = s'échanger une correspondance), le réfléchi (e.g., اِنكَسَرَ [inkasara] = se casser), etc. Les schèmes nominaux sont en revanche très nombreux. En effet, les nombreuses variations pour exprimer une même idée complexifient la possibilité de les dénombrer. Toutefois, il est possible d'en isoler plusieurs dont la valeur est relativement fixe. On retrouve ainsi systématiquement les participes actifs et passifs, les noms d'action, régulièrement les noms d'instruments, de lieu, etc. (Kouloughli, 1994).

2.1.4. Synthèse

Comme constaté précédemment, le système orthographique de l'arabe se caractérise par quelques spécificités. Une des principales caractéristiques de l'arabe provient du fait que la forme de la lettre varie selon sa position dans le mot. La majorité des lettres arabes ont quatre formes différentes selon leur position et sont visuellement similaires quant à leurs formes de base.

Une autre spécificité du système orthographique arabe est la structure bidimensionnelle du mot (racine / schème). Contrairement aux langues européennes où les voyelles font partie de l'alphabet et sont représentées par des lettres spécifiques, les voyelles

⁸ C = consonne de la racine.

sont représentées par des signes diacritiques. Ainsi, la langue arabe repose sur un système qui peut être voyellé ou non. L'écriture voyellée, utilisée au début de l'apprentissage de la lecture, est phonologiquement transparente. L'écriture non voyellée, destinée au lecteur expert, se caractérise en revanche par son opacité liée à l'absence de signes diacritiques. Pour résoudre l'ambiguïté de certains mots, il faut recourir au contexte.

2.2. La reconnaissance des mots en écriture arabe

La particularité morphologique de l'arabe, construite sur la base de racines, a conduit les chercheurs à s'interroger sur le rôle de ces unités morphémiques dans la reconnaissance de mots. Ces études ont débouché sur des résultats qui apparaissent contradictoires.

Dans une étude de cas, Prunet, Béland et Idrissi (2000) ont collecté un ensemble d'erreurs de métathèse⁹ produites par un patient aphasique bilingue (arabe/français). Le patient était natif du Liban, ayant commencé à étudier parallèlement l'arabe et le français à l'âge de 4 ans. À l'âge de 22 ans, il a poursuivi ses études dans une faculté à Montréal où la langue d'enseignement était le français. Après avoir obtenu son diplôme, il travaille pendant huit ans comme ingénieur dans des compagnies où le français était la langue utilisée. L'arabe était cependant la langue d'usage dans le cadre familial et amical. À l'âge de 32 ans, il a subi un accident vasculaire cérébral qui a entraîné une aphasie mixte sévère, c'est-à-dire un déficit linguistique dans lequel la production et la compréhension étaient altérées en arabe et en français. Néanmoins, il ne rencontrait aucune difficulté pour comprendre des mots isolés.

Le sujet était testé dans les deux langues dans des tâches de lecture de mots isolés fréquents à haute voix, de répétition de mots écoutés, de dénomination d'images et de dictée. Dans les quatre tâches, il commettait de fréquentes erreurs de métathèse en arabe, qui consistaient en la modification de l'ordre linéaire des consonnes de la racine (e.g., مجدوه [m₃du:h]- au lieu de - مجهود [m₃hu:d] = effort). En revanche, les schèmes, les voyelles et les affixes consonantiques restaient intacts. En Français, les erreurs de métathèses portaient parfois sur les consonnes (e.g., eplikser - au lieu de - éclipser) et parfois sur les voyelles et même les syllabes (e.g., palodé - au lieu de - pédalo). En plus de ces différences qualitatives, une différence quantitative a été observée : les erreurs de métathèse étaient beaucoup plus nombreuses en arabe qu'en français (119 contre 12) même si le nombre total

⁹ Inversion de l'ordre de phonèmes ou de syllabes dans un mot.

des stimuli utilisés pour le test était quasiment le même (1455 en arabe contre 1498 en Français). Les auteurs avaient conclu à l'évidence de l'existence des racines en tant qu'unités morphémiques qui feraient partie du lexique mental et qui joueraient un rôle crucial dans l'accès lexical.

Le rôle de la racine dans la reconnaissance des mots en arabe a été également étudié en tâche de décision lexicale à l'aide du paradigme d'amorçage. Une expérience de Boudelaa et Marslen-Wilson (2000) menée avec le paradigme d'amorçage intermodal (amorce auditive/cible visuelle) a testé le rôle de la racine dans le traitement de mots. Dans cette expérience, quatre conditions ont été alternées : une condition [+R+S] dans laquelle l'amorce et la cible partageaient la même racine avec une relation sémantique transparente (e.g., إدخال [ʔdxa:l] = insertion et دخول [dxu:l] = entrée) ; une condition [+R-S] identique à la première, sauf que la relation sémantique entre amorce et cible était opaque (e.g., مداخلة [mda:xlh] = intervention orale et دخول [dxu:l] = entrée) ; une condition [-R+S] dans laquelle l'amorce et la cible étaient reliées sémantiquement mais pas morphologiquement (e.g., منفذ [mnfð] = sortie et دخول [dxu:l] = entrée) ; une condition contrôle dans laquelle l'amorce et la cible n'étaient reliées ni morphologiquement ni sémantiquement (e.g., قهوة [qhwh] = café et دخول [dxu:l] = entrée). Un effet significatif d'amorçage a été obtenu avec des amorces partageant la même racine que la cible, indépendamment de la nature de leur relation sémantique (transparente/opaque), ce qui révèle un traitement morphologique des mots basé sur la racine, indépendamment des relations sémantiques existant entre les mots de la même famille.

Cette sensibilité à la racine a été reconfirmée dans une autre étude de Boudelaa et Marslen-Wilson (2005) en amorçage incrémental masqué. Ces auteurs ont montré qu'avec les quatre asynchronies (SOAs) (32, 48, 64, et 80 ms), la présentation d'une amorce qui partageait la racine du mot cible accélérail les temps de décision lexicale quel que soit le degré de transparence sémantique entre l'amorce et la cible. Par ailleurs, il a été révélé que la détection de cette unité morphémique différait de celle du schème puisque la facilitation par amorçage d'un schème est notée seulement à SOAs 48 et 64ms pour les noms et à SOAs 48ms pour les verbes, ce qui montre que les temps nécessaires à l'extraction de l'information de la racine et du schème n'ont pas le même poids (la racine est plus rapidement repérable dès 32 ms de SOA) étayant de ce fait le rôle que peuvent jouer les racines dans la construction des représentations lexicales.

Grainger, Dichy, El-Halfaoui et Bamhamed (2003) dans une série de deux expériences d'amorçage subliminal, manipulant dans la première le nombre de lettres partagées entre l'amorce et le mot cible (de trois consonnes), et dans la seconde la fréquence de la racine par rapport à celle la forme orthographique du mot, avaient conclu à une absence d'effet de l'amorçage orthographique avec des mots arabes tri-consonantiques. En revanche, ils ont trouvé un effet facilitateur de la fréquence de la racine sur les temps de réaction. Cet effet étant plus important pour les cibles de basse fréquence lexicale que pour les cibles de haute fréquence lexicale.

L'hypothèse du rôle déterminant de la racine dans l'accès lexical a été pourtant infirmée dans une autre étude de Abu-Rabia et Awwad (2004). Dans une tâche de décision lexicale avec amorçage masqué, les auteurs n'ont pas observé d'effet facilitateur de la préexposition à la racine (e.g., **صعد** [s'aʔida] (*monter*) / **مصعد** [misʔad] (*ascenseur*); la racine étant [sʔd]) par rapport à une condition orthographique où la cible partageait les trois consonnes de l'amorce, celle-ci n'étant cependant pas une racine (e.g., **قول** [qu:l] (*proverbe*) / **مقتول** [mqtu:l] (*tué*), suggérant que les mots en arabe ne seraient pas représentés en tant qu'unités morphémiques décomposables (racines et schèmes) mais en tant qu'unités orthographiques. L'effet d'amorçage orthographique obtenu dans d'autres études semble confirmer ce constat. Par exemple chez Perea, Abu Mallouh et Carreiras (2014), des temps de réaction plus rapides ont été enregistrés lorsque l'amorce était reliée orthographiquement à une cible de 5 lettres (e.g., **كخابة** [kxa:ʁh] (*pseudo-mot*) / **كتابة** [kita:baʁh] (*écriture*); la racine étant [ktb]) sans pour autant partager toutes les consonnes de la racine, comparés à une condition non reliée (e.g., **دائلي** [da:θli] (*pseudo-mot*) / **كتابة** [kita:baʁh] (*écriture*)). Dans une autre étude de Perea, Abu Mallouh, Mohammed, Khalifa et Carreiras (2018) avec des cibles de 4 lettres, trois conditions ont été comparées : une condition identique (e.g., **خدمة** [xɗmh] – **خدمة** [xɗmh] = *service*); une condition où l'amorce pseudo-mot était reliée orthographiquement à la cible avec substitution d'une seule consonne de la racine (e.g., **فدمة** - [fɗmh] – **خدمة** [xɗmh]; la racine étant [xɗm]), et une condition créée à partir de pseudo-mots non reliés orthographiquement à la cible (e.g., **خدمة - سكاب** [ska:b] – [xɗmh]). L'effet d'amorçage observé pour la condition reliée était plus important par rapport à la condition non reliée et moins important pour la condition identique.

La littérature ne fait cependant pas l'objet d'un consensus quant à la nature des représentations des mots en arabe. Néanmoins, certaines recherches suggèrent que la

fréquence serait un facteur affectant les représentations lexicales (Bybee, 1995; Katz, Rexer & Lukatela, 1991, cité dans Abu-Rabia & Awwad, 2004). Selon ces auteurs, la décomposition morphologique serait nécessaire à l'accès lexical essentiellement dans le traitement de nouveaux mots ou de mots de basse fréquence. Les mots familiers et de haute fréquence ne nécessiteraient pas une décomposition morphologique. Cet effet de fréquence pourrait expliquer l'absence d'effet morphologique en arabe observé par Abu-Rabia et Awwad (2004) qui n'ont utilisé que des mots de haute fréquence dans leur étude. Il semble également cohérent avec les résultats de Grainger et al. (2003) démontrant un fort effet de la fréquence de la racine pour les mots de basse fréquence, cet effet n'étant plus apparent pour les mots de haute fréquence. Dans les deux études de Boudelaa et Marslen-Wilson (2000, 2005), aucune information sur la fréquence lexicale n'a été fournie, laissant supposer que la présence d'effet d'amorçage par la racine serait due à l'utilisation de mots moins fréquents.

2.3. Le traitement orthographique du mot écrit en arabe

2.3.1. La reconnaissance des lettres

Comme vu au Chapitre 1, les mots ne sont pas traités dans leur globalité mais par un traitement en parallèle des lettres qui les constituent (Pelli, Farell & Moore, 2003 ; Perea & Rosa, 2002 ; Rayner, McConkie, & Zola, 1980 ; Grainger & Jacobs, 1996). Les recherches en écriture latine ont révélé que les lettres sont reconnues via leurs traits visuels distinctifs (lignes horizontales, verticales, diagonales ; courbes ; intersections, etc.) (e.g., Jacobs, Nazir & Heller, 1989 ; Solomon & Pelli, 1994). Les indices de hauteur (e.g., la lettre " l " par rapport à la lettre " n ") et de largeur (e.g., la lettre " H " par rapport à la lettre " i ") sont également des caractéristiques fondamentales pour la discrimination et la détermination de la forme de la lettre (Bouma, 1971). Il a été démontré également que l'effet d'amorçage était équivalent pour les paires de lettres (minuscule/majuscule) qui ont des traits visuels similaires (e.g., c-C), et les paires qui ont des traits dissimilaires (e.g., g-G) (Arguin & Bub, 1995 ; Bowers, Vigliocco & Haan, 1998 ; Ziegler, Ferrand, Jacobs, Rey & Grainger, 2000 ; Kinoshita & Kaplan, 2008). Ce qui converge avec le postulat de l'existence d'une représentation abstraite pour les lettres, indépendamment de la casse.

En arabe, les traits visuels des lettres présentent des similitudes. La majorité des lettres partagent une même forme de base avec un seul trait distinctif, le nombre et l'emplacement

des points diacritiques (e.g., ح , خ , ج) (cf. Tableau 2). Il en résulte que certains mots semblent visuellement identiques, comme dans le cas des deux mots بحار ([bħa:r] = mers) et بخار ([bxa:r] = vapeur) qui ne possèdent qu'un seul trait distinctif, le point sur la deuxième lettre.

Savoir si cette ressemblance visuelle des lettres en arabe affecte la reconnaissance de mots est une question qui a donné lieu à quelques travaux. Dans une tâche de décision lexicale avec amorçage masqué, Perea, Abu Mallouh, Mohammed, Khalifa et Carreiras (2016) ont alterné trois conditions : une première condition dans laquelle l'amorce était identique à la cible (e.g., صحفية - صحفية [s^hħfjh] = journaliste) ; une deuxième condition où l'amorce et la cible ne se distinguaient que par une seule lettre ayant la même forme que la lettre critique (e.g., صحفية [s^xħfjh] - صحفية [s^hħfjh]), et une troisième condition où l'amorce et la cible ne se distinguaient que par une seule lettre ayant une forme différente que celle de la lettre critique (e.g., صحفية [s^kħfjh] - صحفية [s^hħfjh]). Les résultats ont montré que les deux conditions de substitution produisaient des temps de réaction équivalents mais supérieurs à la condition identique. (Voir Perea et al., 2018 pour des résultats similaires). Le constat tiré de ces deux études est que les traits (nombre et emplacement des points diacritiques) distinguant les lettres à forme de base commune sont rapidement traités.

Connaitre comment les lettres sont représentées en mémoire chez l'arabophone (représentation respectant les caractéristiques physiques vs. représentation abstraite comme en écriture latine) est une question qui a fait l'objet de quelques travaux. Comme vu précédemment, le système orthographique arabe ne comporte ni majuscules ni minuscules, seulement des variations allographiques (cf. Tableau 1). Par exemple, la lettre ع [ʕ] peut prendre quatre formes (allographes) selon qu'elle est au début de mot comme dans علوم ([ʕlu:m] = sciences) ; au milieu comme dans سعيد ([sʕi:d] = heureux) ; à la fin, connectée à la lettre précédente, comme dans رابع ([ra:bʕ] = quatrième) ; ou à la fin, en forme isolée, comme dans وداع ([wada:ʕ] = adieu). Cependant, quelques lettres peuvent garder quasiment la même forme visuelle quelle que soit leur position (e.g., ط , ط , ط , [t^ʕ]). Pour vérifier si un effet d'amorçage s'opère aussi bien sur des paires d'allographes visuellement similaires que sur des paires d'allographes visuellement dissimilaires, Carreiras et al. (2012) ont conduit une expérience avec une tâche *identique-différent*, dans laquelle le sujet était amené à décider si l'allographe cible était identique ou différent de l'allographe de référence (e.g., ك [k] / ك [k] comme paire d'allographes visuellement dissimilaires ; ف [f] / ف [f] comme paire d'allographes visuellement similaires). Les résultats ont révélé un effet d'amorçage équivalent à la fois pour

les paires d'allographes visuellement similaires et dissimilaires, ce qui montre que les variations allographiques n'impactent pas la reconnaissance visuelle du mot, et suppose que la représentation de l'identité abstraite des lettres n'est pas spécifique aux langues à alphabet latin mais pourrait refléter un mécanisme universel de traitement de lettres.

2.3.2. Le codage de la position des lettres en arabe

Nous avons vu au Chapitre 1 que les recherches en écriture latine portant sur le codage des positions des lettres dans le mot ont mis en évidence que les non-mots créés par transposition de deux lettres internes activaient les représentations lexicales de leurs mots de base (Perea & Lupker, 2003a, 2003b ; Perea & Carreiras, 2006 ; Schoonbaert & Grainger, 2004). Perea et Lupker (2003) ont montré par exemple que le non-mot *jugde* peut être perçu similairement que son mot de base *judge*. La robustesse de cet effet de transposition de lettres a contredit les modèles de la reconnaissance de mots écrits basés sur la position absolue des lettres dans le mot (e.g., IAM, McClelland & Rumelhart, 1981) et a conforté les modèles plus flexibles à codage relatif de la position des lettres (e.g., *Open-Bigram Model*, Grainger & van Heuven, 2003 ; *SERIOR Model*, Whitney, 2001 ; *SOLAR Model*, Davis, 1999 ; *Overlap Model*, Gomez, Ratcliff & Perea, 2008).

L'arabe, tout comme le français ou l'anglais, est une langue basée sur un système alphabétique où les lettres représentent des phonèmes. En revanche, elle a un système morphologique où la plupart des mots sont construits sur la base de deux morphèmes (racine et schème). Contrairement aux langues européennes à morphologie concaténative, la racine et le schème en arabe ne sont pas combinés linéairement, mais entrelacés. Par exemple le mot مفتاح [mifta:h] = une clé, est construit de la racine فتح [fth] qui fait référence au concept d'ouverture et du schème /miC1C2a:C3/ signifiant le nom de l'instrument (les lettres C indiquent la position de chaque consonne de la racine). La même racine peut se combiner à un autre schème (e.g., /maC1a:C2i:C3/ qui signifie la forme plurielle) pour former le mot مفاتيح [mafa:ti:h] = des clés.

Bien que l'arabe soit une langue alphabétique, l'effet de transposition de lettres y semble difficile à obtenir. Les travaux qui ont porté sur le rôle des consonnes de la racine dans l'accès lexical (e.g., Boudelaa & Marslen-Wilson 2000, 2005 ; Mahfoudhi 2007) ont conclu qu'une racine présentée en amorce facilitait la reconnaissance d'une cible dérivée de cette même racine, ce qui laisse supposer que le codage orthographique en arabe ne porte pas sur

toutes les lettres du mot mais principalement sur les consonnes de la racine. Ces résultats pourraient expliquer l'inflexibilité du codage de la position des lettres en arabe (Pollatsek & Treiman, 2015). Plusieurs racines partagent les mêmes consonnes, ainsi l'altération de l'ordre des consonnes des 4858 racines trilitères et quadrilatères les plus utilisées donne à 54% des cas une racine existante comparé à l'anglais où ces cas ne dépassent pas les 7% (Boudelaa et al., 2019). Par exemple l'altération des consonnes de la racine سبَح [sbħ] = nager, peut produire les racines حَسَب [ħsb] = compter ; حَبَس [ħbs] = emprisonner et سَحَب [shb] = retirer.

La première démonstration de l'inflexibilité du codage de la position des lettres en arabe a été démontrée par Perea, Abu Mallouh et Carreiras (2010). Dans leur étude, deux expériences avec amorçage dans une tâche de décision lexicale ont été conduites : une première expérience dans laquelle la cible (e.g, بعيد [bʕi:d] = loin) était précédée d'une amorce dont la transposition de lettres impliquaient deux lettres de la racine (e.g, عبيد [ʕbi:d] = esclaves) , et une deuxième expérience dans laquelle la cible (e.g, قياس [qja:s] = mesure) était précédée d'une amorce dont la transposition de lettres impliquaient une lettre de la racine, et une lettre du schème, sans que l'ordre des consonnes de la racine soit altéré (e.g, يقاس [jqa:s] = être mesuré). Les résultats ont montré que l'effet d'amorçage était plus robuste quand l'amorce maintenait l'ordre correcte des consonnes de la racine (Expérience 2) que quand elle ne le maintenait pas (Expérience 1). En paradigme des frontières (Boundary paradigm), Hermena, Juma et Aljassmi (2021) ont mis également en évidence un effet de facilitation morphologique quand la cible et l'aperçu parafovéal partageaient les mêmes lettres de la racine dans le bon ordre (e.g., تبديل [tbdi:l] et بدائل [bda:ʔl] ; la racine étant [bdl]). Cet effet disparaissait quand les lettres de la racine étaient dans un ordre différent (e.g., تبديل [tbdi:l] et تديب [tdi:b]). Ceci suggère que l'ordre des consonnes de la racine est important dans la reconnaissance de mots en arabe.

Une des principales caractéristiques de l'arabe est le changement de forme des lettres en fonction de leurs positions dans le mot (initiale, médiane, finale ou isolée) (Tableau 1). Ainsi, la position détermine lequel des allographes utiliser pour telle lettre. Une autre caractéristique est les ligatures entre lettres : vingt-deux lettres de l'alphabet arabe se relient aux lettres qui les précèdent et qui les suivent, et possèdent quatre allographes, alors que six lettres ne se relient qu'aux lettres précédentes et ne possèdent que deux allographes (ا و ز ر) (ذ د). L'allographie en arabe est extrêmement informative à l'égard du codage de la position des lettres. D'une part, parce que chaque allographe est déterminé par sa position dans le

mot. Par exemple l'allographe ع [ʕ] ne peut être ni au début ni à la fin, mais bien au milieu du mot, alors que l'allographe ح [ħ], ne peut être qu'en fin de mot. D'autre part, parce que les ligatures affectent l'emplacement de l'espace à l'intérieur du mot dans le cas des lettres qui ne se relient pas aux lettres suivantes. Par exemple, l'espace dans le mot تجربة (= expérience), apparaît après la troisième lettre, mais en transposant deux lettres internes, l'espace apparaît plutôt après la deuxième lettre dans le pseudo-mot ترجمة.

L'étude de Perea et al. (2010) a été contestée pour deux raisons. D'abord parce que les transpositions de lettres n'ont pas pris en compte les variations allographiques (Kinoshita, Norris & Siegelman, 2012) : les lettres transposées dans les amorces impliquaient des variations allographiques par rapport à la cible (e.g., بعيد [bʕi:d] = loin, et عبيد [ʕbi:d] = esclaves). Ensuite, les amorces étant bien des racines de mots existants, l'absence d'effet de transposition serait dû à la compétition de ces deux racines (Boudelaa, Norris, Mahfoudi & Kinoshita, 2019).

Dans l'étude de Boudelaa et al. (2019), ces deux variables ont été contrôlées. Avec le même paradigme d'amorçage, deux expériences ont été menées. Une en tâche de décision lexicale, l'autre en tâche *identique-différent*. Dans les deux tâches, la transposition des deux consonnes de la racine impliquait dans une première condition une variation allographique de l'amorce par rapport à la cible (e.g., يسعدون [jsʕdu:n] – يسعدون [jʕsdu:n] (= être heureux)) alors que dans une deuxième condition, aucune variation allographique entre l'amorce et la cible n'était impliquée (e.g., يعسدون [jʕsdu:n] – يسعدون [jsʕdu:n] (= être heureux)). Les résultats n'ont révélé aucun effet de transposition de lettres en décision lexicale, en revanche, un effet robuste d'amorçage modulé par les variations allographiques a été révélé dans la tâche *identique-différent* : la condition à variations allographiques était moins facilitatrice par rapport à la condition où aucune variation allographique n'était opérée, supposant que ces variations impactent le processus de traitement orthographique, et fournissent d'importantes informations sur l'ordre des lettres dans le mot. Un deuxième constat tiré de cette étude est que dans la tâche *identique-différent*, qui ne requiert pas un accès lexical, l'effet de transposition de lettres ne semble pas être affecté par la structure morphologique (*i.e.*, l'altération des consonnes de la racine) (Kinoshita et al., 2012), contrairement à la tâche de décision lexicale où le traitement est influencé par la structure morphologique du mot, et l'intégration des informations concernant la racine semble cruciale pour faire une décision lexicale.

