

Identification du mot écrit et de son apprentissage

Marie-Christine CAZALY

Maitrise des langages 2018

Sommaire

Introduction	5
Chapitre 1 : Identification du mot écrit et de son apprentissage	7
A : L'identification du mot écrit chez le lecteur expert : l'hypothèse des deux voies de lecture	7
1. Mise en évidence de deux voies de lecture par les recherches en neuropsychologie.....	7
2. Le modèle à deux voies de lecture Coltheart 1978.....	8
3. Evolution du modèle à deux voies de lecture.....	10
3.1 Le modèle DRC de Coltheart et al 2001,	
3.2 Le modèle connexionniste CDP+ de Perry, Ziegler et Zorzi (2007)	
4. Les différents effets de lecture.....	13
4.1 Les effets de fréquence	
4.2 Les effets de lexicalité	
4.3 Les effets de régularité	
B : Mécanismes d'apprentissage d'identification du mot écrit et de ses difficultés	17
1. Modèles développementaux	17
1.1 Le modèle en stades de Frith (1985)	
1.2 Le modèle à "double fondation" de Seymour (1997)	
2. Le rôle prépondérant du décodage pour la construction du lexique orthographique. ..	22
3. La lecture par analogie	23
4. Comprendre certaines difficultés d'identification du mot écrit	24
4.1 L'apport des études inter-langues	
4.1.1 Transparence et complexité	
4.1.2 Effet de consistance	
4.1.3 Théorie de la granularité	
4.2 Les effets de la fréquence du voisinage orthographique et syllabique	
4.2.1 L'effet des voisins orthographiques	
4.2.2 L'effet des voisins syllabiques	

4.3 Les effets de familiarité

4.4. La situation du français langue seconde

4.5 La dyslexie

4.5.1 La dyslexie de surface

4.5.2 La dyslexie phonologique

Résumé Chapitre 1.....31

Bibliographie

Introduction

La lecture n'est pas une activité naturelle car le cerveau n'est pas programmé pour lire comme il peut l'être pour le langage oral. L'écrit est une invention de l'homme tout comme les activités de comptage et suppose qu'il y ait apprentissage. Certains chercheurs, comme Ludovic Ferrand ont qualifié la lecture de talent cognitif prodigieux, automatique, mais son apprentissage peut devenir laborieux pour un certain nombre d'enfants. C'est le cas, en Nouvelle Calédonie, où le nombre d'élèves en difficulté de lecture représente plus de 20% de la population scolaire. Parmi ceux-ci, beaucoup ont des difficultés d'identification du mot écrit et n'accèdent pas à une lecture fluide. Le grand défi de l'école est alors de réussir à déterminer ce qui peut être source de difficulté afin de concevoir des aides appropriées. C'est notamment le rôle des enseignants spécialisés qui à partir d'investigations menées, essaient d'en dégager des profils de lecteurs. En effet, plus l'analyse du profil de lecteur sera fine, plus les aides apportées seront efficaces.

C'est pourquoi, à travers la présentation d'un ensemble de recherches les plus adaptées à la lecture à voix haute, cette étude se proposera dans un premier temps de faire le point sur les connaissances dont nous disposons sur les mécanismes d'identification du mot écrit pour le lecteur expert. Puis, d'étudier l'apprentissage de ces mécanismes tels qu'ils se développent chez l'enfant. Et enfin, nous nous intéresserons particulièrement aux différentes sources explicatives des difficultés de reconnaissance du mot écrit.

Chapitre 1 :

Les modèles théoriques de l'identification du mot écrit et de son apprentissage

Que sait-on de la lecture experte ?

La finalité de la lecture est avant tout la compréhension de ce qui vient d'être lu. Pour Ecalle et Magnan, 2010, cette finalité résulte de deux composantes essentielles : l'identification du mot écrit et la compréhension de phrases ou de textes. Ainsi, la reconnaissance des mots écrits conditionne toute l'activité de lecture car elle est un préalable à la compréhension. Pour être efficaces, les mécanismes d'identification du mot écrit doivent être automatisés. Mais quels sont ces mécanismes ? Comment se traduisent-ils chez le lecteur expert ? C'est pourquoi, avant de décrire son apprentissage et ses difficultés nous exposerons dans cette première partie, les théories explicatives de la lecture experte.

A. L'identification du mot écrit chez le lecteur expert : L'hypothèse de deux voies de lecture

1. Mise en évidence de deux voies de lecture par les recherches en neuropsychologie

Deux voies de lecture ont été mises en évidence par Marshall et Newcombe (1966,1973) à partir de l'analyse des erreurs de patients ayant subi des lésions.

Dans le cas de l'alexie périphérique, il existe des erreurs de substitution ou de réarrangement des lettres : APPLE qui est lu ABLE. La personne change une ou plusieurs lettres et commet une erreur de lecture. L'alexie profonde correspond à des erreurs de substitution sémantique. Le patient remplace le mot « little » par un mot sémantiquement correct comme « small ». Enfin, l'alexie de surface qui s'explique par des erreurs de lecture de phonèmes, donc à une sur-régularisation des mots lus telle que 'lite' pour 'lit'. L'alexie de surface provoque de nombreux néologismes au cours de la lecture.

Ces observations démontrent qu'il existe certainement deux routes principales pour la lecture : Une qui conduit à une adresse phonologique et l'autre à une adresse sémantique, les deux combinées permettent la prononciation du mot lu.

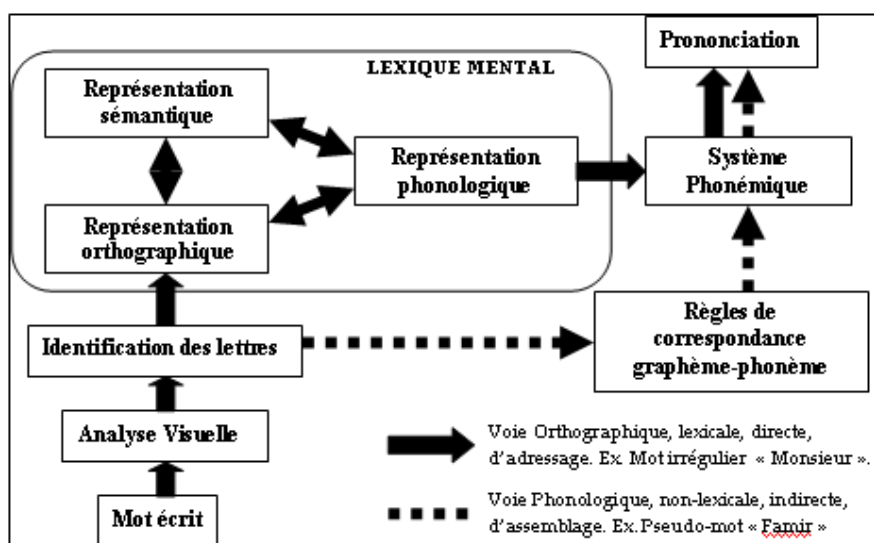
Il existe différentes catégories de mots qui conduisent à l'utilisation de voies différentes :

- Les pseudo-mots comme « sabifo » qui ne peuvent être lus que par la voie d'assemblage c'est-à-dire la conversion de graphèmes en phonèmes
- Les mots irréguliers comme « femme » qui ne peuvent être lus que par la voie d'adressage, c'est-à-dire la reconnaissance visuelle immédiate du mot stocké dans le lexique mental.
- Les mots réguliers comme « vélo » dont la lecture peut utiliser les deux voies précédentes.

Mais ces constats font débat à propos de différents facteurs qui peuvent intervenir lors de la lecture tels que la fréquence des mots, l'ambiguïté, l'homophonie ainsi que la concrétude (mots plus ou moins concrets). Mais avant de s'intéresser à ces différents facteurs, il nous semble essentiel de décrire les modèles qui tentent de décrire la lecture experte, dont le plus connu est le modèle à deux voies de lecture de Coltheart, 1978.

2. Le modèle à deux voies de lecture (Dual route model) Coltheart 1978

Les recherches portant sur la lecture experte ont largement démontré l'importance des procédures d'identification des mots écrits isolés. Parmi ces recherches, le modèle à deux voies de Coltheart (1978) s'est révélé très fertile et a servi de point de référence à de nombreuses recherches dans le domaine de l'identification du mot écrit chez le lecteur expert. D'après ce modèle, pour lire un mot, le lecteur expert peut utiliser deux procédures : 1) La voie directe, dite aussi orthographique, lexicale ou d'adressage et 2) La voie indirecte, dite aussi phonologique, sub-lexicale, décodage ou d'assemblage.



Modèle de la double-voie (Coltheart, 1978)

La voie directe, d'adressage permet de récupérer en mémoire la représentation phonologique lexicale directement. En quelque sorte, le mot est reconnu d'emblée sans avoir besoin de le déchiffrer. Cette voie serait la plus utilisée par le lecteur expert. Elle est également la seule utilisable pour la lecture de mots irréguliers. Cette route est sensible à la fréquence des mots. En effet, plus un mot est fréquent et plus vite il est reconnu et lu. Les mots fréquents sont identifiés plus rapidement que les mots rares. Cet effet de fréquence montre l'existence même de la voie directe. Elle est révélée aussi par le fait qu'il est impossible de lire un mot nouveau ou pseudo-mot (mot qui n'existe pas) directement. Le lecteur est obligé d'assembler les phonèmes et les graphèmes afin d'aboutir à la prononciation du mot.

