



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



MEMOIRE présenté pour l'obtention du
CERTIFICAT DE CAPACITE D'ORTHOPHONISTE

Par

DELEUZE Anaïs
MURIAN Élodie

ERREURS ORTHOGRAPHIQUES ET PARAMÈTRES
EN TEMPS RÉEL :

Analyse de textes d'élèves de CM2, 5^e et 3^e

Maître de Mémoire

FAYOL Michel

Membres du Jury

BALDY MOULINER Florence

CHOSSON Christelle

GUILLON-INVERNIZZI Fanny

Date de Soutenance

30 Juin 2011

ORGANIGRAMMES

1. Université Claude Bernard Lyon1

Président
Pr. BONMARTIN Alain

Vice-président DEVU
Pr. SIMON Daniel

Vice-président CA
Pr. ANNAT Guy

Vice-président CS
Pr. MORNEX Jean-François

Directeur Général des Services
M. GAY Gilles

1.1 Secteur Santé :

U.F.R. de Médecine Lyon Est
Directeur **Pr. ETIENNE Jérôme**

U.F.R d'Odontologie
Directeur **Pr. BOURGEOIS Denis**

U.F.R de Médecine Lyon-Sud
Charles Mérieux
Directeur **Pr. GILLY François
Noël**

Institut des Sciences Pharmaceutiques
et Biologiques
Directeur **Pr. LOCHER François**

Institut des Sciences et Techniques de
Réadaptation
Directeur **Pr. MATILLON Yves**

Comité de Coordination des
Etudes Médicales (C.C.E.M.)
Pr. GILLY François Noël

Département de Formation et Centre
de Recherche en Biologie Humaine
Directeur **Pr. FARGE Pierre**

1.2 Secteur Sciences et Technologies :

U.F.R. de Sciences et Technologies
Directeur **Pr GIERES François**

IUFM
Directeur **M. BERNARD Régis**

U.F.R. de Sciences et Techniques
des Activités Physiques et
Sportives (S.T.A.P.S.)
Directeur **Pr. COLLIGNON Claude**

Ecole Polytechnique Universitaire de
Lyon (EPUL)
Directeur **M. FOURNIER Pascal**

Institut des Sciences Financières et
d'Assurance (I.S.F.A.)
Directeur **Pr. AUGROS Jean-Claude**

Ecole Supérieure de Chimie Physique
Electronique de Lyon (CPE)
Directeur **M. PIGNAULT Gérard**

Observatoire Astronomique de
Lyon **M. GUIDERDONI Bruno**

IUT LYON 1
Directeurs **M. COULET Christian et
Pr. LAMARTINE Roger**

2. Institut Sciences et Techniques de Réadaptation FORMATION ORTHOPHONIE

Directeur ISTR
Pr. MATILLON Yves

Directeur de la formation
Pr. TRUY Eric

Directeur des études
BO Agnès

Directeur de la recherche
Dr. WITKO Agnès

Responsables de la formation clinique
THEROND Béatrice
GUILLON Fanny

Chargée du concours d'entrée
PEILLON Anne

Secrétariat de direction et de scolarité
BADIOU Stéphanie
CLERGET Corinne

REMERCIEMENTS

Nous remercions tout d'abord nos lectrices, Florence Baldy-Moulinier et Christelle Chosson, pour la justesse et la précision de leurs remarques. Leur présence à notre soutenance, ainsi que celle de Fanny Guillon-Invernizzi, nous honore.

Agnès Witko, dont l'énergie et la disponibilité valorisent nos travaux de recherche, reçoit toute notre gratitude, tout comme son « messenger » au sein de notre promotion, notre future consœur Angélique Raffara.

Nous remercions également Florence Chenu et Bernard Lété qui, dans les premiers temps de notre travail, ont su nous initier avec patience et pédagogie à certains aspects techniques.

L'élaboration de notre réflexion et l'analyse de nos résultats n'auraient pu se faire sans l'aide inestimable de Séverine Maggio, doctorante en psychologie à l'Université Blaise Pascal de Clermont-Ferrand. En témoignent nos 183 courriels échangés... Elle sera c'est certain une digne représentante de la recherche et une excellente professeure. Notre collaboration restera un très bon souvenir. Jusqu'à la prochaine ?

Ce parcours initiatique, jalonné par ce mémoire, qu'auront été nos études, n'aurait pu se faire sans soutien. Nous remercions nos familles et nos amis, français comme québécois, pour leur présence et leur patience sans faille.

Notre année d'études à l'Université de Montréal et notre découverte du Québec auront su tisser les liens d'une amitié indéfectible. Les heures de réflexion, d'échanges et de rédaction sont autant d'heures passées à se découvrir et à grandir, avec ce beau Québec en commun.

Élodie, ces deux années, charnières de nos vies personnelle et professionnelle, n'auraient pas eu la même valeur sans ta présence et ton amitié. Aie confiance en toi, ta valeur est certaine.

Anaïs, écrire, grandir, effacer, partager, réécrire, collaborer, apprendre de toi et sentir que rien ne sera plus comme avant notre rencontre... Liées même au-delà d'un océan. Pour tout ça et bien plus encore, merci pour tout.

Enfin, que dire qui n'aurait déjà été évoqué à propos de Michel Fayol ? Qu'il est un chercheur reconnu ? Tout le monde le sait. Un excellent professeur ? Quelle évidence ! Un maître de mémoire exemplaire ? Cela a été dit cent fois... Mais avant tout, Monsieur Fayol, vous aurez été pour nous un guide attentif, respectueux et valorisant, un maître encadrant et présent, un amoureux du Québec tout comme nous. Travailler avec vous aura été un véritable plaisir et un grand honneur.

« Les meilleurs professeurs sont ceux qui savent se transformer en ponts et qui invitent leurs élèves à les franchir. »

Nikos Kazantzakis

SOMMAIRE

| | |
|---|-----------|
| ORGANIGRAMMES | 2 |
| 1. <i>Université Claude Bernard Lyon I</i> | 2 |
| 1.1 <i>Secteur Santé</i> : | 2 |
| 1.2 <i>Secteur Sciences et Technologies</i> : | 2 |
| 2. <i>Institut Sciences et Techniques de Réadaptation FORMATION ORTHOPHONIE</i> | 3 |
| REMERCIEMENTS..... | 4 |
| SOMMAIRE..... | 5 |
| INTRODUCTION..... | 7 |
| PARTIE THEORIQUE | 8 |
| I. LES ERREURS ORTHOGRAPHIQUES | 9 |
| 1. <i>Acquisition de l'orthographe</i> | 9 |
| 2. <i>Statut de l'erreur</i> | 12 |
| 3. <i>Étude des erreurs orthographiques</i> | 14 |
| II. LA PRODUCTION ECRITE DE TEXTE..... | 16 |
| 1. <i>Description des processus en jeu dans la production écrite</i> | 16 |
| 2. <i>Interactions entre les processus rédactionnels</i> | 18 |
| 3. <i>Analyse en temps réel de la production écrite</i> | 20 |
| PROBLEMATIQUE ET HYPOTHESES..... | 23 |
| I. PROBLEMATIQUE | 24 |
| II. HYPOTHESES..... | 25 |
| 1. <i>Hypothèse générale</i> | 25 |
| 2. <i>Hypothèses opérationnelles</i> | 25 |
| PARTIE EXPERIMENTALE | 26 |
| I. METHODE DE RECUEIL DES DONNEES..... | 27 |
| 1. <i>Participants</i> | 27 |
| 2. <i>Matériel</i> | 27 |
| 3. <i>Procédure expérimentale</i> | 28 |
| II. CONSTITUTION DU CORPUS..... | 29 |
| 1. <i>Corpus général</i> | 29 |
| 2. <i>Corpus des erreurs orthographiques : constitution des critères d'analyse</i> | 30 |
| PRESENTATION DES RESULTATS | 33 |
| I. DONNEES DESCRIPTIVES DU CORPUS..... | 34 |
| II. COMPARAISON DES MOTS AVEC OU SANS ERREUR(S) ORTHOGRAPHIQUE(S) | 35 |
| 1. <i>Débit d'écriture</i> | 35 |
| 2. <i>Pauses</i> | 36 |
| 3. <i>Paramètres en temps réel et nature des mots</i> | 39 |
| III. COMPARAISON DES DIFFERENTS TYPES D'ERREURS ORTHOGRAPHIQUES | 39 |
| 1. <i>Données descriptives</i> | 39 |
| 2. <i>Débit d'écriture</i> | 41 |
| 3. <i>Pauses</i> | 42 |
| 4. <i>Révisions</i> | 45 |
| DISCUSSION DES RESULTATS..... | 47 |
| I. RAPPELS DES DONNEES DE L'ETUDE | 48 |
| 1. <i>Objectif</i> | 48 |
| 2. <i>Hypothèses</i> | 48 |
| 3. <i>Méthode</i> | 48 |
| II. INTERPRETATION DES RESULTATS..... | 48 |
| 1. <i>Données générales : fluence, pourcentages d'erreurs et pourcentages de types d'erreurs par niveau de classe</i> | 48 |
| 2. <i>Comparaison des pourcentages de types d'erreurs par niveau scolaire</i> | 49 |

| | | |
|------|--|-----------|
| 3. | <i>Paramètres en temps réel et erreur orthographique</i> | 51 |
| III. | BIAIS ET LIMITES DE L'ETUDE | 57 |
| 1. | <i>Population</i> | 57 |
| 2. | <i>Création d'une typologie</i> | 57 |
| 3. | <i>Analyse</i> | 58 |
| IV. | PERSPECTIVES, APPORTS PROFESSIONNELS ET PERSONNELS | 58 |
| 1. | <i>Perspectives de recherche</i> | 58 |
| 2. | <i>Apports pour la profession d'orthophoniste</i> | 59 |
| 3. | <i>Apports personnels</i> | 60 |
| | CONCLUSION | 61 |
| | BIBLIOGRAPHIE | 63 |
| | GLOSSAIRE | 68 |
| | ANNEXES | 69 |
| | LISTE DES ANNEXES | 70 |
| | ANNEXE I : INFORMATIONS DISPONIBLES AU SEIN DES TABLEAUX D'ANALYSE..... | 71 |
| 1. | <i>Données concernant le mot N</i> | 71 |
| 2. | <i>Données concernant le mot N+1</i> | 73 |
| 3. | <i>Données concernant le mot N-1</i> | 73 |
| | ANNEXE II : PRODUCTIONS ET ANALYSE DES ERREURS ORTHOGRAPHIQUES D'UN ELEVE DE CM2 | 74 |
| 1. | <i>Brouillon du texte expositif</i> | 74 |
| 2. | <i>Brouillon du texte narratif</i> | 74 |
| 3. | <i>Analyse des erreurs</i> | 75 |
| | ANNEXE III : PRODUCTIONS ET ANALYSE DES ERREURS ORTHOGRAPHIQUES D'UN ELEVE DE 5 ^E | 76 |
| 1. | <i>Brouillon du texte expositif</i> | 76 |
| 2. | <i>Brouillon du texte narratif</i> | 76 |
| 3. | <i>Analyse des erreurs</i> | 77 |
| | ANNEXE IV : PRODUCTIONS ET ANALYSE DES ERREURS ORTHOGRAPHIQUES D'UN ELEVE DE 3 ^E | 79 |
| 1. | <i>Brouillon du texte expositif</i> | 79 |
| 2. | <i>Brouillon du texte narratif</i> | 80 |
| 3. | <i>Analyse des erreurs</i> | 81 |
| | TABLE DES ILLUSTRATIONS | 83 |
| | TABLE DES MATIERES | 85 |

INTRODUCTION

Le langage écrit, à la fois dépendant d'aptitudes générales (les mêmes qui régissent le langage oral) et spécifiques, représente un enjeu crucial des apprentissages scolaires. Son acquisition oblige à la maîtrise combinée de nombreuses compétences, dont celle de l'orthographe.