2.4. Les traitements visuo-attentionnels et la position optimale de fixation (*Optimal Viewing Position*) en écriture arabe

2.4.1. Traitements visuo-attentionnels

Les expériences utilisant le paradigme de la fenêtre mobile (*movig window*) ont prouvé que l'empan perceptif s'étend plus à droite qu'à gauche du point de fixation (McConkie & Rayner, 1975, 1976 ; Rayner & Bertera, 1979). Ce qui suggère que l'asymétrie de l'empan perceptif dépend des facteurs attentionnels liés aux habitudes de lecture. Lorsque nous lisons de gauche à droite (dans les langues se lisant dans ce sens), les traitements visuo-attentionnels se portent surtout à droite du point de fixation car c'est dans cette partie du champ visuel que se situe l'information utile pour la suite de la lecture. Cette hypothèse est confirmée par le fait que l'asymétrie de l'empan s'inverse dans les langues sémitiques se lisant de droite à gauche. Lorsque des bilingues anglais-hébreu lisent de l'hébreu, les traitements visuo-attentionnels se portent plus à gauche du point de fixation qu'à droite (Pollatsek, Bolozky, Well, & Rayner, 1981). De même, les bilingues anglais-arabe montrent des performances de lecture supérieures pour les fenêtres attentionnelles s'étendant vers la gauche lorsqu'ils lisent en arabe, alors qu'ils montrent un avantage pour les fenêtres s'étendant vers la droite lorsqu'ils lisent en anglais (Jordan, et al., 2014).

2.4.2. Position optimale de fixation

Si la courbe de l'OVP est caractérisée par une asymétrie gauche-droite (*J* inversé) et de meilleures performances pour une position imposée au centre gauche du mot pour des langues lues de gauche à droite (Holmes & O'Regan, 1987 ; Nazir, 1993 ; Nazir et al., 1998 ; O'Regan & Jacobs, 1992 ; O'Regan et al., 1984 ; Vitu, 1991 ; Vitu et al., 1990), l'asymétrie de la courbe de la position optimale de fixation est moins apparente dans les langues lues de droite à gauche (Brysbart & Nazir, 2005). Dans une étude de Farid et Grainger (1996), il a été montré que l'OVP en arabe dépend de la structure morphologique du mot. Dans cette étude, des mots préfixés et suffixés ont révélé des asymétries divergentes : les pourcentages de reconnaissance des mots suffixés étaient plus élevés lorsque les fixations étaient situées au début des mots que lorsqu'elles étaient situées à la fin ; des performances inverses étaient obtenues sur des mots préfixés mettant en évidence une asymétrie inversée, ce qui rejoint les prédictions de l'hypothèse des contraintes lexicales (cf.1.3.2.3)

Une courbe symétrique (en forme de U inversé) a été obtenue cependant dans une étude de Jordan, Almabruk, McGowan et Paterson (2011) où les participants étaient amenés à faire des décisions lexicales sur des mots (et pseudo-mot) de cinq lettres en fonction de six positions imposées (cf. Figure 9) : à droite (position 1), à gauche (position 6), ou à l'une des quatre positions interlettres (2-3-4-5). Les résultats ont révélé des performances faibles pour les positions 1 et 6, moyennes pour les positions 2 et 5, et supérieures pour les positions 3 et 4, c'est-à-dire au centre du mot (cf. Figure 10).

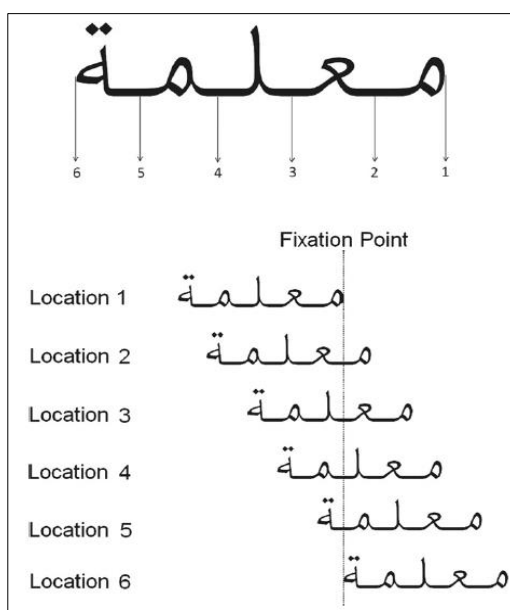


Figure 9. Jordan et al. (2011). Positions (*locations*) des six fixations utilisées pour chaque stimulus.

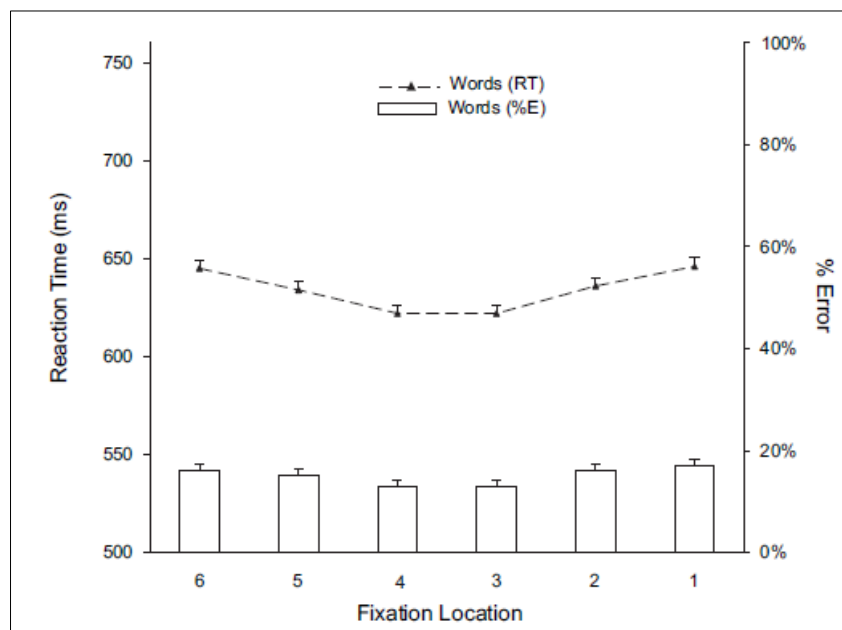


Figure 10. Jordan et al. (2011). Temps de réaction et pourcentage d'erreurs pour chaque position de fixation. Les barres correspondent à 95% d'intervalles de confiance.

La même courbe symétrique (U) a été observée dans une autre étude de Ganayim (2016). Sur des mots de trois lettres aucun effet d'OVP n'a été trouvé. En revanche, sur des mots de quatre et cinq lettres, l'OVP avait tendance à se situer au centre du mot : les temps de réaction étaient similaires quand la deuxième et troisième lettres étaient fixées, et plus élevés lors des fixations des autres lettres (1^{ère} ; 4^{ème} ; et 5^{ème}).

3. Chapitre 3 : Reconnaissance des mots écrits et apprentissage de la lecture en langue arabe

Les recherches portant sur l'apprentissage de la lecture en arabe sont peu nombreuses dans la littérature scientifique en langue anglaise. Elles ont essentiellement exploré trois champs de recherche : d'abord, le problème de la diglossie lié à la distance linguistique entre l'arabe standard lu et l'arabe dialectal parlé ; ensuite, l'effet des voyelles diacritiques et du contexte sur la lecture et, enfin, le problème de la perceptibilité des lettres en écriture arabe lié à la similarité visuelle des lettres entre-elles et au changement de la forme des lettres en fonction de leur position dans le mot. Dans la grande majorité des cas, les aspects orthographiques sont peu abordés, voire pas du tout.

3.1. Organisation de l'enseignement primaire au Maroc

Au Maroc, les niveaux de classes sont décrits en années. Le cycle primaire dure six ans : la 1^{ère} année correspond au CP français (6-7 ans), la 6^{ème} année à la première année de collège (6^{ème}, 11 -12 ans). Dans les écoles privées, on y retrouve un enseignement préscolaire (de 4 à 6 ans) et primaire (de 6 à 12 ans) alors que, dans les écoles publiques, la scolarisation ne débute qu'à partir de la première année (CP français, 6 ans).

L'enseignement primaire est structuré en deux cycles (cf. Figure 11). Le 1er cycle (cycle de base) dure deux ans. Le second cycle (cycle intermédiaire) se déroule sur quatre ans. La fin du primaire est sanctionnée par un certificat d'études primaires.

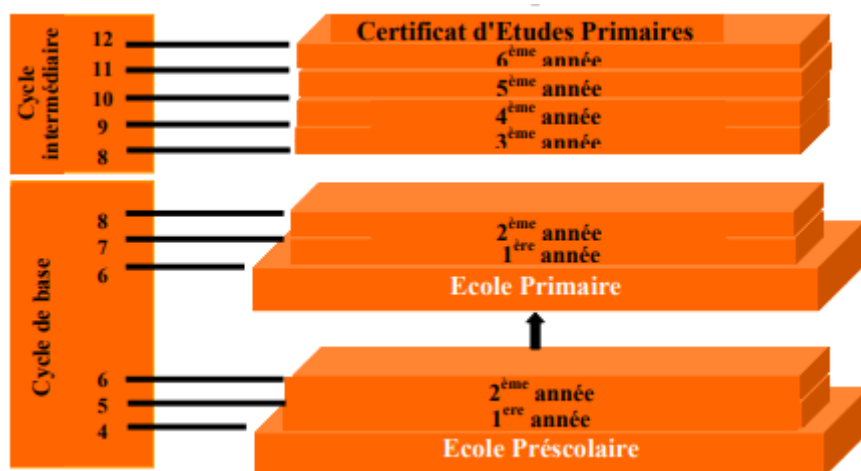


Figure 11 - Structuration de l'enseignement primaire marocain (Ministère de l'Éducation Nationale, de l'Enseignement Supérieur, de la Formation des Cadres de la Recherche Scientifique - MENESCFRCRS, 2007).

Ces deux cycles sont présents dans le système éducatif privé et public. Cependant, la 1^{ère} et 2^{ème} année du premier cycle de base de l'école préscolaire privée, ne sont généralement pas présentes dans les écoles publiques.

Dans l'enseignement privé, l'apprentissage de l'écrit en arabe et en français débute simultanément en 2^{ème} année du préscolaire (Grand Section de Maternelle en France, 5-6 ans). Dans les écoles publiques, l'apprentissage du français ne débute qu'à partir du second cycle, soit en 3^{ème} année (CE2 français, 8-9 ans), alors que l'apprentissage de la langue arabe démarre plus tôt, dès la première année du 1er cycle du primaire (CP français, 6-7 ans).

3.2. Le problème de la diglossie pour l'apprenti-lecteur arabe

La situation de diglossie est définie par Ferguson (1959) comme une situation où sont en usage deux langues structurellement apparentées. En arabe, deux registres coexistent : l'arabe dialectal, pratiqué dans les conversations quotidiennes, est essentiellement oral ; l'arabe standard, pratiqué dans les situations officielles et formelles, est généralement réservé à l'écrit. Des différences importantes existent entre les deux registres aux niveaux phonologique, lexical et syntaxique. En effet, l'arabe dialectal diffère de l'arabe standard par une phonologie altérée, un lexique riche en emprunts lexicaux et une syntaxe simplifiée.

Cette situation sociolinguistique particulière rend l'apprentissage de la lecture en arabe plus difficile. En effet, dans la plupart des systèmes orthographiques, la lecture constitue un prolongement des connaissances linguistiques déjà acquises à l'oral. L'apprenti-lecteur débute son apprentissage avec des connaissances phonologiques assez développées qui devront être exploitées dans l'apprentissage des conversions graphophonologiques. En revanche, l'apprenti-lecteur arabophone aborde l'apprentissage de la lecture dans une langue arabe standard différente de l'arabe dialectal utilisé pour l'apprentissage de sa langue maternelle. Ainsi, un même référent (unité lexicale) est représenté par deux formes phonologiques (arabe dialectal et arabe standard) et par une seule forme orthographique en arabe standard. Par exemple, le mot " beau" se prononce [zwi:n] en arabe dialectal, [zami:l] en arabe standard et s'écrit جميل. L'apprenti-lecteur arabophone doit donc apprendre une autre langue. C'est pourquoi certains chercheurs (Baccouche, 2003 ; Ibrahim, 2006 ; Saiegh-Haddad, 2004, 2005) parlent de bilinguisme pour caractériser l'apprentissage de la lecture

d'un élève arabophone. Cette diglossie donne lieu également à une variabilité phonologique entre l'arabe standard et l'arabe dialectal, marquée principalement par la chute des voyelles courtes. Au niveau des verbes, la chute touche systématiquement la première voyelle et la voyelle finale (e.g., كتب [kataba] en arabe standard qui signifie "il a écrit" se prononce [ktab] en arabe dialectal). Au niveau des noms féminins, la chute de la voyelle courte ne concerne pas la voyelle finale comme pour les verbes (celle-ci est conservée pour marquer le genre) mais l'avant dernière voyelle (e.g., شجرة [ʃaʒara] en arabe standard qui signifie "arbre" se prononce [ʃaʒra] en arabe dialectal). Au niveau de certains adjectifs, la chute des voyelles courtes concerne systématiquement la première voyelle (e.g., كريم [kari:m] en arabe standard qui signifie "généreux" se prononce [kri:m] en arabe dialectal). Ce phénomène répandu au niveau des structures des mots lors du passage du standard au dialectal entraîne une modification de la structure syllabique, compactant à l'écrit des unités lexicales qui tendent, pour certaines, vers le monosyllabisme (Mzoughi, 2015).

Cette distance phonologique entre les deux registres est souvent perçue comme étant à l'origine de nombreuses difficultés d'apprentissage de la lecture (Ammar, 1997 ; Ayari, 1996 ; Belajouza & Snoussi, 2006 ; Feitelson, Goldstein, Iraqi & Share, 1993 ; Saiegh-Haddad, 2003, 2004 ; Snoussi, 2007). Par exemple, dans une épreuve d'identification de mots chez des élèves de 2^{ème} et 3^{ème} année primaire, Ammar (1997) a observé un effet important de la diglossie : plusieurs élèves remplaçaient la structure phonologique du mot cible en arabe standard par celle d'un mot en arabe dialectal. Ce phénomène d'interférence phonologique entre les deux registres lexicaux s'observait essentiellement chez les élèves de 2^{ème} année, il était plus rare en 3^{ème} année. Najem (2006) considère que lors du passage de l'arabe dialectal à l'arabe standard, les arabophones bénéficient de l'invariance du morphème consonantique (*i.e.* la structure consonantique des mots présente dans le lexique des deux registres) mais sont pénalisés par les modifications qui touchent surtout la structure phonologique représentée par les voyelles courtes.

3.3. Le traitement des signes diacritiques

D'autres facteurs liés à la spécificité du système orthographique arabe impactent l'apprentissage de la lecture.

L'arabe possède un système orthographique qui se présente sous deux formes : une forme non vocalisée, notant uniquement le squelette du mot, composée de consonnes et de voyelles longues (ا و ي /i: u: a:), et une forme vocalisée où les voyelles courtes sont ajoutées par les signes diacritiques superposés au squelette du mot (َ ُ ِ). Ce système produit des relations graphème-phonème transparentes, ce qui fait qu'il est tout à fait adapté à des apprentis-lecteurs. Pendant les deux premières années d'apprentissage, tous les textes des manuels scolaires sont vocalisés. Ayant mémorisé un ensemble de formes orthographiques et de formes phonologiques qui lui sont associées, l'apprenti-lecteur commence à s'exercer à une lecture sans voyelles à partir de la 3^{ème} année seulement (vers l'âge de 8 ans). En effet, pour lire un mot sans voyelles, il faut en principe l'avoir déjà correctement vocalisé et l'avoir mémorisé. Ce passage progressif de la forme vocalisée à la forme non vocalisée au cours de l'apprentissage de la lecture est considéré comme une transition d'une phase de codage phonologique vers une phase de codage orthographique (Taouk & Coltheart, 2004).

Ces signes diacritiques représentant les voyelles courtes sont les premières caractéristiques orthographiques de l'arabe ayant attiré l'attention des chercheurs. Leur importance dans l'apprentissage de la lecture a été confirmée par une série d'expériences réalisées par Abu Rabia (1996, 1997a, 1997b, 1998, 2001 ; Abu Rabia & Siegel, 1995). Les résultats obtenus dans la lecture de paragraphes ou de phrases et de mots isolés (vocalisés, partiellement vocalisés et non vocalisés) indiquent que la présence des voyelles facilite l'identification des mots aussi bien chez les bons lecteurs que les faibles lecteurs (Abu Rabia, 1997b). Cependant, ces voyelles, bien qu'utiles pour réduire l'ambiguïté phonologique, peuvent exiger un traitement visuo-attentionnel plus important. C'est pourquoi elles peuvent perturber la fluidité de la lecture même chez les lecteurs expérimentés qui n'ont plus besoin de la présence de voyelles pour identifier correctement le mot (Saiegh-Haddad & Schiff, 2016). Ce constat a été démontré dans une étude de Ibrahim (2013) menée auprès d'élèves de 8^{ème} année. Dans une tâche de lecture de mots, les résultats ont montré que les mots non voyellés étaient lus plus rapidement que les mots voyellés, les mots voyellés requérant une charge perceptuelle additionnelle. Al-Samarraie, Sarsam, Alzahrani et Alalwan (2020), avec des données EEG collectées auprès d'universitaires lors d'une tâche de lecture de textes dans deux conditions (avec et sans voyelles), ont montré un impact significatif des voyelles sur l'activité corticale, suggérant que les voyelles augmentent le niveau d'attention qu'un individu accorde au texte.

En l'absence de voyelles diacritiques (forme non vocalisée), les mots sont des homographes phonologiquement et sémantiquement ambigus, ce qui rend difficile l'apprentissage de la lecture en arabe du fait que la seule issue pour assurer une identification de mots correcte est de s'appuyer sur le contexte. En version non vocalisée, le lecteur ne peut pas par exemple identifier sans ambiguïté la flexion casuelle d'un verbe (actif ou passif) dans une phrase sans recourir aux éléments qui le suivent, ce qui se traduit par des mouvements oculaires plus importants sur les derniers éléments de la phrase (Roman, Pavard & Asselah, 1985). Cette nécessité de recourir au contexte pour lever certaines ambiguïtés pénalisent les faibles lecteurs (Ammar, 1997). Ceux-ci se retrouvent dans une situation de surcharge cognitive qui ne leur permet pas de traiter à la fois l'information orthographique émanant du squelette consonantique du mot et l'information contextuelle.

3.4. La perceptibilité des lettres en écriture arabe

Pour mémoire, comme indiqué en partie 2.1.2, l'écriture arabe est semi-cursive, 22 lettres s'écrivent reliées (ligaturées) entre elles dans le mot, et 6 lettres (ا و ز ر د) ne se relient ni entre elles, ni à la lettre qui les suit dans le mot, mais peuvent se relier à la lettre qui les précède. Chaque lettre est représentée par différentes formes (allographes) en fonction de la position qu'elle occupe dans le mot (initiale, médiane ou finale). Une grande similarité visuelle est observée entre les lettres du fait que la majorité d'entre elles partagent la même forme de base et ne se distinguent que par le nombre et la position des points diacritiques (e.g., ث ت ب /b, t, θ /). Ces caractéristiques font de l'arabe l'un des systèmes d'écriture les plus complexes, après le chinois, le japonais, le kannada et le coréen (Verhoeven & Perfetti, 2021).

Certaines études (e.g., Abdelhadi, Ibrahim & Eviatar, 2011 ; Eviatar & Ibrahim, 2004, 2007, 2014 ; Eviatar, Ibrahim & Ganayim, (2004) ; Ibrahim & Eviatar, 2009; Ibrahim, Eviatar & Aharon-Peretz, 2002) ont exploré l'effet de la complexité de l'écriture arabe sur la lecture, aussi bien chez l'adulte que chez l'apprenti-lecteur.

Dans une tâche de dénomination de trois lettres constituant un pseudo-mot, Eviatar et Ibrahim (2004) ont mesuré le temps d'affichage des stimuli à partir duquel un taux d'erreur de 50% était obtenu. Les participants étaient des étudiants arabophones, hébreophones ou anglophones. Les anglophones dépassaient 50% d'erreur lorsque les stimuli étaient présentés très rapidement (50% d'erreur à 35 ms) ; les hébreophones, lorsque les stimuli étaient

présentés deux fois plus longtemps (50% d'erreur à 70 ms) et les arabophones dépassaient 50% d'erreur lorsque les stimuli étaient présentés également deux fois plus longtemps (50% d'erreur à 140 ms). Dans une autre étude menée auprès d'étudiants bilingues arabophones et hébreophones (Eviatar et al., 2004), les auteurs ont présenté aux participants des paires de lettres en hébreu et en arabe afin de décider si elles étaient identiques ou non. Les temps de réaction ont montré que les réponses sur les paires de lettres en arabe étaient plus lentes comparées à celles sur les paires de lettres en hébreux (+30 ms).

L'étude de Ibrahim et al. (2002), chez des sujets bilingues (10^{ème} grade) avec l'arabe comme langue maternelle et l'hébreu comme langue seconde, a révélé qu'ils traitaient les lettres plus rapidement en hébreu qu'en arabe. D'autres résultats également de Abdelhadi, Ibrahim et Eviatar (2011) ont révélé que des appreni-lecteur bilingues de 3^{ème} et 6^{ème} grade détectaient les voyelles diacritiques dans des mots en hébreux plus rapidement que dans des mots en arabe. Les auteurs expliquent cette difficulté par la plus grande similarité visuelle des lettres en arabe qu'en hébreu, ce qui augmente la charge cognitive et ralentit les processus de reconnaissance des mots.

Dans la même veine, une étude de Dai, Ibrahim et Share (2013) auprès d'enfants de 3^{ème} grade a testé l'effet des points diacritiques sur l'apprentissage de la lecture. Les résultats ont montré que ces derniers ralentissaient le temps de lecture de plus d'une demi-seconde par mot. Pour les auteurs, il semble que les points diacritiques constituent une autre source de difficulté pour les jeunes lecteurs arabophones. On peut cependant relativiser cette conclusion en considérant que les signes diacritiques permettent un meilleur traitement de la structure phonologique des mots impliquée dans la lecture à voix haute. Le ralentissement observé peut donc être vu comme plus bénéfique que préjudiciable pour l'enfant car il lui permet de renforcer les représentations lexicales des mots.

Finalement, comme l'écriture arabe est semi-cursive, la structure interne du mot peut se présenter sous trois formes : une forme totalement connectée (e.g., سليم (sain)) ; une forme partiellement connectée (e.g., سائق (conducteur)) et une forme non connectée (e.g., أرز (riz)). Le lecteur se trouve par conséquent en présence d'espaces à l'intérieur même du mot qu'il lit, alors qu'ils n'ont pas le rôle de séparateur entre les mots mais qu'ils font partie inhérente de la structure orthographique des mots en arabe. Cette particularité a conduit à s'interroger sur son impact sur le traitement orthographique. Dans ce contexte, Abdelhadi et al. (2011) ont montré que des sujets de 3^{ème} et 6^{ème} grade (CE2 et 6^{ème} français) détectaient

plus rapidement les voyelles diacritiques qui portaient sur des mots avec des lettres connectées plutôt que non connectées. Chez des adultes, Taha, Ibrahim et Khateb (2013) ont examiné l'effet de la connectivité orthographique entre les lettres sur la vitesse de réponse électrophysiologique lors d'une tâche de décision lexicale de deux catégories de mots : avec des lettres totalement connectées ou partiellement connectées. Les résultats ont révélé des temps de réactions plus élevés et de faibles précisions de réponses pour les mots partiellement connectés, par rapport aux mots totalement connectés. Ces résultats montrent que la connectivité orthographique des lettres en arabe semble avoir un impact sur la reconnaissance des mots aussi bien chez les enfants que chez les adultes. Pour Abdelhadi et al. (2011), ce résultat pourrait refléter un simple effet de fréquence. En effet, la majorité des mots en arabe sont principalement composés de lettres complètement connectées, les mots composés de lettres non connectées étant plus rares.