La voie indirecte, d'assemblage mène à l'élaboration de cette représentation phonologique par assemblage des phonèmes et des graphèmes composant le mot. Procédure aussi appelée 'médiation phonologique'. Elle est principalement utilisée par le lecteur expert pour les mots nouveaux ou les pseudo-mots. De plus, elle présente un effet de régularité. C'est-à-dire que les mots réguliers sont mieux lus que les mots irréguliers, par le fait que la correspondance des graphèmes phonèmes ne permet pas d'identifier le mot irrégulier comme « femme » correctement. Il se produit alors une erreur de régularisation qui aboutit à la prononciation « fame ». Cet effet de régularité démontre la présence de la voie indirecte au cours de l'identification des mots écrits. Mais également, l'effet de lexicalité, c'est-à-dire la différence entre la lecture de mots et de pseudo-mots, où les mots sont mieux lus que les pseudo-mots, montre que deux procédures sont utilisées.

Dans ce modèle à deux routes, pour le lecteur expert, l'utilisation de la voie indirecte est facultative alors que la voie directe, plus rapide, joue le premier rôle. Les deux voies s'activent en parallèle. Il est donc possible d'évaluer l'identification des mots écrits en prononciation (à voix haute) en utilisant des tâches avec effet de fréquence, de lexicalité et de régularité. Dans ce cas, on mesure les réponses exactes ou inexactes ainsi que les temps de latence des réponses correctes.

Ainsi le modèle à deux voies est à l'origine de nombreuses recherches et a fait l'objet de modélisation informatique qui permet de rendre compte des différents effets de lecture et d'affiner la compréhension du processus de lecture experte. Le modèle DRC en fait partie.

3. Evolution du modèle à deux routes de lecture

3.1 Le modèle DRC (Modèle à deux routes en cascade) de Coltheart, Rastle, Perry, Langdon et Ziegler (2001)

Dans les années 2000, le modèle initial à deux routes de Coltheart (1978) à évolué vers une modélisation par ordinateur : Le modèle à deux routes en cascade par Coltheart, Rastle, Perry, Langdon et Ziegler (2001). Le principe est le suivant :

L'information se déplace en cascade de niveau en niveau par activation systématique d'un niveau à l'autre. Par exemple : l'activation de l'unité lettres active à son tour les autres niveaux. Les unités interagissent par des connexions excitatrices ou inhibitrices qui se répandent de façon progressive. Trois routes composées de différents niveaux en interaction constituent le modèle DRC :

- La route lexicale / sémantique (non implémentée)
- La route lexicale / non sémantique
- Une route non lexicale (règles de conversions graphèmes -> phonèmes)

Chaque niveau comporte soit des mots, des lettres, des phonèmes ou des traits visuels. Le modèle contient 26 lettres, 43 phonèmes et 7981 mots monosyllabiques.

La route lexicale non sémantique fonctionne de la façon suivante : (Chaque niveau activant le suivant). (-> = activation).

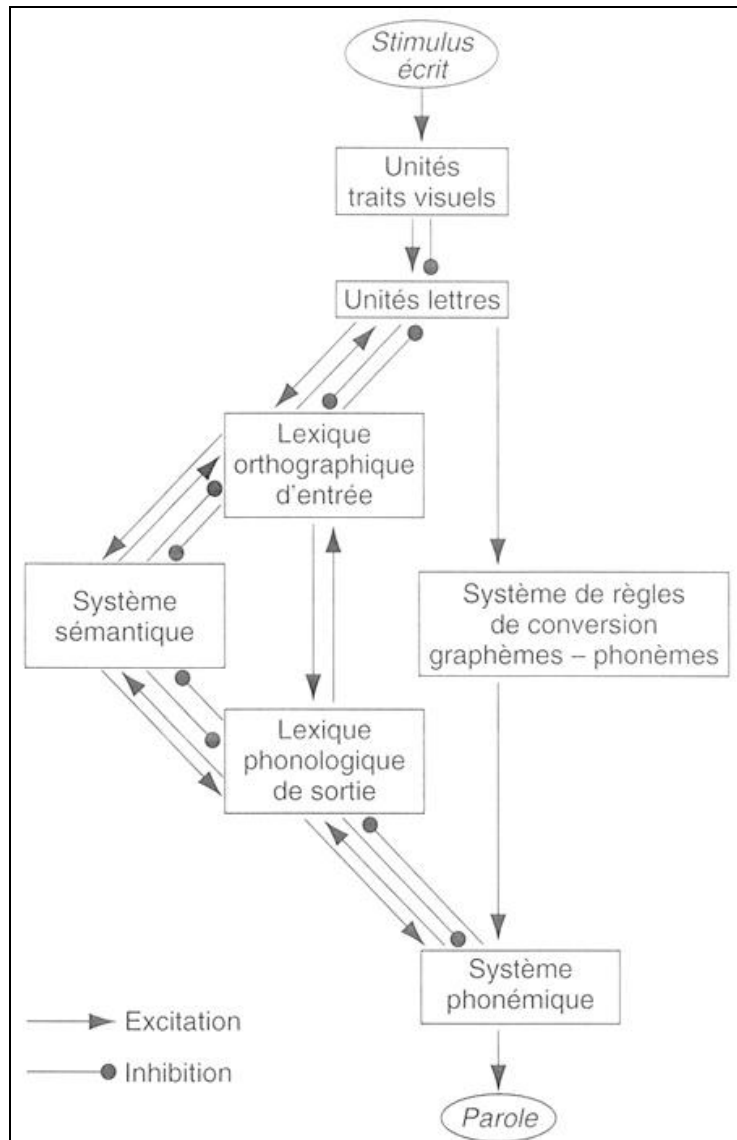
Mot en présentation -> traits visuels correspondant -> lettres -> unité orthographique -> unité phonologique -> phonèmes correspondant.

Les mots contenus dans le lexique orthographique par fréquence d'occurrence pour les mots les plus fréquents sont plus rapidement activés que les mots moins fréquents.

La route non lexicale fonctionne grâce à des règles de conversions graphèmes -> phonèmes. Ces règles ont été choisies uniquement de façon statistique. C'est-à-dire que pour un graphème donné ; c'est le phonème le plus fréquent qui lui est assigné. La lecture est sérielle, lettres après lettres et de gauche à droite.

Performance du modèle : Sur les 7981 mots du modèle, 98,9 % sont lus correctement ; les erreurs portent sur les homographes hétérophones (mots qui s'écrivent pareils mais ne se prononcent pas pareils) et sur des sur-régularisations en prononciation. Sur les 7000 non-mots du modèle, 98,9 % sont bien lus. Le modèle DRC simule correctement les effets repérés

pour des humains, c'est-à-dire les effets de fréquence, de régularité, de pseudo homophonie, de fréquence des pseudo-homophones, de taille du voisinage pour les mots et les non mots et de répétition. DRC est aussi capable de simuler la dyslexie de surface et phonologique. Ces performances montrent bien la force de ce modèle pour la compréhension des mécanismes sous-jacents à l'identification des mots écrits.



Modèle de la « DRC » par Coltheart et al. (2001). *Figure tirée de Ferrand (2007).*

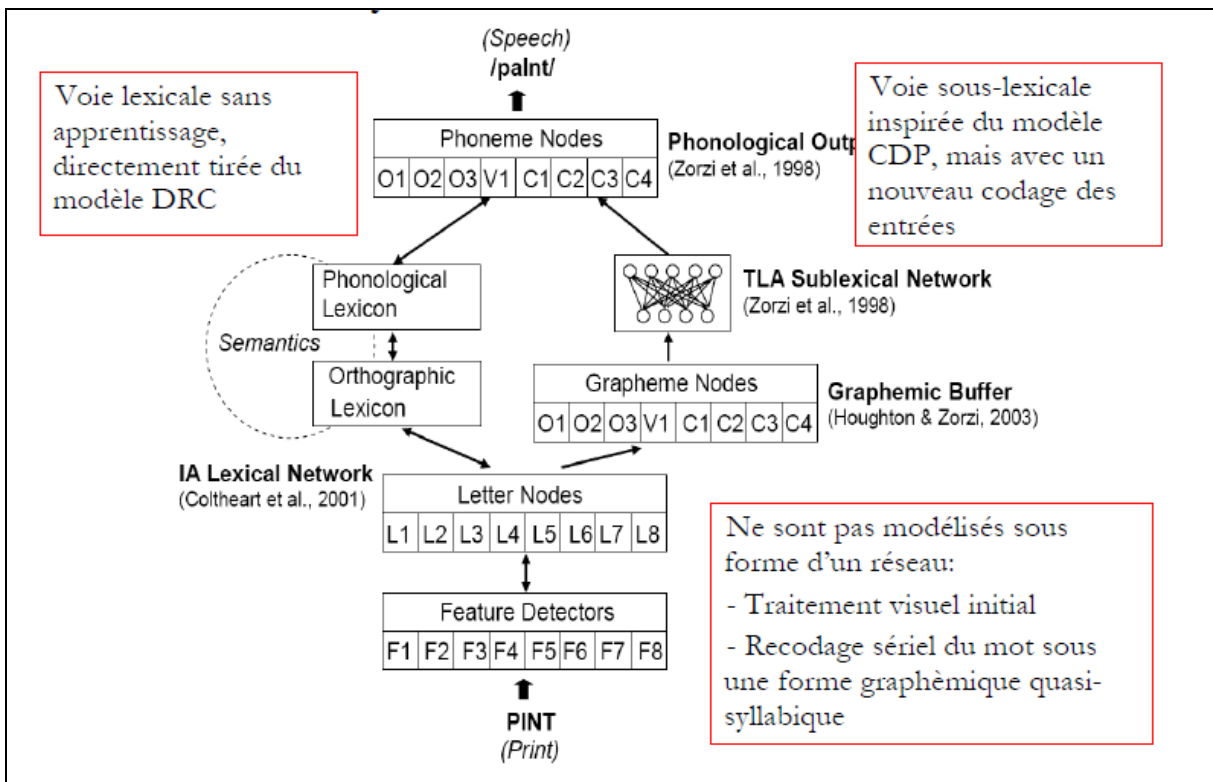
Limite du modèle : Le modèle ne peut pas simuler les effets d'amorçage rapide avec masquage et ne peut pas lire les mots de plus d'une syllabe. Bien qu'il soit adapté à des langues alphabétiques telles que l'allemand (Ziegler, Perry, Coltheart, 2000) et au français (Ziegler, Perry, Coltheart, 2003), langues contenant des voyelles, il ne peut prétendre au principe universel d'un codage phonologique dans la lecture silencieuse.