En français, l'apprenti-scripteur ne peut s'appuyer sur la seule application du principe alphabétique pour orthographier tous les mots de la langue. Ainsi, l'inconsistance des correspondances phonèmes-graphèmes oblige à la mobilisation d'autres compétences. Il doit donc faire appel à des connaissances lexicales (régularités graphotactiques et unités du lexique orthographique) et morphologiques (flexionnelles et dérivationnelles). Les erreurs orthographiques (généralement catégorisées en erreurs phonologiques, lexicales et morphologiques) sont le reflet de ces apprentissages, de leur intrication et de leurs conflits.

L'orthographe n'est en fait qu'un sous-processus de l'activité complexe qu'est la production écrite. Le scripteur doit en effet gérer de nombreuses contraintes afin de planifier, mettre en texte et réviser son énoncé. L'automatisation progressive et partielle de certains processus (orthographe et graphisme) permet le développement d'autres compétences, dites de haut niveau (telle la planification) qui ne pourraient se dérouler si les premières étaient trop coûteuses.

L'utilisation récente des tablettes graphiques (et de logiciels d'analyse, tel Eye and Pen) au sein de la recherche sur la production écrite a permis de matérialiser, du moins en partie, la dynamique de celle-ci. Les interactions entre les processus rédactionnels sont ainsi mieux décrites. L'analyse des paramètres en temps réel de la production écrite que sont le débit d'écriture et les pauses constituent en effet autant d'indices des coûts cognitifs associés à ces procédures.

Une analyse dynamique de la production orthographique, c'est-à-dire étudiée à la lumière des paramètres temporels qui lui sont associés, pourrait permettre de définir, de façon plus écologique, les processus en jeu et leur coût cognitif.

Nous nous intéressons ici aux caractéristiques des paramètres temporels de la production écrite de texte (débit et pauses) associées aux erreurs orthographiques. La finalité de cette étude est de déterminer si et dans quelle mesure, ces paramètres en temps réel reflètent le coût associé au traitement de l'erreur orthographique.

Après avoir étudié les étapes de l'acquisition de l'orthographe et le statut de l'erreur, nous présenterons un état des données concernant la production écrite, sa dynamique et son analyse.

Au regard de ces connaissances théoriques, nous présenterons l'étude menée ici, qui a consisté en l'analyse des erreurs orthographiques d'élèves de CM2, 5^e et 3^e produites lors de la rédaction libre de textes. Classées selon une typologie créée pour les besoins de cette recherche, ces erreurs ont ensuite été analysées au regard des paramètres en temps réel leur étant associés.

Chapitre I
PARTIE THEORIQUE

I. Les erreurs orthographiques

L'orthographe fait l'objet de nombreux débats, depuis la création de l'Académie française jusqu'aux différentes tentatives de réforme. À ce propos, Jaffré (2006) explique que :

« L'orthographe présente une nature conservatrice [...] d'ailleurs renforcé[e] par le statut social de la norme orthographique qui rend la plupart des usagers réfractaires à toute modification. Cela d'autant plus que, dans les sociétés à tradition écrite, la maîtrise réelle ou supposée de cette norme confère du pouvoir à celui qui en dispose. ».

Déjà, Catach, dans son ouvrage Les délires de l'orthographe (1989), déplorait le trop grand intérêt porté aux normes formelles de l'orthographe, allant même jusqu'à proposer de s'opposer à la philologie (étude de la grammaire et de la linguistique des textes) pour s'adonner à la « philographie », le plaisir libéré d'utiliser et de transformer la langue écrite.

Ainsi, le scripteur, qu'il soit novice ou expert, doit se conformer à des normes orthographiques, se voit confronté au quotidien au doute et risque d'être stigmatisé au regard de son niveau en orthographe (Rey, Pacton & Perruchet, 2005).

En ce sens, l'orthographe est, depuis longtemps, un objet de recherche. Des étapes de son acquisition ont notamment été décrites. Toute étude visant à l'analyse de productions orthographiques, et tout particulièrement d'enfants, ne saurait faire l'économie de ces connaissances.

1. Acquisition de l'orthographe

De nombreux modèles descriptifs de l'acquisition de la lecture et de l'orthographe ont été élaborés (David, 2003 ; Frith, 1980, 1985 ; Gentry, 1978, 1982). Même si le découpage des étapes (ou stratégies) peut varier, Content et Zesiger (1999, p.193) remarquent que ces modèles s'accordent sur le fait qu'il existe « une séquence impliquant une phase d'orthographe « pré-communicative » ou « symbolique », suivie d'une phase de développement de l'orthographe phonétique impliquant la médiation phonologique puis d'une phase d'acquisition de l'orthographe conventionnelle (orthographe lexicale) ».

Fayol (2010) distingue quant à lui trois grands types d'acquisitions indispensables à l'accès à une orthographe experte : l'apprentissage du principe alphabétique, les apprentissages dits « orthographiques » et les apprentissages morphologiques.

1.1. Apprentissage du principe alphabétique

L'apprentissage des langues alphabétiques, tel le français, implique la compréhension d'un principe majeur, le principe alphabétique, selon lequel les séquences

de lettres (constituées de graphèmes) entretiennent des correspondances régulières (mais parfois complexes) avec les séquences sonores (constituées de phonèmes).

Au sein d'un système alphabétique totalement transparent, où à chaque graphème correspondrait un phonème, la seule intégration du principe alphabétique permettrait idéalement de lire et d'écrire tous les mots de la langue. Ainsi, pour certaines langues, dont l'orthographe est dite « transparente » (comme l'italien, l'espagnol ou encore le finnois), la procédure de simple appariement graphèmes-phonèmes s'avère généralement efficace (Jaffré, 2008, p.89). Sa connaissance suffit à assurer la réussite en lecture et en écriture.

La situation de la langue française est particulière, puisque là où les relations graphèmes-phonèmes (lecture) sont relativement régulières les relations phonèmes-graphèmes (orthographe) sont, à l'inverse, plus inconsistantes (irrégulières). Ainsi, en français, dont l'orthographe est dite « opaque » (Jaffré, 2008, p.89), la seule utilisation d'une procédure de conversion s'avère insuffisante et ne permettrait, selon Véronis (1988), de transcrire qu'un mot sur deux de la langue.

La capacité de conversion phonèmes-graphèmes permet l'accès à une orthographe dite « phonologique », première étape, imparfaite, de l'acquisition de l'orthographe. Elle s'avère pourtant indispensable puisqu'elle marque l'acquisition de régularités, auxquelles viennent ensuite s'ajouter les connaissances propres aux irrégularités de la langue. Des apprentissages d'ordre orthographique et d'autres d'ordre morphologique permettent d'orthographier les mots irréguliers, là où la médiation phonologique seule s'avère insuffisante (Fayol, 2010).

1.2. Apprentissages orthographiques

1.2.1. Acquisition des connaissances orthographiques générales : les régularités graphotactiques

Au contact de l'écrit, l'apprenti scripto-lecteur repère et mémorise inconsciemment nombre de régularités propres à cette modalité. L'enfant devient ainsi implicitement (c'est-à-dire intuitivement et inconsciemment) sensible aux probabilités, plus ou moins fortes, qu'à des séquences de lettres figées correspondent des configurations sonores précises. Il ne se limite plus à la compréhension du principe alphabétique pour orthographier, mais repère et applique certaines associations de graphèmes (pouvant inclure des lettres muettes et des irrégularités) dont l'orthographe est stable et qui sont assez fréquentes pour être mémorisées comme des unités à part entière (comme les suffixes -able, -age, -ienne, etc.).

Ces connaissances implicites permettent à l'enfant (et à l'adulte) de préférer instinctivement certaines configurations de lettres ou positions à d'autres statistiquement peu fréquentes ou absentes en français (Pacton, Perruchet, Fayol & Cleeremans, 2001, cité par Fayol, 2010 ; Pacton, Fayol & Perruchet, 2005). Elles permettent aussi d'orthographier de façon plausible des mots, même nouveaux.

1.2.2. Apprentissages des connaissances lexicales spécifiques : constitution du lexique orthographique

L'exposition toujours plus grande de l'enfant apprenti-scripteur à l'écrit lui permet de repérer des invariants orthographiques, mais aussi de constituer, peu à peu, son lexique orthographique. L'existence d'un effet de fréquence (plus les mots ont été rencontrés souvent, plus leur orthographe sera mémorisée) et d'un effet d'analogie (un mot nouveau peut s'orthographier bien souvent par référence à des mots déjà connus) en attestent l'existence et la construction progressive.

Selon les modèles à double voie (Baron, 1973 ; Coltheart, 1978 ; Forster, 1976 ; Paap & Noël, 1991 ; Seidenberg, 1985 ; cités par Ecalle & Magnan, 2002), l'utilisation de la voie d'adressage, permettant une récupération directe de la forme orthographique, s'avère plus automatique, rapide et précise que l'application séquentielle des règles de conversion phonèmes-graphèmes (voie d'assemblage). Sprenger-Charolles et Casalis (1996) précisent à ce propos qu'à la fin de la première année du primaire, la procédure par médiation phonologique coexiste avec le traitement orthographique, aussi bien en lecture qu'en écriture.

1.3. Apprentissages morphologiques

La morphologie, branche de la linguistique étudiant la façon dont les morphèmes, unités minimales de sens, se combinent pour former des mots (ou plus précisément des lemmes), est classiquement scindée en deux dimensions : la morphologie dérivationnelle et la morphologie flexionnelle.

1.3.1. Acquisition de la morphologie dérivationnelle

La morphologie dérivationnelle consiste en la combinaison de morphèmes racines et d'affixes (suffixes et préfixes) dans le but de produire des mots nouveaux. Elle est en cela un moyen d'accroître le lexique.

En français, son utilisation permet également souvent de guider le scripteur qui, grâce à ses connaissances morphologiques, pourra décider de la terminaison d'un mot (ex. lourd prend un « d » final parce qu'il appartient à la même famille de mots que lourdeur et lourdement). Pourtant, même si Sénéchal, Basque et Leclaire (2006, cités par Fayol, 2010) ont montré un impact de la morphologie sur la production orthographique dès le CM1 (9-10 ans), ces résultats sont à pondérer dans la mesure où il est souvent difficile de faire la part entre application de connaissances morphologiques et recours à des régularités graphotactiques (ex. des différentes graphies de /o/ ; Pacton et al., 2005). De plus, l'utilisation de connaissances morphologiques peut parfois mener à des surgénéralisations orthographiques (ex. numérot* et abrit* ; Pacton & Casalis, 2006).