4. Chapitre 4 : Synthèse théorique

Contrairement aux langues latines à construction morphologique linéaire (ajout d'un préfixe ou d'un suffixe à une base), l'arabe se caractérise par une construction morphologique non-linéaire, procédant par l'entrelacement d'un schème (composé soit d'une séquence de voyelles, soit d'une séquence de voyelles et de consonnes) à une racine, généralement composée de trois consonnes. L'entrelacement s'effectue par le placement des éléments du schème entre les consonnes de la racine, ce qui conduit à une relation non contiguë entre les unités morphémiques (racine/schème).

Dans cette perspective, l'idée largement soutenue par les chercheurs dans le domaine de la reconnaissance du mot écrit en arabe est que l'accès lexical nécessite prioritairement l'extraction des consonnes de la racine, qui joueraient un rôle important dans l'activation des représentations lexicales (Prunet et al., 2000). Comme nous l'avons vu, cette hypothèse fait pourtant controverse : l'effet facilitateur en amorçage morphologique (par la racine) étant observé dans certains travaux (Boudelaa & Marslen-Wilson, 2000, 2005 ; Grainger et al., 2003) mais pas dans d'autres (Abu-Rabia & Awwad, 2004), suggérant que les mots en arabe ne seraient pas cognitivement représentés en tant qu'unités morphémiques décomposables (racines et schèmes) mais en tant qu'unités orthographiques. Une telle hypothèse est confortée par les effets d'amorçage orthographique obtenus par Perea et al. (2014, 2018) avec des amorces créées à partir de substitution d'une seule lettre de la racine du mot de base (e.g., كخابة [kxa:bh] / كتابة [kita:bah] (*écriture*) ; la racine étant [ktb]).

Pour ce qui concerne le traitement orthographique, la plupart des auteurs s'accordent sur le fait que la lettre représente l'unité de base de la perception des mots écrits, y compris chez l'apprenti-lecteur (e.g., McClelland & Rumelhart, 1981 ; Grainger, Lété, Bertrand, Dufau, & Ziegler, 2012). Ils avancent l'hypothèse d'une première étape de codage visuel des traits impliquant que la reconnaissance d'une lettre dépend du nombre de traits visuels la constituant, du nombre de traits partagés avec d'autres lettres, et de la présence de traits distinctifs (e.g., "I" et "L" ; "O" et "Q"). En arabe, la présence de traits partagés entre les lettres est plus systématique qu'en écriture à alphabet latin étant donné leurs similitudes visuelles (e.g., ت – ث – ب). Néanmoins, les données issues des études de Perea et al. (2016, 2018) ont révélé que les lettres à forme de base identique étaient traitées différemment, suggérant une sensibilité du lecteur aux traits distinctifs (points diacritiques) malgré la similarité visuelle

entre les lettres. D'un autre côté, les études en écriture à alphabet latin portant sur la nature de la représentation des lettres ont manipulé la casse (minuscule/MAJUSCULE) dans des tâches *Same-Different Matching* avec amorçage masqué (e.g., Arguin & Bub, 1995 ; Bowers, Vigliocco & Haan, 1998 ; Kinoshita & Kaplan, 2008 ; Ziegler, Ferrand, Jacobs, Rey, & Grainger, 2000). À l'instar de ces études, Carreiras et al. (2012) ont manipulé des amorces et des cibles présentant différentes formes dépendant de leur position dans le mot (médiane/finale), du fait que le système alphabétique arabe ne comporte pas de minuscules et de majuscules mais des allographes (différentes graphies de la même lettre). Ils ont observé un effet d'amorçage équivalent pour des paires d'allographes visuellement similaires (e.g., ف / ف [f]) et des paires visuellement dissimilaires (e.g., ك / ك [k]). Ces données laissent donc penser que le traitement orthographique en écriture latine et arabe suit le même principe, avec une première étape de codage des traits visuels suivie d'une étape d'identification abstraite des lettres.

Le traitement orthographique ne repose pas uniquement sur le codage des traits et de l'identité des lettres mais également sur leur codage positionnel. Nous avons vu que les effets d'amorçage orthographique obtenus en manipulant la position des lettres partagées par l'amorce et la cible (e.g., Acha & Perea, 2008 ; Peressotti & Grainger, 1999 ; Schoonbaert & Grainger, 2004) ont conduit à nuancer les conceptions initiales des modèles de traitement orthographique à codage en position absolue (e.g., McClelland & Rumelhart, 1981), ouvrant la voie vers des modèles plus flexibles comme le modèle de Grainger et Van-Heuven (2003) qui propose un codage relatif des lettres en terme de bigrammes ouverts. Le rôle des bigrammes dans le traitement orthographique a été confirmé en paradigme *Flanking letter lexical decision* en lecture experte (Dare & Schillcock, 2013 ; Grainger et al., 2014) et également chez les apprenti-lecteurs (Snell et al., 2020). Le paradigme a permis également de montrer un biais attentionnel lié au sens de la lecture, semblable à celui observé en situation naturelle de lecture (Snell et Grainger, 2018).

Pour leur part, les données rapportées en arabe sur le codage positionnel des lettres (Boudelaa et al., 2019 ; Hermena et al., 2021 ; Perea et al., 2010) ne confirment que partiellement l'idée d'un codage relatif vu l'inconstance de l'effet de transposition des lettres de la racine dépendant de la tâche utilisée. Toutefois, l'effet de transposition impliquant une lettre de la racine et une lettre du schème reste robuste (Perea et al., 2010).

Le premier objectif de ce travail sera précisément de compléter ces travaux sur le codage de la position des lettres en arabe en utilisant le paradigme FLLD. Il s'agira en

particulier de tester le modèle de Grainger et Van-Heuven (2003) en arabe et de vérifier si on peut observer un codage de la position des lettres en bigrammes ouverts.

Les données recueillies auprès d'enfants permettent de souligner deux aspects majeurs distinguant l'apprentissage de l'arabe des langues latines. Le premier aspect est lié à la diglossie. La situation d'apprentissage de la lecture en arabe standard ne s'inscrit pas, comme en langues latines, dans la continuité des connaissances phonologiques acquises en langue maternelle (l'arabe dialectal), étant donné les différences notables entre les deux registres (dialectal/standard) sur le plan phonologique, lexical et syntaxique. L'apprentissage de l'arabe standard est à ce titre considéré comme un apprentissage d'une seconde langue (e.g., Saiegh-Haddad, 2004, 2005).

Le second aspect est lié aux caractéristiques orthographiques de l'arabe. L'arabe possède deux formes d'écriture : une forme notant les signes diacritiques (forme voyellée) et une forme ne notant que le squelette consonantique (forme non voyellée). La forme voyellée est transparente, elle est destinée à l'enfant en phase initiale d'apprentissage. L'importance des signes diacritiques dans la facilitation de l'identification des mots chez l'apprenti-lecteur a été confirmée dans une série d'expériences (Abu Rabia, 1996, 1997a, 1997b, 1998, 2001 ; Abu Rabia & Siegel, 1995). Progressivement, l'apprenti-lecteur s'habitue à lire dans la forme non voyellée. L'écriture arabe dans sa forme non voyellée est considérée comme l'une des plus opaques, en raison de l'ambiguïté phonologique qu'elle présente : selon le contexte phrastique, un squelette consonantique peut se lire différemment, ce qui fait que le recours au contexte est la principale caractéristique de la lecture en arabe.

La complexité de l'orthographe arabe est considérée également comme étant une caractéristique pénalisant le traitement perceptif chez l'apprenti-lecteur. Les études comparant souvent l'arabe à l'hébreu ou à d'autres langues indo-européennes comme l'anglais (e.g., Abdelhadi, Ibrahim & Eviatar, 2011 ; Eviatar & Ibrahim, 2004, 2007, 2014 ; Eviatar, Ibrahim & Ganayim, 2004 ; Ibrahim & Eviatar, 2009 ; Ibrahim, Eviatar & Aharon-Peretz, 2002) ont ainsi mis en évidence des différences notoires dans la rapidité d'apprentissage de la lecture. En raison de ces particularités orthographiques, l'arabe fait l'objet d'un apprentissage de la lecture plus lent et plus coûteux. Néanmoins, ces données recueillies en langue arabe auprès d'apprenti-lecteurs semblent négliger les aspects orthographiques, alors que cet axe de recherche présente un enjeu majeur dans la compréhension des mécanismes

cognitifs chez l'apprenti-lecteur, notamment dans une perspective développementale, permettant ainsi une meilleure appréhension de la mise en place progressive de l'expertise en lecture en écriture arabe. Cela sera précisément le deuxième objectif de ce travail de thèse.

Enfin, pour ce qui est de la question relative aux traitements visuo-attentionnels et la direction de la lecture en arabe, les données de la littérature révèlent un biais attentionnel à gauche (dans le sens de la lecture) en situation de lecture de phrases (Jordan et al., 2014). Cependant, les études sur cette question restent trop peu nombreuses pour pouvoir émettre une conclusion avérée. Le dernier objectif de ce travail sera donc de compléter les données existantes à l'aide du paradigme FLLD pour étudier l'effet de la direction de la lecture sur l'intégration spatiale de l'information orthographique lors de la lecture de mots en langue arabe.

5. Chapitre 5 : Expérience 1. Étude de l'effet des flankers chez l'adulte

5.1. Objectif

À l'aide du paradigme FLLD, l'objectif de cette expérience était d'étudier l'effet de la présentation de flankers bigrammes sur le codage orthographique d'un mot cible chez l'adulte arabophone. Il s'agissait en particulier de vérifier si l'effet de la permutation des bigrammes observé dans les écritures latines était avéré également en arabe (i.e., RO ROCK CK = CK ROCK RO comparés à DA ROCK SH ; les deux bigrammes non reliés constituant un autre mot : DASH).

5.2. Méthode

5.2.1. Participants

Cent-soixante participants ont été testés pour cette première expérience issus de deux échantillons :

- ✓ Un groupe de 80 élèves du lycée public Alkotrobi de la ville de Ouazzane (région de Tanger au nord-ouest du Maroc) de niveaux 1^{ère} année (N = 43 ; 31 garçons) et 2^{ème} année (N = 37 ; 32 garçons) du cycle secondaire qualifiant (respectivement 2 ans et 1 an avant le Baccalauréat) de moyenne d'âge de 17 ans 3 mois (étendue = 14 ans 11mois à 24 ans ; écart-type = 1 an 6 mois) ;
- ✓ Un groupe de 80 étudiants de la faculté des Lettres et Sciences Humaines de Dhar-El-Mehrez de Fès de niveaux Master (N = 55 ; 29 garçons) et Licence (N = 25 ; 10 garçons) de moyenne d'âge de 26 ans 11 mois (étendue = 18 ans 4 mois à 48 ans 5 mois ; écart-type = 6 ans 2 mois).

Tous les participants avaient une vue normale ou corrigée, étaient natifs du Maroc et avaient appris à lire la langue arabe dès la première année du primaire.

5.2.2. Matériel

Le matériel était constitué de deux ensembles d'items :

- ✓ Un ensemble d'**items cibles** constitué de 112 mots et 112 pseudo-mots de 4 lettres toutes différentes (aucune lettre répétée) et sans signes diacritiques ;
- ✓ Pour créer les flankers en condition Bigrammes Différents, un ensemble d'**items flankers** constitué de 112 mots et 112 pseudo-mots de 4 lettres, constitué selon

les mêmes critères que les items cibles (quatre lettres non répétées et sans signes diacritiques).

Chaque item flanker a été apparié à un item cible de façon à ce qu'aucune lettre ne soit partagée entre l'item cible et l'item flanker.

Les mots cibles et les mots flankers ont été sélectionnés à partir de la base de données lexicales ARALEX (Boudelaa & Marslen-Wilson, 2010) en fonction de leur fréquence lexicale et de la fréquence de leur bigramme initial, interne et final qui a également été contrôlée entre les quatre catégories d'items (i.e., items cibles et items flankers pour les mots et les pseudo-mots). Certains items ont été pré-sélectionnés à partir du matériel de Perea et al. (2018).

Les statistiques détaillées du matériel sont présentées au Tableau 6. La totalité du matériel est présentée en Annexe.

Tableau 6. Statistiques du matériel expérimental de l'Expérience 1.

Type d'items	N	Sélection Perea & al., 2018/ARALEX)	Fréquence lexicale ^a	Fréquence du bigramme initial ^b	Fréquence du bigramme interne ^b	Fréquence du bigramme final ^b
Mots cibles	112	72 / 40	34 [0.03 – 275.3]	3.1	3.2	3.0
Mots flankers	112	72 / 40	8.3 [0.03 – 200.7]	2.9	3.2	3.1
Pseudo-mots cibles	112	112 / 00	-	2.8	3.1	2.9
Pseudo-mots flankers	112	112 / 00	-	2.7	3.1	2.9

Notes.^a moyenne de la fréquence par million et étendue ; ^b moyenne du log de la fréquence par million.

Quatre types de flankers ont été manipulés (cf. Tableau 7). Dans la condition Cible isolée (ci-après CIB), la cible était présentée isolément sans flankers. Dans la condition Bigrammes adjacents (ADJ), la cible était entourée de ses deux bigrammes constitutifs, initial et final, en conservant leur ordre droite-gauche (43 21). Dans la condition Bigrammes inversés (INV), la cible était entourée de ses deux bigrammes constitutifs, initial et final, mais leur ordre droite-gauche était inversé (21 43). Dans la condition Bigrammes différents (DIF), la cible était entourée des deux bigrammes de son flanker différent. L'ordre droite-gauche des bigrammes était conservé (87 65).

Il est à noter que la forme visuelle des lettres dans les bigrammes a été maintenue telle qu'elle apparaissait dans l'item cible dans les conditions ADJ et INV, et telle qu'elle apparaissait dans l'item flanker dans la condition DIF.

Tableau 7. Conditions expérimentales de présentation des flankers pour l'Expérience 1.

CIBLES	Cible isolée (CIB)	FLANKERS																		
		Bigrammes adjacents (ADJ)			Bigrammes inversés (INV)			Bigrammes différents (DIF)												
Mot	خيار 4321	ار	خيار	ار	خيار	ار	خيار	ار	خيار	ار	خيار	ار	خيار	ار	خيار	ار	خيار	ار	خيار	ار
Pseudo-mot	مئس 4321	س	مئس	س	مئس	س	مئس	س	مئس	س	مئس	س	مئس	س	مئس	س	مئس	س	مئس	س

5.2.3. Procédure

La présentation des stimuli et l'enregistrement des temps de réaction et des réponses étaient pilotés par le logiciel DMDX (Forster & Forster, 2003) sous Windows XP, sur des portables FUJITSU SIEMENS (écran 15") avec une résolution d'écran de 1280*800 pixels et un taux de rafraîchissement de 60 Hz.

Les 896 stimuli ((112 mots cibles + 112 pseudo-mots cibles) x 4 Type de Flankers) ont été présentés en format image bitmap de 1284 x 793 pixels. Les stimuli étaient écrits en blanc sur fond noir en police Courier New 17.5 ppp. À 50 cm de distance, le stimuli (la cible et ses deux flankers) correspondait à 7° d'angle visuel.

Quatre listes de passation ont été constituées de façon à ce que chaque cible soit présentée dans chaque condition de présentation à travers les participants.

Le sujet était assis devant un écran d'ordinateur à une distance de 50 cm. Chaque essai débutait par la présentation d'une croix de fixation durant 800 ms au centre de l'écran, remplacée ensuite par le stimulus qui restait affiché jusqu'à la réponse du participant. Une durée maximale de 3000 ms était permise pour répondre avant que l'item suivant soit présenté. Il était indiqué au participant qu'il devait focaliser son attention sur la suite de lettres au milieu de l'écran (entre les deux espaces) sans se préoccuper des autres lettres qui

l'entouraient. Il devait décider le plus rapidement et le plus exactement possible si cette suite de lettres constituait un mot en appuyant sur la touche *Shift droite* ou un pseudo-mot en appuyant sur la touche *Shift gauche* (les touches étaient inversées pour les participants gauchers de façon à ce que la réponse "mot" corresponde à la main dominante du participant). Les deux touches étaient marquées par des pastilles adhésives de couleur rouge. Aucun feedback positif ou négatif était affiché après la réponse du participant.

Les 224 stimuli étaient présentés en 8 blocs expérimentaux de 28 stimuli (14 mots et 14 pseudo-mots), avec des pauses entre chaque bloc. L'ordre des blocs à l'intérieur des listes et l'ordre des stimuli à l'intérieur des blocs étaient aléatorisés. Les blocs expérimentaux étaient précédés d'un bloc d'entraînement de 20 essais (10 mots et 10 pseudo-mots) différents de ceux utilisés dans la session expérimentale, permettant ainsi la familiarisation des participants avec le matériel. Au total, la passation durait 15 minutes.

5.3. Résultats

Les temps de réaction (TR) aberrants inférieurs à 300 ms et supérieurs à 3000 ms ont d'abord été retirés des analyses. Puis un score z a été calculé pour chaque TR basé sur la moyenne et l'écart-type des TR de chaque participant sur chaque type de cible (mot vs pseudo-mot). Les TR avec un score z inférieur ou supérieur à 2.5 écarts-types (outliers) ont été retirés des analyses. Au total, ces deux critères ont écarté 3% des TR sur les mots (3% pour les conditions CIB et INV ; 2% pour la condition ADJ ; 4% pour la condition DIF) ; et 3% sur les pseudo-mots (2% pour les conditions CIB et ADJ ; et 3% pour les conditions INV et DIF).

Les pourcentages d'erreur ont été calculés sur les essais restants.

Les analyses statistiques ont été réalisées sous JASP. Pour chaque type de cible (mot et pseudo-mot), une ANOVA a été conduite sur les TR et les pourcentages d'erreurs en prenant, d'une part, les participants comme facteur aléatoire (F_1) selon le plan S_{160} * Type de Flankers₄ et, d'autre part, les items comme facteur aléatoire (F_2) selon le plan I_{112} * Type de Flankers₄.

5.3.1. Analyse sur les mots

Les résultats sur les TR et les pourcentages d'erreurs apparaissent au Tableau 8.

Tableau 8. Moyennes des TR et des pourcentages d'erreurs (écarts types entre parenthèses) selon le type de flankers pour les cibles mots.

Conditions	Cible Isolée	Bigrammes adjacents	Bigrammes inversés	Bigrammes différents
TR	749 (145)	767 (145)	772 (152)	786 (148)
% d'erreurs	5 (5)	5 (6)	6 (6)	6 (8)

L'ANOVA à mesures répétées, effectuée sur les TR, a révélé un effet significatif du type de flankers, $F_1(3, 477) = 27.98, p < .001$; $F_2(3, 333) = 19.26, p < .001$. Les comparaisons post-hoc effectuées sur les participants avec correction de Holm (cf. Figure 12) ont montré que la condition Cible isolée ($M = 749, SD = 145$) était significativement plus rapide que les trois autres conditions : Bigrammes adjacents ($M = 767, SD = 145$), $t(159) = -4.98, p < .001$; Bigrammes inversés ($M = 772, SD = 152$), $t(159) = -5.72, p < .001$; Bigrammes différents ($M = 786, SD = 148$), $t(159) = -8.97, p < .001$. Les comparaisons indiquaient également des TR significativement plus rapides pour la condition Bigrammes adjacents ($M = 767, SD = 145$) par rapport à la condition Bigrammes différents ($M = 786, SD = 148$), $t(159) = -4.56, p < .001$; et des TR significativement plus rapides pour la condition Bigrammes inversés ($M = 772, SD = 152$) par rapport à la condition Bigrammes différents ($M = 786, SD = 148$), $t(159) = -3.34, p < .01$. Aucune différence significative n'était observée entre les deux conditions Bigrammes adjacents ($M = 767, SD = 145$) et Bigrammes inversés ($M = 772, SD = 152$), $t(159) = -1.20, p = .23$.

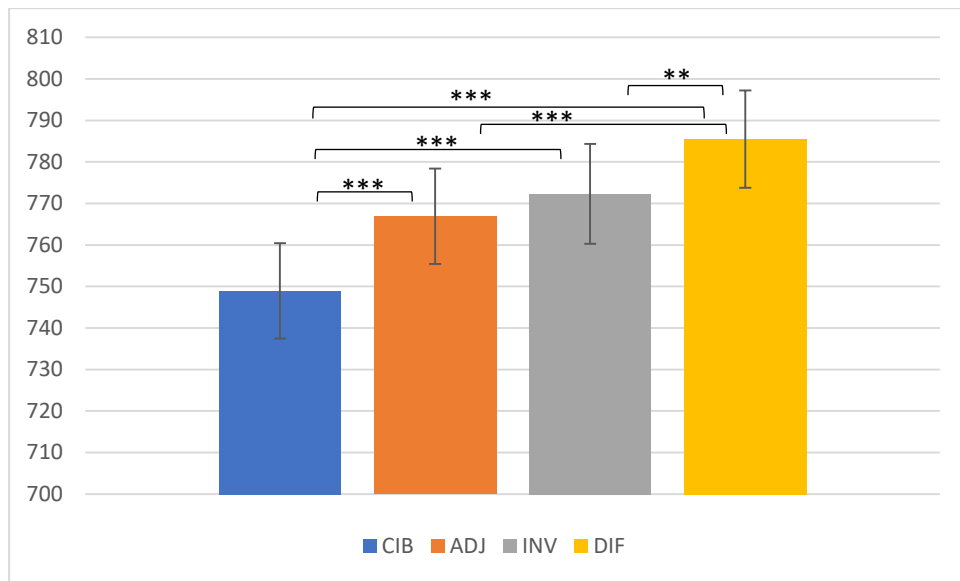


Figure 12. Moyennes des TR en ms selon le type de flankers. Les barres d'erreurs représentent l'erreur standard. (** $p < .01$ – *** $p < .001$)

L'ANOVA effectuée sur les pourcentages d'erreurs a révélé un effet principal du type de flankers avec des pourcentages d'erreurs plus importants pour les conditions Bigrammes Inversés et Bigrammes Différents, $F_1(3, 477) = 4.39, p < .01$; $F_2(3, 333) = 7.08, p < .001$.

5.3.2. Analyse sur les pseudo-mots

Les résultats sur les TR et les pourcentages d'erreurs apparaissent au Tableau 9.

Tableau 9. Moyennes des TR et des pourcentages d'erreurs (écarts types entre parenthèses) selon le type de flankers pour les cibles pseudo-mots.

Conditions	Cible Isolée	Bigrammes Adjacents	Bigrammes Inversés	Bigrammes Différents
TR	979 (262)	1013 (280)	1013 (272)	1012 (281)
% d'erreurs	10 (9)	9 (9)	9 (9)	8 (8)

Sur les pseudo-mots, l'ANOVA à mesures répétées a révélé un effet significatif du type de flankers, $F_1(3, 477) = 15.80, p < .001$; $F_2(3, 333) = 7.86, p < .001$. Les tests post-hoc effectués sur les participants avec correction de Holm (cf. Figure 13) ont montré que la condition Cible isolée ($M = 979, SD = 262$) était significativement plus rapide que les trois autres conditions : Bigrammes adjacents ($M = 1013, SD = 280$), $t(159) = -5.24, p < .001$; Bigrammes inversés ($M = 1013, SD = 272$), $t(159) = -6.17, p < .001$; Bigrammes différents ($M = 1012,$

$SD = 281$), $t(159) = -5.58$, $p < .001$. En revanche aucune différence significative n'était observée entre les conditions Bigrammes adjacents ($M = 1013$, $SD = 280$) et Bigrammes inversés ($M = 1013$, $SD = 272$), $t(159) = -0.09$, $p = 1.00$; Bigrammes adjacents ($M = 1013$, $SD = 280$) et Bigrammes différents ($M = 1012$, $SD = 281$), $t(159) = 0.16$, $p = 1.00$; Bigrammes inversés ($M = 1013$, $SD = 272$) et Bigrammes différents ($M = 1012$, $SD = 281$), $t(159) = 0.28$, $p = 1.00$.