Encore plus récemment, le modèle connexionniste à deux processus CDP+ de Perry, Ziegler et Zorzi (2007) est capable de simuler beaucoup plus d'effets et se montre encore plus puissant et prometteur.

3.2 Le modèle connexionniste à deux processus CDP+ de Perry, Ziegler et Zorzi (2007)

Ce modèle est intéressant dans le cadre de ce mémoire car il est capable de simuler un très grand nombre d'effets observés chez le lecteur humain en prononciation (à voix haute). Il est très détaillé mais très complexe dans sa composition car sa structure est entièrement connexionniste, c'est-à-dire qu'il est composé de réseaux d'unités interconnectées. Les auteurs de ce modèle (Perry et al 2007) font l'hypothèse qu'il y a bien deux voies distinctes de lecture.

Il associe de façon incrémentale d'autres modèles antérieurs et en reprend les meilleurs aspects : le DRC de Coltheart et al (2001) présenté dans la section précédente (3.1) pour la voie de gauche, lexicale et le modèle CDP de Zorzi, Houghton et Butterworth (1998) non présenté ici pour la voie de droite, dite sous-lexicale.



Le modèle de Perry, Ziegler et Zorzi (2007) : CDP+ = synthèse des modèles CDP et DRC

Schéma issu du cours N°4 Collège de France – Stanislas DEHAENE – Mécanismes cérébraux de la lecture.

La voie sous-lexicale (de droite) comprend un système de phonologie assemblée à deux niveaux et une mémoire tampon pour les graphèmes ; le buffer graphémique, qui n'apparaissait pas dans le modèle CDP de Zorzi et al (1998). Les graphèmes des mots en entrée sont extraits, traités et recodés lettre par lettre, de façon sérielle pour reconstituer grâce à la centration sur les voyelles, la structure syllabique du mot. Les étapes de prétraitement visuel sont programmées sur ordinateur et facilitent le travail des réseaux de neurones du modèle. La voie lexicale fonctionne comme le modèle DRC.

Ce modèle est très prometteur dans le sens où il est capable de lire un très grand nombre de mots et de pseudo-mots d'une syllabe et la comparaison entre les temps de réaction simulés et les temps de réaction observés montrent une grande similitude. Par contre, le problème de ce modèle tient principalement au fait que le modèle ne lit que des mots mono-syllabiques en anglais alors que le français ou l'anglais comportent des mots pluri-syllabiques. L'extension à d'autres langues alphabétiques comme le français est également attendue.

Mais quels sont les différents effets que l'on peut observer chez le lecteur humain ?

La section suivante présente ces différents effets à partir du tableau de Perry et al (2007) reprenant les tests expérimentaux de référence par les principaux modèles de lecture à voix haute.

4. Les différents effets de lecture

L'étude des différents effets de lecture revêt un rôle important dans la compréhension des mécanismes de lecture du lecteur expert et en difficulté ainsi que pour le développement des processus d'identification du mot écrit. C'est pourquoi, nous présentons ci-après les principaux effets de lecture observés chez le lecteur expert.

D'autres effets, plus spécifiques comme par exemple l'effet de consistance seront vus dans la partie B - Apprentissage des mécanismes d'identification des mots écrits et de ses difficultés.

Liste des effets expérimentaux simulés (+) ou non simulés(-) par les principaux modèles de lecture à voix haute : DRC (Coltheart et al, 2001) et CDP+ (Perry et al, 2007). D'après Ziegler et Zorzi (2007).

Tableau issu du livre de L.Ferrand, 2007, Psychologie cognitive de la lecture (p.290)

Effets	Expériences	Description	Modèle DRC	Modèle CDP+
Fréquence	Jared (2002, Exp 2) Weekes (1997)	Les mots de haute fréquence prononcés plus rapidement / correctement que les mots de basse fréquence	+	+
Lexicalité	McCann & Besner (1987) Weekes (1997)	Les mots sont prononcés plus rapidement / mieux que les non-mots	+	+
Fréquence X régularité	Paap & Noel (1991) Jared (2002, Exp 2)	Les mots irréguliers sont prononcés plus lentement / moins bien que les mots réguliers. Pas d'interaction avec la fréquence (selon Jared, 2002)	+	+
Consistance (mots)	Jared (2002, Exp 1)	Les mots inconsistants sont prononcés plus lentement / moins bien que les mots consistants. La taille de l'effet dépend du taux amis/ennemis	-	+
Consistance (non-mots)	Andrews & Scarrat (1998)	La prononciation des non-mots montre des effets graduels de la consistance : les gens n'utilisent pas toujours les correspondances graphèmes-phonèmes les plus communes	-	+
Longueur X régularité	Weekes (1997) Ziegler et al (2001)	Les latences de prononciation augmentent linéairement avec chaque lettre additionnelle	+	+
Position de l'irrégularité	Rastle & Coltheart (1999)	La taille de l'effet de régularité est plus grande pour les mots avec une irrégularité en position initiale (ex. Chef) qu'en seconde ou troisième position.	+	+
Voisinage Rime	Ziegler et al (2001)	Les mots avec beaucoup de voisins-rimes sont prononcés plus rapidement/correctement que ceux ayant peu de voisins-rimes	-	+
Amorçage Onset	Forster & Davis (1991)	Les mots précédés par une amorce partageant l'attaque sont prononcés plus rapidement/correctement que des mots précédés par des amorces non reliées.	+	+
Avantage pseudo-homophonique	McCann & Besner (1987) Reynolds & Besner (2005)	Les non-mots se prononçant comme des mots réels (e.g. bloo) sont prononcés plus rapidement/correctement que des contrôles orthographiques	+	+
Dyslexie de surface	Patterson & Behrmann (1997)	Patient MP : trouble spécifique de la prononciation de mots irréguliers, modulé par le taux de consistance des mots	-	+
Dyslexie phonologique	Derouesné & Beauvois (1985)	Patient LB : Trouble spécifique de la prononciation de non-mots, réduit lorsque les non-mots sont des pseudo-homophones similaires orthographiquement	+	+
Large bases de Données	Spieler & Balota (1997) Balota & Spieler (1998)	Les latences de prononciation du modèle subissent une régression sur la moyenne des latences de chaque item pris dans les larges bases de données contenant plusieurs milliers d'items	-	+

4.1 Les effets de fréquence :

La définition générale de la fréquence correspond au nombre de fois où une unité linguistique apparaît dans un énoncé. En lecture, elle signifie le nombre de fois qu'un

lecteur a rencontré un mot donné. La fréquence des mots en français est contenue dans un large corpus, LEXIQUE (New, Pallier, Ferrand & Matos, 2001). Cette base de données permet de connaître la fréquence d'occurrence de la plupart des mots du français, donc de savoir si un mot est plus ou moins fréquent.

En lecture, l'effet de fréquence signifie que plus un mot est fréquent, plus vite il est lu par rapport à un mot peu fréquent. On suppose que le lecteur a rencontré ce mot un grand nombre de fois et qu'il le récupère beaucoup plus vite qu'un mot moins souvent rencontré. Cet effet de fréquence est robuste et démontré par un grand nombre de recherches sur des tâches :

- de décision lexicale (Allen, Mc Neal & Kvak, 1992 par exemple)
- d'identification perceptive (Grainger & Jacob, 1990)
- de prononciation (Andrews & Heathcote, 2001)
- d'enregistrement des mouvements des yeux (Inhoff & Rayner, 1986)
- de catégorisation sémantique (Andrews & Heathcote, 2001)

De plus, cet effet de fréquence a été constaté pour plusieurs langues dont l'anglais, le français, l'hébreu, le chinois, l'espagnol, le serbo-croate et révélé pour la lecture silencieuse en décision lexicale. Il est par contre sous estimé pour la lecture en prononciation car la route non lexicale utilise des règles de correspondances graphèmes-phonèmes qui ralentissent la lecture. Il sera donc amoindri pour les langues qui utilisent principalement la voie non lexicale, comme le Serbo-Croate (Frost et al, 1987), langue très régulière.