1.3.2. Acquisition de la morphologie flexionnelle

La morphologie flexionnelle concerne quant à elle les variations de la forme des unités lexicales en fonction de facteurs grammaticaux. Ces variations sont traditionnellement catégorisées en deux types de flexions : la déclinaison (phénomènes d'accord en nombre et en genre des noms, adjectifs, pronoms et verbes) et la conjugaison (marques des modes et temps des verbes).

Comme c'est le cas pour la morphologie dérivationnelle, l'impact de la morphologie flexionnelle sur l'acquisition des compétences orthographiques est à relativiser, puisqu'elle se trouve bien souvent en concurrence avec d'autres stratégies, notamment graphotactiques. Son utilisation, sa connaissance et son enseignement explicite prennent en revanche tout leur sens pour ce qui est des marques flexionnelles (marques du pluriel) dont les séquences graphémiques ne trouvent pas de correspondances phonémiques (ex. l'enfant court / les enfants courent).

L'application de la morphologie peut se faire, en français, de deux façons. D'une part, la récupération directe de formes orthographiques mémorisées fluidifient la rédaction (Cousin, Largy & Fayol, 2002 ; Largy, Cousin, Bryant & Fayol, 2007, cités par Fayol, 2010 ; Largy, Fayol & Lemaire, 1996). D'autre part, l'application explicite de procédures (application des règles enseignées), bien que coûteuse en attention, constitue au besoin une procédure de recours (Fayol, Thévenin, Jarousse & Totereau, 1999) lorsque la compétition entre procédure et récupération (Logan, 1988) induit des erreurs.

2. Statut de l'erreur

2.1. Statut social et pédagogique de l'erreur

Aujourd'hui encore, le terme de « faute » reste couramment utilisé, dans l'enseignement comme dans le langage courant. Or, celui-ci renvoie étymologiquement en latin à l'action de faillir et originellement « au fait de manquer aux prescriptions d'une religion [...] ou à la règle morale » (Rey, 1995, p.782). Considérant ce terme comme emprunt d'une gravité excessive, les pédagogues préfèrent aujourd'hui utiliser le terme d'erreur, qui renvoie quant à lui à l'« action d'errer çà et là » (Rey, 1995, p.717) et semble moins connoté négativement.

Cette évolution s'inscrit dans une réflexion plus générale de la recherche en pédagogie sur la place de l'évaluation et donc de la prise en compte des erreurs dans le processus d'apprentissage. Astolfi (1997) comme Reboul (1980) reconnaissent ainsi le caractère constructif et dynamisant des erreurs dans l'apprentissage. Rey et al. (2005, p.103) affirment en outre que « l'erreur permet [...] de déterminer les procédures et stratégies déjà acquises et celles qui demeurent encore instables ou en cours d'acquisition ».

En ce sens, les textes officiels de l'Éducation Nationale préconisent désormais l'usage d'une correction constructive « donnant [des] réponses ou éclaircissements, valorisant [le] travail et [les] efforts [de] l'enfant [...] [et] se démarquant très nettement

d'une correction-sanction (celle qui comptabilise les « fautes » et les observations négatives) » (Babin, 2003, p.80).

L'enseignant comme le rééducateur, tel l'orthophoniste, est donc amené à relever, analyser et utiliser les erreurs (notamment d'orthographe) comme autant d'éléments permettant l'élaboration d'une démarche d'aide soutenant l'enfant dans son apprentissage.

2.2. Intérêt et analyse de l'erreur orthographique dans la recherche scientifique

Parce qu'elles présentent des régularités et qu'elles peuvent rendre compte de différents niveaux de traitement cognitif, les erreurs de production verbale constituent une source d'information intéressante (Bonin, 2003). L'un des principaux postulats de la recherche dans le domaine de l'orthographe est que les erreurs orthographiques commises sont le reflet des difficultés de conversion phonèmes-graphèmes, de la connaissance des formes orthographiques figées de la langue et des règles d'accord (Fayol, 2008).

L'analyse des erreurs d'orthographe suppose leur recueil et leur classification. En 1956, Simon (cité par Bellone, 2003) propose une épreuve de dictée et choisit d'étudier les erreurs orthographiques selon une typologie dite « R.U.P » :

- « règles » : erreurs d'accord, concernant des règles de grammaire ou portant sur des exceptions ;
- « usage » : erreurs sur le radical des mots n'entraînant pas d'altération phonétique ;
- « phonétique » : erreurs sur le radical entraînant une altération phonétique.

Des auteurs ont depuis élaboré des typologies dont la structure globale s'avère comparable à celle de Simon, sans pour autant qu'émerge un consensus en ce qui concerne la terminologie choisie ou le nombre de catégories et de sous-catégories distinguées (Catach, Gruaz & Duprez, 1996 ; Pothier, 1994 ; Moats, 1995 ; Estienne, 1999). Les données du rapport de l'Inserm « Dyslexie, Dysorthographie, Dyscalculie » (2007) ont plus récemment proposé une typologie qui, inspirée des données développementales, distingue quatre catégories d'erreurs orthographiques :

- erreurs de type phonétique ;
- « erreurs phonétiquement acceptables mais ne respectant pas l'orthographe conventionnelle » (p.278) ;
- « erreurs pouvant être prévenues si des connaissances morphologiques sont disponibles et appliquées » (p.279) ;
- erreurs d'accord, en genre et en nombre.

Considérant l'analyse de corpus d'erreurs comme une démarche cruciale dans la compréhension des mécanismes qui sous-tendent l'apprentissage et l'utilisation de l'orthographe, Fayol (2008) propose de dépasser l'idée d'un résultat (le nombre d'erreurs commises) pour se poser la question de leur origine. C'est en effet à la seule condition

d'en connaître les facteurs explicatifs qu'il est possible de remédier aux erreurs orthographiques. Or, ce type d'analyse reste encore rare.

3. Étude des erreurs orthographiques

3.1. Origines des erreurs orthographiques

D'après Fayol et Miret (2005), les erreurs de production verbale (et notamment orthographiques) peuvent, selon l'hypothèse d'une triple détermination, être imputées à :

1. « **une surcharge de coût** » (Fayol & Miret, 2005, p.394) ne permettant pas une implication homogène de la charge attentionnelle disponible sur tous les processus en jeu ;
2. « **la (relative) fragilité du fonctionnement d'une composante** » (Fayol & Miret, 2005, p.395) qui affecte la mise en œuvre de celle-ci ;
3. « **l'objectif poursuivi [...] qui focalise l'attention sur une composante aux dépens des autres** » (Fayol & Miret, 2005, p.395).

Foulin (1997) a proposé de répertorier les principaux facteurs en jeu dans la réussite en orthographe, permettant ainsi de différencier les « bons » des « mauvais orthographes » chez les adultes et chez les enfants. Ainsi, en référence à Fayol et Miret (2005), la fragilité d'un ou plusieurs de ces facteurs pourrai(en)t être responsable(s) des erreurs orthographiques.

Chez l'enfant-apprenant, Foulin (1997) reconnaît comme facteurs de réussite :

- l'acquisition de la procédure phono-graphémique, via l'étape alphabétique déjà largement définie (Fayol, 2010) et par extension, des connaissances stables en correspondances graphèmes-phonèmes ;
- la conscience phonologique ;
- la mémoire de travail phonologique ;
- le développement du lexique orthographique (via l'étape lexicale).

De plus, la conscience morphologique décrite par Carlisle (1995) joue elle aussi un rôle essentiel dans le développement de l'orthographe.

Chez l'adulte (ou plutôt chez « l'orthographe expert »), Foulin (1997) reconnaît l'importance de l'efficacité, de la qualité et de l'intégrité :

- du lexique mental orthographique ;
- des connaissances orthographiques (régularités graphotactiques) ;
- de la stratégie d'analogie ;
- d'une lecture permettant une analyse complète des mots (pour éviter une mémorisation orthographique des mots basée sur une analyse trop parcellaire).

Jaffré et Fayol parlent également de « zones de fragilité » (1997, p.102) pour désigner les mots comportant des variations graphiques ne correspondant pas strictement à des variations sonores (diacritiques, ambiguïtés de transcription). Ils reconnaissent ainsi ces « zones » comme plus susceptibles d'entraîner des erreurs orthographiques.

Finalement, les causes à l'origine d'erreurs orthographiques peuvent donc tenir au scripteur (niveau de conscience phonologique, d'attention, etc.) ou à la langue (opacité, « zones de fragilité », etc.).

3.2. Distribution des erreurs orthographiques au cours de l'apprentissage

Martinet et Valdois (1999) constatent que le nombre total d'erreurs orthographiques tend à diminuer entre les classes de CP et de CM2 et ce, quelle que soit la complexité des mots considérée.

L'expertise collective de l'Inserm (2007) propose un aperçu de la progression des compétences orthographiques des scripteurs selon leur âge et le type d'erreurs orthographiques.

Ainsi, les erreurs de type phonétique (omissions, substitutions ou modifications de l'ordre des graphèmes : ex. « pidon » pour « bidon » p.278) prédominent en début d'apprentissage puis se raréfient. Martinet et Valdois (1999) précisent qu'au cours de l'apprentissage, le pourcentage d'erreurs phonologiquement plausibles augmente là où celui des erreurs non phonologiquement plausibles baisse.

Les « erreurs phonétiquement acceptables mais ne respectant pas l'orthographe conventionnelle » (Inserm, 2007, p.278), portant sur des mots irréguliers (que même la stratégie d'analogie ne permet pas d'orthographier) ou sur des régularités graphotactiques, restent fréquentes même une fois le niveau d'« expert » atteint. Les séquences dites illégales de lettres sont par contre amenées à disparaître quasi complètement dès le CM2 (Martinet & Valdois, 1999).

« Les erreurs pouvant être prévenues si des connaissances morphologiques sont disponibles et appliquées » (Inserm, 2007, p.279) interviennent lorsque le scripteur, indépendamment de son âge, présente un défaut d'anticipation d'une dérivation (ex. retar*, p.279) ou au contraire opère des sur-généralisations (ex. numérot*, p.279).

Enfin, « les erreurs d'accord, en genre et en nombre » (omissions, substitutions) s'avèrent fréquentes en début d'apprentissage et peuvent subsister, à moindre fréquence, chez l'adulte (notamment concernant les homophones verbaux du type « timbre » ou « ferme » ; Fayol, Largy & Lemaire, 1994 ; Largy, Fayol & Lemaire, 1996).

II. La production écrite de texte

La production écrite est une activité complexe impliquant la gestion de nombreuses tâches cognitives. La difficulté est, pour les adultes comme pour les enfants, de coordonner en temps réel toutes les composantes en jeu, compte tenu de la capacité de traitement cognitif limitée d'un point de vue attentionnel et mnésique.