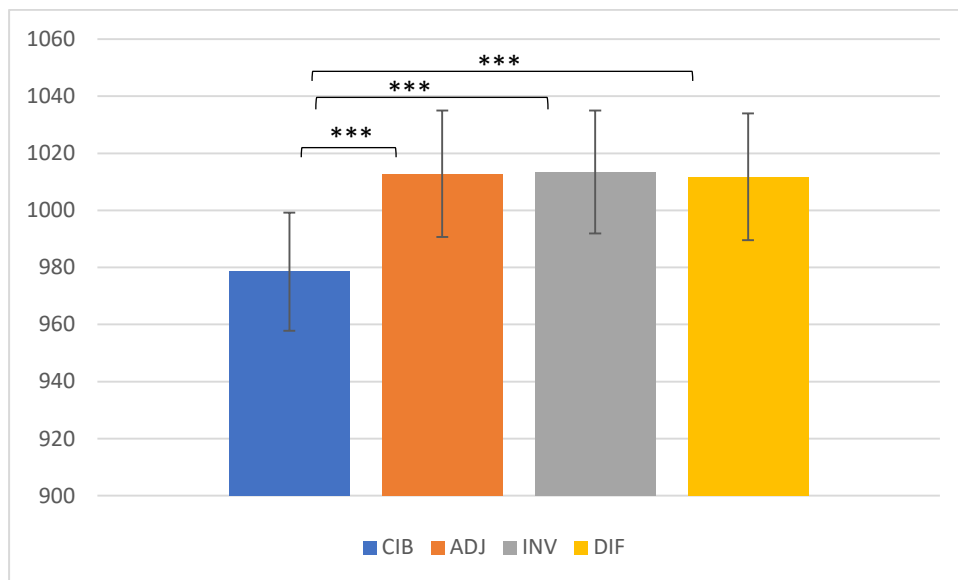


Figure 13. Moyennes des TR en ms selon le type de flankers. Les barres d'erreurs représentent l'erreur standard. (***) $p < .001$

L'ANOVA effectuée sur les pourcentages d'erreurs a révélé un effet marginal du type de flankers, $F_1(3, 477) = 2.35$, $p = .07$; $F_2(3, 333) = 2.36$, $p = .07$.

5.4. Synthèse de l'expérience 1

Dans leur étude, Dare et Shillcock (2013) ont introduit un nouveau paradigme nommé *Flanking Letter Lexical Decision* (FLLD), dans lequel la cible est entourée à gauche et à droite de bigrammes. Comparativement à une condition sans recouvrement orthographique, ils ont montré que la décision lexicale était facilitée quand les bigrammes étaient reliés orthographiquement à la cible, quel que soit leur ordre de présentation. Ainsi, la reconnaissance de la cible [ROCK] était autant facilitée par les bigrammes adjacents [RO ROCK CK] que par les mêmes bigrammes mais inversés [CK ROCK RO], comparativement aux bigrammes [DA ROCK SH]. Des effets similaires ont été obtenus par Grainger et al. (2014). Les résultats de ces études ont étayé le modèle de traitement orthographique de Grainger et Van

Heuven (2003), impliquant un codage contigu et non-contigu de combinaisons non ordonnées de paires de lettres, appelées "bigrammes ouverts".

Notre première expérience a été conduite afin de vérifier si cet effet de la permutation des bigrammes observé dans les écritures latines était avéré également en arabe. Nos résultats confirment bien cet effet dans cette écriture. Il apparaît avec beaucoup de robustesse uniquement sur les cibles mots, les décisions sur les pseudo-mots ne bénéficiant pas de l'information orthographique portée par les flankers adjacents. De plus, il apparaît que l'intégration de l'information orthographique pénalise le traitement de la cible, les cibles sans flankers étant reconnues plus rapidement que celles avec flankers.

6. Chapitre 6 : Expérience 2. Étude de l'effet des flankers chez l'apprenti-lecteur

6.1. Objectif

Toujours à l'aide du paradigme FLLD, l'objectif de cette expérience était d'étudier l'effet de la présentation de flankers bigrammes sur le codage orthographique d'un mot cible chez des apprentis-lecteurs arabophones. Du fait que l'effet était avéré chez l'adulte arabophone (Expérience 1), il s'agissait en particulier de révéler à quel moment de l'apprentissage de la lecture cette expertise se met en place. Snell et al. (2020) ont montré chez des élèves français du CP à la 6^{ème} des effets significatifs des flankers dès la fin de la première année d'apprentissage. On peut suspecter que, compte-tenu des contraintes imposées par le système orthographique arabe recensées au Chapitre 3, qui rendent l'apprentissage de la lecture plus difficile qu'en français, cet effet soit retrouvé plus tardivement chez l'élève arabophone.

6.2. Méthode

6.2.1. Participants

Trois groupes de participants issus de trois écoles primaires publiques de la ville de Ouazzane (École Abi Hayan Attawhidi ; École Abou Moussa Alachaari ; École Moulay Elmahdi Louazzani), ont été testés pour cette deuxième expérience :

- ✓ Un groupe de 60 élèves de niveau 2^{ème} année primaire (25 garçons) de moyenne d'âge de 7 ans 3 mois (étendue = 6 ans 1 mois à 9 ans 5 mois ; écart-type = 7 mois) ;
- ✓ Un groupe de 60 élèves de niveau 4^{ème} année primaire (38 garçons) de moyenne d'âge de 9 ans 7 mois (étendue = 8 ans 1 mois à 12 ans 1 mois ; écart-type = 8 mois) ;
- ✓ Un groupe de 64 élèves de niveau 6^{ème} année primaire (33 garçons) de moyenne d'âge de 11 ans 5 mois (étendue = 9 ans 1 mois à 14 ans ; écart-type = 8 mois).

Tous les participants avaient une vue normale ou corrigée, étaient natifs du Maroc et avaient appris à lire la langue arabe dès la première année du primaire.

6.2.2. Matériel

Le matériel était le même que celui utilisé pour l'expérience 1. Il était adapté à une population d'élèves de primaire du fait que la sélection des cibles/ flankers a été faite sur la base de la fréquence lexicale : les mots plus fréquents ont été placés en cible, et les mots moins fréquents en flankers différents.

6.2.3. Procédure

La même procédure que celle de l'expérience 1 a été utilisée.

6.3. Résultats

Les temps de réaction (TR) aberrants inférieurs à 300 ms et supérieurs à 3000 ms ont d'abord été retirés des analyses. Puis un score z a été calculé pour chaque TR basé sur la moyenne et l'écart-type des TR de chaque participant sur chaque type de cible (mot vs pseudo-mot). Les TR avec un score z inférieur ou supérieur à 2.5 écarts-types (outliers) ont été retirés des analyses. Au total, ces deux critères ont écarté :

- ✓ Pour le niveau 2^{ème} année : 42% des TR sur les mots (42% pour les conditions CIB et INV ; 44% pour les conditions ADJ et DIF) et 51% sur les pseudo-mots (53% pour les conditions CIB et INV ; et 49% pour les conditions ADJ et DIF) ;
- ✓ Pour le niveau 4^{ème} année : 12% des TR sur les mots (11% pour la condition CIB ; 12% pour les conditions ADJ, INV et DIF) et 19% sur les pseudo-mots (18% pour les conditions CIB et ADJ ; 19% pour la condition INV ; 21% pour la condition DIF) ;
- ✓ Pour le niveau 6^{ème} année : 5% des TR sur les mots (3% pour la condition CIB ; 6% pour les conditions ADJ et INV ; 5% pour la condition DIF) et 7% sur les pseudo-mots (6% pour la condition CIB ; 7% pour les conditions ADJ et DIF ; 8% pour la condition INV).

Les pourcentages d'erreur ont été calculés sur les essais restants.

Les analyses statistiques ont été réalisées sous JASP. Pour chaque type de cible (mot et pseudo-mot), une ANOVA a été conduite sur les TR et les pourcentages d'erreurs en prenant, d'une part, les participants comme facteur aléatoire (F_1) selon le plan $S_{60} * \text{Type de Flankers}_4$ et, d'autre part, les items comme facteur aléatoire (F_2) selon le plan $I_{112} * \text{Type de Flankers}_4$.

6.3.1. Analyse sur les mots

Les résultats sur les TR et les pourcentages d'erreurs selon le niveau d'étude apparaissent au Tableau 10.

Tableau 10. Moyennes des TR et des pourcentages d'erreurs (écarts types entre parenthèses) selon le niveau d'étude et le type de flankers pour les cibles mots.

Niveau		Cible isolée	Bigrammes adjacents	Bigrammes inversés	Bigrammes différents
2 ^{ème} année (CE1 français)	TR	1964 (395)	1964 (408)	1965 (432)	1935 (403)
	% d'erreurs	33 (22)	32 (22)	31 (23)	31 (22)
4 ^{ème} année (CM1 français)	TR	1551 (401)	1596 (409)	1575 (381)	1619 (374)
	% d'erreurs	19 (15)	17 (14)	19 (15)	19 (14)
6 ^{ème} année (6 ^{ème} français)	TR	1209 (323)	1236 (333)	1230 (324)	1270 (339)
	% d'erreurs	11 (9)	11 (8)	12 (9)	12 (10)

Dans le groupe de niveau 2^{ème} année, l'ANOVA à mesures répétées effectuée sur les TR n'a pas révélé d'effet significatif du type de flankers, F_1 et $F_2 < 1$ (cf. Figure 14).

Dans le groupe de niveau 4^{ème} année, l'ANOVA à mesures répétées a révélé un effet significatif du type de flankers, $F_1(3, 177) = 4.57, p = .004$; $F_2(3,333) = 5.33, p = .001$. Les comparaisons post-hoc avec correction de Holm effectuées sur les participants (cf. Figure 14) ont montré que la condition Cible isolée ($M = 1551, SD = 401$) était significativement plus rapide que la condition Bigrammes différents ($M = 1619, SD = 374$), $t(59) = -3.10, p = .02$. Un effet marginal était observé entre la condition Cible isolée ($M = 1551, SD = 401$) et la condition Bigrammes adjacents ($M = 1596, SD = 409$), $t(59) = -2.43, p = .09$. Cependant, aucune différence significative n'était observée entre la condition Cible isolée ($M = 1551, SD = 401$) et la condition Bigrammes inversés ($M = 1575, SD = 381$), $t(59) = -1.17, p = .49$. Aucune différence n'était observée également entre la condition Bigrammes adjacents ($M = 1596, SD = 409$) et les deux conditions : Bigrammes inversés ($M = 1575, SD = 381$), $t(59) = 1.16, p = .49$, et Bigrammes différents ($M = 1619, SD = 374$), $t(59) = -1.44, p = .45$, et aucune différence entre la condition Bigrammes inversés ($M = 1575, SD = 381$) et Bigrammes différents ($M = 1619, SD = 374$), $t(59) = -2.20, p = .12$.

Dans le groupe de niveau 6^{ème} année, l'ANOVA à mesures répétées a révélé un effet significatif du type de flankers, $F_1(3,189) = 7.26, p < .001$; $F_2(3,333) = 4.13, p = .007$. Les comparaisons post-hoc avec correction de Holm effectuées sur les participants (cf. Figure 14) ont montré que la condition Cible isolée ($M = 1209, SD = 323$) était significativement plus

rapide que la condition Bigrammes différents ($M = 1270, SD = 339$), $t(63) = -4.00, p = .001$. La condition Bigrammes inversés ($M = 1230, SD = 324$) également était significativement plus rapide que la condition Bigrammes différents ($M = 1270, SD = 339$), $t(63) = -2.86, p = .02$. Un effet marginal était observé entre la condition Cible isolée ($M = 1209, SD = 323$) et la condition Bigrammes adjacents ($M = 1236, SD = 333$), $t(63) = -2.31, p = .07$ et entre la condition Bigrammes adjacents ($M = 1236, SD = 333$) et la condition Bigrammes différents ($M = 1270, SD = 339$), $t(63) = -2.44, p = .07$. En revanche, aucune différence significative n'était observée entre la condition Cible isolée ($M = 1209, SD = 323$) et la condition Bigrammes inversés ($M = 1230, SD = 324$), $t(63) = -1.86, p = .13$ et entre la condition Bigrammes adjacents ($M = 1236, SD = 333$) et la condition Bigrammes inversés ($M = 1230, SD = 324$), $t(63) = 0.46, p = .64$.

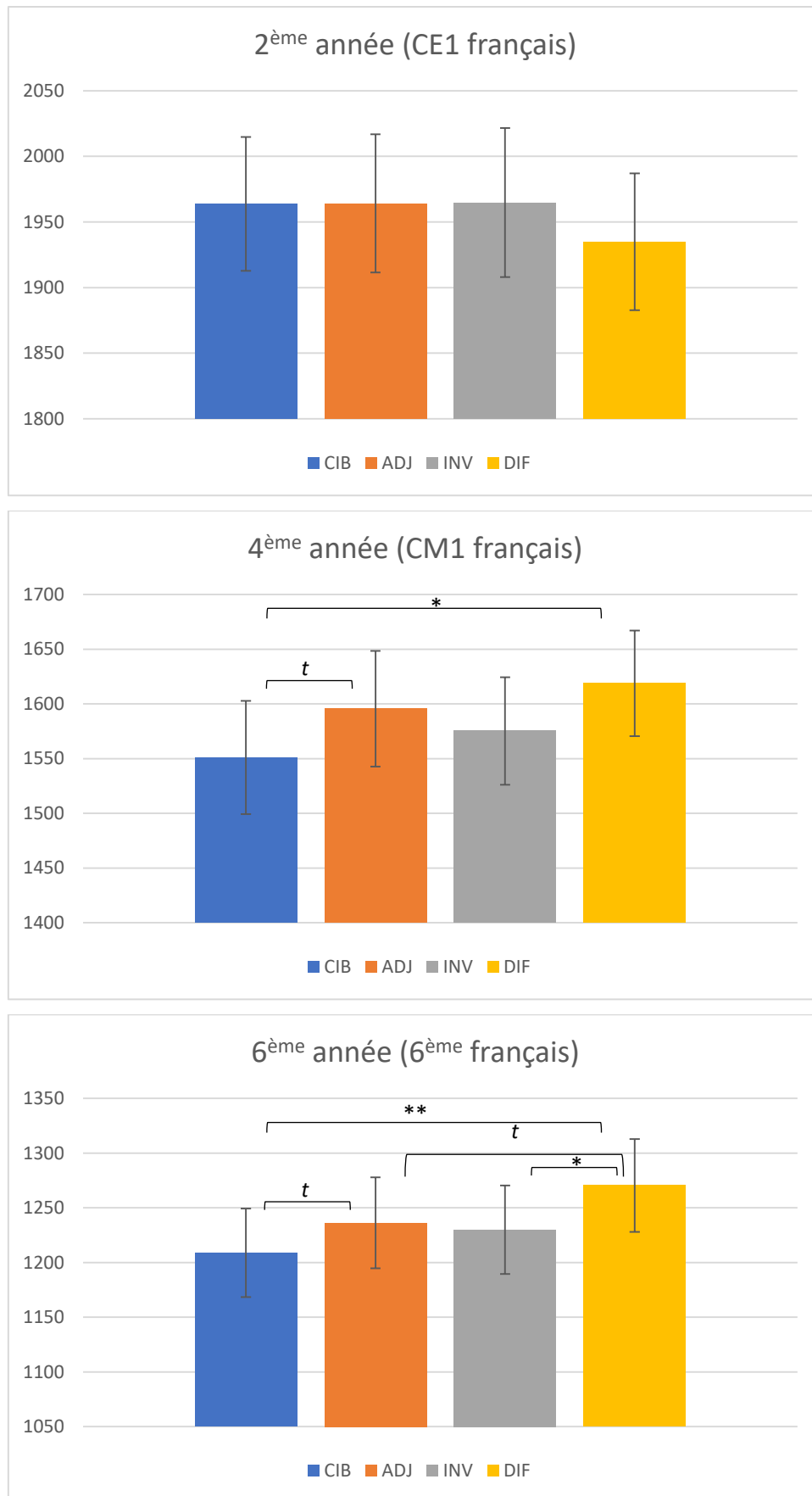


Figure 14. Moyennes des TR en ms selon le niveau d'étude et le type de flankers. Les barres d'erreurs représentent l'erreur standard. ($t = .05 < p < .10$ - * $p < .05$ - ** $p < .01$)

L'ANOVA effectuée sur les pourcentages d'erreurs n'a pas révélé d'effet significatif du type de flankers pour le niveau 2^{ème} année, $F_1 < 1$; $F_2(3, 333) = 1.43, p = .23$; le niveau 4^{ème} année, $F_1(3, 177) = 1.39, p = .24$; $F_2 < 1$ et le niveau 6^{ème} année, F_1 et $F_2 < 1$.

6.3.2. Analyse sur les pseudo-mots

Les résultats sur les TR et les pourcentages d'erreurs selon le niveau d'étude apparaissent au Tableau 11.

Tableau 11. Moyennes des TR et des pourcentages d'erreurs (écarts types entre parenthèses) selon le niveau d'étude et le type de flankers pour les cibles pseudo-mots.

Niveau		Cible isolée	Bigrammes adjacents	Bigrammes inversés	Bigrammes différents
2 ^{ème} année (CE1 français)	TR	2178 (446)	2110 (442)	2133 (471)	2052 (516)
	% d'erreurs	39 (30)	35 (25)	38 (28)	35 (28)
4 ^{ème} année (CM1 français)	TR	1949 (412)	1947 (397)	1984 (392)	1938 (388)
	% d'erreurs	25 (21)	22 (22)	25 (23)	23 (21)
6 ^{ème} année (6 ^{ème} français)	TR	1549 (403)	1595 (401)	1592 (429)	1580 (408)
	% d'erreurs	15 (14)	13 (14)	15 (15)	13 (12)

Sur les pseudo-mots, l'ANOVA à mesures répétées n'a pas révélé d'effet significatif du type de flankers pour le groupe de 2^{ème} année, F_1 et $F_2 < 1$, et le groupe de 4^{ème} année, $F_1(3, 177) = 1.59, p = .19$; $F_2(3, 333) = 1.78, p = .15$ (cf. Figure 15).

Pour le groupe de 6^{ème} année, l'ANOVA a révélé un effet significatif du type de flankers sur les participants, $F_1(3, 189) = 3.86, p = .01$, mais pas sur les items, $F_2(3,333) = 1.46, p = .22$. Les comparaisons post-hoc avec correction de Holm effectuées sur les participants (cf. Figure 15) ont montré que la condition Cible isolée ($M = 1549, SD = 403$) était significativement plus rapide que la condition Bigrammes adjacents ($M = 1595, SD = 401$), $t(63) = -3.22, p = .01$. Cependant, aucune différence n'était observée entre la condition Cible isolée ($M = 1549, SD = 403$) et les deux autres conditions : Bigrammes inversés ($M = 1592, SD = 429$), $t(63) = -2.36, p = .11$, et Bigrammes différents ($M = 1580, SD = 408$), $t(63) = -2.33, p = .11$. Aucune différence n'était observée également entre la condition Bigrammes adjacents

($M = 1595$, $SD = 401$) et les deux conditions Bigrammes inversés ($M = 1592$, $SD = 429$), $t(63) = 0.18$, $p = .92$, et Bigrammes différents ($M = 1580$, $SD = 408$), $t(63) = 1.14$, $p = .77$, et aucune différence entre la condition Bigrammes inversés ($M = 1592$, $SD = 429$) et Bigrammes différents ($M = 1580$, $SD = 408$), $t(63) = 0.74$, $p = .92$.

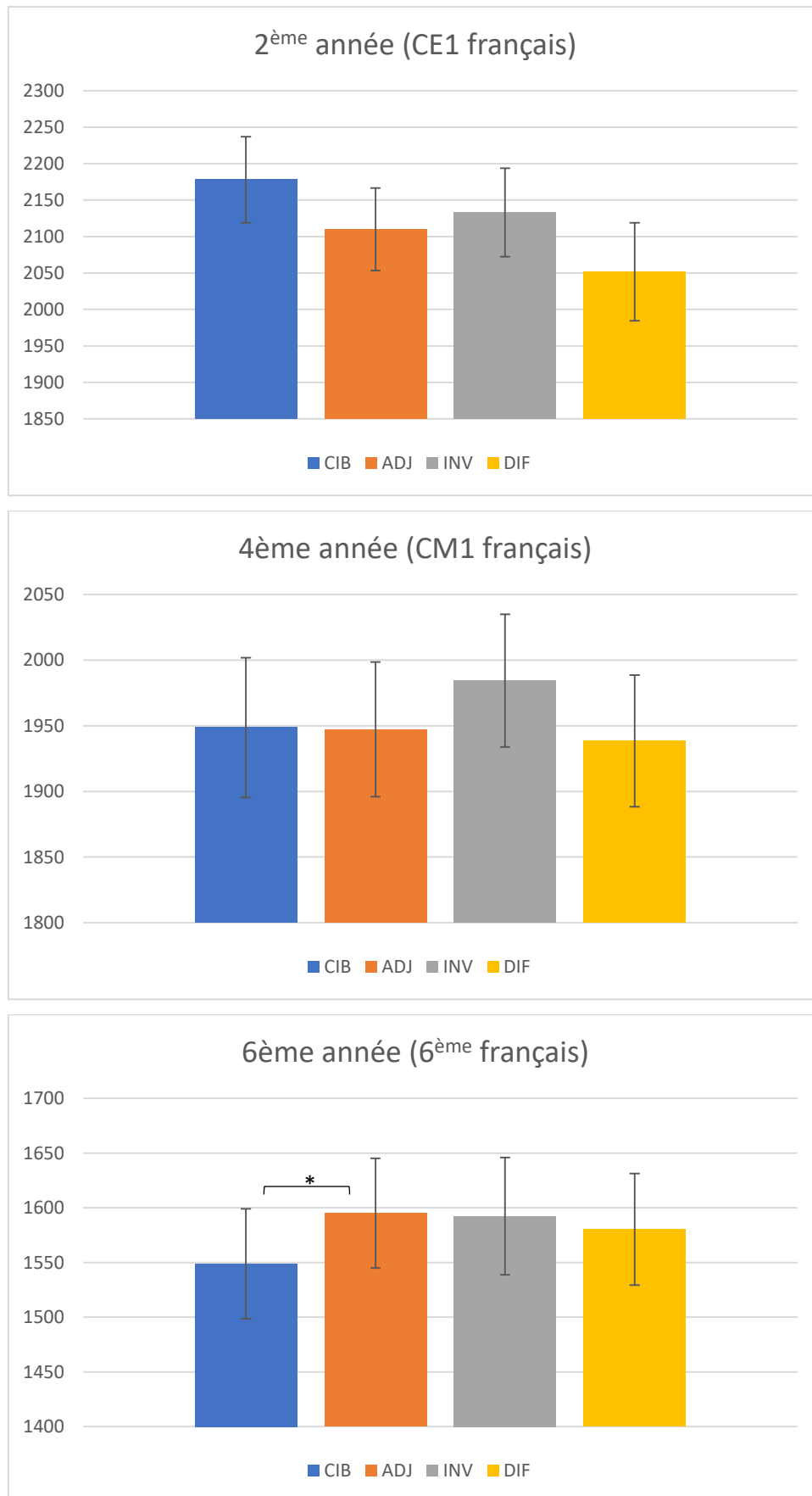


Figure 15. Moyennes des TR en ms selon le niveau d'étude et le type de flankers. Les barres d'erreurs représentent l'erreur standard. (* $p < .05$)

L'ANOVA effectuée sur les pourcentages d'erreurs n'a pas révélé d'effet significatif du type de flankers pour le niveau 2^{ème} année, F_1 et $F_2 < 1$; le niveau 4^{ème} année, $F_1(3, 177) = 1.23$, $p = .30$; $F_2 < 1$ et le niveau 6^{ème} année, $F_1(3, 189) = 1.54$, $p = .20$; $F_2(3, 333) = 1.25$, $p = .29$.