Dans le modèle à Traitement Parallèle Distribué, Seidenberg & McClelland (1989) font l'hypothèse qu'à chaque fois que le lecteur rencontre un mot, sa représentation orthographique est renforcée. Par voie de conséquence, les autres unités phonologiques et sémantiques sont également renforcées. Les mots les plus fréquents sont donc lus plus vite que les mots moins fréquents. Cette explication est également valable pour l'orthographe.

4.2 Effet de lexicalité :

L'effet de lexicalité s'explique par le fait qu'un mot sera lu plus rapidement qu'un non-mot. La logique sous-jacente correspond à l'hypothèse des deux voies de lecture. En effet, un non-mot ne peut être lu que par la voie indirecte c'est-à-dire par assemblage des

graphèmes-phonèmes et sa lecture sera plus lente qu'un mot qui a été rencontré plusieurs fois (fréquent) qui lui, sera lu par la voie directe, beaucoup plus rapide.

4.3 Effet de régularité :

Un mot régulier est un mot qui peut être lu par assemblage graphèmes-phonèmes classique, donc totalement décodé. A l'inverse, un mot irrégulier, qui ne peut être décodé normalement car certains graphèmes, doit être lu différemment. C'est le cas pour le mot « oignon » où le graphème « oi » correspond au phonème « o », conversion très exceptionnelle. L'effet de régularité indique alors qu'un mot régulier sera lu plus rapidement et mieux qu'un mot irrégulier.

En résumé, les principaux modèles d'identification du mot écrit chez le lecteur expert que nous venons d'exposer, tentent de retracer le déroulement du processus d'identification d'un mot écrit. Le premier correspond au modèle à double voie de Coltheart, 1978 qui décrit deux procédures distinctes et parallèles. C'est-à-dire, d'une part, la voie indirecte, dite d'assemblage, permettant de décoder les mots nouveaux et les non-mots et d'autre part la voie directe, dite d'adressage qui permet l'accession et la récupération directe d'un mot, comme c'est le cas pour la lecture de mots irréguliers. Elle représente la voie la plus utilisée par le lecteur expert. Le second type de modèle, correspond aux modèles connexionnistes (DRC, CDP+), souvent modélisés par informatique, qui décrivent une seule procédure. Celle-ci fonctionne par distribution sur deux axes, des unités phonologiques et orthographiques, qui en fin de compte permet la lecture de tous les mots. Ces modélisations sont issues directement du modèle à double voie de Coltheart et permettent de rendre compte et de simuler différents effets de lecture constatés chez le lecteur expert tels que la fréquence, la régularité et la lexicalité pour les plus robustes. Ces effets représentent la marque de ces procédures. Même si ces modèles sont très prometteurs quant à la compréhension de la lecture experte, ils ne décrivent pas comment se mettent en place ces processus. C'est pourquoi, nous exposerons dans cette seconde partie théorique l'apprentissage des mécanismes d'identification du mot écrit chez l'enfant et de ses difficultés.

B. Mécanismes d'apprentissage d'identification du mot écrit

Mais que sait-on de l'apprentissage des procédures d'identification du mot écrit ? Le mot est l'unité de base de l'écrit et se situe au cœur des processus de bas et de haut niveau.

L'identification du mot écrit suppose que le lecteur établisse une correspondance entre ce qu'il voit, un mot écrit, et sa représentation mentale pour aboutir à sa compréhension. Pour être efficace, un certain nombre de procédures doivent être automatisées afin d'aboutir à une lecture experte. Les méthodes d'apprentissage de la lecture s'appuient pour la plupart d'entre elles sur une conception développementale de l'apprentissage, notamment celle en stades dont le modèle le plus connue est celui d'Uta FRITH (1985).

1. Modèles développementaux

Les modèles développementaux ont permis d'infirmer l'hypothèse de la lecture globale, thèse qui, à un moment donné, était largement répandue chez les enseignants. En effet, Bruce McCandliss (2006) a pu démontrer que les modèles stadistes sont directement inspirés du modèle à deux voies de lecture de Coltheart (1978) qui prennent pour hypothèse que le lecteur expert utilise bien deux routes distinctes pour accéder au lexique mental.

Une question fondamentale de l'apprentissage des mécanismes d'identification du mot écrit s'est posée à partir de l'observation des erreurs commises par des lecteurs apprenants. Parmi ces modèles, le schéma de référence reste encore aujourd'hui le modèle d'Uta Frith (1985) car il permet de comprendre le fonctionnement normal de l'apprentissage de la lecture et de ses difficultés.

1.1 Le modèle en stades de FRITH (1985)

Le modèle de FRITH (1985) décrit des stades d'apprentissage qui ne sont pas divisés mais qui se recouvrent dans le temps. Selon ce modèle, l'apprentissage de la lecture procède en trois grandes étapes :

➤ Le stade Logographique ou pictural :

Il s'agit en fait d'une pseudo-lecture sur la base de traits graphiques saillants et d'indices contextuels. Par exemple, le logo caractéristique de « Coca Cola » est reconnu d'emblée grâce à quelques traits particuliers ou même la couleur (rouge ici). Gouch et Juel (1989) ont

montré qu'un indice visuel associé à un mot permet la reconnaissance de ce mot ultérieurement.

La base de cette lecture est variable et ne peut être généralisée. L'ordre des lettres n'est pas pertinent. La reconnaissance du mot se fait sans traitement linguistique mais plutôt par l'utilisation d'un système sémantique pictural (Morton, 1990). La reconnaissance est globale, sous forme de devinette, de par cœur et se montre peu économique et surtout limitée. Ce stade n'est pas non plus prédictif d'un bon apprentissage de la lecture.

➤ **Le stade Phonologique ou alphabétique :**

Ce stade correspond à la mise en place et à l'utilisation progressive de la voie de conversion graphèmes-phonèmes (médiation phonologique), où l'ordre des lettres est pertinent, et dans le même temps, à l'émergence de la conscience phonologique essentielle à la réussite de cet apprentissage. La conscience phonologique, c'est à dire comprendre de façon explicite que la parole se décompose en unités plus petites que sont les mots, les syllabes et les phonèmes. La conscience phonologique est fortement réduite chez les illettrés voir inexistante. Morais, Cary, Alegria et Bertelson (1979) ont pu montrer sur des adultes analphabètes que la manipulation des sons du langage était difficile, comme par exemple l'ajout ou la suppression de syllabes ou de phonèmes. La conscience phonologique est à la fois cause et conséquence de l'apprentissage de la lecture, autrement dit de la mise en place de la voie phonologique. Cette étape alphabétique se caractérise également par des effets de longueur des mots, c'est-à-dire qu'un mot long est plus difficilement lu qu'un mot court. Cet effet diminue au cours de l'apprentissage. Il est également constaté que les syllabes complexes ralentissent considérablement le décodage car elles sont composées d'un grand nombre de phonèmes. Par exemple, le mot « strict » est particulièrement difficile à lire par rapport au mot « bol » qui ne comprend que 3 phonèmes. La longueur des mots et la complexité des syllabes sont sources de difficulté pour les lecteurs apprenants. La phase alphabétique est sensible à la régularité des mots, et la lecture des mots irréguliers conduit à des erreurs de régularisation (par exemple : « femme » lu « fame »). Mais il n'y a pas d'effet de fréquence, ni de lexicalité. C'est-à-dire que les mots courants ne sont pas forcément mieux lus que les pseudo-mots.

Enfin, cette voie phonologique qui nécessite de connaître l'alphabet, doit faire l'objet d'un apprentissage explicite où toutes les correspondances graphèmes-phonèmes doivent être

enseignées, apprises. Bruce McCandliss, chercheur américain, a pu démontrer, grâce à ses travaux d'imagerie fonctionnelle, l'inefficacité d'un enseignement par une méthode globale. Cette étape est donc cruciale et prédictive du niveau de lecture de l'enfant.

➤ **Le stade orthographique :**

Le stade orthographique est la phase de l'accès à l'expertise en lecture, à l'automatisation, et correspond à la reconnaissance directe des mots ou des morphèmes sans nécessité de conversion grapho-phonémique. L'effet de régularité est alors moins important mais les effets de fréquence et de lexicalité sont constatés (voir expérience de Sprenger-Charolles, 1992, ci-dessous). C'est-à-dire qu'à cette étape, les mots fréquents sont mieux lus que les mots rares et les mots sont lus plus rapidement et mieux que les non-mots. On constate également la disparition de l'effet de longueur des mots. C'est également la phase où une certaine expertise de lecture des bi-grammes, des préfixes et des suffixes se met en place.

Tableau 2 : *Modèle en six étapes des habiletés en lecture et en écriture (Frith, 1986)*

Etape	Lecture	Ecriture
1a	<i>logographique 1</i>	(symbolique)
1b	logographique 2	logographique 2
2a	logographique 3	<i>alphabétique 1</i>
2b	alphabétique 2	alphabétique 2
3a	<i>orthographique 1</i>	alphabétique 3
3b	orthographique 2	orthographique 2

Note : "1" signifie un très faible niveau d'expertise ; "2" et "3" une expertise de plus en plus importante

Ce modèle décrit le développement progressif des deux voies de lecture des mots écrits, la voie indirecte et la voie directe, décrit dans le modèle à deux voies de Coltheart (1978).