1. Description des processus en jeu dans la production écrite

1.1. Modèle princeps d'Hayes et Flower

Flower et Hayes (1980) considèrent qu'écrire consiste à jongler entre différentes contraintes. Dans leur article princeps (1981), ces auteurs proposent un modèle rédactionnel en trois dimensions (Figure I.1):

- **l'environnement de la tâche**, qui comprend toutes les caractéristiques liées à la tâche elle-même, c'est-à-dire les thématiques abordées, le destinataire, la motivation qu'elle suscite auprès du scripteur. Il comprend également le texte déjà produit ;
- **les connaissances en mémoire à long terme**, qui incluent les connaissances déclaratives (ce que le scripteur sait et veut dire), mais aussi procédurales (accès au lexique, élaboration des formes syntaxiques, réalisation des accords orthographiques, réalisation d'un plan, autoévaluation de ses compétences) ;
- **l'ensemble des processus rédactionnels** que sont :
 - **la planification (planning)**, qui consiste à générer des idées en les récupérant en mémoire à long terme et à les organiser. Ce processus est perpétuellement influencé par le but rhétorique poursuivi par le scripteur ;
 - **la formulation**, également nommée **mise en texte (translating)**, consistant à traduire ces idées en langage écrit, tout en respectant les conventions ;
 - et **la révision (reviewing)**, qui correspond à l'évaluation du texte déjà produit et à sa modification si le scripteur le juge nécessaire.

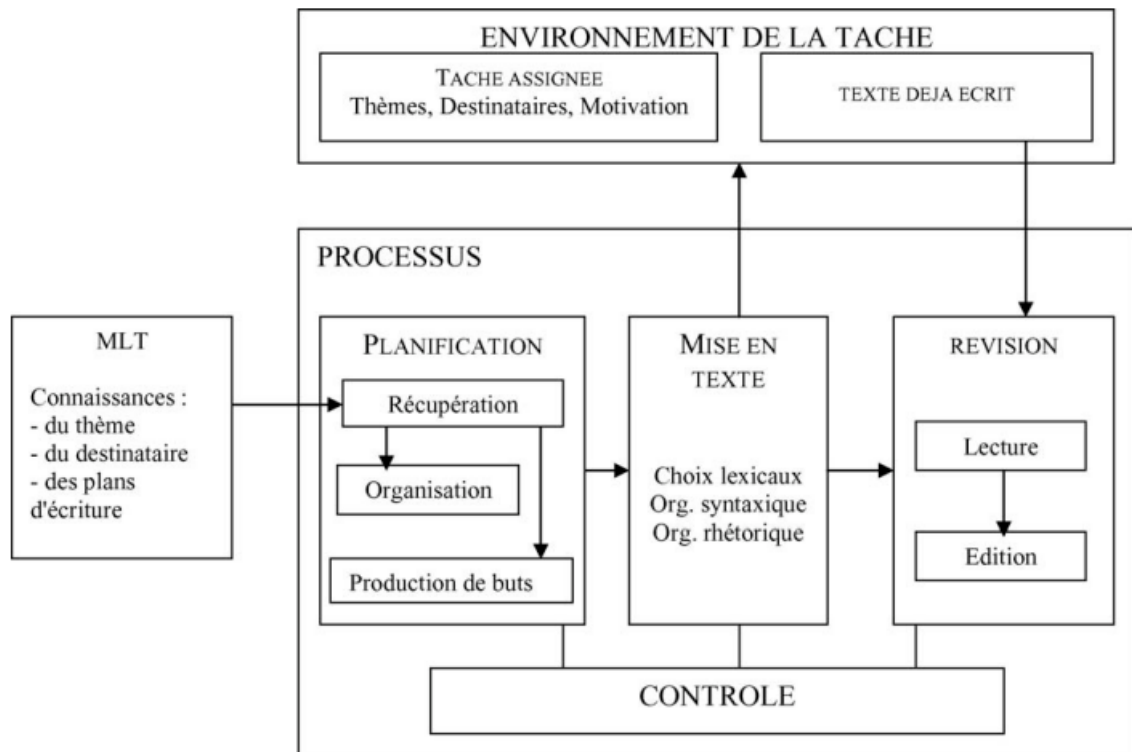


Figure I.1 : Les composants cognitifs de la production écrite selon Flower et Hayes (1980), schéma extrait de Favart et Olive (2005)

Cette première description du processus de production écrite a montré depuis certaines limites. En effet, selon Alamargot et Fayol (2009), il s'agit d'un modèle de la production écrite valable chez l'adulte mais ne mettant pas en évidence les aspects développementaux. Ces mêmes auteurs constatent également que le processus de formulation est peu décrit par Flower et Hayes (1980 et 1981).

1.2. Processus de formulation

Le processus de formulation (ou de mise en texte) est, des trois processus décrits originellement par Flower et Hayes (1981), celui qui a connu le plus de développement, notamment parce qu'il a suscité, chez les chercheurs, de nombreux questionnements quant à son acquisition. Berninger, Fuller et Whitaker (1996) ont ainsi proposé une modélisation au sein de laquelle la mise en texte comprend deux sous-processus :

- **la génération de texte**, qui correspond à la mise en langage de l'idée au sein de la mémoire de travail. Il s'agit d'un processus linéaire (Fayol, 1997, cité par Fayol, Foulin, Maggio, & Lété, in press), commun aux modalités orales et écrites. Ce sous-processus implique selon Levelt (1989) une sélection lexicale, un agencement syntaxique et une organisation rhétorique des éléments. Bien que concernant en premier lieu la modalité orale, ces résultats semblent applicables à la production écrite de textes (Bock & Levelt, 1994) ;
- **la transcription**, spécifique à la modalité écrite, qui équivaut à la transformation en symboles écrits du résultat de la génération de texte, c'est-à-dire au codage orthographique, à la segmentation du texte (ponctuation et cohésion) et à

l'exécution motrice (Alamargot & Fayol, 2009). Graham, Berninger, Abbott, Abbott et Whitaker (1997) parlent à ce propos de processus mécaniques.

2. Interactions entre les processus rédactionnels

Au-delà d'une description statique des processus rédactionnels en jeu dans la production écrite, s'intéresser aux interactions qui déterminent leur efficacité semble un enjeu majeur de la recherche actuelle dans le domaine.

2.1. Temporalité de l'écriture : partage attentionnel et théorie de la capacité

Just et Carpenter (1992) décrivent que le caractère limité des ressources cognitives nécessaires à la réalisation d'une tâche oblige à une compétition entre les processus en jeu. Ainsi, des processus non-automatisés pourraient subir les conséquences de ce partage de ressources au point de perdre en efficacité et en fluidité, tout se passant « comme si chacun des traitements contrôlés était effectué de façon séquentielle, autonome et (seulement en apparence) encapsulée » (Chanquoy & Alamargot, 2002).

D'abord appliquée aux tâches de lecture et de compréhension de textes (Just & Carpenter, 1992), la théorie de la capacité a été depuis reprise et appliquée à la production de texte. Ainsi, McCutchen (1994, 1996) affirme que l'exigence de gestion simultanée en mémoire de travail des trois processus rédactionnels (planification, formulation, révision) au cours de la production d'un texte affecterait la qualité de rédaction de l'apprenti-scripteur.

Le sous-processus de transcription (graphisme et orthographe) concentre, chez les plus jeunes, la majorité du coût cognitif total impliqué dans la production écrite (Alamargot & Fayol, 2009). C'est l'automatisation de ces processus de bas niveau qui permet de dégager l'attention nécessaire aux processus dits de haut niveau que sont la planification, la révision et la génération de texte.

Plane, Alamargot et Lebrave (2010, p.11) décrivent ainsi la production écrite comme « un ensemble de plaques correspondant à chacun des phénomènes en jeu et glissant l'une sur l'autre à des vitesses différentes ».

2.2. Implications développementales

2.2.1. Modèles développementaux

Selon Alamargot et Fayol (2009), un modèle développemental de la production écrite devrait permettre à la fois de décrire les relations entre les processus à un âge donné et les caractéristiques de la production finale. Bien que la recherche ait su modéliser certains aspects du développement de la production écrite, elle ne semble actuellement pas offrir une description complète.

Bereiter et Scardamalia (1987) ont distingué deux types de stratégies rédactionnelles selon le niveau d'expertise du scripteur. Ainsi, de 9 à 16 ans, le rédacteur applique une « stratégie des connaissances rapportées ou racontées », qui consiste en une transcription mots pour mots, au fur et à mesure de la récupération en mémoire à long terme. McCutchen (1996) parle à ce propos de « cycles répétitifs de récupération et de formulation « Think it-Write it » » (Chanquoy & Alamargot, 2002, p.28-29). Plus tard, l'automatisation de certains processus permet au scripteur expert de mobiliser simultanément et de façon fluide le contenu conceptuel et les formes linguistiques adaptées aux objectifs communicatifs, appliquant ainsi une stratégie dite « des connaissances transformées ». McCutchen (1996) précise que c'est principalement l'automatisation de la transcription, notamment graphomotrice, qui permet l'accès à cette expertise.

S'inspirant des travaux de McCutchen, Berninger et Swanson (1994) ont élaboré un modèle développemental au sein duquel l'accès au niveau d'expert se fait par stades et où chaque stade correspond à un type de contraintes imposé à la mémoire de travail lors de la production écrite. Ainsi, lors des années d'apprentissage de l'écrit (du CP au CE2), la transcription (orthographe et graphisme) concentre la majorité des ressources cognitives de l'enfant, alors que du CM1 à la 6^e, grâce à l'automatisation progressive de la transcription, il peut davantage investir ses ressources sur l'activité linguistique. Enfin, de la 5^e à la 3^e, ce sont les processus de planification et de révision qui recueillent le plus de ressources.

2.2.2. Facteurs développementaux de la production écrite

Le développement de la production écrite modélisé par ces auteurs dépend à la fois de la maturation cognitive et de l'expérience du scripteur (Alamargot & Fayol, 2009).

Selon Kellogg (2008 ; cité par Alamargot & Fayol, 2009), la maturation cognitive permet l'augmentation de l'empan de mémoire de travail et l'acquisition de compétences métacognitives et métalinguistiques. À ce propos, Case (1985 ; cité par Alamargot & Fayol, 2009) explique l'amélioration des capacités mnésiques par l'automatisation de processus, faisant de ce phénomène une conséquence indirecte. La métacognition correspond quant à elle aux connaissances de l'individu sur ses états et ses processus cognitifs et au contrôle qu'il applique volontairement sur ces derniers en fonction du but à atteindre (Gombert, 1990). Dans le cadre de la production écrite, métacognition et métalinguistique permettent l'accession au stade d'expert par le raisonnement, l'analyse réflexive, la gestion et la coordination des processus rédactionnels ainsi que par la vérification de l'adéquation entre intention et texte produit. Les effets de la maturation cognitive sont particulièrement marqués chez les plus jeunes (automatisation progressive des processus de bas niveau). Ils sont aussi présents plus tard, bien que moins étudiés (Alamargot & Fayol, 2009). En effet, vers 14-15 ans, le passage de la « stratégie des connaissances rapportées » à la « stratégie des connaissances transformées » (amélioration des processus de planification et de révision) coïnciderait avec cette maturation (Bereiter & Scardamalia, 1987).

La confrontation du scripteur aux activités de production écrite constitue une expérience indispensable à l'amélioration de ses compétences par l'apprentissage implicite et explicite qui en résulte. Selon Alamargot et Fayol (2009), la part de

l'apprentissage explicite est d'ailleurs négligée dans la recherche, bien que ses effets aient été mis en évidence à plusieurs reprises. L'enseignement peut prendre la forme :

- de tâches de soutien à la production écrite. Le scripteur est amené à améliorer ses performances par la mobilisation de connaissances, l'application de stratégies (e.g. faire un plan), la modification de l'environnement de la tâche, etc. ;
- d'entraînements permettant l'automatisation de processus pris isolément (orthographe et graphisme généralement, e.g. entraîner la motricité fine en vue de l'amélioration du graphisme) ou l'acquisition de nouvelles compétences.