6.4. Synthèse de l'expérience 2

L'objectif de cette deuxième expérience était d'étudier avec le paradigme FLLD la trajectoire développementale du traitement orthographique en arabe chez des enfants d'école primaire de trois grade (2^{ème}, 4^{ème} et 6^{ème}), afin d'appréhender la transition d'un traitement séquentiel des lettres dans les mots chez les lecteurs débutants vers un traitement parallèle prenant en compte les informations orthographiques environnantes. On s'attendait à observer à un certain niveau chez l'apprenti-lecteur, un effet de facilitation orthographique des bigrammes reliés à la cible, quel que soit leur ordre, par rapport aux bigrammes différents.

Les données recueillies aux trois niveaux confirment l'hypothèse d'un traitement orthographique parallèle des lettres, se mettant en place progressivement au fur et à mesure du développement des aptitudes en lecture et des capacités visuo-attentionnelles.

Chez les apprentis-lecteurs de 2^{ème} niveau, la reconnaissance des mots cibles ne semble pas être sensible à la présence des bigrammes adjacents. Les temps de réaction restent équivalents dans toutes les conditions, mettant en avant l'idée d'un traitement purement séquentiel des mots, sans prise en compte des informations orthographiques environnantes.

Les apprentis-lecteurs de 4^{ème} niveau ne présentent pas un bénéfice pour les bigrammes reliés orthographiquement par rapport aux bigrammes différents. Néanmoins, ils reconnaissent plus rapidement les mots présentés isolément, comparativement aux mots entourés de bigrammes différents. Cet effet qui perturbe leur décision lexicale semble traduire les prémices d'une légère sensibilité aux informations environnantes.

Les apprentis-lecteurs de 6^{ème} niveau présentent des performances similaires à celles des apprentis-lecteurs de 4^{ème} niveau en ce qui concerne le traitement des mots présentés isolément. Néanmoins, ils manifestent une sensibilité plus importante aux types de bigrammes. On observe un patron de résultat quasiment similaire à celui observé chez l'adulte, mais celui-ci manque de robustesse à l'analyse statistique. Il existe un effet avéré pour les bigrammes inversés par rapport aux bigrammes différents. L'effet entre les bigrammes adjacents et les bigrammes différents est statistiquement marginal.

Nos données recueillies auprès d'apprenti-lecteurs arabophones permettent d'avancer que le traitement parallèle des informations orthographiques est un processus très lent en langue arabe qui ne commence à se manifester qu'à un stade avancé de l'apprentissage de la lecture (6^{ème} niveau seulement), comparativement aux langues latines où il se met en place assez précocement (Khelifi et al., 2012, 2015 ; Reichle et al., 2013 ; Snell et al., 2020).

7. Chapitre 7 : Étude de l'effet de la direction de la lecture sur le traitement des mots chez l'adulte

7.1. Objectif

L'objectif de ces deux expériences était d'étudier l'effet de la direction de la lecture sur les traitements visuo-attentionnels chez l'adulte, à l'aide du paradigme FLLD. Étant donné le sens de la lecture en langue arabe (de droite à gauche), on s'attendait à observer un biais attentionnel à gauche, à l'opposé de celui observé en français par Snell et Grainger (2018). Une première expérience a été menée en présentation unilatérale de bigrammes. Devant l'absence de résultats probants, une seconde expérience a été menée cette fois-ci en présentation bilatérale de mots.

7.2. Expérience 3 : Direction de la lecture – Présentation unilatérale de bigrammes

7.2.1. Méthode

7.2.1.1. Participants

L'échantillon était composé de 80 élèves du lycée public Alkotrobi de la ville de Ouazzane (région de Tanger au nord-ouest du Maroc) de niveaux 2^{ème} année (N = 41 ; 29 garçons) et 3^{ème} année (N = 39 ; 34 garçons) du cycle secondaire qualifiant (respectivement le Baccalauréat, et 1 an avant le Baccalauréat) de moyenne d'âge de 17 ans 4 mois (étendue = 16 ans 1 mois à 20 ans ; écart-type = 1 an 7 mois).

Tous les participants avaient une vue normale ou corrigée, étaient natifs du Maroc et avaient appris à lire la langue arabe dès la première année du primaire.

7.2.1.2. Matériel

Le matériel était le même que celui utilisé pour l'expérience 1. Les images des deux conditions (Bigrammes adjacents et Bigrammes inversés) de l'expérience 1 ont été retraitées sous Python de façon à garder soit le bigramme de gauche soit celui de droite afin de créer les quatre types de flankers en présentation unilatérale (cf. Tableau 12). Dans la condition Bigramme adjacent gauche (ci-après ADJ gauche), le bigramme final de la cible était présenté

à gauche ; dans la condition Bigramme adjacent droite (ADJ droite) le bigramme initial de la cible était présenté à droite ; dans la condition Bigramme inversé gauche (INV gauche) le bigramme initial de la cible était présenté à gauche ; dans la condition Bigramme inversé droite (INV droite) le bigramme final de la cible était présenté à droite.

Tableau 12. Conditions expérimentales de présentation unilatérale des flankers pour l'expérience 3.

		CIBLES	
		Mot	Pseudo-mot
FLANKERS	Bigramme adjacent gauche (ADJ gauche)	خيار ار 43 4321	مائس ئس 43 4321
	Bigramme adjacent droite (ADJ droite)	خيد خيار 4321 21	ما مائس 4321 21
	Bigramme inversé gauche (INV gauche)	خيار خيد 21 4321	مائس ما 21 4321
	Bigramme inversé droite (INV droite)	ار خيار 4321 43	ئس مائس 4321 43

7.2.1.3. Procédure

La même procédure que celle de l'expérience 1 a été utilisée.

7.2.2. Résultats

Les temps de réaction (TR) aberrants inférieurs à 300 ms et supérieurs à 3000 ms ont d'abord été retirés des analyses. Puis un score z a été calculé pour chaque TR basé sur la moyenne et l'écart-type des TR de chaque participant sur chaque type de cible (mot vs pseudo-mot). Les TR avec un score z inférieur ou supérieur à 2.5 écarts-types (outliers) ont été retirés des analyses. Au total, ces deux critères ont écarté 3% des TR sur les mots (3% pour les conditions ADJ gauche et ADJ droite ; 2% pour la condition INV gauche ; 4% pour la condition INV droite) et 2% sur les pseudo-mots (2% pour les conditions ADJ gauche, ADJ droite et INV gauche ; 3% pour la condition INV droite).

Les pourcentages d'erreur ont été calculés sur les essais restants.

Les analyses statistiques ont été réalisées sous JASP. Pour chaque type de cible (mot et non-mot), une ANOVA a été conduite sur les temps de réaction et les pourcentages d'erreurs en prenant, d'une part, les participants comme facteur aléatoire (F_1) selon le plan $S_{80} * \text{Type de Flankers}_4$ et, d'autre part, les items comme facteur aléatoire (F_2) selon le plan $I_{112} * \text{Type de Flankers}_4$.

7.2.2.1. Analyse sur les mots

Les résultats sur les TR et les pourcentages d'erreurs apparaissent au Tableau 13.

Tableau 13. Moyennes des TR et des pourcentages d'erreurs (écarts types entre parenthèses) selon le type de flankers pour les cibles mots.

Conditions	Bigramme adjacent gauche	Bigramme adjacent droite	Bigramme inversé gauche	Bigramme inversé droite
TR	821 (199)	829 (218)	827 (195)	823 (198)
% d'erreurs	7 (6)	7 (11)	7 (6)	8 (11)

L'ANOVA à mesures répétées, effectuée sur les TR, n'a révélé aucun effet significatif du type de flankers, F_1 et $F_2 < 1$ (cf. Figure 16).

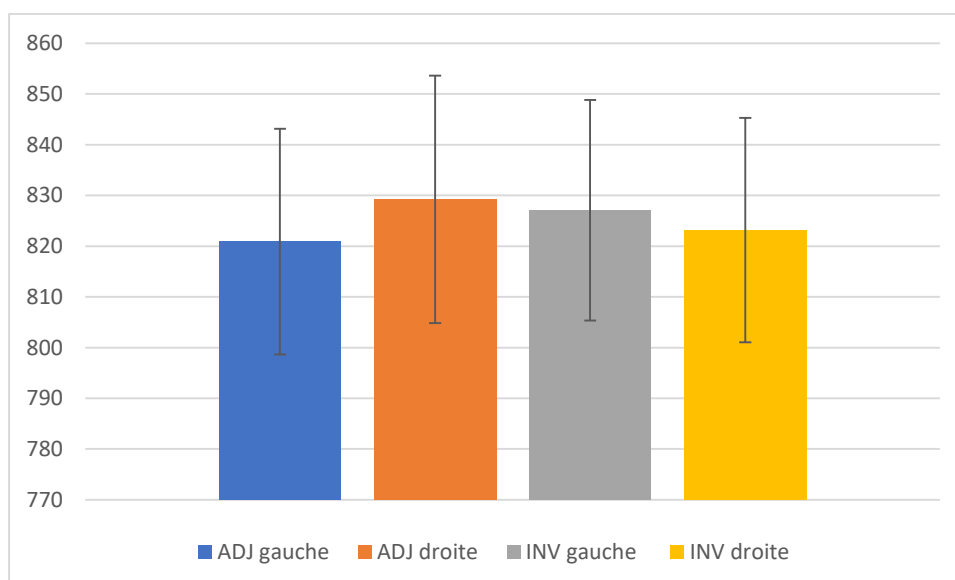


Figure 16. Moyennes des TR en ms selon le type de flankers. Les barres d'erreurs représentent l'erreur standard.

L'ANOVA effectuée sur les pourcentages d'erreurs n'a pas révélé d'effet significatif du type de flankers, $F_1 < 1$; $F_2(3, 333) = 1.40, p = .24$.

7.2.2.2. Analyse sur les pseudo-mots

Les résultats sur les TR et les pourcentages d'erreurs apparaissent au Tableau 14.

Tableau 14. Moyennes des TR et des pourcentages d'erreurs (écarts types entre parenthèses) selon le type de flankers pour les cibles pseudo-mots.

Conditions	Bigramme adjacent gauche	Bigramme adjacent droite	Bigramme inversé gauche	Bigramme inversé droite
TR	1048 (324)	1066 (344)	1059 (342)	1063 (335)
% d'erreurs	9 (8)	8 (7)	8 (8)	10 (8)

Sur les pseudo-mots, l'ANOVA à mesures répétées n'a pas révélé d'effet significatif du type de flankers, $F_1(3, 237) = 1.39, p = .25$; $F_2 < 1$ (cf. Figure 17).

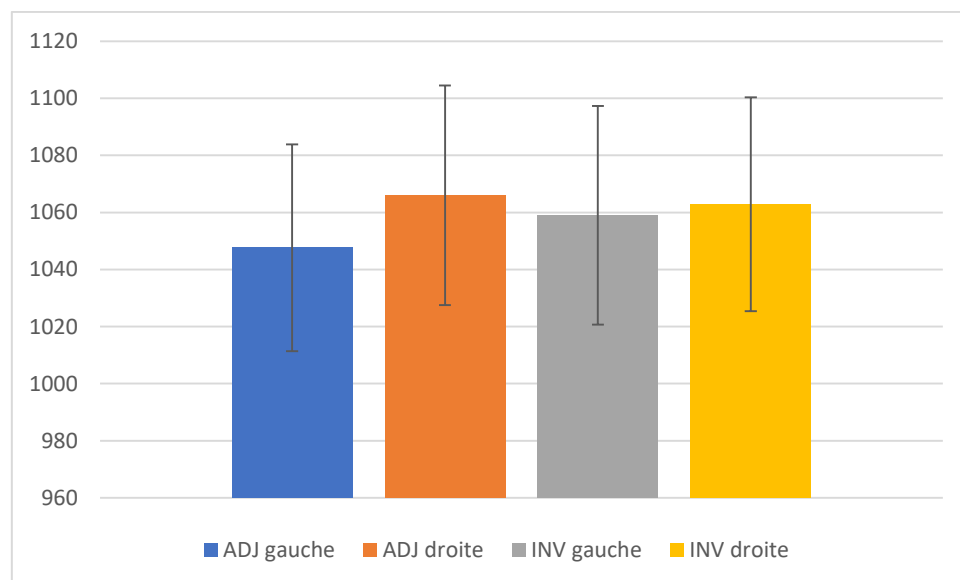


Figure 17. Moyennes des TR en ms selon le type de flankers. Les barres d'erreurs représentent l'erreur standard.

L'ANOVA effectuée sur les pourcentages d'erreurs n'a pas révélé d'effet significatif du type de flankers, $F_1(3, 237) = 1.68, p = .17$; $F_2(3, 333) = 1.57, p = .20$.

7.3. Expérience 4 : Direction de la lecture – Présentation bilatérale de mots

7.3.1. Méthode

7.3.1.1. Participants

Un groupe de 80 élèves du lycée public Alkotrobi de la ville de Ouazzane (région de Tanger, nord-ouest du Maroc) a été testé pour cette quatrième expérience, de niveaux 1^{ère} année (N = 25 ; 17 garçons) et 2^{ème} année (N = 55 ; 39 garçons) du cycle secondaire qualifiant (respectivement 2 ans et 1 an avant le Baccalauréat) de moyenne d'âge de 17 ans 2 mois (étendue = 14 ans 11mois à 20 ans 10 mois ; écart-type = 1 an 4 mois).

Tous les participants avaient une vue normale ou corrigée, étaient natifs du Maroc et avaient appris à lire la langue arabe dès la première année du primaire.

7.3.1.2. Matériel

Le matériel était le même que celui utilisé pour l'expérience 1.

Quatre types de flankers en présentation bilatérale ont été manipulés (cf. Tableau 15). Dans la condition Flankers répétés (ci-après REP), les flankers présentés à droite et à gauche étaient identiques à la cible. Dans la condition Flankers différents (DIF), les flankers présentés à droite et à gauche étaient différents orthographiquement et sémantiquement de la cible. Dans la condition Mixte flanker répété à gauche (M-REP gauche), le flanker présenté à gauche était identique à la cible et celui de droite était différent. Dans la condition Mixte flanker répété à droite (M-REP droite), le flanker présenté à droite était identique à la cible et celui de gauche était différent.

Tableau 15. Conditions expérimentales de présentation bilatérale des flankers pour l'expérience 4.

		CIBLES					
		Mot			Pseudo-mot		
		خيار 4321	خيار 4321	خيار 4321	مائس 4321	مائس 4321	مائس 4321
FLANKERS	Flankers répétés (REP)	خيار 4321	خيار 4321	خيار 4321	مائس 4321	مائس 4321	مائس 4321
	Flankers différents (DIF)	نغمة 8765	خيار 4321	نغمة 8765	روتج 8765	مائس 4321	روتج 8765
	Mixte flanker répété à gauche (M-REP gauche)	خيار 4321	خيار 4321	نغمة 8765	مائس 4321	مائس 4321	روتج 8765
	Mixte flanker répété à droite (M-REP droite)	نغمة 8765	خيار 4321	خيار 4321	روتج 8765	مائس 4321	مائس 4321

7.3.1.3. Procédure

La même procédure que celle de l'expérience 1 a été utilisée à l'exception qu'une durée d'affichage du stimulus pendant 170 ms a été imposée. Le but était de garantir que le participant ne fasse pas un mouvement oculaire sur les flankers. Les stimuli ont été présentés en format texte, écrits en blanc sur fond noir en police Courier New 12 ppp. À 50 cm de distance, le stimuli (la cible et ses deux flankers) correspondait à 4.6° d'angle visuel.

7.3.2. Résultats

Les temps de réaction (TR) aberrants inférieurs à 300 ms et supérieurs à 3000 ms ont d'abord été retirés des analyses. Puis un score z a été calculé pour chaque TR basé sur la moyenne et l'écart-type des TR de chaque participant sur chaque type de cible (mot vs pseudo-mot). Les TR avec un score z inférieur ou supérieur à 2.5 écarts-types (outliers) ont été retirés des analyses. Au total, ces deux critères ont écarté 3% des TR sur les mots (2% pour la condition REP ; et 3% pour les conditions DIF, M-REP gauche et M-REP droite) et 2% sur les pseudo-mots (3% pour la condition REP ; et 2% pour les conditions DIF, M-REP gauche et M-REP droite).

Les pourcentages d'erreur ont été calculés sur les essais restants.

Les analyses statistiques ont été réalisées sous JASP.

Pour chaque type de cible (mot et non-mot), une ANOVA a été conduite sur les temps de réaction et les pourcentages d'erreurs en prenant, d'une part, les participants comme

facteur aléatoire (F_1) selon le plan S_{80} * Type de Flankers₄ et, d'autre part, les items comme facteur aléatoire (F_2) selon le plan I_{112} * Type de Flankers₄.

7.3.2.1. Analyse sur les mots

Les résultats sur les TR et les pourcentages d'erreurs apparaissent au Tableau 16.

Tableau 16. Moyennes des TR et des pourcentages d'erreurs (écarts types entre parenthèses) selon le type flankers pour les cibles mots.

Conditions	Flankers répétés	Flankers différents	Mixte flanker répété à gauche	Mixte flanker répété à droite
TR	748 (153)	765 (160)	765 (172)	758 (159)
% d'erreurs	9 (6)	10 (8)	9 (6)	9 (7)

L'ANOVA à mesures répétées a révélé un effet significatif du type de flankers sur les participants, $F_1(3, 237) = 3.71, p = .01$, mais pas sur les items, $F_2(3, 333) = 1.66, p = .17$. Les comparaisons post-hoc effectuées sur les participants avec correction de Holm (cf. Figure 18) ont montré que la condition Flankers répétés ($M = 748, SD = 153$) étaient significativement plus rapide que la condition Flankers différents ($M = 765, SD = 160$), $t(79) = -2.94, p = .03$ et la condition Mixte flanker répété à gauche ($M = 765, SD = 172$), $t(79) = -2.75, p = .04$. Cependant, entre la condition Flankers répétés ($M = 748, SD = 153$) et la condition Mixte flanker répété à droite ($M = 758, SD = 159$), la différence n'était pas significative, $t(79) = -1.81, p = .30$. Aucune différence significative n'était observée également entre les conditions Flankers différents ($M = 765, SD = 160$) et Mixte flanker répété à gauche ($M = 765, SD = 172$), $t(79) = 0.10, p = .92$; Flankers différents ($M = 765, SD = 160$) et Mixte flanker répété à droite ($M = 758, SD = 159$), $t(79) = 1.20, p = .70$; Mixte flanker répété à gauche ($M = 765, SD = 172$) et Mixte flanker répété à droite ($M = 758, SD = 159$), $t(79) = 0.98, p = .70$.

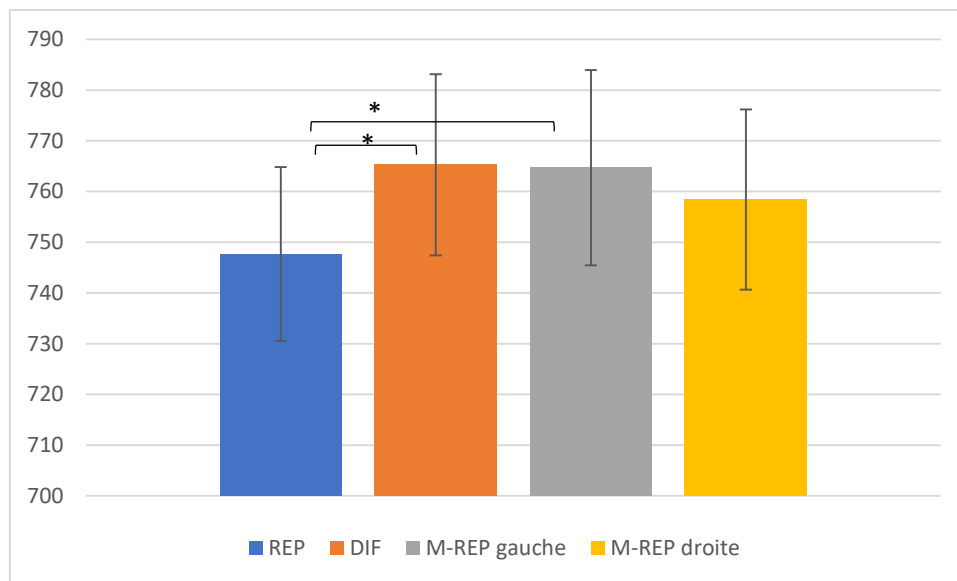


Figure 18. Moyennes des TR en ms selon le type de flankers. Les barres d'erreurs représentent l'erreur standard. (* $p < .05$)

L'ANOVA effectuée sur les pourcentages d'erreurs n'a pas révélé d'effet significatif du type de flankers, F_1 et $F_2 < 1$.

7.3.2.2. Analyse sur les pseudo-mots

Les résultats sur les TR et les pourcentages d'erreurs apparaissent au Tableau 17.

Tableau 17. Moyennes des TR et des pourcentages d'erreurs (écarts types entre parenthèses) selon le type de flankers pour les cibles pseudo-mots.

Conditions	Flankers répétés	Flankers différents	Mixte flanker répété à gauche	Mixte flanker répété à droite
TR	973 (260)	957 (252)	967 (261)	966 (261)
% d'erreurs	14 (12)	13 (11)	14 (11)	14 (11)

Sur les pseudo-mots, l'ANOVA à mesures répétées n'a pas révélé d'effet significatif du type de flankers, F_1 et $F_2 < 1$ (cf. Figure 19).

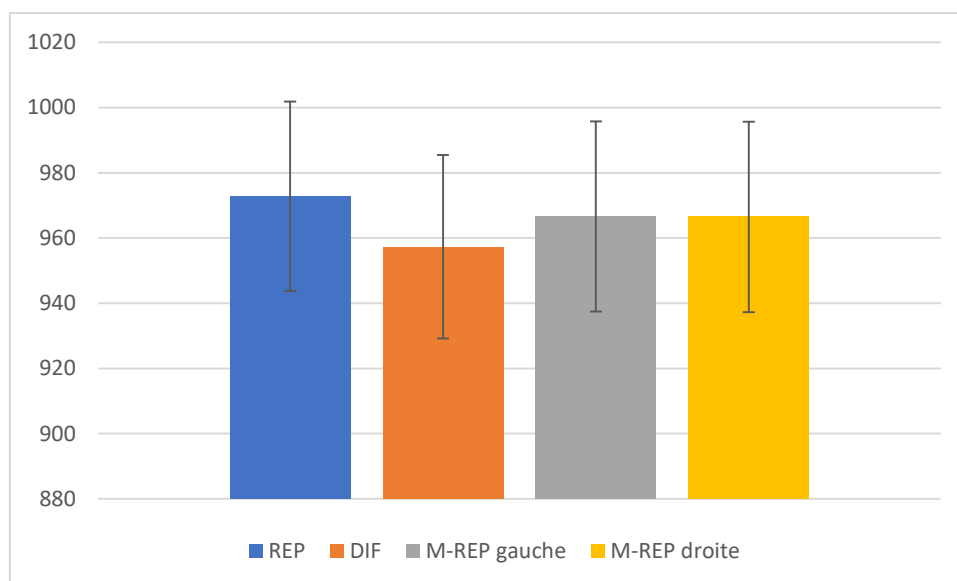


Figure 19. Moyennes des TR en ms selon le type de flankers. Les barres d'erreurs représentent l'erreur standard.