Comme le démontre ce graphique, l'installation des deux voies de lecture conduit également à l'apprentissage de l'orthographe où l'attention se porte petit à petit sur des groupes de lettres qui sont porteurs de sens, les morphèmes.

L'expérience de Sprenger-Charolles en 1992, exposée ci-dessous, montre cette évolution de l'installation des deux voies de lecture et le déroulement de l'apprentissage, et plus exactement le passage du stade alphabétique au stade orthographique. L'apparition des effets de lexicalité et de fréquence, montre le passage au stade orthographique des élèves

de CE1, puisque contrairement aux enfants de CP, les enfants de CE1 commencent à utiliser préférentiellement la voie orthographique, démontrée par les mots mieux lus que les non-mots (effet de lexicalité) et mieux lus que les mots irréguliers (effet de fréquence).

Les stades alphabétique & orthographique : expérience princeps de Sprenger-Charolles (1992)

- sujets : enfants de CP et CE1
- tâche : lire à voix haute différents items linguistiques écrits.
- matériel : listes d'items dans lesquelles on fait varier:
 - la *lexicalité*: mots *versus* non-mots ;
 - la *fréquence*: mots fréquents *versus* mots rares;
 - la *régularité*: mots réguliers *versus* mots irréguliers (par exemple, "*porte*" *versus*. "*pays*").
- Points observés : exactitude de la lecture (erreurs) et le temps de latence des réponses
- résultats : CP: Effet important de la régularité des mots. Pas d'effet de la lexicalité, ni de la fréquence.
CE1: baisse de l'effet de régularité. Apparition d'effets de lexicalité et de fréquence.

Le modèle en stade de Frith (1985) est à l'origine de nombreux tests d'évaluation de la lecture tels que le LMC-R de Khomsj, 1999 ou l'EVALEC de Sprenger-Charolles et al, 2005.

Enfin, ce modèle en stades représente bien le côté explicite de l'apprentissage de la lecture, mais son architecture est séquentielle et peu dynamique. Le modèle à double fondation de Seymour (1990,1997) ajoute l'hypothèse d'une certaine interactivité des processus de lecture tout au long de l'apprentissage.

1.2 Le modèle à « double fondation » de SEYMOUR (1997)

Selon Seymour, l'identification du mot écrit se fait à partir de deux processeurs (qui représentent la double fondation) interagissant avec la conscience linguistique :

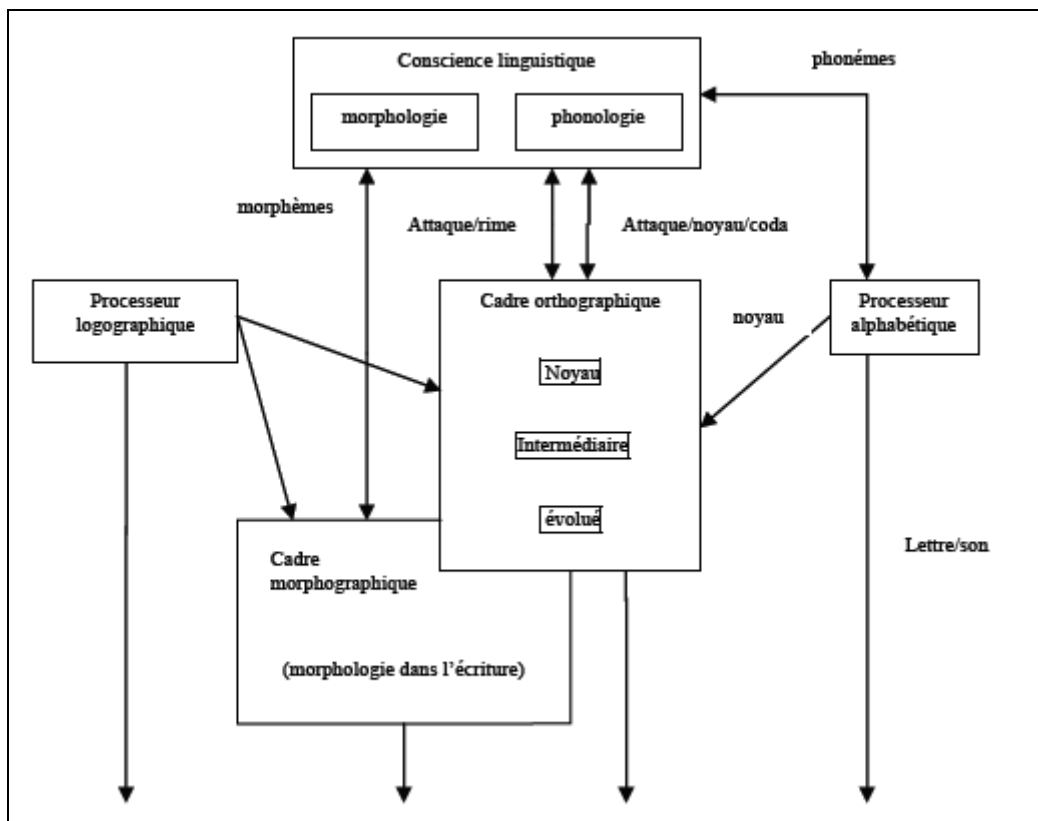
- le processeur logographique et le processeur alphabétique.

La mise en correspondance de l'oral et de l'écrit, les interactions entre le processeur alphabétique et la conscience linguistique permettent de faire émerger le cadre orthographique.

La conscience linguistique définit toutes les connaissances linguistiques du lecteur apprenant sur les différentes dimensions du mot telles que la phonologie, la morphologie etc. Le processeur alphabétique alimente la conscience linguistique et inversement.

Toutes ces interactions permettent la construction progressive du cadre orthographique qui lui-même continue à alimenter la conscience linguistique. Ce modèle fait donc apparaître clairement le côté dynamique de l'apprentissage de la lecture.

Le cadre orthographique se construit peu à peu en commençant par le stockage de petites unités orthographiques comme la syllabe (noyau): par exemple le « peu » de « chapeau » avec une certaine importance de la rime des mots. Puis atteint un niveau intermédiaire qui permet de lire des mots à deux syllabes, des homophones ou quelques mots irréguliers puis atteint le stade évolué vers le CE2 qui permet de lire des mots de plus longs et complexes.



Le modèle à « double fondation » de Seymour (1997),
Schéma issu du cours de Carine Royer 2012 – Master EERBEP UE1.

D'après ce modèle, le cadre morphologique, qui commence en écriture des mots, apparaît quand le cadre orthographique est suffisamment développé. A ce stade, l'enfant va prendre en compte la morphologie des mots, c'est dire qu'il est capable de repérer des morphèmes, porteurs de sens comme par exemple, comprendre que la fleuriste est quelqu'un qui vend des fleurs car il reconnaît le morphème « fleur » dans le mot « fleuriste ». Ces connaissances morphologiques sont d'ailleurs maintenant utilisées pour aider les dyslexiques.

Tous les cadres se construisent en interaction constante avec les connaissances linguistiques. Les uns nourrissant les autres.

Mais la principale critique de ce modèle porte sur le cadre morphologique qui d'après Seymour apparaît uniquement quand le cadre orthographique est suffisamment développé. Or, les recherches actuelles montrent que les enfants ont des connaissances morphologiques très tôt qui peuvent être utilisés plus précocement dans l'enseignement de la lecture à l'école.

De plus, les recherches n'ont pas pu démontrer le rôle du processeur logographique dans l'apprentissage de la lecture et donc au développement du cadre orthographique. C'est également le cas pour le modèle de Frith. Par contre, les études de Masonheimer et coll.1984, in Sprenger-Charolles et Colé, 2003, montrent que le processeur logographique n'est pas obligatoire ni généralisable.

En résumé, ces modèles en stades nous permettent de mieux comprendre le déroulement de l'apprentissage des mécanismes d'identification du mot écrit chez l'enfant. En effet, on perçoit un peu plus l'importance de chaque procédure pour lire un mot. En effet, chez l'apprenti lecteur, la procédure indirecte, permet de décoder les mots nouveaux par assemblage des graphèmes et des phonèmes. Mais il est essentiel de comprendre son rôle tenu dans le développement du lexique orthographique. C'est pourquoi nous continuons cet exposé par le rôle prépondérant du décodage dans ce développement.

2 Le rôle prépondérant du décodage pour la construction du lexique orthographique

Beaucoup d'études ont démontré l'importance de la procédure phonologique (voie indirecte) pour le développement du capital orthographique. En effet, cette voie permet de lire tous les mots réguliers mais aussi en partie les mots irréguliers. La lecture des mots irréguliers aboutit à des erreurs de régularisation comme par exemple le mot « aquarium » lu « acarium ». Ces erreurs peuvent être rectifiées par le lecteur par la récupération inférentielle d'un mot ressemblant dans le lexique oral.

On sait également que la lecture répétée d'un mot permet de l'inscrire dans le capital lexical orthographique et de le lire ultérieurement rapidement par la voie directe, Share, 1999.