S'ajoutent à ces facteurs influençant le niveau d'expertise du scripteur d'autres éléments tels, selon Plane et al. (2010, p.2) « [la] familiarité [du scripteur] avec le thème ou avec le genre du texte, ou encore [ses] états psychologiques ou physiologiques particuliers (pathologies du langage, attrition, dégénérescence sénile, etc.) ».

3. Analyse en temps réel de la production écrite

3.1. Méthodes d'analyse de la production écrite

Selon Favart et Olive (2005), la production écrite peut être étudiée de deux manières distinctes, à l'aide de méthodes dites a posteriori ou de méthodes en temps réel. Une analyse a posteriori s'intéresse au résultat de la production écrite (mot ou texte) de façon subjective (avis d'un lecteur sur la production), ou plus objective (analyse de caractéristiques textuelles).

Une analyse en temps réel, au contraire, s'intéresse à la dynamique de la production écrite. Classiquement, les cognitivistes font appel au paradigme expérimental de double-tâche afin de montrer la répartition des ressources à l'œuvre pendant la production écrite.

Dès 1980, Flower et Hayes utilisent des protocoles verbaux qui consistent pour les scripteurs à exprimer à haute voix tout ce qu'ils pensent au moment de la rédaction, dans le but de mettre en évidence les processus en jeu.

Suite aux avancées de la recherche concernant le partage attentionnel entre les processus rédactionnels, de nouveaux protocoles d'analyse en temps réel ont vu le jour. Ainsi, récemment, et en lien avec le développement de nouvelles technologies, il est possible de prendre en compte les paramètres en temps réel (débit d'écriture et temps de pause) et de repérer les révisions à l'aide de traitements de textes et de tablettes graphiques. Cette méthodologie permet de contourner les biais liés au procédé de verbalisation. Ces techniques de recueil de données, moins coûteuses en traitement et applicables à de larges effectifs, ont permis de dépasser les limites méthodologiques des protocoles d'enregistrement vidéo de l'écriture, utilisés jusqu'alors (Matsushashi, 1981; Chanquoy, Foulin & Fayol, 1993; Foulin, Chanquoy & Fayol, 1989 ; Foulin, 1993 ; cités par Chesnet, Guillabert & Espéret, 1994).

L'utilisation de tablettes graphiques digitalisées et de leur logiciel d'analyse, G-Studio originellement (Chesnet et al., 1994) et Eye and Pen plus récemment (Chesnet & Alamargot, 2005) permet de recueillir la production écrite manuscrite de façon

écologique et fiable, via des données objectives traitées informatiquement et facilement analysables statistiquement.

3.2. Paramètres en temps réel et processus rédactionnels

3.2.1. Débit d'écriture et temps de pause

Selon Fayol (en préparation), la production écrite consiste en une succession de séquences d'écriture et de pauses. Les nouvelles méthodes d'analyse (Eye and Pen principalement) permettent d'isoler des éléments rendant compte de cette dynamique temporelle : les paramètres en temps réel (débit d'écriture et pauses).

Ainsi, le débit d'écriture correspond au temps (en millisecondes) alloué au tracé d'un caractère et les pauses au temps (en millisecondes) pendant lequel aucune trace graphique n'est enregistrée.

Le postulat de la recherche à ce sujet est que ces paramètres temporels constituent une trace tangible des processus rédactionnels en jeu dans la production écrite. Plus précisément, « les contraintes capacitaires [...] conduisent à moduler le décours de la rédaction en provoquant une variation observable de la vitesse d'écriture d'une part et la durée des pauses d'écriture d'autre part » (Chanquoy, Foulin & Fayol, 1990, cité par Plane et al. 2010).

Les paramètres en temps réel seraient donc le reflet des coûts attentionnels associés à chaque processus (planification, formulation, révision). Plane et al. (2010) rapportent d'ailleurs qu'une surcharge attentionnelle entraîne un ralentissement du débit d'écriture et un allongement des pauses dans le but de privilégier les processus prioritaires.

3.2.2. Révisions

Il semble possible d'étudier le processus de révision et sa part dans l'acte de production écrite à travers les différentes « rétroactions » (Plane et al., 2010) enregistrées par la tablette graphique, c'est-à-dire les retours du scripteur sur des unités présentes en amont. En effet, ceci rend compte de la dynamique temporelle de l'écriture.

Le processus de révision correspond, selon Piolat et Roussey (1991-1992 ; cité par Largy, 2001, p.4), aux « activités de relecture du texte déjà produit ainsi [qu'aux] procédures mises en œuvre pour le corriger ». Scardamalia et Bereiter (1983) ont ainsi proposé une décomposition du processus de révision en trois étapes successives « CDO » (« compare, diagnose, operate ») :

1. le scripteur constate, dans un premier temps, une inadéquation entre son intention et sa production ;
2. il caractérise cette inadéquation (graphique, orthographique ou textuelle) ;
3. il apporte la ou les modifications qu'il juge nécessaire(s).

Les opérations cognitives associées à chacune de ces étapes ont depuis été décrites (Flower, Hayes, Carey, Schriver & Stratman, 1986 ; Hayes, Flower, Schriver, Stratman & Carey, 1987 ; cités par Ganier, 2006), les auteurs précisant par ailleurs qu'il existe deux types de révisions, l'une modifiant la production en surface (orthographe, ponctuation, etc.), l'autre en profondeur (signification).

Plus récemment, Butterfield, Hacker et Albertson (1996 ; cité par Ganier, 2006) ont distingué deux composantes en jeu au sein du processus de révision, l'environnement du scripteur (dimensions rhétoriques et pragmatiques de la production) et les systèmes de traitement de l'information. Ceux-ci interviennent selon les auteurs à des niveaux cognitifs (connaissances stockées en mémoire à long terme) et métacognitifs (sélection et mise en œuvre de stratégies).

Ces révisions peuvent porter sur des unités de toute taille (de la structure globale du texte aux unités infragraphiques tels les accents), en autant que le scripteur considère qu'il existe une inadéquation entre son intention et sa production. Elles peuvent intervenir à différents moments de la production écrite : immédiatement lors de la production, en différé ou au cours d'une relecture finale.

Bien que les révisions ne soient pas des paramètres en temps réel, elles n'en sont pas moins des témoins-clés de la dynamique textuelle. Les prendre en compte dans une analyse portant sur la temporalité de la production écrite paraît donc nécessaire.

Chapitre II
PROBLEMATIQUE ET HYPOTHESES

I. Problématique

La quasi-totalité des travaux ayant porté sur l'orthographe, son usage et son acquisition, reposent sur deux approches complémentaires. L'une étudie les erreurs orthographiques en les décrivant et en les classant dans des catégories plus ou moins nombreuses et précises. Elle s'attache donc aux produits et infère à partir de la nature des erreurs commises les processus susceptibles d'intervenir. La seconde analyse la production de mots, presque toujours isolés, en prenant en considération certains paramètres temporels : généralement la pause avant (i.e. la latence), plus rarement le débit d'écriture. Les erreurs commises sont écartées du fait que l'interprétation des temps y pose problème. C'est la dynamique des processus qui est visée et abordée à partir des variations de durées interprétées en fonction de modèles théoriques (Flower & Hayes, 1981 ; McCutchen, 1994, 1996 par exemple). Jusqu'alors, aucune recherche n'a pris en compte les erreurs et les paramètres temporels qui les caractérisent. Tel sera notre objectif.

Pour la première fois, nous allons essayer de décrire les relations entre les performances orthographiques et les paramètres temporels qui leur sont associés (pauses précédant les mots, pauses internes aux mots et débit d'écriture). Les erreurs sont ainsi classées en fonction de leur nature et en suivant des critères inédits de catégorisation. Nous nous appuyons pour cela sur des travaux récents cherchant à étudier la dynamique de la production écrite de textes. Les mots n'y sont plus isolés mais intégrés dans des textes (récits et descriptions). Notre objectif est essentiellement descriptif : même si nous sommes en mesure, à partir des travaux antérieurs, de proposer des hypothèses, celles-ci sont peu reliées à la situation particulière que nous analysons : la production de mots erronés dans des textes.

Le cadre général de notre interprétation est celui du coût cognitif des processus mis en jeu lors de la production (Fayol, 1999). Chez l'adulte expert, le coût des traitements orthographiques est faible. L'orthographe lexicale est en effet connue pour la plupart des mots, ce qui permet un accès direct en mémoire, peu coûteux et dépendant uniquement de la fréquence (et de la longueur, les deux variables étant difficiles à dissocier). Les accords, quant à eux, sont réalisés rapidement et exactement, hormis dans quelques situations critiques (Fayol et al., 1994 ; Largy et al., 1996). Dès qu'une difficulté surgit, elle provoque une augmentation du coût, laquelle se traduit selon les sites concernés par un accroissement des durées de pauses (initiales ou internes aux mots) et/ou un ralentissement du débit. Ce sont donc les variations de ces paramètres, mais cette fois chez l'enfant (de CM2, 5^e et 3^e), que nous allons décrire et analyser.

Notre objectif est donc de déterminer si et dans quelle mesure les paramètres en temps réel reflètent le coût associé au traitement de l'erreur orthographique.

De cette problématique découlent les hypothèses suivantes.

II. Hypothèses

1. Hypothèse générale

La présence d'erreurs orthographiques influe sur les valeurs des paramètres en temps réel (débit d'écriture et pauses).

2. Hypothèses opérationnelles

Cette hypothèse donne naissance à deux hypothèses opérationnelles.

2.1. Première hypothèse

Prendre en compte le coût cognitif associé à la production orthographique amène nécessairement à comparer les mots correctement orthographiés et ceux qui ne le sont pas. Ainsi, nous posons l'hypothèse que **les valeurs des paramètres d'analyse en temps réel diffèrent selon que les mots comportent ou non des erreurs orthographiques.**

Pour les mots contenant une ou plusieurs erreur(s) orthographique(s) :

- le débit d'écriture est plus lent
- les pauses (avant et à l'intérieur des mots) sont plus longues.

2.2. Deuxième hypothèse

Les compétences nécessaires à une orthographe correcte étant diverses et ne faisant pas intervenir le même type de connaissances, nous faisons l'hypothèse que **les valeurs des paramètres d'analyse en temps réel diffèrent selon la nature des erreurs commises** (phonologique seule, phono-lexicale, lexicale seule, phono-morphosyntaxique ou morphosyntaxique seule). Dans la mesure où une telle analyse n'a jamais été menée auparavant, nous ne pouvons nous appuyer sur la littérature afin de développer davantage cette hypothèse. Celle-ci est également difficile à déterminer en faisant appel à nos seules intuitions, connaissances ou raisonnements.