L'ANOVA effectuée sur les pourcentages d'erreurs n'a pas révélé d'effet significatif du type de flankers, F_1 et $F_2 < 1$.

7.4. Synthèse des expériences 3 et 4

Un biais attentionnel à droite du point de fixation, semblable à celui observé en lecture de phrases avec le paradigme de mouvements oculaires, a été mis en évidence en français à l'aide du paradigme FLLD par Snell et Grainger (2018). Leur étude a révélé des temps de réaction plus courts pour les conditions Flanker répété à droite (e.g., rock rock) et Flankers mixtes répétés à droite (e.g., step rock rock), par rapport aux conditions Flanker répété à gauche (e.g., rock rock) et flankers mixtes répétés à gauche (e.g., rock rock step).

L'objectif de ces deux expériences était d'étudier, avec le même paradigme, l'effet de la direction de la lecture sur les traitements visuo-attentionnels chez l'adulte arabophone, en cherchant à produire l'effet inverse obtenu par Snell et Grainger (2018), étant donné le sens de la lecture de la langue arabe (de droite à gauche).

En présentation unilatérale de bigrammes (Expérience 3), aucun effet lié à la direction de la lecture n'a été observé, les cibles étant traitées à la même vitesse qu'elles soient entourées à gauche ou à droite de leurs bigrammes adjacents ou inversés.

L'absence d'effet de direction de la lecture en présentation unilatérale de bigrammes nous a conduit à basculer vers une présentation bilatérale de mots (Expérience 4), en

présentant des mots à gauche et à droite de la cible. L'idée était de se rapprocher d'une situation de lecture naturelle, dans laquelle il existe une contiguïté d'unités lexicales. On s'attendait à observer que la décision lexicale soit facilitée plus par les flankers répétés à gauche de la cible que par les flankers répétés à droite de la cible.

Contrairement à nos attentes, les résultats n'ont pas révélé d'effet lié à la direction de la lecture. Les participants ne bénéficient pas plus des flankers répétés à gauche de la cible que de ceux à droite de la cible. Les temps de réaction de ces deux conditions apparaissent comme identiques à ceux de la condition Flankers différents. Il apparaît donc que, pour qu'un effet de répétition orthographique soit observé en arabe, les flankers droite et gauche doivent participer en même temps au processus d'intégration orthographique du mot cible.

Les données de ces deux expériences ne mettent pas en évidence un biais attentionnel asymétrique à gauche en arabe, et ne confirment pas de ce fait l'hypothèse d'un traitement visuo-attentionnel inverse à celui observé chez des lecteurs français. Du fait que l'effet de répétition orthographique est avéré uniquement en présence des deux flankers, il apparaît que l'attention semble se distribuer symétriquement de part et d'autre du mot cible chez le lecteur arabophone.

8. Chapitre 8 : Discussion générale et perspectives de recherche

8.1. Rappel des objectifs et synthèse des résultats

L'objectif général de cette thèse était d'étudier les mécanismes d'intégration spatiale de l'information orthographique lors de la lecture de mots en écriture arabe chez l'adulte et l'apprenti-lecteur, à l'aide du paradigme FLLD (Dare & Shillcock, 2013). Pour cela, trois études ont été menées. Les données collectées ont été comparées avec celles recueillies en langues latines avec le même paradigme (Dare & Shillcock, 2013 ; Grainger et al., 2014 ; Snell & Grainger, 2018 ; Snell et al., 2020). La question générale était de savoir si les effets observés en écriture latine peuvent refléter des mécanismes cognitifs universels sous-tendant la lecture ou s'il s'agit de mécanismes spécifiques sous-tendus par les caractéristiques du système orthographique.

Trois questions spécifiques ont été soulevées. La première concernait la nature du codage orthographique utilisée dans le système d'écriture de l'arabe. Il s'agissait de vérifier si le modèle à bigrammes ouverts de Grainger et Van Heuven (2003) expliquant le processus du traitement orthographique en écriture latine, pouvait être également étendu aux processus de codage orthographique impliqués dans la reconnaissance de mots écrits en arabe. La deuxième question concernait la trajectoire développementale des processus d'intégration spatiale des informations orthographiques au cours de l'apprentissage de la lecture en arabe en cycle primaire. Il s'agissait d'étudier si le traitement parallèle des informations orthographiques observé chez l'apprenti-lecteur français (Snell et al., 2020) s'opérait à la même période chez l'apprenti-lecteur arabophone. La troisième question concernait l'impact de la direction de la lecture en arabe sur les mécanismes d'intégration de l'information orthographique. Il s'agissait de déterminer si le biais attentionnel révélé dans un système orthographique qui se lit et s'écrit de gauche à droite (Snell et al., 2020) s'observe également dans un système orthographique qui se lit et s'écrit de droite à gauche.

Concernant la première question, nos résultats (Expérience 1, Chapitre 5) sont similaires à ceux obtenus par Dare et Shillcock (2013) et Grainger et al. (2014) : les bigrammes adjacents à la cible impactent le traitement orthographique de la cible qu'ils soient dans l'ordre correct droite-gauche ou inversés. Par exemple, la reconnaissance de la cible **كاتب** était facilitée au même degré par les bigrammes **نبا**, **كاتب**, et les bigrammes **بكا**,

comparativement aux bigrammes *رو كاتب تح*. Ces effets obtenus en écriture arabe étayent le modèle à bigrammes ouverts de Grainger et Van Heuven (2003). Ce modèle comporte trois niveaux de traitement : un premier niveau traite en parallèle l'identité de toutes les lettres d'un mot. Cette information est ensuite transmise au niveau de la carte des positions relatives des lettres qui mobilise un ensemble de représentations utilisant des bigrammes dits "ouverts" (i.e., représentations codant l'ordre des lettres dans le sens de la lecture). L'activation de cet ensemble non ordonné de bigrammes ouverts fournit l'ordre correct des lettres composant le mot. Les bigrammes ouverts activés à ce deuxième niveau envoient de l'activation au niveau des représentations orthographiques compatibles avec ces bigrammes. Ainsi, selon le principe du modèle, le mot *كاتب* serait codé par l'ensemble suivant de bigrammes ouverts : *تب , اب , ات , كب , كد , كا*.

Ces résultats suggèrent que, à l'instar des langues latines, la position relative des lettres en écriture arabe serait codée en bigrammes ouverts et que la nature du traitement orthographique des langues alphabétiques pourrait refléter des mécanismes cognitifs universels sous-tendant la lecture.

En ce qui concerne la deuxième question, la littérature sur les trajectoires développementales de l'apprentissage de la lecture a mis en évidence l'influence du niveau de lecture sur le développement de la taille de l'empan perceptif et sur la quantité d'information orthographique parafovéale pouvant être extraite lors d'une fixation. L'empan perceptif des apprenti-lecteurs étant plus restreint, ils accordent plus d'attention au traitement fovéal, leur permettant seulement un traitement orthographique séquentiel des mots fixés (Häikiö et al., 2009 ; Rayner, 1986). Au cours de l'apprentissage de la lecture, on constate une progression de la taille de l'empan perceptif chez les apprentis lecteurs qui devient similaire à celle observée chez le lecteur expert, permettant ainsi un traitement parallèle des informations orthographiques fovéales et parafovéales en une seule fixation. Il a été montré par exemple que les enfants de grade 1 et 2 prétrahaient peu, voire pas, d'informations en parafovéa (Reichle et al., 2013). D'autres données en français ont pu révéler que le bénéfice de l'information orthographique parafovéale pouvait être observé dès le grade 3 (Khelifi et al., 2012, 2015) ou même dès la fin du grade 1 (Snell et al., 2020). En effet, dans leur étude, Snell et al. (2020) ont obtenu à l'aide du paradigme FLLD le même effet observé chez le lecteur expert (Dare & Shillcock, 2013 ; Grainger et al., 2014) : dès le grade 1, les

apprentis-lecteurs répondaient plus rapidement aux cibles entourées de bigrammes reliés comparativement aux cibles entourées de bigrammes différents, suggérant que les bigrammes sont traités parallèlement avec la cible.

Nos résultats obtenus chez des apprentis-lecteurs arabophones de trois niveaux du cycle primaire (2^{ème}, 4^{ème} et 6^{ème}) montrent que le bénéfice pour les bigrammes reliés orthographiquement à la cible n'est observable qu'à partir du 6^{ème} niveau (12 ans). À partir de ce niveau, l'apprenti-lecteur arabophone commence à manifester un patron similaire à celui d'un lecteur expert arabophone. La mise en place progressive de l'expertise en lecture, traduite par un traitement parallèle des informations orthographiques, se manifeste donc assez tardivement au cours de l'apprentissage de la lecture en arabe, comparativement aux langues latines où elle se met en place plus précocement. Ce résultat met en perspective l'effet pénalisant de la complexité du système orthographique arabe chez l'apprenti-lecteur arabophone (Abdelhadi et al., 2011 ; Eviatar & Ibrahim, 2004, 2007, 2014 ; Eviatar et al., 2004 ; Ibrahim et al., 2002 ; Ibrahim & Eviatar, 2009). Un déploiement plus important des ressources attentionnelles dédiées au traitement des informations fovéales semble s'effectuer au détriment du traitement des informations parafovéales. Il faut plusieurs années de pratique de lecture avant que l'apprenti-lecteur arabophone ne se familiarise avec la complexité des traits visuels de l'écriture arabe et de son système linguistique et acquiert les ressources attentionnelles nécessaires pour pouvoir traiter les informations orthographiques parafovéales.

S'agissant de la troisième question relative à l'impact de la direction de la lecture, la revue de littérature a montré que lors de la lecture, une partie de l'attention est portée sur le mot fixé en fovéa mais également en parafovéa, sans pour autant impliquer un déplacement des yeux grâce à un mécanisme d'attention endogène. Les données recueillies en mouvements oculaires ont mis en évidence le caractère asymétrique de la distribution de l'attention, lié à la direction de la lecture. Dans les langues à alphabet latin, les traitements visuo-attentionnels se portent surtout à droite du point de fixation, car c'est dans cette partie du champ visuel que le lecteur extrait l'information (McConkie & Rayner, 1975, 1976 ; Rayner & Bertera, 1979). Ce biais attentionnel à droite a été mis en évidence également à l'aide du paradigme FLLD par Snell et Grainger (2018). Leur étude a révélé que les participants bénéficiaient plus des flankers répétés à droite que ceux répétés à gauche. En revanche, dans

les langues sémitiques se lisant de droite à gauche, la distribution de l'attention s'inverse pour devenir asymétrique à gauche du point de fixation (e.g., l'arabe, Jordan et al., 2014 ; l'hébreu, Pollatsek, Bolozky, Well, & Rayner, 1981). Cet effet a été obtenu à l'aide du recueil de mouvements oculaires. Aucune étude à ce jour n'a testé cet impact de la direction de la lecture en FLLD. Nos résultats, obtenus chez de jeunes adultes au lycée (M = 17 ans), ne révèlent pas de biais attentionnel vers la gauche, ceci tant en présentation unilatérale de bigrammes que bilatérale de mots.

Deux explications peuvent être avancées pour rendre compte de ce résultat. La première est liée au paradigme utilisé. Du fait que, dans nos expériences, l'effet de répétition orthographique est avéré uniquement en présence des deux flankers, il apparaît que l'attention semble se distribuer symétriquement de part et d'autre du mot cible. Il est donc possible que le biais attentionnel en arabe soit plus difficile à mettre en évidence en lecture de mots isolés avec le paradigme FLLD qu'en lecture de phrases (Jordan, et al., 2014) quand le lecteur est engagé dans un processus de lecture naturelle.

La seconde explication est liée à l'expertise en lecture. En effet, les seules données en arabe proviennent de lecteurs experts de niveau universitaire (Jordan, et al., 2014). Il est donc possible que le biais attentionnel soit davantage présent chez des lecteurs expérimentés que chez des lecteurs plus jeunes (lycéens). Nous avons en effet montré que l'intégration spatiale de l'information orthographique (en parafovée) était beaucoup plus lente à se mettre en place chez le lecteur arabophone. Une période relativement longue de pratique de la lecture semble nécessaire pour la voir s'installer au même niveau que le lecteur français ou anglais.

8.2. Perspectives de recherche

Nous avons choisi d'explorer dans cette dernière partie trois axes de perspectives de recherche que nous souhaitons mener avec le paradigme FLLD.

1) La question de la direction de la lecture.

Notre expérience 3 (Chapitre 7) sur l'impact de la direction de la lecture sur l'intégration spatiale de l'information orthographique réalisée avec des lycéens n'a pas révélé de biais attentionnel dans le sens de la lecture dans cette tranche d'âge. Il serait donc intéressant de rechercher cet effet chez des jeunes adultes de niveau universitaire supposés être d'un meilleur niveau d'expertise. Un effet avéré dans cette population montrerait que le biais attentionnel survient tardivement dans un système mature conformément à ce qui est observé par Jordan et al. (2014) dans le paradigme de fenêtre mobile. Cependant, la comparaison des effets obtenus entre les deux systèmes d'écriture reste à étudier dans le paradigme FLLD étant donné, qu'à notre connaissance, aucune étude à ce jour n'a étudié dans ce paradigme cet effet de direction chez des lecteurs non experts. Un effet non avéré dans le paradigme FLLD montrerait que ce biais attentionnel n'existe pas dans le système d'écriture arabe, ou qu'il est aussi important dans le sens de la lecture que dans l'autre sens. À ce stade, les tentatives d'explications de ce résultat restent très spéculatives. L'interprétation la plus plausible serait liée à la symétrie de la position optimale de regard en arabe, située juste au milieu du mot, faisant que le lecteur arabophone traiterait les informations orthographiques aussi bien à gauche qu'à droite du point de fixation, contrairement au lecteur occidental pour lequel la position optimale se situe au centre gauche du mot, déployant de ce fait plus d'attention à droite afin d'optimiser la visibilité de l'ensemble des lettres composant le mot fixé.

2) La question de l'impact des variations allographiques des lettres.

Dans nos expériences, nous avons maintenu les formes des lettres des bigrammes telles qu'elles apparaissaient dans la cible, n'autorisant pas de variations allographiques. Dans notre matériel, pour le stimulus 43 4321 21, les variations allographiques observées concernent la lettre 2 du bigramme de droite (à cause de l'espace suivant) et la lettre 3 du bigramme de gauche (à cause de l'espace qui précède) dans 54% des cas (61 cibles sur 112).

Sur ces 61 cibles, la variation allographique est systématiquement plus importante pour la lettre 2 du bigramme de droite (i.e., la modification de la forme de la lettre est très importante perceptivement, e.g., *مِكْتَبِ تَب* ; *مِكْ مِكْتَبِ طَح* ; *أَسِ أَسْطَحِ طَح*) que pour la lettre 3 du bigramme de gauche (i.e., la modification de la forme de la lettre est très peu visible, e.g., *مِكْ مِكْتَبِ تَب* ; *مِكْ مِكْتَبِ تَب* ; *أَسِ أَسْطَحِ طَح*).

L'idée serait de rechercher si les effets orthographiques des flankers observés dans notre expérience 1 en conditions flankers "adjacents" et "inversés" peuvent être répliqués en utilisant les variations allographiques des lettres. Pour ce faire, quatre conditions seraient créées : deux conditions avec lien orthographique entre cible et flankers impliquant soit une variation allographique des lettres des flankers (e.g., *مِكْ مِكْتَبِ تَب*) soit aucune variation (mêmes formes des lettres que dans le mot cible ; matériel utilisé dans le présent travail) (e.g., *مِكْ مِكْتَبِ تَب*) ; et deux conditions sans lien orthographique (conditions contrôle) impliquant soit une variation allographique des lettres des flankers (e.g., *أَسِ مَكْتَبِ طَح*) soit aucune variation (e.g., *أَسِ مَكْتَبِ طَح*). Quatre autres conditions pourraient être manipulées pour les flankers inversés. Si l'effet de similarité orthographique des flankers est obtenu également en condition de variation allographique, la participation de l'identité abstraite des lettres pourrait être invoquée sans équivoque dans cet effet. Cette implication des représentations abstraites des lettres pourrait donc être démontrée plus facilement en écriture arabe qu'en écriture latine où les variations allographiques n'existent pas.

3) La question de l'intégration de l'information morphologique.

Cauchi, Beyersmann, Lété et Grainger (soumis 2021) ont récemment étudié cet effet de facilitation de l'information morphologique des flankers sur le traitement de la cible auprès d'élèves d'école primaire (3^{ème}, 4^{ème} et 5^{ème} grade) et auprès de jeunes adultes de niveau universitaire. Deux conditions morphologiques ont été créées : dans la première, la cible était une racine et les flankers étaient une dérivation suffixée de la cible (e.g., *danseur danse danseur*) ; dans la seconde, la cible était le dérivé et les flankers étaient la racine (e.g., *danse danseur danse*). L'effet morphologique des flankers de ces deux conditions était comparé à l'effet de répétition orthographique dans deux conditions : condition racine en cible et en flankers (e.g., *danse danse danse*) et condition mots suffixés en cible et en flankers (e.g., *danseur danseur danseur*). Ces quatre conditions reliées étaient comparées respectivement à quatre conditions de flankers différents : condition racine en cible et mots suffixés en flankers

(e.g., toiture danse toiture) ; condition mots suffixés en cible et racine en flankers (e.g., toit danseur toit) ; condition racine en cible et en flankers (e.g., toit danse toit), et condition mots suffixés en cible et en flankers (e.g., toiture danseur toiture). Les données de cette étude n'ont pas révélé un effet de facilitation morphologique des flankers sur le traitement de la cible chez les apprentis-lecteurs. L'effet a été observé chez l'adulte mais uniquement dans la condition cible-dérivé et flanker-racine (i.e., *danse danseur danse*) suggérant que l'intégration de l'information morphologique est un effet qui se développe tardivement en français.

À l'opposé des langues latines, la morphologie de l'arabe fonctionne essentiellement sur le croisement de racines (généralement de trois consonnes) et de schèmes constitués par une suite de voyelles et/ou d'autres consonnes qui viendront former :

- soit un préfixe : par exemple, l'ajout de la consonne م [m] à la racine كتب [ktb] (désignant le concept d'écriture) forme le mot مكتب [mktb] qui signifie "bureau" ;

- soit un infixe : par exemple, l'ajout de la voyelle longue ا [a:] à la racine [ktb] forme le mot كاتب [ka:tb] qui signifie "écrivain" ;

- soit les deux : par exemple, les ajouts de la consonne م [m] et de la voyelle longue و [u:] à la racine [ktb] forme le mot مكتوب [mktu:b] qui signifie "écrit".

Les suffixes dérivationnels, très nombreux en français, sont très rares en arabe. La voyelle longue ا [a:] et la consonne ن [n] servent par exemple à dériver quelques formes adjectivales comme عطشان [ʔtʕʔa:n] signifiant "assoiffé", dérivée de la racine عطش [ʔtʕʔ] (désignant le concept de soif). La plupart des suffixes servent à construire des formes flexionnelles (genre, nombre, personne, etc.).

Le matériel de l'étude serait donc constitué de racines et de dérivés à partir de préfixes et/ou d'infixes vu que la dérivation à partir de suffixes est très limitée en arabe. Les mêmes conditions que Cauchi et al. (2021) seraient manipulées :

- Conditions morphologiques : a) racine en cible / dérivé en flankers (e.g., كاتب كتب كاتب [ka:tb ktb ka:tb]) ; b) dérivé en cible/ racine en flankers (e.g., كتب كاتب كتب [ktb ka:tb ktb]).

- Conditions Répétitions : c) répétition de racine en cible et en flankers (e.g., كتب كتب كتب [ktb ktb ktb]) ; d) répétition de dérivé en cible et en flankers (e.g., كاتب كاتب كاتب [ka:tb ka:tb ka:tb]).

- Conditions contrôle flankers différents, respectivement : a) racine en cible et dérivé en flankers (e.g., خروج كتب خروج [xru:ʒ ktb xru:ʒ]) ; b) dérivé en cible et racine en flankers (e.g.,

خرج كاتب خرج [xrʒ ka:tb xrʒ] ; c) racine en cible et en flankers (e.g., خرج كتب خرج [xrʒ ktb xrʒ]) ; d) dérivé en cible et en flankers (e.g., خروج كاتب خروج [xru:ʒ ka:tb xru:ʒ]).

L'avantage que présente cette étude en écriture arabe est qu'elle permet d'étudier un effet purement morphologique sans intervention d'un effet de recouvrement orthographique (i.e. nombre de lettres partagées entre racine (cible) et dérivé (flankers) ; (e.g., كاتب كتب كاتب [ka:tb ktb ka:tb])), chose qui est difficile à contrôler en langues latines étant donné que dans la plupart des cas les lettres des suffixes du mot dérivé (flankers) font partie de celles de la racine (cible) (e.g., amit*î*é am*î* amit*î*é ; haut*ai*n haut haut*ai*n).

Compte-tenu de cette spécificité morphologique des mots arabes (basés sur des racines), l'intégration de l'information morphologique pourrait apparaître plus précocement chez le lecteur arabophone que chez le lecteur non arabophone.

Conclusion

Les recherches en psychologie cognitive de la lecture et les modèles explicatifs qui en découlent nous fournissent des données émanant en grande partie des langues à alphabet latin. Toutefois, un enjeu majeur des recherches dans ce domaine est de recueillir des données interlangues permettant d'élaborer des théories générales sur les mécanismes de lecture dans les langues alphabétiques. Notre travail apporte des éléments de réponse à cet enjeu scientifique grâce aux effets orthographiques que nous avons pu répliquer, et qui confortent le modèle à bigrammes ouverts de Grainger et Van Heuven (2003).

Le paradigme FLLD s'est révélé être un outil adéquat pour examiner les mécanismes de l'intégration de l'information orthographique en arabe. Nos données peuvent être considérées comme un complément important aux études antérieures en langues à alphabet latin. Nous espérons avoir contribué à améliorer notre compréhension des processus du traitement orthographique en arabe et apporter des informations supplémentaires sur la façon dont s'articulent les aspects universaux et spécifiques des mécanismes de la lecture. D'autres études restent à mener avec le même paradigme pour approfondir les résultats que nous avons obtenus et explorer d'autres questions.