Une étude longitudinale de Byrne & al, 1992, fait clairement apparaître que les lecteurs qui obtiennent au commencement de bons résultats en lecture de pseudo-mots sont ceux qui progressent le plus par la suite en lecture de mots réguliers et irréguliers.

Une autre étude à partir du français de Sprenger-Charolles & al, 1998b, indique que les lecteurs apprenants qui font un grand nombre d'erreurs de régularisation sur les mots irréguliers étaient ceux qui obtenaient le plus de réponses correctes et inversement. En effet, ces enfants utilisent essentiellement au début de l'apprentissage, la voie indirecte pour accéder aux mots. Cette étude met donc en relief le rôle crucial de la bonne maîtrise du décodage pour l'apprentissage de la lecture. En effet, le lecteur apprenant augmente petit à petit son capital lexical orthographique grâce à l'utilisation du décodage. Cet aspect va à l'encontre de stades séquentiels où le passage est déterminé par l'acquisition complète du stade antérieur. Le modèle de Seymour se montre plus pertinent quant aux interactions constantes entre le processeur alphabétique et l'augmentation du capital lexical. Ces recherches montrent bien l'absolue nécessité d'apprentissage des règles de conversions graphèmes-phonèmes.

Mais dans la construction de cette voie de lecture, un autre procédé est essentiel à l'élaboration de ce lexique orthographique aussi bien pour l'écriture que pour la lecture. Il s'agit de l'hypothèse de Goswami (1999) qui décrit la lecture par analogie, procédé fortement sollicité par les enseignants pour l'apprentissage de la lecture.

3. La lecture par analogie

Goswami, 1999, a développé l'hypothèse de la lecture par analogie lors de l'apprentissage de la lecture et de l'écriture. C'est-à-dire, l'utilisation de la rime pour lire et écrire. En effet, la rime représente l'unité naturelle la plus facile dans le développement de la conscience phonologique. De fait, le lecteur apprenant commence par utiliser des unités larges comme la syllabe pour écrire ou lire un autre mot. Par exemple, le «teau » de château.

Puis, dans un second temps, il sera capable de segmenter les rimes pour établir une correspondance entre la rime et sa représentation orthographique (le « eau » de château pour écrire « râteau ». Cette capacité à segmenter les rimes se révèle être le moteur dans le développement du capital lexical orthographique car il permet à l'apprenti lecteur d'être de plus en plus compétent pour lire ou écrire les mots. En effet, il y a une véritable réciprocity entre la lecture et l'écriture car par exemple, après avoir écrit le mot râteau par analogie avec le « eau » de château, l'enfant sera capable de reconnaître le « eau » dans n'importe quel autre mot.

L'effet moteur de la structuration progressive des représentations phonologiques les plus larges aux plus petites est le contraire de la construction du cadre orthographique du modèle à double fondation de Seymour, 1997.

En résumé, l'apprentissage la lecture dépend de la mise en place de procédures de lecture, et pour devenir lecteur habile, l'enfant doit posséder des mécanismes d'identification du mot écrit en élaborant notamment, lors de l'apprentissage, un assembleur phonologique efficace qui permet de lire tout ce qui peut être compris oralement. Cela suppose qu'il puisse faire des connexions entre les mots écrits et leur signification stockée en mémoire. Ces connexions sont possibles car les graphèmes simples ou complexes correspondent à des phonèmes ou à des syllabes. Mais, on constate d'énormes différences individuelles quant à l'apprentissage de la lecture entre les enfants. Il y a ceux qui n'éprouvent pas de difficulté particulières et ceux pour qui cet apprentissage est problématique. En conséquence, la compréhension de certaines difficultés d'identification du mot écrit est essentielle d'autant que ces difficultés peuvent aussi concerner des élèves de cycle 3. C'est pourquoi, nous essaierons d'effectuer une revue des différents apports de chercheurs nous permettant de mieux appréhender les difficultés qu'éprouvent certains élèves.

4. Comprendre certaines difficultés d'identification du mot écrit

La littérature et les recherches dans l'analyse et la compréhension des difficultés d'identifications du mot écrit sont nombreuses. Mais parmi celles-ci, les études comparatives entre différentes langues alphabétiques apportent un éclairage nouveau et peuvent fournir des informations essentielles à la compréhension de certaines difficultés de lecture.

4.1 L'apport des études inter-langues

4.1.1 Transparence et Complexité

Beaucoup de langues utilisent le système alphabétique, c'est-à-dire qu'elles sont basées sur un alphabet, pourtant on constate des grandes différences dans les codes grapho-phonologiques de ces langues. L'anglais et le français ont un système de conversion graphèmes-phonèmes ou phonèmes-graphèmes complexe et peu transparent. Par exemple,

en français, la rime « ille » se lit différemment dans « bille » ou dans « ville ». En anglais, la rime « ave » se prononce différemment dans « save » et « have ».

Par contre, l'italien et l'allemand ont un système de conversion graphèmes-phonèmes beaucoup plus simple. C'est-à-dire que chaque graphème a pour correspondance un phonème et un seul (dans la plupart des cas). On dit alors que ces langues sont soit transparente, régulière (italien et allemand), soit opaques ou irrégulières (anglais, français). L'apprentissage de la lecture suivant ces langues est par conséquent plus ou moins aisé. Le manque de transparence de la langue anglaise provoque des difficultés d'apprentissage pour les enfants anglais, beaucoup plus que pour les enfants italiens qui ont à apprendre à lire dans une langue dite régulière.

Seymour et al, 2003, ont présenté une étude qui consistait à vérifier le pourcentage de mots correctement lus à la fin de la première année d'apprentissage de la lecture. Les résultats sont les suivants : 1° Allemand 98% - 2° Français 79% - 3° Anglais 34%. Il n'y a pas de différences dues à l'âge de début d'instruction sur la lecture des mots et des pseudo-mots mais la différence est due au degré de transparence des différents codes orthographiques.

D'après Van den Bosch, Content, Daelemans et Gelder, 1994, le degré d'opacité d'un système orthographique se mesure suivant trois critères : 1) La complexité des correspondances entre les graphèmes et les phonèmes, 2) la diversité des graphèmes et la complexité qu'il y a à les segmenter au sein d'un mot, 3) l'étendue avec laquelle l'information morphologique et syntaxique est représentée au niveau de l'orthographe mais non de la phonologie.

En français, d'après Goetry 2002, pour la lecture des mots fréquents, on constate 130 graphèmes pour 35 phonèmes et dans une langue plus transparente comme le néerlandais, il y a 35 graphèmes pour 35 phonèmes.

Le français présente un code relativement transparent pour la lecture mais relativement opaque pour l'écriture. Mais malgré tout, il existe de nombreuses ambiguïtés qui peuvent considérablement gêner l'apprentissage. En effet, comment lire « ch » dans « hacher » et dans « orchidée ». En écriture, comment transcrire /e/ : ai, é, er,... ? Ces ambiguïtés peuvent être partiellement compensées par l'intervention d'autres régularités, notamment de type morphologique. Par exemple, « baigne » et « baignoire ». L'enfant doit également faire face au fait que le même graphème (-s-) ne correspond pas toujours au même phonème (bise, mis, anse) et inversement, le même phonème ne se réalise pas toujours avec le même

graphème (/o/ se transcrit par o, au, eau, ...). Ces ambiguïtés font l'objet de nombreuses recherches donc celles de Ziegler et al (1997) sur les effets de consistances orthographiques.

4.1.2 Effet de consistance

Ziegler et al, 1997, lors de recherches à partir du français ont pu mettre en évidence un **effet dit de consistance orthographique**, la consistance et à l'opposé l'inconsistance. L'observation des rimes montre que certaines sont consistantes comme le /yb/ qui s'écrit toujours -ube (tube, cube) et d'autres inconsistantes comme la rime /am/ qui peut s'écrire de plusieurs façons (dame, flamme, âme, femme). En lecture, l'ambiguïté de certains mots inconsistants peut provoquer des erreurs de lecture et surtout un fort ralentissement d'accès à la prononciation, ce qui le cas de beaucoup de mots irréguliers. En conclusion, les temps de lecture sont plus longs pour les mots inconsistants que pour les mots consistants. Ziegler et al 1996, montrent également que le français est relativement consistant sur le mode lecture (87,6%) et très inconsistant dans le mode écriture (79,1%).

Mais Ziegler et Al, 2005, ont développé une nouvelle théorie pour tenter d'expliquer les difficultés d'apprentissage de lecture plus importantes pour les enfants anglais par rapport à d'autres enfants des pays européens : la théorie de la granularité.

4.1.3 Théorie de la granularité (Ziegler et Goswami, 2005)

Pour ces auteurs, deux éléments peuvent expliquer ces difficultés : la taille des unités orthographiques et leur régularité. En effet, les unités de petites tailles telles que les phonèmes et les graphèmes sont plus ou moins régulières suivant les langues.