Chapitre III
PARTIE EXPERIMENTALE

I. Méthode de recueil des données

Notre étude consiste en l'analyse d'un corpus de textes déjà recueilli, en 2006, par l'équipe « Acquisition et cognition de l'enfant à l'adulte » du laboratoire Dynamique Du Langage (UMR 5596, Université Lumière Lyon II), en collaboration avec le laboratoire LAPSCO (UMR 6024, Université Blaise Pascal, Clermont-Ferrand). Ce recueil s'inscrivait dans le cadre d'un réseau international et dans le programme de recherche « Développement syntaxique et discursif tardif » de ces équipes. Ainsi, nous profitons de la pertinence et de la fiabilité de ce corpus pour répondre à notre problématique.

1. Participants

139 élèves âgés de 10 à 15 ans et issus de trois groupes d'individus de niveaux scolaires différents ont participé à cette étude. Chaque niveau scolaire comprenait deux groupes selon que les élèves avaient des difficultés scolaires ou non. Ces groupes étaient constitués de :

- 46 élèves (22 filles et 24 garçons) de cours moyen deuxième année (CM2, équivalent au niveau G5 dans le système éducatif anglophone) issus des écoles primaires Sablons, Prainet II et Soie de Décines-Charpieu, proches de Lyon ;
- 47 élèves (23 filles et 24 garçons) de 5^e (G7) du collège G. Brassens de Décines-Charpieu ;
- 46 élèves (22 filles et 24 garçons) de 3^e (G9) du collège G. Brassens de Décines-Charpieu.

Les participants étaient tous francophones natifs mais pouvaient être en contact avec d'autres langues. Tous appartenaient, selon les professions et catégories socioprofessionnelles de leurs tuteurs légaux, à la classe socioculturelle défavorisée et tous étaient scolarisés en ZEP (zone d'éducation prioritaire) ou REP (réseau d'éducation prioritaire).

Le niveau de performances scolaires des participants a été déterminé sur la base de leurs notes en français (moyenne inférieure à 9,5/20 : avec difficultés ; moyenne supérieure à 10,5/20 : sans difficulté), tout en veillant à ce qu'ils aient un niveau similaire dans les matières principales telles que les mathématiques ou l'histoire-géographie et en tenant compte de l'avis des enseignants.

2. Matériel

Le recueil des données a été permis par l'utilisation du dispositif Eye and Pen, reliant des tablettes graphiques à un ordinateur portable. Grâce à l'utilisation des tablettes graphiques, différentes informations spatiales et temporelles ont pu être recueillies : pression du stylo, débit d'écriture, pauses, révisions.



Figure III.1: Dispositif Eye and Pen (www.eyepandpen.net)

Les participants visionnaient en petit groupe une vidéo présentant des situations de vie quotidienne à l'école (exclusion, solidarité entre élèves, vandalisme...).

Par la suite, installés devant une tablette graphique, ils devaient produire un brouillon puis une version finale d'un texte narratif et d'un texte expositif (discuter, exposer ses connaissances, donner son avis) en lien avec la vidéo visionnée.

3. Procédure expérimentale

L'expérimentation a eu lieu sur deux sessions séparées d'une semaine. Il a été demandé aux élèves de produire un texte. La moitié des élèves produisaient un brouillon et une version finale d'un texte narratif à la première session, puis un brouillon et une version finale d'un texte expositif à la deuxième session. Le protocole expérimental était inversé pour l'autre moitié des élèves. Pour les textes narratifs, il était précisé de ne pas résumer la vidéo visionnée au préalable, mais de raconter un événement personnel suggéré par celle-ci. Ainsi, chaque individu a rédigé quatre textes :

- un brouillon de texte narratif ;
- un brouillon de texte expositif ;
- une version finale de texte narratif ;
- une version finale de texte expositif.

Notre étude ne tient pas compte des versions finales des textes rédigés par les participants. Une analyse de données dite « en temps réel » nécessite de ne tenir compte que d'écrits de premier jet. Ainsi, seuls les brouillons des textes rédigés par les participants ont été analysés. Ils rendent compte de l'ensemble des processus de production écrite là où un texte final n'est qu'une copie d'un texte déjà produit. À titre d'exemples, les deux brouillons d'un élève de chaque niveau scolaire sont proposés en annexes II, III et IV.

II. Constitution du corpus

Tous les mots produits par les participants ont été isolés et classés au sein de tableaux Excel afin d'en faciliter l'analyse. De cette façon, chaque mot correspond à une entrée et est individuellement analysé. Cela permet une décomposition précise de la production écrite.

Deux tableaux ont été construits :

- un tableau contenant tous les mots produits par les participants (corpus général) ;
- un autre contenant tous les mots comportant au moins une erreur orthographique (corpus des erreurs orthographiques).

1. Corpus général

Ces tableaux rassemblent des informations permettant de replacer chaque mot :

- dans son **contexte de production** : clé unique du mot, numéro de ligne, niveau scolaire (G5, G7, G9), numéro du scripteur, sexe, type de texte, ordre de passation (descriptif ou expositif produit en premier ou en second), avec ou sans difficulté scolaire (d = avec, w = sans), position du mot dans le texte, numéro de page (lorsque l'élève en produit plusieurs) et localisation du mot dans le texte ;

Tableau III.1 : Exemples d'entrées du tableau rassemblant les données du corpus.

| Clé | N°de Ligne | level | subject_index | sex | texttype | order | expression | page | Position |
|------|------------|-------|---------------|-----|----------|--------|------------|------|----------|
| 3441 | 12 | G7 | 1 | M | E | ECRIT2 | D | | 10 |
| 3442 | 51 | G7 | 1 | M | E | ECRIT2 | D | | 23 |
| 3443 | 52 | G7 | 1 | M | E | ECRIT2 | D | | 24 |
| 3444 | 63 | G7 | 1 | M | E | ECRIT2 | D | | 31 |

- des **informations linguistiques sur le mot** : indices des bases de données Cordial et Manulex à propos du type de mot, de sa fonction dans la phrase, de sa fréquence dans la langue, de sa composition (nombre de lettres, de phonèmes), etc. ;
- les **informations dites « de temps réel »** recueillies par le logiciel Eye and Pen via la tablette graphique :
 - débit (en millisecondes par caractère, la durée de la pause interne au mot ayant été soustraite) ;
 - pauses (durée en millisecondes, position, nombre à l'intérieur d'un mot, somme des durées de pauses à l'intérieur d'un mot, etc.) ;
 - révision.

Pour chaque entrée du tableau, l'ensemble des informations concernant les mots précédant et suivant est indiqué.

2. Corpus des erreurs orthographiques : constitution des critères d'analyse

L'analyse du corpus des erreurs orthographiques s'est faite en deux temps. Elle a d'abord consisté en un classement et une comptabilisation des erreurs selon les critères de classification choisis. Ces données ont ensuite été analysées statistiquement afin d'objectiver les liens et influences entre les différents paramètres considérés (types d'erreurs, niveaux scolaires et paramètres en temps réel).

Les mots comportant au moins une erreur orthographique ont été extraits du tableau général, soit 5252 items. Ce nouveau tableau a ensuite été scindé selon le niveau scolaire des participants (G5, G7, et G9). Six colonnes ont été ajoutées afin de traiter les erreurs relevées dans chaque mot, en fonction de la typologie élaborée dans le cadre de cette étude :

- **erreurs phonologiques seules** (« phono seule ») ;
- **erreurs phono-lexicales** (« phono-lex ») ;
- **erreurs lexicales seules** (« lex seule ») ;
- **erreurs phono-morphosyntaxiques** (« phono-morpho ») ;
- **erreurs morphosyntaxiques seules** (« morpho seule ») ;
- **total d'erreurs** (« total ») ;
- **combinaisons d'erreurs dans chaque mot** (« combinaisons »). Exemple (clé 3467 du tableau G7) : « bagarer* » pour « bagarré » : « lex seule + morpho seule ».

L'annexe I présente une liste des informations disponibles au sein des tableaux d'analyse. L'analyse des erreurs d'un élève de chaque niveau scolaire est proposée à titre d'exemple en annexes II, III et IV.

2.1. « Erreurs phono-lexicales »

Sont considérées comme « erreurs phono-lexicales » les parties de mots pour lesquelles la forme finale n'est pas phonologiquement plausible. Il s'agit de productions dont la forme ne respecte ni l'orthographe ni la forme phonologique de l'item cible.

Exemple (clé 7 du tableau G5) : « gos* » pour « gros ».

2.2. « Erreurs lexicales seules »

Les « erreurs lexicales seules » correspondent quant à elles aux erreurs phonologiquement plausibles ou de régularisation, c'est-à-dire aux productions dont la forme ne respecte pas l'orthographe de l'item cible tout en en préservant la forme phonologique.

Exemple (clé 4 du tableau G5) : « avvait* » pour « avait ».

2.3. « Erreurs morphosyntaxiques seules »

Les erreurs « morphosyntaxiques seules » correspondent aux critères suivants, sans altération phonologique :

- erreurs d'accord en genre des noms, adjectifs, pronoms et verbes (clé 66 du tableau G5: « ami » pour « amie ») ;
- erreurs d'accord en nombre des noms, adjectifs, pronoms et verbes (clé 55 du tableau G5 : « personne » pour « personnes ») ;
- erreurs sur la catégorie des mots (clé 19 du tableau G5 : « arriver » pour « arrivé ») ;
- confusion entre homophones hétérographes (ou logogrammes selon Catach et al., 1980) (clé 16 du tableau G5 : « sa » pour « ça ») ;
- erreurs de segmentation (clé 79 du tableau G5 : « quil » pour « qu'ils »).

2.4. « Erreurs phono-morphosyntaxiques »

Sont considérées comme des erreurs « phono-morphosyntaxiques » les sections de mots erronées correspondant aux critères suivants et entraînant une altération phonologique :

- erreurs d'accord en genre des noms, adjectifs, pronoms et verbes (clé 4772 du tableau G7: « permis » pour « permise ») ;
- erreurs d'accord en nombre des noms, adjectifs, pronoms et verbes (clé 4765 du tableau G7 : « tout » pour « tous ») ;
- erreurs portant sur le radical et consistant en une mauvaise connaissance de la conjugaison (clé 857 du tableau G5 : « croivent » pour « croient ») ;
- erreurs de mode (clé 951 du tableau G5 : « aimera » pour « aimerait »).

Ainsi, il s'agit de productions à l'écrit de formes orales erronées ou de formes résultant d'un mauvais transcodage oral/écrit (méconnaître par exemple la nécessité du « e » final pour transcoder le son /z/ dans le mot « permise »).

2.5. « Erreurs phonologiques seules »

L'« erreur phonologique seule » consiste en une altération phonologique consécutive à la présence d'une erreur sur une autre section du mot. Elle est donc systématiquement combinée à un autre type d'erreur.

Exemple (clé 4367 du tableau G7) : « aves* » pour « avec ». La substitution du « s » au « c » entraîne en premier lieu une erreur « phono-lex » car le /k/ n'est pas transcrit. La conséquence indirecte de cette erreur est que le /ɛ/ ne se prononce plus alors que la graphie « e » est respectée.