Références

- Abdelhadi, S., Ibrahim, R., & Eviatar, Z. (2011). Perceptual load in the reading of Arabic: Effects of orthographic visual complexity on detection. *Writing Systems Research, 3*, 117-127.
- Abu-Rabia, S. (1996). The role of vowels and context in reading of highly skilled native Arabic readers. *Journal of Psycholinguistic Research, 25*, 629-641.
- Abu-Rabia, S. (1997a). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers. *Reading and Writing: An Interdisciplinary Journal, 9*, 65-78.
- Abu-Rabia, S. (1997b). Reading in Arabic orthography: The effect of vowels and context on reading accuracy of poor and skilled native Arabic readers in reading paragraphs, sentences and isolated words. *Journal of Psycholinguistic Research, 26*, 465-482.
- Abu-Rabia, S. (1998). Reading Arabic texts: Effects of text type, reader type and vowelization. *Reading and Writing: An Interdisciplinary Journal, 10*, 105-119.
- Abu-Rabia, S. (2001). The role of vowels in reading Semitic scripts: Data from Arabic and Hebrew. *Reading and Writing 14*, 39–59.
- Abu-Rabia, S., & Awwad, Y. (2004) Morphological structures in visual word recognition: Evidence from Arabic. *Journal of Research in Reading, 27*, 321-336.
- Abu-Rabia, S., Share, D., & Mansour, M. S. (2003). Word recognition and basic cognitive processes among reading-disabled and normal readers in Arabic. *Reading and Writing: An Interdisciplinary Journal, 16(5)*, 423–442.
- Abu-Rabia, S., & Siegel, L. S. (1995). Different orthographies different context effects: The effects of Arabic sentence context in skilled and poor readers. *Reading Psychology: An International Quarterly, 16*, 1-19.
- 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, 245-64*.
- Acha, J. & Perea, M. (2008). The effect of neighborhood frequency in reading: Evidence with transposed-letter neighbors. *Cognition, 290-300*.
- Adelman, J. S., Marquis, S. J., & Sabatos-DeVito, M. G. (2010). Letters in words are read simultaneously, not in left-to-right sequence. *Psychological Science, 21(12)*, 1799–1801.
- AlJassmi, M. A., Hermena, E. W., & Paterson, K. B. (2021). Eye movements in Arabic reading. In D. Ntelitheos and T. Tsz-Cheung Leung (Eds.), *Experimental Arabic Linguistics* (pp. 86–108). UAE: John Benjamins Publishing Company.
- Al-Samarraie, H., Sarsam, S. M., Alzahrani, A. I., & Alalwan, N. (2020). Reading text with and without diacritics alters brain activation: The case of Arabic. *Current Psychology: A Journal for Diverse Perspectives on Diverse Psychological Issues, 39(4)*, 1189–1198.
- Ammar, M. (1997). *Les stratégies d'identification de mots écrits en arabe*, Thèse de doctorat, Université de Nantes.
- Andrews, S. (1996). Lexical retrieval and selection processes: Effects of transposed-Letter confusability. *Journal of Memory and Language, 775-800*.
- Arguin, M., & Bub, D. (1995). Priming and response selection processes in letter classification and identification tasks. *Journal of Experimental Psychology: Human Perception and Performance, 21(5)*, 1199–1219.
- Awadh, R. (2016). *Caractérisation et rôle de l'empan visuo-attentionnel chez les lecteurs arabophones adultes et enfants (experts et dyslexiques développementales)*. Thèse de doctorat de psychologie, Université Grenoble Alpes.

- Ayari, S. (1996). Diglossia and illiteracy in the Arab world. *Language, Culture and Curriculum*, 9, 243-253.
- Baccouche, T. 2003. La Langue arabe : spécificités et évolution. In Nouri Romdhane, M., Gombert, J. E., & Belajouza, M. (Eds.), *L'apprentissage de la lecture : Perspectives comparatives. Presses universitaires de Rennes*.
- Belajouza, M., & Snoussi, F. (2006). Apprentissage de la lecture en langue arabe et contraintes externes. *Muqabasat intersignes*, 1, 91-113.
- Blanchard, H. E., Pollatsek, A., & Rayner, K. (1989). The acquisition of parafoveal word information in reading. *Perception & Psychophysics*, 46(1), 85–94.
- Bouanani, M. (2015). الإنجاز اللغوي العربي ومسارات النفاذ الى المعجم الذهني: بين اللساني والمعرفي [Production de la langue arabe et processus d'accès au lexique mental: entre linguistique et cognitif]. In Bouanani, M., Zarhouch, B., & Alaoui, I. (Eds), *المعجم الذهني واللغة العربية - Langue arabe et lexique mental* (pp. 13-44). Fès: Publications du laboratoire des sciences cognitives (LASCO).
- Boudelaa, S., & Marslen-Wilson, W. D. (2000). Non-concatenative morphemes in language processing: Evidence from Modern Standard Arabic. *Proceedings of the Workshop on Spoken Word Access Processes*, 23–26.
- Boudelaa, S., & Marslen-Wilson, W. D. (2005). Discontinuous morphology in time: Incremental masked priming in Arabic. *Language and Cognitive Processes*, 20(1), 207–260.
- Boudelaa, S., Norris, D., Mahfoudhi, A., & Kinoshita, S. (2019). Transposed letter priming effects and allographic variation in Arabic: Insights from lexical decision and the same-different task. *Journal of experimental psychology. Human perception and performance*, 45(6), 729–757.
- Bouma, H. (1971). Visual recognition of isolated lower-case letters. *Vision Research*, 11, 459-474.
- Bowers, J. S. (2000). In defense of abstractionist theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, 7(1), 83–99.
- Bowers, J. S., Vigliocco, G., & Haan, R. (1998). Orthographic, phonological, and articulatory contributions to masked letter and word priming. *Journal of Experimental Psychology: Human Perception and Performance*, 24(6), 1705–1719.
- Brysbaert, M., & Nazir, T. (2005). Visual constraints in written word recognition: Evidence from the optimal viewing-position effect. *Journal of Research in Reading*, 28(3), 216–228.
- Brysbaert, M., Vitu, F., & Schroyens, W. (1996). The right visual field advantage and the optimal viewing position effect: On the relation between foveal and parafoveal word recognition. *Neuropsychology*, 10(3), 385–395.
- Bybee, J. (1995). Diachronic and typological properties of morphology and their implication for representation. In L.B. Feldman (Eds.), *Morphological aspects of language processing*. (pp. 225–246). Hillsdale, NJ: Erlbaum.
- Carreiras, M., Perea, M. & Mallouh, R.A. (2012). Priming of abstract letter representations may be universal: The case of Arabic. *Psychon Bull Rev*, 19, 685–690.
- Carver, R. P. (1990). Reading rate: A review of research and theory. *Academic Press*.
- Cattell, J.M. (1886). The time taken up by cerebral operations. *Mind*, 11,220-242,377-392,524-538.
- Cauchi, C., Beyersmann, E., Lété, B., & Grainger, J. (soumis 2021). A developmental perspective on morphological processing in the flankers task. *Journal of Experimental Child Psychology*.
- Cauchi, C., Lété, B., & Grainger, J. (2020). Orthographic and phonological contributions to flanker effects. *Attention, Perception, & Psychophysics*, 82, 3571-3580.
- Cohen, L., Dehaene, S., Vinckier, F., Jobert, A., & Montavont, A. (2008). Reading normal and degraded words: contribution of the dorsal and ventral visual pathways. *Neuroimage*, 40(1), 353-366.

- Cohen, L., Martinaud, O., Lemer, C., Lehericy, S., Samson, Y., Obadia, M., et al. (2003). Visual word recognition in the left and right hemispheres: anatomical and functional correlates of peripheral alexias. *Cerebral Cortex*, *13*(12), 1313-1333.
- Coltheart, M., & Rastle, K. (1994). Serial processing in reading aloud: Evidence for dual-route models of reading. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(6), 1197–1211.
- Coltheart, M., Rastle, K., Perry, C., Langdon, R., & Ziegler, J. (2001). DRC: a dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, *108*(1), 204-56.
- Content, A., & Peereman, R. (1999). La reconnaissance des mots écrits In Rondal, J.A. & Seron, X. (1999). *Troubles du langage. Bases théoriques, diagnostic et rééducation*. Liège: Mardaga.
- Dai, J., Ibrahim, R., & Share, D.L. (2013) The influence of orthographic structure on printed word learning in Arabic. *Writing Systems Research*, *5*(2), 189-213.
- Dare, N., & Shillcock, R. (2013). Serial and parallel processing in reading: Investigating the effects of parafoveal orthographic information on nonisolated word recognition. *The Quarterly Journal of Experimental Psychology*, *66*(3), 487–504.
- Davis, C. J. (1999). The self-organising lexical acquisition and recognition (SOLAR) model of visual word recognition. *Dissertation Abstracts International*, *62*, 594.
- Davis, C. J., & Bowers, J. S. (2006). Contrasting five different theories of letter position coding: Evidence from orthographic similarity effects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(3), 535–557.
- Den Buurman, R., Roersema, T., & Gerrissen, J. (1981). Eye movements and the perceptual span in reading. *Reading Research Quarterly*, *16*(2), 227-235.
- Deutsch, A., & Rayner, K. (1999). Initial fixation location effects in reading Hebrew words. *Language and Cognitive Processes*, *14*(4), 393–421.
- Drieghe, D., Rayner, K., & Pollatsek, A. (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance*, *31*(5), 954–969.
- Elbeheri, G., & Everatt, J. (2007). Literacy ability and phonological processing skills amongst dyslexic and non-dyslexic speakers of Arabic. *Reading and Writing*, *20*, 273-294.
- Evett, L. J., & Humphreys, G. W. (1981). The use of abstract graphemic information in lexical access. *Quarterly Journal of Experimental Psychology*, *33A*(4), 325-350.
- Eviatar, Z., & Ibrahim, R. (2004). Morphological and orthographic effects on hemispheric. *Reading and Writing*, *17*, 691–705.
- Eviatar, Z., & Ibrahim, R. (2007). Morphological structure and hemispheric functioning: The contribution of the right hemisphere to reading in different languages. *Neuropsychology*, *21*(4), 470–484.
- Eviatar Z., Ibrahim R. (2014) Why is it hard to read arabic?. In: Saiegh-Haddad E., Joshi R. (eds) *Handbook of Arabic Literacy. Literacy Studies (Perspectives from Cognitive Neurosciences, Linguistics, Psychology and Education)*, vol 9. Springer, Dordrecht.
- Eviatar, Z., Ibrahim, R., & Ganayim, D. (2004). Orthography and the hemispheres: visual and linguistic aspects of letter processing. *Neuropsychology*, *18*(1), 174–184.
- Farid, M., & Grainger, J. (1996). How initial fixation position influences visual word recognition: A comparison of French and Arabic. *Brain and Language*, *53*(3), 351–368.
- Feitelson, D., Goldstein, Z., Iraqi, J., & Share, D. L. (1993). Effects of listening to story reading on aspects of literacy acquisition in a diglossic situation. *Reading Research Quarterly*, *28*(1), 71–79.
- Ferguson, Ch. (1959). Diglossia. *Word*, *15*, 325–340.

- Ferrand, L. (2007). *Psychologie cognitive de la lecture. Reconnaissance des mots écrits chez l'adulte*. Ouvertures psychologiques. De Boeck.
- Ferrand, L., & New, B. (2004). Semantic and associative priming in the mental lexicon. In P. Bonin (Ed.), *Mental lexicon: "Some words to talk about words"* (p. 25–43). Nova Science Publishers.
- Ganayim, D. (2016). Letter visibility and the optimal viewing position effect of isolated connected and un-connected letters in Arabic. *Psychology of Language and Communication*, 19(3), 174-200.
- Gomez, P., Ratcliff, R., & Perea, M. (2008). The overlap model: a model of letter position coding. *Psychological Review*, 115(3), 577–600.
- Grainger, J. (2008). Cracking the orthographic code: An introduction. *Language and Cognitive Processes*, 23(1), 1–35.
- Grainger, J., Dichy, J., El-Halfaoui, M., & Bamhamed, M. (2003). Approche expérimentale de la reconnaissance du mot écrit en arabe. *Faits De Langues*, 77-86.
- 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. (1996). Orthographic processing in visual word recognition: A multiple read-out model. *Psychological Review*, 103, 518–565.
- Grainger, J., Lété, B., Bertand, D., Dufau, S., & Ziegler, J. C. (2012). Evidence for multiple routes in learning to read. *Cognition*, 123(2), 280–292.
- Grainger, J., Mathôt, S., & Vitu, F. (2014). Tests of a model of multi-word reading: Effects of parafoveal flanking letters on foveal word recognition. *Acta Psychologica*, 146, 35-40 .
- Grainger, J., & Van Heuven, W. J. B. (2003). Modeling letter position coding in printed word perception. In P. Bonin (Ed.), *Mental lexicon: "Some words to talk about words"* , (pp. 1-23). Nova Science Publishers.
- Häikiö, T., Bertram, R., & Hyönä, J. (2010). Development of parafoveal processing within and across words in reading: Evidence from the boundary paradigm. *Quarterly Journal of Experimental Psychology*, 63(10), 1982–1998.
- Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology*, 102(2), 167–181.
- Harm, M. W., & Seidenberg, M. S. (2001). Are there orthographic impairments in phonological dyslexia? *Cognitive Neuropsychology*, 18(1), 71–92.
- Harm, M. W., & Seidenberg, M. S. (2004). Computing the meanings of words in reading: cooperative division of labor between visual and phonological processes. *Psychological Review*, 111(3), 662–720.
- Henderson, J. M., & Ferreira, F. (1990). Effects of foveal processing difficulty on the perceptual span in reading: Implications for attention and eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(3), 417–429.
- Hermena, E. W., Juma, E. J., & AlJassmi, M. A. (2021). Parafoveal processing of orthographic, morphological, and semantic information during reading Arabic: A boundary paradigm investigation. *PLoS one*, 16(8), e0254745.
- Holmes, V.M. & O'Regan, J.K. (1987) Decomposing french words. In: *Eye movements: From physiology to cognition*, (Ed.) J.K. O'Regan & A. Lévy-Schoen. North Holland.
- Holmes, V. M., & O'Regan, J. K. (1992). Reading derivationally affixed French words. *Language and Cognitive Processes*, 7(2), 163–192.
- Humphreys, G. W., Evett, L. J., & Quinlan, P. T. (1990). Orthographic processing in visual word identification. *Cognitive Psychology*, 22(4), 517–560.

- Ibrahim, R. (2006). Morpho-Phonemic similarity within and between languages: A factor to be considered in processing Arabic and Hebrew. *Reading and Writing, 19*(6), 563-586.
- Ibrahim, R. (2013). Reading in Arabic: New evidence for the role of vowel signs. *Creative Education, 4*, 248-253.
- Ibrahim, R., & Eviatar, Z. (2009). Language status and hemispheric involvement in reading: Evidence from trilingual Arabic speakers tested in Arabic, Hebrew, and English. *Neuropsychology, 23*(2), 240–254.
- Ibrahim, R., Eviatar, Z., & Aharon-Peretz, J. (2002). The characteristics of Arabic orthography slow its processing. *Neuropsychology, 16*(3), 322–326.
- Ikeda, M., & Saida, S. (1978). Span of recognition in reading. *Vision Research, 18*(1), 83–88.
- Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics, 40*(6), 431–439.
- Inhoff, A. W., & Liu, W. (1998). The perceptual span and oculomotor activity during the reading of Chinese sentences. *Journal of Experimental Psychology: Human Perception and Performance, 24*(1), 20–34.
- Jacobs, A.M., Nazir, T.A. & Heller, O. (1989). Perception of lowercase letters in peripheral vision: A discrimination matrix based on saccade latencies. *Perception & Psychophysics 46*, 95–102.
- Jeon, S. T., Hamid, J., Maurer, D., & Lewis, T. L. (2010). Developmental changes during childhood in single-letter acuity and its crowding by surrounding contours. *Journal of Experimental Child Psychology, 107*(4), 423–437.
- Johnson, R. L., Perea, M., & Rayner, K. (2007). Transposed-letter effects in reading: Evidence from eye movements and parafoveal preview. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 209–229.
- Jordan, T., Almabruk, A.A., Gadalla, E., McGowan, V.A., White, S., Abedipour, L., & Paterson, K. (2014). Reading direction and the central perceptual span: Evidence from Arabic and English. *Psychonomic Bulletin & Review, 21*, 505-511.
- Jordan, T. R., Almabruk, A. A. A., McGowan, V. A., & Paterson, K. B. (2011). Evaluating hemispheric divisions in processing fixated words: The evidence from Arabic. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior, 47*(8), 992–997.
- 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.
- Katz, L. & Frost, R. (1992). The reading process is different for different orthographies: The orthographic depth hypothesis. In Frost, R. & Katz, L., (Eds.). *Orthography, Phonology, Morphology, and Meaning*, pp. 67-84. Amsterdam: Elsevier North Holland Press.
- Katz, L., Rexer, K., & Lukatela, G. (1991). The processing of inflected words. *Psychological Research, 53*(1), 25–32.
- Kennison, S. M., & Clifton, C. (1995). Determinants of parafoveal preview benefit in high and low working memory capacity readers: Implications for eye movement control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*(1), 68–81.
- Khelifi, R., Sparrow, L., & Casalis, S. (2012). Sensitivity to lateral information on a perceptual word identification task in French third and fifth graders. *Brain and Cognition, 78*(2), 123–132.
- Khelifi, R., Sparrow, L., & Casalis, S. (2015). Third and fifth graders' processing of parafoveal information in reading: A study in single-word recognition. *Journal of Experimental Child Psychology, 139*, 1–17.
- Kinoshita, S., Norris, D., & Siegelman, N. (2012). Transposed-letter priming effect in Hebrew in the same-different task. *Quarterly Journal of Experimental Psychology, 65*(7), 1296–1305.

- Kinoshita, S., & Kaplan, L. (2008). Priming of abstract letter identities in the letter match task. *The Quarterly Journal of Experimental Psychology*, *61*(12), 1873–1885.
- Kouloughli D. E. , (1994), *Grammaire de l'arabe d'aujourd'hui*, Paris, Pocket - Langues pour Tous.
- Layes, S., & Rebaï, M. (2015). Le rôle du traitement visuo-attentionnel dans la reconnaissance des mots chez les enfants dyslexiques : Effet de la connectivité interne (totale versus partielle) des mots en arabe. *AL-Lisaniyyat*, *21*, (29-48).
- Leibnitz, L., Grainger, J., Muneaux, M. & Ducrot, S. (2016). Processus visuo-attentionnels et lecture : une synthèse. *L'Année Psychologique*, *116*(4), 597-622.
- Liberman, A.M., Fitch, H.L., Halwes, T., & Erickson, D.M.(1980). Perceptual equivalence of two acoustic cues for stop consonant manner. *Perception and Psychophysics*, *27*, 343-350.
- Mahfoudhi, A. (2007). Roots and patterns in Arabic lexical processing. In E. Benmamoun (Ed.), *Perspectives on Arabic linguistics XIX* (pp. 123–150). Amsterdam: John Benjamins
- Mason, M. (1982). Recognition time for letters and nonletters: Effects of serial position, array size, and processing order. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(5), 724–738.
- Mason, M., & Katz, L. (1976). Visual processing of nonlinguistic strings: redundancy effects and reading ability. *Journal of Experimental Psychology. General*, *105*(4), 338–348.
- Marzouki, Y., & Grainger, J. (2008). Effects of prime and target eccentricity on masked repetition priming. *Psychonomic bulletin & review*, *15*(1), 141–148.
- Marzouki, Y., & Grainger, J. (2014). Effects of stimulus duration and inter-letter spacing on letter-in-string identification. *Acta Psychologica*, *148*, 49–55.
- Marx, C., Hutzler, F., Schuster, S., & Hawelka, S. (2016). On the development of parafoveal preprocessing: Evidence from the incremental boundary paradigm. *Frontiers in Psychology*, *7*, 514.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: I. An account of basic findings. *Psychological Review*, *88*, 375–407.
- McConkie, G. W., & Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Perception & Psychophysics*, *17*(6), 578–586.
- McConkie, G. W., & Rayner, K. (1976). Asymmetry of the perceptual span in reading. *Bulletin of the Psychonomic Society*, *8*(5), 365–368.
- McConkie, G. W., & Zola, D. (1979). Is visual information integrated across successive fixations in reading? *Perception & Psychophysics*, *25*(3), 221–224.
- Montant, M., Nazir, T. A., & Poncet, M. (1998). Pure alexia and the viewing position effect in printed words. *Cognitive Neuropsychology*, *15*(1-2), 93–140.
- Morris, R. K., Rayner, K., & Pollatsek, A. (1990). Eye movement guidance in reading: The role of parafoveal letter and space information. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(2), 268–281.
- Mzoughi, I. (2015). *Intégration des emprunts lexicaux au français en arabe dialectal tunisien*. Thèse de doctorat en linguistique. Université de Cergy Pontoise.
- Najem, A. (2006). Stratégies d'identification des mots en arabe standard par des arabophones et des non arabophones: rôle de la diglossie. In Colloque International: *Approche cognitive de l'apprentissage de la langue écrite*. Université Rennes 2 - Rennes, Haute Bretagne, France.
- Nazir, T. A. (1993). On the relation between the optimal and the preferred viewing position in words during reading. In G. d'Ydewalle & J. van Rensbergen (Eds.), *Perception and cognition: Advances in eye movement research* (pp. 349–361). Amsterdam: North-Holland.
- Nazir, T. A., Jacobs, A. M., & O'Regan, J. K. (1998). Letter legibility and visual word recognition. *Memory & Cognition*, *26*(4), 810-821.

- Nazir, T. A., O'Regan, J. K., & Jacobs, A. M. (1991). On words and their letters. *Bulletin of the Psychonomic Society*, 29(2), 171–174.
- O'Regan J. K. (1990). Eye movements and reading. *Reviews of oculomotor research*, 4, 395–453.
- O'Regan, J. K. (1992). Solving the "real" mysteries of visual perception: The world as an outside memory. *Canadian Journal of Psychology/Revue canadienne de psychologie*, 46(3), 461–488.
- 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 & Performance*, 18, 185–197.
- O'Regan, J. K., & Lévy-Schoen, A. (1987). Eye-movement strategy and tactics in word recognition and reading. In M. Coltheart (Ed.), *Attention and performance 12: The psychology of reading* (p. 363–383). Lawrence Erlbaum Associates, Inc.
- O'Regan, J. K., Lévy-Schoen, A., Pynte, J., & Brugailière, B. (1984). Convenient fixation location within isolated words of different length and structure. *Journal of Experimental Psychology: Human Perception and Performance*, 10(3), 393.
- Paap, K. R., Newsome, S. L., McDonald, J. E., & Schvaneveldt, R. W. (1982). An activation–verification model for letter and word recognition: The word-superiority effect. *Psychological Review*, 89, 573–594.
- Paap, K. R., Newsome, S. L., & Noel, R. W. (1984). Word shape's in poor shape for the race to the lexicon. *Journal of experimental psychology. Human perception and performance*, 10(3), 413–428.
- Pelli, D. G., Farell, B., & Moore, D. C. (2003). The remarkable inefficiency of word recognition. *Nature*, 423(6941), 752–756.
- Perea, M., Abu Mallouh, R., & Carreiras, M. (2010). The search for an input-coding scheme: Transposed-letter priming in Arabic. *Psychonomic Bulletin & Review*, 17(3), 375–380.
- Perea, M., Abu Mallouh, R., Mohammed, A., Khalifa, B., & Carreiras, M. (2016). Do diacritical marks play a role at the early stages of word recognition in Arabic? *Frontiers in Psychology*, 7, 1255
- Perea, M., Abu Mallouh, R., Mohammed, A., Khalifa, B., & Carreiras, M. (2018). Does visual letter similarity modulate masked form priming in young readers of Arabic? *Journal of experimental child psychology*, 169, 110–117.
- Perea, M., & Carreiras, M. (2006). Do transposed-letter similarity effects occur at a prelexical phonological level? *Quarterly Journal of Experimental Psychology*, 59(9), 1600–1613.
- Perea, M., & Fraga, I. (2006). Transposed-letter and laterality effects in lexical decision. *Brain and Language*, 97(1), 102–109.
- Perea, M., & Lupker, S. J. (2003a). Does jugde activate COURT? Transposed-letter similarity effects in masked associative priming. *Memory and Cognition*, 31, 829–841.
- Perea, M., & Lupker, S. J. (2003b). Transposed-letter confusability effects in masked form priming. In S. Kinoshita & S. J. Lupker (Eds.), *Masked priming: State of the art* (pp. 97–120). Hove, UK: Psychology Press.
- Perea, M., & Lupker, S. J. (2004). Can CANISO activate CASINO? Transposed-letter similarity effects with nonadjacent letter positions. *Journal of Memory and Language*, 51, 231–246.
- Perea, M., Mallouh, R. A., & Carreiras, M. (2014). Are root letters compulsory for lexical access in Semitic languages? The case of masked form-priming in Arabic. *Cognition*, 132(3), 491–500.
- Perea, M., Rosa, E. Does “whole-word shape” play a role in visual word recognition? *Perception & Psychophysics* 64, 785–794 (2002).
- Perea, M., Rosa, E., & Gómez, C. (2005). The frequency effect for pseudowords in the lexical decision task. *Perception and Psychophysics*, 67, 301–314.