Dans le cas d'une langue très régulière comme l'allemand, ces petites unités le sont aussi (c'est-à-dire que les graphèmes correspondent exactement aux phonèmes) et réduisent alors considérablement le nombre de correspondances à apprendre (50 environ pour l'allemand). Mais dans le cas d'une langue très irrégulière comme l'anglais ou un peu moins comme le français, les petites unités ont des correspondances beaucoup moins régulières donc beaucoup plus nombreuses et leurs apprentissages deviennent très difficiles. C'est pourquoi, les enfants anglais ne peuvent pas décoder les mots en utilisant ces petites unités et sont obligés d'utiliser les unités beaucoup grandes comme les rimes, les syllabes ou les mots. En anglais, le nombre de correspondances se monte à environ 500, ce qui a pour conséquence un ralentissement des mécanismes de l'identification du mot écrit.

Cette théorie de la granularité, appelée aussi théorie du grain linguistique, pourrait permettre de comprendre pourquoi, les enfants anglais sont ceux qui décodent le moins bien en fin de première année d'apprentissage de la lecture.

En résumé, les unités telles que la syllabes jouent un rôle prépondérant dans l'apprentissage de la lecture dans une langue relativement irrégulière comme le français. Mais dans le cas de l'effet de fréquence, elle peut être source d'erreurs comme nous allons le voir maintenant en étudiant les effets de la fréquence du voisinage des mots à lire.

4.2 Les effets de la fréquence du voisinage orthographique et syllabique

Effets d'inhibition des voisins orthographiques et syllabiques :

4.2.1 L'effet des voisins orthographiques :

Le voisin orthographique est un mot qui partage toutes les lettres d'un mot à lire, sauf une lettre, positionnée à la même place. Par exemple, le mot chapeau est un voisin orthographique du mot « chameau » où seul la lettre « m » est changée par « P ».

D'après le modèle à activation interactive de McClelland et Rumelhart 1981, l'identification d'un mot écrit se réfère à l'activation de sa représentation mais aussi de ses voisins orthographiques et met en jeu un processus d'inhibition réciproque. En effet, plus les mots se ressemblent sur le plan orthographique, plus le degré d'activation de ces mots augmente. De plus, les mots fréquents ont des degrés d'activation plus robustes que les mots de basse fréquence. En conséquence, les voisins orthographiques plus fréquents que le mot cible seront plus fortement activés que le mot à lire et vont inhiber la reconnaissance de ce mot, rendant sa lecture plus difficile et plus lente.

4.2.2 L'effet des voisins syllabiques

Un voisin syllabique est un mot qui partage une syllabe avec le mot à lire comme par exemple « bateau » et « bataille ».

De façon équivalente, les voisins syllabiques fréquents d'un mot cible jouent un rôle inhibiteur sur la lecture de ce mot, ou plus exactement le nombre de voisins syllabiques plus fréquent que le mot à lire exerce un effet inhibiteur sur le mot cible. Dans notre exemple, tous les mots courants de deux syllabes commençant par « ba » produisent alors un effet de compétition entre les mots voisins qui partagent la même syllabe, surtout en début de mot. Ce niveau de compétition dépend du degré de fréquence du mot à lire et des

mots voisins. En conséquence, la prononciation de ce mot peut être erronée lorsque celui-ci n'est pas fréquent.

Mais un autre effet semble tout aussi essentiel pour tenter d'expliquer des erreurs de lecture de mots peu fréquents, il s'agit de l'effet de familiarité des mots.

4.3 Les effets de familiarité

Les bases de données de la fréquence lexicale sont basées sur les écrits et non pas sur la parole. Or, il est incontestable que certains mots fréquents ne sont pas familiers des enfants. C'est encore plus vrai pour ceux issus de milieux défavorisés ou quand le français est la langue de scolarisation ou langue seconde. En effet, le vécu de chacun de ces enfants est à l'origine des mots qui leur sont familiers. Le lieu de vie détermine également la familiarité des mots. Par exemple, la « murène » est un mot familier pour un enfant de Nouvelle Calédonie et beaucoup d'enfants connaissent ce poisson dangereux, pourtant le mot n'existe pas dans le dictionnaire de windows. On peut alors comprendre que les mots fréquents issus de l'écrit ne sont pas forcément familiers pour les enfants qui apprennent à lire. Les travaux de Gernsbacher, 1984 ont établi qu'un mot familier (surtout de basse fréquence) pour une personne ne le sera pas forcément pour une autre personne (exemple, le mot « étayage » plus familier pour un enseignant que pour d'autres corps de métier). Cette recherche indique donc que l'effet de familiarité prend plus d'importance que l'effet de fréquence dans la lecture de mots écrits chez le lecteur débutant.

Mais une autre situation particulière est à prendre en compte dans la compréhension des difficultés d'identification du mot écrit : il s'agit du français en tant que langue de scolarisation ou langue seconde. Sans entrer dans le détail des recherches existantes qui sont peu nombreuses, il est tout de même intéressant de citer une recherche qui éclaire un peu cette situation particulière.

4.4 La situation du français langue seconde

Les travaux de Geva & Siegel, 2000 évoquent la situation d'apprentissage de la lecture dans un milieu bilingue. En effet, le degré de complexité de chacune des deux langues est importante surtout si la langue seconde a un code orthographique plus régulière que la première. Geva et Siegel précisent que l'apprentissage de la lecture sera facilité dans le cas

d'une langue seconde régulière même si l'enfant a des compétences linguistiques insuffisamment développées dans la langue seconde.

De même, lors d'une étude d'enfants de la communauté française de Belgique, placés dans un contexte d'immersion en néerlandais, Philippe Mousty, Katia Lecocq, Régine Kolinsky, José Morais, & Jesus Alegria, 2007, avaient pour objectif de déterminer s'il y avait une influence de la langue seconde sur le développement des habiletés d'identification du mot écrit dans chacune des langues (français et néerlandais) apprises par ces enfants.

Ils ont montré que l'apprentissage de la lecture dans une langue seconde régulière comme le néerlandais, facilite l'apprentissage des processus de lecture en français. Cet effet positif a lieu uniquement si la langue première est moins régulière que la seconde. Il serait donc préférable d'apprendre à lire dans une des deux langues la plus régulière. Les compétences acquises dans cette langue sont transférables à la première.

Les résultats de cette recherche sont intéressants car l'épreuve de lecture des pseudo-mots permet d'évaluer si l'enfant est performant dans sa capacité à utiliser ou non la voie indirecte, d'assemblage, puisqu'il est impossible de les lire par la voie directe (mots jamais rencontrés ne faisant donc pas partie du capital lexical orthographique de l'enfant). Les enfants qui ont appris à lire en néerlandais ont donc mis en place une procédure indirecte dans cette langue presque aussi facilement que les enfants ne parlant que le néerlandais malgré l'insuffisance de développement des compétences linguistiques dans cette langue. Simplement, les auteurs de cette recherche montrent qu'il s'effectue un certain ralentissement d'apprentissage des habiletés d'identification du mot écrit la première année mais qui néanmoins aboutit à la mise en place de cette procédure lors de la deuxième année d'apprentissage. Par contre la lecture de mots indique que les enfants apprenant à lire en néerlandais ont plus de mal que les monolingues à développer la procédure d'adressage, c'est-à-dire la voie directe. Les auteurs émettent l'hypothèse que cette difficulté provient du manque de vocabulaire dans cette langue seconde (le néerlandais), difficulté corroborée par un effet de fréquence moins marqué que celui constaté dans les autres groupes évalués (enfants monolingues français ou néerlandais). Le manque de vocabulaire provoque un ralentissement de la lecture puisque les enfants ne peuvent récupérer le mot directement en mémoire lexical.

Mais les auteurs nuancent ces conclusions en disant que :

« Ces constats optimistes doivent toutefois être nuancés car, bien qu'ils ne soient pas moins bons que les enfants francophones, les enfants du groupe d'immersion ayant appris à lire d'abord en néerlandais ont plus de difficultés que les deux autres groupes francophones à traiter les graphèmes spécifiques du français. »

Philippe Mousty, Katia Lecocq, Régine Kolinsky, José Morais, & Jesus Alegria, 2007,
http://www.enseignement.be/download.php?do_id=5113&do_check=)

Toutes ces recherches que nous venons d'exposer concernent des difficultés pour l'apprenti lecteur tout venant, mais certains enfants ont plus de difficulté encore pour apprendre à lire. Il s'agit des enfants ayant un trouble spécifique du langage écrit : les dyslexiques.

4.5 La dyslexie

4.5.1 La dyslexie de surface

La dyslexie de surface ou morphémique (les mots) se définit par un dysfonctionnement de la voie directe. Elle renvoie à une lecture lente des mots mais aussi des pseudo-mots avec la présence de nombreuses erreurs de régularisation sur les mots irréguliers. Dans ce cas, la lecture s'effectue par la voie d'assemblage qui n'est pas encore tout à fait efficiente. Pour Frith, 1978, elle serait le résultat d'un blocage au niveau de la procédure alphabétique qui entraverait l'accession au stade orthographique.

4.5.2 La dyslexie phonologique

Ce trouble est très caractérisé par des difficultés d'ordre phonologique dominantes. En effet, la lecture des mots rares et des pseudo-mots est difficile et découle, toujours d'après le modèle stadiste de Frith, 1978, d'une difficulté à accéder à la procédure alphabétique d'assemblage.