2.6. Items problématiques

Certains items considérés par l'étude initiale comme erronés mais n'étant pas par définition des erreurs orthographiques ou ne rentrant pas dans le cadre de notre mémoire ont été retirés de l'analyse et donc non retranscrits dans les tableaux. Il s'agit :

- des items non reconnus comme des unités linguistiques par la base de données Manulex (griffonnages, mots incomplets et abréviations) ;
- des mots révisés mais non corrigés : Exemple : clés 3450 et 3451 du tableau G7, l'élève a d'abord produit « sens » puis l'a révisé pour produire « sans » alors que la cible était « sang ». Sa révision ne lui a donc pas permis de corriger le mot. Laisser les deux entrées dans le tableau destiné à l'analyse statistique reviendrait à les considérer comme deux mots différents alors qu'ils n'en sont qu'un.

Il a été également nécessaire de prendre des décisions quant au classement de certaines erreurs récurrentes. Ainsi :

- la distinction /e/-/ɛ/ étant en France floue, puisque dépendante des accents régionaux, il est apparu difficile de les distinguer comme deux phonèmes à part entière. Les erreurs de substitution d'accents « é » / « è » ou encore les terminaisons « ai » / « ais » n'ont donc pas été considérées comme entraînant une altération phonologique ;
- les locutions figées, également appelées nominales, sont des groupements de mots indissociables et lexicalisés par l'usage. En ce sens, les erreurs, même d'apparence morphologiques, ont été comptabilisées comme des erreurs lexicales ou phono-lexicales (ex. clé 239 du tableau G5 « tous le monde » pour « tout le monde ») ;
- pour maintenir l'idée d'un travail sur le texte, replacer certaines erreurs dans leur contexte global de production a été nécessaire. En effet, des erreurs phonologiques isolées ne le sont pas toujours si le mot est réintégré dans la structure sonore de la phrase (ex. clé 1042 du tableau G5 « cette » pour « c'est » constitue une erreur phonologique s'il est analysé isolément. Or, se référer au brouillon du texte permet de constater que l'élève a écrit « cette une fille » pour « c'est une fille », retranscrivant ainsi la liaison). La référence à la production initiale est donc indispensable.

Chapitre IV
PRESENTATION DES RESULTATS

Dans le but de permettre l'application de tests paramétriques (t-Test et ANOVA) plus puissants que les tests non paramétriques, les données brutes ont été transformées afin d'être normalisées selon la formule « $y' = \ln(y) + 4$ » (la constante 4 permettant d'éliminer les valeurs négatives ou nulles). Par souci de clarté et pour en permettre une meilleure lecture, les données citées ont ensuite été retransformées (et donc restituées dans leur unité d'origine). Pour ces analyses, le logiciel PASW Statistics 18 (SPSS Inc.) a été utilisé. Les données chiffrées ont été arrondies et présentées avec deux décimales. L'intervalle de confiance choisi pour décrire la significativité des résultats est de 95%.

I. Données descriptives du corpus

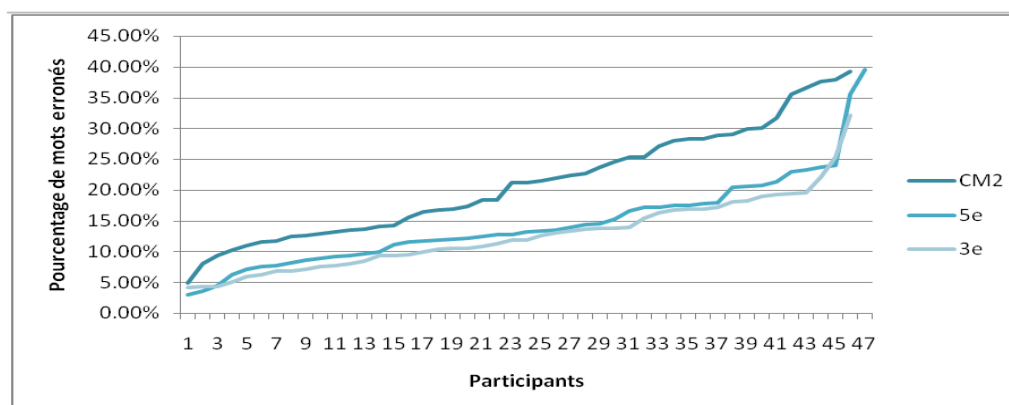
Le tableau IV.1 illustre le rapport entre le nombre total de mots produits et le nombre de mots erronés et ce, à chaque niveau scolaire. Il est à noter que, dans cette rédaction libre de texte, plus le niveau scolaire est élevé et plus l'élève produit de mots.

Ces données montrent également une diminution du pourcentage d'erreurs produites entre les classes (CM2 > 5^e > 3^e). Plus le niveau scolaire est élevé, moins les élèves produisent d'erreurs.

Tableau IV.1 : Nombre de mots (total et erronés) et pourcentage de mots erronés selon le niveau scolaire.

| Niveau scolaire | N | Total de mots produits (avec et sans erreurs) | Production moyenne de mots par participant | Écart-type | Nombre de mots avec erreur(s) orthographique(s) | Pourcentage de mots avec erreur(s) orthographique(s) |
|-----------------|-----|---|--|------------|---|--|
| CM2 | 46 | 5804 | 126.17 | 70.60 | 1122 | 26.96% |
| 5 ^e | 47 | 11146 | 237.15 | 134.35 | 1502 | 18.39% |
| 3 ^e | 46 | 14681 | 319.15 | 154.34 | 1762 | 16.98% |
| Total | 139 | 31631 | | | 4389 | |

Il est important de noter que les performances au sein d'un même niveau scolaire sont hétérogènes (Graphique IV.1). Ainsi, le pourcentage de mots erronés varie de 3% à 40%. Cette dispersion se retrouve dans les trois niveaux scolaires.



Graphique IV.1 : Rapport des mots erronés au nombre total de mots produits par chaque participant selon le niveau scolaire (classement croissant).

II. Comparaison des mots avec ou sans erreur(s) orthographique(s)

Nous avons cherché à mettre en évidence les variations du coût cognitif associées à la production d'erreurs orthographiques à travers l'étude des données de temps réel. Pour chacune des mesures, un t-Test a permis de comparer les mots sans erreur (MSE) à ceux comportant au moins une erreur orthographique (MAE).

Les graphiques présentent les moyennes ainsi que l'erreur standard, donnée plus représentative de la valeur de la moyenne que l'écart-type. Ainsi, plus l'erreur standard est faible, plus la moyenne est fiable et la variance des données réduite.

1. Débit d'écriture

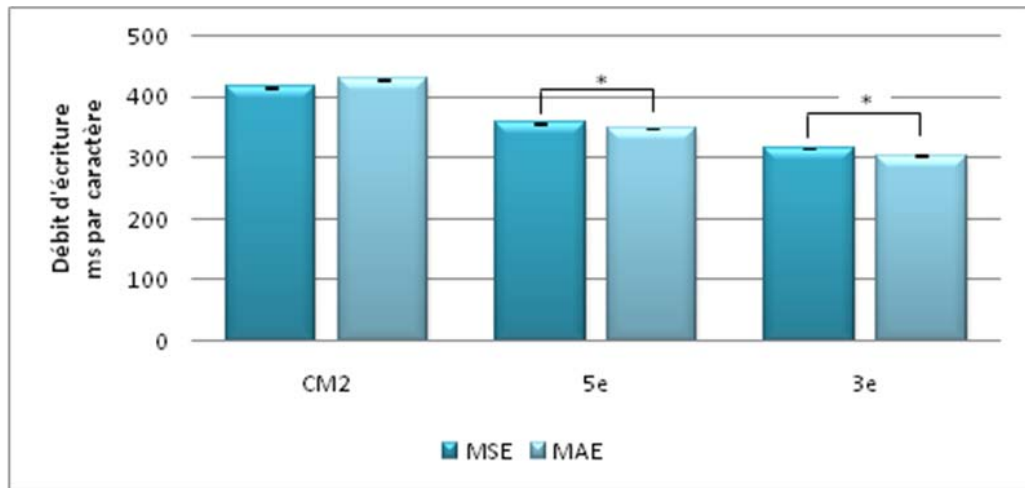
Le tableau IV.2 présente l'ensemble des données concernant les liens entre le débit d'écriture (en millisecondes par caractère) et la présence ou non d'erreurs orthographiques et ce, selon le niveau scolaire. Le graphique IV.2 résume ces données et met en évidence les résultats significatifs ($p < .05$).

Tableau IV.2 : Débit d'écriture (en ms par caractère) pour les mots avec et sans erreurs selon le niveau scolaire.

| Niveau scolaire | Mots | N | Débit | Moyenne | Écart-type | Erreur standard moyenne |
|-----------------|-------------------|------|-------------------|---------------|-------------|-------------------------|
| CM2 | Mots sans erreurs | 3040 | Débit | 6.62 | 0.26 | 0.00 |
| | | | Débit brut | 415.93 | 9.01 | 0.28 |
| | Mots avec erreurs | 693 | Débit | 6.63 | 0.26 | 0.01 |
| | | | Débit brut | 429.12 | 9.30 | 0.28 |
| 5 ^e | MSE | 6666 | Débit | 6.55 | 0.15 | 0.00 |
| | | | Débit brut | 357.50 | 7.75 | 0.24 |
| | MAE | 937 | Débit | 6.54 | 0.16 | 0.01 |
| | | | Débit brut | 349.03 | 7.56 | 0.23 |
| 3 ^e | MSE | 8613 | Débit | 6.50 | 0.14 | 0.00 |
| | | | Débit brut | 316.80 | 6.86 | 0.21 |
| | MAE | 1065 | Débit | 6.48 | 0.14 | 0.00 |
| | | | Débit brut | 304.56 | 6.60 | 0.20 |

L'analyse du débit d'écriture des élèves de 5^e comme de 3^e infirme l'hypothèse selon laquelle la vitesse du tracé est plus lente dès lors que le mot comporte une ou plusieurs erreurs orthographiques. En effet, le temps nécessaire pour la production d'un caractère est significativement moins important au sein des mots erronés qu'à l'intérieur de mots correctement orthographiés ($p = .043$ pour les élèves de 5^e et $p < .001$ pour les élèves de 3^e).

En revanche, le débit d'écriture des élèves de CM2 ne montre aucune différence significative de débit entre mots avec ou sans erreur ($t=ns$).



Graphique IV.2 : Débit d'écriture (en ms par caractère) pour les mots avec et sans erreurs selon le niveau scolaire.

Il est à noter, de plus, que le débit d'écriture est plus rapide d'un niveau scolaire à l'autre ($CM2 > 5^e > 3^e$) et ce, que le mot soit correctement orthographié ou non.

2. Pauses

Les liens entre les temps de pauses et les mots, erronés ou non, ont été étudiés à la fois pour les pauses avant les mots et pour les pauses internes.