- Peressotti, F., & Grainger, J. (1995). Letter-position coding in random consonant arrays. *Perception & Psychophysics*, *57*(6), 875–890.
- Peressotti, F., & Grainger, J. (1999). The role of letter identity and letter position in orthographic priming. *Perception & Psychophysics*, *61*(4), 691–706.
- 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*, 56–115.
- Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language*, *14*(1), 174–180.
- Pollatsek, A., Rayner, K., & Balota, D. A. (1986). Inferences about eye movement control from the perceptual span in reading. *Perception & Psychophysics*, *40*(2), 123–130.
- Pollatsek, A., & Treiman, R. (2015). *The Oxford handbook of reading*. Oxford University Press.
- Prunet, J.F., B eland, R., & Idrissi, A. (2000). The mental representation of Semitic words. *Linguistic Inquiry*, *31*, 609–648.
- Rayner, K. (1975a). Parafoveal identification during a fixation in reading. *Acta Psychologica*, *39*, 272–282.
- Rayner, K. (1975b). The perceptual span and peripheral cues in reading. *Cognitive Psychology*, *7*(1), 65–81.
- Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin*, *85*, 618–660.
- Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology*, *41*(2), 211–236.
- Rayner, K. (1997). Understanding eye movements in reading. *Scientific Studies of Reading*, *1*(4), 317–339.
- 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. *Quarterly Journal of Experimental Psychology*, *62*(8), 1457–1506.
- Rayner, K., & Bertera, J. H. (1979). Reading without a fovea. *Science*, *206*(4417), 468–469.
- Rayner, K., & Liversedge, S. P. (2004). Visual and linguistic processing during eye fixations in reading. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 59–104). Psychology Press.
- Rayner, K., McConkie, G. W., & Ehrlich, S. (1978). Eye movements and integrating information across fixations. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(4), 529–544.
- Rayner, K., McConkie, G. W., & Zola, D. (1980). Integrating information across eye movements. *Cognitive Psychology*, *12*(2), 206–226.
- Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.
- Rayner, K., Slattery, T. J., & B elanger, N. N. (2010). Eye movements, the perceptual span, and reading speed. *Psychonomic bulletin & review*, *17*(6), 834–839.
- Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception & Psychophysics*, *27*(6), 537–544.
- Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception & Psychophysics*, *31*(6), 537–550.
- Reicher, G. M. (1969). Perceptual recognition as a function of meaningfulness of stimulus material. *Journal of Experimental Psychology*, *81*(2), 275–280.

- Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S., White, S. J., & Rayner, K. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review*, *33*(2), 110–149.
- Roman, G., Pavard, B., & Asselah, B. (1985). Traitement perceptif des phrases ambiguës en arabe [Perceptive analysis of ambiguous Arabic sentences]. *Cahiers de Psychologie Cognitive/Current Psychology of Cognition*, *5*(1), 5–22.
- Rumelhart, D. E., & Siple, P. (1974). Process of recognizing tachistoscopically presented words. *Psychological Review*, *81*, 99–118.
- Saiegh-Haddad, E. (2003). Linguistic distance and initial reading acquisition: The case of Arabic diglossia. *Applied Psycholinguistics*, *24*(3), 431–451.
- Saiegh-Haddad, E. (2004). The impact of phonemic and lexical distance on the phonological analysis of word and pseudo words in a diglossic context. *Applied Psycholinguistics*, *25*, 495–512.
- Saiegh-Haddad, E. (2005). Correlates of reading fluency in Arabic: Diglossic and orthographic factors. *Reading and Writing: An Interdisciplinary Journal*, *18*(6), 559–582.
- Saiegh-Haddad, E., & Schiff, R. (2016). The impact of diglossia on vowel and unvowel word reading in Arabic: A developmental study from childhood to adolescence. *Scientific Studies of Reading*, *20*(4), 311–324.
- Schoonbaert, S., & Grainger, J. (2004). Letter position coding in printed word perception: Effects of repeated and transposed letters. *Language and Cognitive Processes*, *19*, 333–367.
- Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception & Psychophysics*, *74*(1), 5–35.
- Schroyens, W., Vitu, F., Brysbaert, M., & d'Ydewalle, G. (1999). Eye movement control during reading: Foveal load and parafoveal processing. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *52A*(4), 1021–1046.
- Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(1), 65–76.
- Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*(4), 523–568.
- Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.
- Share, D. L. (2008). On the Anglocentrism of current reading research and practice: The perils of overreliance on an “outlier orthography”. *Psychological Bulletin*, *134*, 584–615.
- Snell, J., Bertrand, D., & Grainger, J. (2018). Parafoveal letter-position coding in reading. *Memory & Cognition*, *46*(4), 589–599.
- Snell, J., Cauchi, C., Grainger, J., & Lété, B. (2020). Attention extends beyond single words in beginning readers. *Attention, Perception, & Psychophysics*, 1–9.
- Snell, J., & Grainger, J. (2018). Parallel word processing in the flanker paradigm has a rightward bias. *Attention, Perception & Psychophysics*, *80*(6), 1512–1519.
- Snoussi, F. (2007). Effets de la situation de diglossie sur l'apprentissage de la lecture en langue arabe. *Traverse*, *6*, 63–77.
- Solomon, J., Pelli, D. (1994). The visual filter mediating letter identification. *Nature*, *369*, 395–397.
- Sperlich, A., Schad, D. J., & Laubrock, J. (2015). When preview information starts to matter: Development of the perceptual span in German beginning readers. *Journal of Cognitive Psychology*, *27*(5), 511–530.
- Sperlich, A., Meixner, J., & Laubrock, J. (2016). Development of the perceptual span in reading: A longitudinal study. *Journal of Experimental Child Psychology*, *146*, 181–201.

- Taha, H., Ibrahim, R., & Khateb, A. (2013). How does Arabic orthographic connectivity modulate brain activity during visual word recognition: an ERP study. *Brain Topography*, *26*(2), 292–302.
- Taouk, M., & Coltheart, M. (2004). The cognitive processes involved in learning to read in Arabic. *Reading and Writing: An Interdisciplinary Journal*, *17*(1-2), 27–57.
- Townsend, J.T., Taylor, S.G. & Brown, D.R. (1971). Lateral masking for letters with unlimited viewing time. *Perception & Psychophysics*, *10*, 375–378.
- Tydgat, I., & Grainger, J. (2009). Serial position effects in the identification of letters, digits, and symbols. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(2), 480–498.
- Underwood, N. R., & McConkie, G. W. (1985). Perceptual span for letter distinctions during reading. *Reading Research Quarterly*, *20*(2), 153–162.
- Valdois, S., Carbonnel, S., Juphard, A., Baciú, 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), 149-62.
- Verhoeven, L., & Perfetti, C. (2021): Universals in Learning to Read Across Languages and Writing Systems, *Scientific Studies of Reading*, DOI: 10.1080/10888438.2021.1938575
- Vitu, F. (1991a). The existence of a center of gravity effect during reading. *Vision Research*, *31*, 1289–1313.
- Vitu, F. (1991b). The influence of parafoveal preprocessing and linguistic context on the optimal landing position effect. *Perception & Psychophysics*, *50*, 58–75.
- Vitu, F., Brysbaert, M., & Lancelin, D. (2004). A test of parafoveal-on-foveal effects with pairs of orthographically related words. *European Journal of Cognitive Psychology*, *16*(1-2), 154–177.
- Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception & Psychophysics*, *47*, 583–600.
- Weekes, B.S., (1997). Differential effects of number of letters on word and nonword naming latency. *The Quarterly Journal of Experimental Psychology*, *50*(2), 439-456.
- White, S.J. (2007). Foveal load and parafoveal processing: The case of word skipping. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds). *Eye movements: A window on mind and brain*. (pp. 409-424). Oxford: Elsevier.
- White, S. J., Rayner, K., & Liversedge, S. P. (2005). The influence of parafoveal word length and contextual constraint on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review*, *12*(3), 466–471.
- Whitney, C. (2001). How the brain encodes the order of letters in a printed word: The SERIOL model and selective literature review. *Psychonomic Bulletin and Review*, *8*, 221–243.
- Whitney, C. (2008b). Supporting the serial in the SERIOL model. *Language and Cognitive Processes*, *23*(6), 824-865.
- Wolfe, J. M. (2003). Moving towards solutions to some enduring controversies in visual search. *Trends in cognitive sciences*, *7*(2), 70-76.
- Yao-N'dré, M. (2013). *Étude des aspects visuels et oculomoteurs de la lecture en vision centrale et périphérique*. Thèse de Doctorat de neurosciences, Université Aix-Marseille.
- Ziegler, J. C., Ferrand, L., Jacobs, A. M., Rey, A., & Grainger, J. (2000). Visual and phonological codes in letter and word recognition: Evidence from incremental priming. *The Quarterly journal of experimental psychology. A, Human experimental psychology*, *53*(3), 671–692.
- Ziegler, J. C., Goswami, U. & Richardson, U. (2005). The effects of spelling consistency on phonological awareness : A comparison of English and German. *Journal of Experimental Child Psychology*, *92*(4), 345-365.

Annexe

Matériel expérimental utilisé pour les trois études.

a) Liste des mots

Item number	Target	Translit.	Meaning	Unrelated Flanker	Translit.	Meaning
1	خيار	xyAr	choice	نغمة	ngmp	tone
2	حساب	HsAb	calculation	مفتش	mft\$	inspector
3	جناح	jnAH	wing	رشيق	r\$yq	agile
4	حواس	HwAs	senses	مطبخ	mTbx	kitchen
5	سباق	sbaq	race	فصيح	fSyH	eloquent
6	طابع	TAbE	seal	رموش	rmw\$	eyelashes
7	سعيد	sEyD	happy	طالب	TAlb	student
8	طاقة	TAqp	energy	شموخ	\$mwx	greatness
9	بركة	brkp	lake	شيوخ	\$ywx	old
10	سلطة	slTp	power	غاضب	gADb	angry
11	مسجد	msjd	mosque	رخيص	rxyS	cheap
12	نشيد	n\$yd	hymn	صاعق	SAEq	striking
13	قالب	qAlb	mould	ضحكة	DHkp	laugh
14	باحث	bAHv	researcher	مفرش	mfr\$	tablecloth
15	لسان	lsAn	tongue	معجب	mEjb	fan
16	مرطب	mrTb	moisturizing	هادف	hAdf	purposeful
17	ناقص	nAqS	incomplete	ملبس	mlbs	dress
18	نقطة	nqTp	dot	بارد	bArd	cold
19	قادم	qAdm	coming	فرصة	frSp	opportunity
20	معبر	mEbr	passage	فائز	fA}z	winner
21	موظف	mwZf	functionary	كاذب	kA*b	liar
22	جهنم	jhnM	hell	ساحة	sAHp	courtyard
23	فساد	fsAd	corruption	بشرة	b\$rp	skin
24	مطلق	mTlq	absolute	سخاء	sxA'	generosity
25	فكرة	fkrp	idea	شموع	\$mWE	candles
26	شريف	\$ryf	honest	محزن	mHzn	sad
27	نقاش	nqA\$	discussion	صريح	SryH	frank
28	خيام	xyAm	tents	طبله	Tblp	drum
29	مكان	mkAn	place	رسوب	rswb	failure
30	واثق	wAvq	sure	مخدر	mxdr	drug
31	ضفدع	DfdE	frog	نقاب	nqAb	veil
32	طفلة	Tflp	kid	رسام	rsAm	painter
33	حملة	Hmlp	campaign	رهيب	rhyb	frightening

Item number	Target	Translit.	Meaning	Unrelated Flanker	Translit.	Meaning
34	راشد	rA\$d	mature	وصفة	wSfp	recipe
35	خطوة	xTwp	step	مصنف	mSnf	classified
36	ضابط	DAbT	officer	متين	mtyn	resistant
37	شرطي	\$rTy	policeman	مخبز	mxbz	bakery
38	فاشل	fA\$I	loser	مسند	msnd	armrest
39	حركة	Hrkp	movement	مساء	msA'	evening
40	مرشح	mr\$H	candidate	بالغ	bAlg	adult
41	سقوط	sqwT	fall	راحة	rAHp	rest
42	أسطح	>sTH	roofs	منظف	mnZf	cleaner
43	مصنع	mSnE	factory	رافض	rAfD	rejecting
44	متحف	mtHf	museum	راقص	rAqS	dancer
45	ساخن	sAxn	hot	ريشة	ry\$P	feather
46	مصحف	mSHf	(islamic holy book)	يابس	yAbs	dry
47	مرخص	mrXS	authorized	لقاح	lqAH	vaccine
48	مربع	mrBE	square	أسود	>swd	black
49	نشاط	n\$AT	activity	قديم	qdym	ancient
50	قياس	qyAs	measurement	مفكر	mfr	thinker
51	شروق	\$rwq	sunrise	بطلة	bTlp	heroine
52	عنصر	EnSr	element	محبة	mHbp	love
53	حقول	Hqwl	fields	شجاع	\$jAE	courageous
54	فلاح	fIAH	farmer	سمين	smyn	fat
55	غسيل	gsyl	wash	بحار	bHAr	seas
56	خدمة	xdmp	service	قصير	qSyr	short
57	حديث	Hdyv	conversation	نقوش	nqw\$	engravings
58	مدرس	mdrs	teacher	طابق	TAbq	floor
59	بسيط	bsyT	simple	صافح	SAfH	shook hands
60	راتب	rAtb	salary	صحون	SHwn	plates
61	فريق	fryq	team	سهلة	shlp	easy
62	نسيم	nsym	breeze	خضار	xDAr	vegetables
63	حريص	HryS	mindful	جبال	jbal	mountains
64	شعوب	\$Ewb	peoples	فراغ	frAg	void
65	عاصف	EASf	stormy	قرية	qryp	village
66	نجوم	njwm	stars	تفاح	tfAH	apples
67	خارج	xArj	out	سنوي	snwy	annual
68	سرعة	srEp	speed	خصام	xSAm	dispute
69	مد هش	mdh\$	amazing	نساء	nsA'	women
70	قارب	qArb	boat	ملون	mlwn	dye
71	مركز	mrkz	center	بساط	bsAT	carpet
72	ساجد	sAjd	bowing down	زورق	zwrq	dinghy

Item number	Target	Translit.	Meaning	Unrelated Flanker	Translit.	Meaning
73	واضح	wADH	evident	بصمة	bSmp	fingerprint
74	حامل	HAmI	pregnant	بذور	b*wr	seeds
75	بعيد	bEyd	far	فاتح	fAtH	clear
76	مقبل	mQbl	next	دافع	dAfE	motive
77	بريد	bryd	post	أحمق	>Hmq	crazy
78	عنيف	Enyf	violent	بقرة	bqrp	cow
79	تحفة	tHfp	masterpiece	بلوط	blwT	acorn
80	أجمل	>jml	prettiest	بيئة	by}p	environment
81	تيار	tyAr	current	ضحك	DHwk	laughter
82	وحيد	wHyd	alone	ثائر	vA}r	rebel
83	بؤرة	b&rp	hypocenter	ثقيل	vqyl	heavy
84	شكور	\$kwr	thanks	تحدي	tHdy	challenge
85	ثمار	vmAr	fruits	عجيب	Ejyb	wonderful
86	جدول	jdwl	stream	رحيم	rHym	merciful
87	جلوس	jlws	sitting	أرنب	>rnB	rabbit
88	جملة	jmIp	phrase	قدير	qdyr	almighty
89	حياة	HyAp	life	جذور	j*wr	roots
90	جولة	jlwp	ride	سحاب	sHAb	clouds
91	أسفل	>sfI	below	ثرية	vryp	rich
92	جوقة	jwqp	troupe	حافر	HAfr	hoof
93	حرية	Hryp	freedom	دهون	dhwn	fats
94	حلبة	HIbp	arena	أصفر	>Sfr	yellow
95	حصان	HSAn	horse	بخور	bxwr	incense
96	حائط	HA}T	wall	بخيل	bxyl	stingy
97	خبث	xbyv	maleficent	حساء	HsA'	soup
98	حزام	HzAm	belt	بدين	bdyn	obese
99	أبيض	>byD	white	مقطع	mqTE	part
100	خروج	xrwj	exit	حافة	HAfp	edge
101	باقة	bAqp	bouquet	خروف	xrwf	sheep
102	دائم	dA}m	permanent	رجوع	rjwE	return
103	رحلة	rHIp	trip	دفاع	dfAE	defence
104	سائل	sA}l	liquid	رحيق	rHyq	nectar
105	فندق	fndq	hotel	شعير	\$Eyr	barley
106	عجلة	EjIp	wheel	قنص	qnAS	sniper
107	شبكة	\$bKp	net	قناع	qnAE	mask
108	شرطة	\$rTp	police	عميق	Emyq	deep
109	كرسي	krsy	chair	غائب	gA}b	absent
110	مخزن	mxzn	warehouse	كحول	kHwl	alcohol
111	وصفة	wSfp	recipe	نجار	njAr	carpenter
112	مسرح	msrH	theatre	هاتف	hAtf	telephone

b) Liste des pseudo-mots

Item number	Target	Translit.	Unrelated Flanker	Translit.
1	فنبير	fnbr	شثول	\$vwl
2	مائش	mA}\$	طفين	Tfyn
3	فنلة	fnlp	ثميص	vmyS
4	ساكط	sAkT	وصيخ	wSyx
5	ظفام	ZfAm	كحيق	kHyq
6	زسيم	zsym	روتح	rwtH
7	قاهم	qAhm	تريف	tryf
8	نعاط	nEAT	متيغ	mtyg
9	مرثع	mrve	طاحة	TAHp
10	تجاع	tjAE	فصيك	fSyk
11	هادق	hAdq	مطيف	mTbf
12	لظاح	lZAH	ثركة	vrkp
13	رتيق	rtyq	طالط	TAlv
14	أميغ	>myg	قراغ	qrAg
15	كاضر	kADr	سميغ	smyg
16	شموف	\$mwf	ترطي	trTy
17	معنر	mEnr	راتد	rAtd
18	سافن	sAfn	طعيد	TEyd
19	ظصير	ZSyr	لطان	lTAn
20	أطود	>Twd	صريك	Sryk
21	تموع	tmwE	ثالغ	vAlg
22	فدمة	fdmp	خائز	xA}z
23	غشيد	g\$yd	راتث	rAtv
24	نطلة	nTlp	قضار	qDAr
25	مرطث	mrTv	كخلة	kxlp
26	مكزن	mkzn	حواط	HwAT
27	مطلظ	mTlZ	عبصر	EbSr
28	نطيم	nTym	جثال	jtAl
29	طهلة	Thlp	ضانط	DAnT
30	رهيث	rhyv	سفاء	sfA'
31	فطوة	fTwp	ثحار	vHAr
32	مطند	mTnd	كاذث	kA*v
33	معجث	mEjv	ظياس	ZyAs
34	ثارد	vArd	موظخ	mwZx
35	ضخدع	DxdE	مكاغ	mkAg
36	مدهت	mdht	عازخ	EAzx
37	جاغب	jAgb	متسف	mtsf
38	نظاب	nZAb	طهول	Thwl
39	أسخل	>sxl	طرعة	TrEp

Item number	Target	Translit.	Unrelated Flanker	Translit.
40	تمال	tmAl	مفدر	mfdR
41	كريفص	kryS	خساد	xsAd
42	كديث	kdyv	راضف	rADf
43	مركص	mrkS	غاضث	gADv
44	فيار	fyAr	قالث	qAlv
45	ملوغ	mlwg	راكة	rAkp
46	غساء	gsA'	خكرة	xkrp
47	مرسح	mrsH	غطيل	gTyl
48	ظرية	Zryp	مطاء	mTA'
49	مدرط	mdrT	جالظ	jAlZ
50	فماح	fmAH	نضية	nDyp
51	سكام	skAm	خنجب	xnjb
52	ثلمي	vlmy	حاكس	HAks
53	بهاف	bhAf	تيوم	tywM
54	طهرم	Thrm	نبقة	nbqp
55	سحيم	sHym	طختل	Txtl
56	فشاء	f\$A'	صخوس	Sxws
57	جوتب	jwtb	زهرت	zhrv
58	قطيخ	qTyx	صبلح	SblH
59	قطلة	qTlp	حنفث	Hnfv
60	غليج	glyj	صاطر	SATr
61	شامد	\$Amd	جبلق	jblq
62	تناع	tnAE	لهيس	lhys
63	حيلن	Hyln	ثرفة	vrfp
64	سليخ	slyg	طورخ	T&rx
65	مجلظ	mjlZ	تنسع	tnsE
66	جرخب	jrxb	مسيظ	msyZ
67	غنيف	gnyf	هدوث	hdwv
68	أميف	>myf	شخاض	\$xAD
69	مثيخ	mvyx	كضبه	kDbp
70	طزوح	Tzwh	مجخي	mjxy
71	حبشث	Hb\$v	هياخ	hyAx
72	كهنق	khnq	خامف	xAmf
73	شمني	\$mny	زافد	zAfd
74	اخضب	AxDb	دسور	dswr
75	نبقذ	nbq*	خاظم	xAZm
76	خفيغ	xfyg	جيلث	jylv
77	تهضخ	thDx	كعقب	kEqb
78	فعجة	fEjp	وحشف	wH\$f

Item number	Target	Translit.	Unrelated Flanker	Translit.
79	قبحر	qbHr	تلسك	tlsk
80	طحتل	THtl	شهوغ	\$hwg
81	زئيت	z}yt	حافظم	HAZm
82	خثقل	xvql	مهجط	mhjT
83	ظحاء	ZHA'	خبطش	xbT\$
84	سحام	sHAm	كعيف	kEbf
85	جوشب	jw\$b	فضسل	fDsl
86	زاشد	zA\$d	تبسع	tbsE
87	سنيي	snby	شوعه	\$wEp
88	برفة	brfp	مائس	mA}s
89	سيفب	syfb	تهضح	thDH
90	تمخس	tmxs	دشور	d\$wr
91	مرضس	mrDs	أشبن	>\$bn
92	زشيم	z\$ym	هفبة	hfbp
93	سلفن	slfn	خماح	xmAH
94	حاقم	HAqm	طفيب	Tfyb
95	شبول	\$bwl	مجي	mjHy
96	قبحر	qbxr	وصيح	wSyH
97	كمنس	kmsn	فطلة	fTlp
98	زئيش	z}y\$	عتاخ	EtAx
99	بهاخ	bhAx	محيق	mHyq
100	شمبي	\$mby	سزوح	szwH
101	كليف	klyf	شهوچ	\$hwH
102	طيفر	Tyfr	خامق	xAmq
103	قعجة	qEjp	سبار	sbAr
104	شعبر	\$Ebr	صهيج	Shyx
105	اخضن	AxDn	شبيق	\$byq
106	ضميس	Dmys	شهون	\$hwn
107	سهرم	shrm	بيتف	bytf
108	خاقم	xAqm	تطيس	tTys
109	ضميس	Dmys	روشح	rw\$H
110	ضمبر	Dmbr	سختل	sxtl
111	حمهر	Hmhr	فائل	fAvl
112	جرقب	jrqb	مئيج	mvyH