Les hypothèses sur l'origine des troubles de l'apprentissage de la lecture sont nombreuses et J. ALEGRIA et P.MOUSTY, 2004, dans un article intitulé « Les troubles phonologiques et métaphonologiques chez l'enfant dyslexique » se sont intéressés plus particulièrement à l'hypothèse phonologique (Snowling, 2000). C'est-à-dire que les difficultés d'apprentissage de la lecture prendraient leurs origines dans le traitement phonologique et constituerait le déficit « source » de la dyslexie. En effet, l'ensemble des enfants dyslexiques ont des troubles qui portent sur un ou plusieurs aspects de la phonologie non spécifiques à l'apprentissage de la lecture telles que la dénomination d'images et la fluidité verbale.

L'empan faible en mémoire de travail phonologique à des répercussions négatives sur les réussites futures de l'enfant dans un grand nombre d'activités cognitives, spécialement en lecture (traitements de l'information réduites, difficultés dans le traitement de phrases et dans l'assemblage phonologique intervenant dans l'identification des mots écrits).

Mais d'après les recherches en Imagerie cérébrale, notamment celles de Stanislas DEHAENE, (chair de psychologie cognitive expérimentale au collègue de France), la dyslexie serait une anomalie précoce et généralement d'origine génétique. Pour **DEHAENE, 2003** « **la reconversion neuronale** » pourrait étayer l'hypothèse phonologique. Les acquisitions culturelles ne sont possibles que dans la mesure où elles s'inscrivent dans les possibilités de notre organisation cérébrale, en reconvertissant à un autre usage des prédispositions cérébrales déjà présentes (Dehaene 2003, 2008). Dans son cours sur les mécanismes cérébraux (Collège de France, page 302), Stanislas DEHAENE explique :

« Dans le cas de la lecture, trois grands ensembles de circuits cérébraux sont mis en jeu. Les réseaux de la reconnaissance visuelle invariante interviennent afin d'identifier la chaîne de caractères — ce sont ces circuits qui sont aujourd'hui les mieux connus sur le plan cérébral, et auxquels le modèle du recyclage neuronal s'applique le plus directement. Viennent ensuite la conversion des caractères écrits en une représentation phonologique et, en parallèle, l'accès au lexique et au sens des mots et des phrases. Chacun de ces réseaux se met en place avec l'apprentissage de la lecture, dépend en partie de l'organisation de l'écriture et de l'orthographe de la langue, et peut fait l'objet de déficits cognitifs chez l'adulte comme chez l'enfant. »

Mais également, **l'hypothèse d'un continuum** décrite par Ramus (2005) qui donne une origine génétique à la dyslexie, selon laquelle les différents troubles constatés sont explicités par des anomalies morphologiques. En effet, on constate que les aides basées sur la recherche de morphèmes connus dans un mot facilite l'accès à l'identification du mot chez l'enfant dyslexique, par exemple, le fait de reconnaître le mot « fleur » dans le mot « fleuriste » permet d'accéder à l'identification du mot « fleuriste ».

Résumé de la partie théorique

Nous avons vu dans un premier temps, d'un point de vue théorique, comment le lecteur expert procédait pour identifier les mots à la lumière de modèles robustes, puis dans un deuxième temps, comment se déroulait l'installation des mécanismes d'identification du mot écrit et des difficultés qui pouvaient en découler. Les travaux que nous avons cités, ont clairement démontré que le lecteur utilise deux voies pour reconnaître un mot, c'est-à-dire la voie d'assemblage qui permet de décoder tous les mots et la voie d'adressage qui permet d'accéder directement au mot. De plus, les études inter-langues (européennes) donnent un éclairage particulier sur certaines difficultés d'apprentissage que rencontrent les enfants. En effet, plus une langue est irrégulière, plus l'apprentissage sera long et difficile. Cela s'explique par l'analyse de différents effets. Parmi ceux-ci, nous avons choisi, en rapport avec l'objet de notre étude, les effets de transparence, de consistance, de l'influence des voisins orthographiques et syllabiques ainsi que l'effet de familiarité. Cependant, comme notre étude porte sur des enfants originaires de Nouvelle Calédonie où la langue française est soit première soit langue de scolarisation, nous avons présenté des travaux décrivant l'apprentissage de la lecture d'enfants en immersion dans la langue française. Ces travaux montrent que si la langue d'apprentissage est plus transparente que la langue première, l'apprentissage en sera facilité. Enfin, les troubles spécifiques de la langue écrite nous ont paru nécessaires à la présentation des difficultés rencontrés par certains enfants.

Bibliographie

Alegria J. et Mousty P. (2004). Les troubles phonologiques et métaphonologiques chez l'enfant dyslexique. *Enfance*, N° 3, P. 259 à 271.

Bonin P., Collay S et Fayol M. (2008). La consistance orthographique en production verbale écrite : une brève synthèse. *L'année Psychologique* 108- P. 517 à 546.

Dehaene S. (2007). *Les Neurones de la lecture*. Paris. Odile Jacob

Dehaene S. (2011). *Apprendre à lire – Des sciences cognitives à la salle de classe*. Odile Jacob Sciences.

Dehaene S. (2007). *L'apprentissage de la lecture et ses perturbations - Collège de France*. Cours N° 1.

<http://www.college-de-france.fr/site/stanislas-dehaene/#|m=course|q=../stanislas-dehaene/course-2006-2007.htm|p=../stanislas-dehaene/course-2007-01-08-09h30.htm>

Dehaene S. (2007). *L'apprentissage de la lecture et ses perturbations - Collège de France*. Cours N° 4.

<http://www.college-de-france.fr/site/stanislas-dehaene/#|m=course|q=../stanislas-dehaene/course-2006-2007.htm|p=../stanislas-dehaene/course-2007-01-29-09h30.htm>

Dehaene S. (2007). *L'apprentissage de la lecture et ses perturbations - Collège de France*. Cours N° 6

<http://www.college-de-france.fr/site/stanislas-dehaene/#|m=course|q=../stanislas-dehaene/course-2006-2007.htm|p=../stanislas-dehaene/course-2007-02-12-09h30.htm>

Delahaie M., Sprenger-Charolles L. Serniclaes W. (2007). Effet de lexicalité chez les faibles et très faibles lecteurs comparativement à des normolecteurs de même âge et de même niveau de lecture. *L'année psychologique* 107. P.361 à 396.

Ecalte J. (2010). L'évaluation de la lecture et des compétences associées. *Revue Française de linguistique appliquée* 15.1. P.105 à 129.

Ecalte J. et Magnan A. (2010). *L'apprentissage de la lecture et ses difficultés*. Paris. Dunod.

Ferrand L. (2007). *Psychologie cognitive de la lecture. Reconnaissance des mots écrits chez l'adulte*. Ouvertures psychologiques. De Boeck.

Gombert J.E. , Colé P., Valdois S., Goigoux R., Mousty P. et Fayol M. (2002). *Enseigner la lecture au cycle 2*. Nathan Pédagogie.

Jamet E. (1997). Lecture et réussite scolaire. Paris. Dunod.

Jacquier-Roux, M., Valdois, S. & Zorman, M. (2005). *ODEDYS : Outil de dépistage des dyslexies*. Grenoble, Laboratoire Cogni-Sciences.

Khomsi, A. (1999). *Epreuve d'évaluation de la compétence en lecture*. Paris, ECPA

Khomsi A. (1997) ECS III Evaluation des compétences scolaires au cycle III. Paris. ECPA

<http://www.ecpa.fr/orthophonie/test.asp?id=1623>

Lecocq K. , Mousty P., Kolinsky R. , Goetry V. Morais J., Alegria J. (2006). Évaluation de programmes d'immersion en Communauté française: une étude longitudinale comparative du développement de compétences linguistiques d'enfants francophones immergés en néerlandais – Université libre de Bruxelles.

www.enseignement.be/download.php?do_id=1906&do_check=

Mousty, P., Leybaert, J., Alegria, J., Content, A. & Morais, J. (1994). BELEC : une batterie d'évaluation du langage écrit et de ses troubles. In J. Grégoire & B. Piérart (éds), *Evaluer les troubles de la lecture. Les nouveaux modèles théoriques et leurs implications diagnostiques*, Bruxelles, De Boeck.

National Reading Panel. (2000). Teaching children to read : An evidence-based Assesment of the scientific Research literature on reading and its implications for reading instruction.

<http://www.nationalreadingpanel.org/Publications/publications.htm>

Ramus F. (2008). Les troubles d'apprentissage de la lecture ? Les synergies et partenariats créés, l'apport de l'internet. De Boeck Supérieur. Perspectives en Education et formation. Chapitre 9. P. 171 à 192.

Sprenger-Charolles L. (1994). L'acquisition de la lecture en français : étude longitudinale de la première à la seconde année du primaire. *L'Année psychologique*, Année 1994, Volume 94, Numéro 4. p. 553 – 574

Sprenger-Charolles L. et Serniclaes W. (2003) Acquisition de la lecture et de l'écriture et dyslexie : Revue de littérature. *Revue française de linguistique appliquée* 2003/1 - Vol. VIII pages 63 à 90 - ISSN 1386-

Van Grunderbeeck N. (1998). Les difficultés en lecture, diagnostic et pistes d'intervention. Gaétan Morin.

Persée. Portail de revues scientifiques en sciences humaines www.persee.fr/

Cairn. Revues de sciences humaines et sociales en texte intégral www.cairn.info/