2.1. Pauses précédant les mots

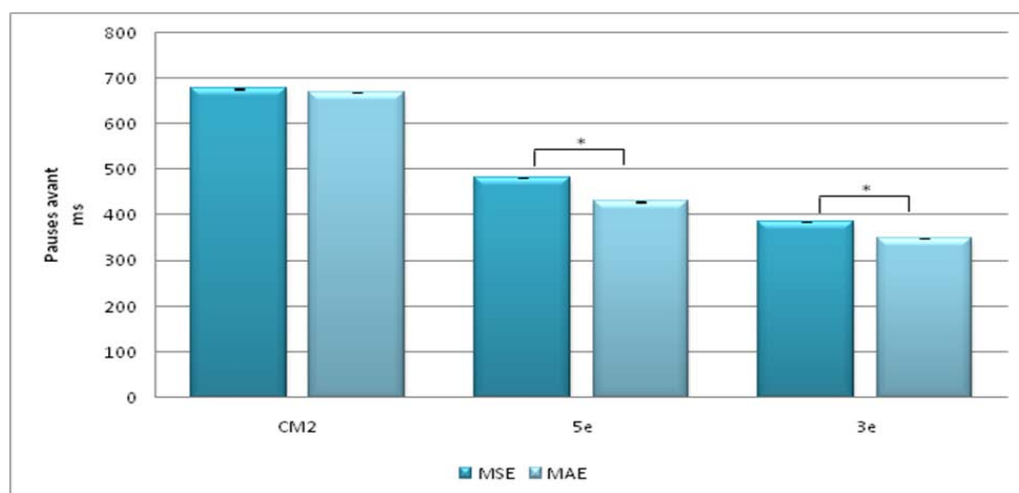
Le tableau IV.3 présente l'ensemble des données concernant les liens entre les pauses précédant les mots (en millisecondes) et la présence ou non d'erreurs orthographiques pour chaque niveau scolaire. Le graphique IV.3 résume ces données et met en évidence les résultats significatifs ($p < .05$).

Tableau IV. 3: Pauses (en ms) avant les mots avec et sans erreurs selon le niveau scolaire.

| Niveau Scolaire | Mots | N | Pauses avant | Moyenne | Écart-type | Erreur standard Moyenne |
|-----------------|------|------|------------------|---------------|--------------|-------------------------|
| CM2 | MSE | 3040 | P_AVA | 6.83 | 0.42 | 0.01 |
| | | | P_AVAbrut | 677.33 | 41.39 | 0.75 |
| CM2 | MAE | 814 | P_AVA | 6.83 | 0.40 | 0.01 |
| | | | P_AVAbrut | 669.40 | 38.90 | 1.36 |
| 5 ^e | MSE | 6666 | P_AVA | 6.68 | 0.45 | 0.01 |
| | | | P_AVAbrut | 481.45 | 32.52 | 0.40 |
| 5 ^e | MAE | 1083 | P_AVA | 6.63 | 0.40 | 0.01 |
| | | | P_AVAbrut | 429.64 | 26.00 | 0.79 |
| 3 ^e | MSE | 8613 | P_AVA | 6.59 | 0.46 | 0.00 |
| | | | P_AVAbrut | 385.15 | 26.66 | 0.01 |
| 3 ^e | MAE | 1229 | P_AVA | 6.54 | 0.40 | 0.01 |
| | | | P_AVAbrut | 349.86 | 21.26 | 0.61 |

L'observation de la durée des pauses avant les mots ne montre aucune différence significative entre les mots correctement orthographiés et les mots erronés dans la classe de CM2 ($t = ns$).

Les résultats obtenus dans les classes de 5^e et 3^e infirment l'hypothèse selon laquelle un scripteur aurait besoin de plus de temps pour traiter l'orthographe d'un mot erroné : les pauses sont significativement plus importantes lorsqu'elles précèdent des mots sans erreurs ($p = .001$ pour la classe de 5^e et $p = .002$ pour la classe de 3^e).



Graphique IV.3 : Pauses (en ms) avant les mots avec et sans erreurs selon le niveau scolaire.

Il est à noter, de plus, que la durée des pauses précédant les mots diminue d'un niveau scolaire à l'autre ($CM2 > 5^e > 3^e$) et ce, que le mot soit correctement orthographié ou non.

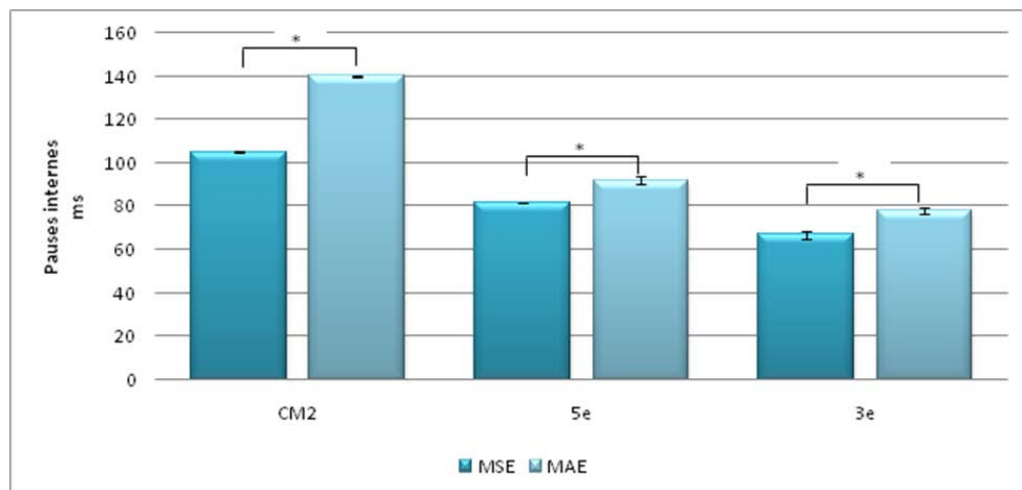
2.2. Pauses internes aux mots

Le tableau IV.4 présente l'ensemble des données concernant les liens entre les pauses internes aux mots (en millisecondes) et la présence ou non d'erreurs orthographiques et ce, selon le niveau scolaire. Le graphique IV.4 résume ces données et met en évidence les résultats significatifs ($p < .05$).

Tableau IV.4 : Pauses (en ms) internes aux mots avec et sans erreurs selon le niveau scolaire.

| Niveau scolaire | Mots | N | Pauses internes | Moyenne | Écart-type | Erreur standard moyenne |
|-----------------|------|------|-------------------|---------------|-------------|-------------------------|
| CM2 | MSE | 2502 | P_Int | 6.02 | 0.37 | 0.01 |
| | | | P_Int brut | 105.11 | 6.47 | 0.13 |
| CM2 | MAE | 629 | P_Int | 6.15 | 0.43 | 0.02 |
| | | | P_Int brut | 140.16 | 9.70 | 0.39 |
| 5 ^e | MSE | 5385 | P_Int | 5.91 | 0.34 | 0.00 |
| | | | P_Int brut | 81.42 | 4.71 | 0.06 |
| 5 ^e | MAE | 828 | P_Int | 5.96 | 0.34 | 0.01 |
| | | | P_Int brut | 91.82 | 5.66 | 1.85 |
| 3 ^e | MSE | 7162 | P_Int | 5.83 | 0.36 | 0.00 |
| | | | P_Int brut | 66.86 | 4.12 | 1.84 |
| 3 ^e | MAE | 941 | P_Int | 5.89 | 0.36 | 0.01 |
| | | | P_Int brut | 77.87 | 4.80 | 1.57 |

Les pauses à l'intérieur des mots produits par les élèves, quel que soit le niveau scolaire, sont significativement plus importantes lorsque le mot est erroné que lorsqu'il est correctement orthographié ($p < .001$ pour les trois classes).



Graphique IV.4 : Pauses (en ms) internes aux mots avec et sans erreurs selon le niveau scolaire.

Il est à noter, de plus, que la durée des pauses internes aux mots diminue d'un niveau scolaire à l'autre (CM2 > 5^e > 3^e) et ce, que le mot soit correctement orthographié ou non.

3. Paramètres en temps réel et nature des mots

Grâce aux données linguistiques extraites du corpus (nature des mots, fréquence, consistance, etc.), il a été possible d'explorer l'existence d'un éventuel lien entre paramètres en temps réel (pauses et débit d'écriture) et nature des mots (adjectif, adverbe, conjonction, déterminant, nom commun, nom propre, préposition, pronom, verbe) selon qu'ils contiennent ou non au moins une erreur orthographique.

Un seul des résultats paraît facilement analysable, là où les autres mériteraient une analyse plus précise qui n'est pas l'objet de ce mémoire. Ainsi, quel que soit le niveau scolaire, les pauses précédant les verbes sont significativement plus longues pour ceux contenant une ou plusieurs erreurs orthographiques que pour les mots sans erreurs ($p=.001$ en CM2, $p<.001$ en 5^e, $p=.002$ en 3^e). Il semble donc que la difficulté à orthographier un verbe soit anticipée par l'élève. Contrairement à l'orthographe d'usage qui atteint chez le scripteur expert un degré certain d'automatisation, l'orthographe grammaticale mérite en effet toujours de l'attention et demande notamment une analyse des liens de sens dans la phrase. Cette disparité pourrait expliquer ces résultats.

III. Comparaison des différents types d'erreurs orthographiques

L'objectif de cette analyse est de déterminer s'il existe un effet du type d'erreurs orthographiques sur les paramètres en temps réel, autrement dit si le coût cognitif diffère selon les erreurs produites.

Une analyse de variance (ANOVA univariée) a été effectuée avec le facteur Type d'erreurs orthographiques comportant 4 modalités (erreurs lex. seule vs morpho. seule vs phono-lex vs phono-morpho). La modalité « phono. seule » n'a pu être prise en compte, le nombre d'unités étant trop faible pour être traité. Un Tukey test a été administré afin de préciser les relations significatives.

Les graphiques présentent les moyennes ainsi que les écart-types des données, le logiciel utilisé pour effectuer les ANOVA ne fournissant pas dans ce cas les erreurs standards à la moyenne. Ainsi, plus l'écart-type est faible, moins les données sont dispersées.

1. Données descriptives

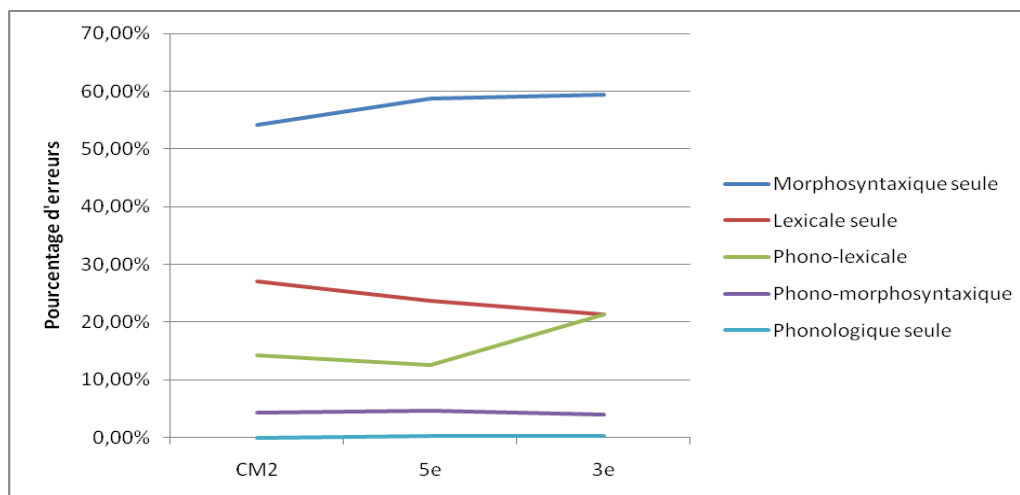
Le tableau IV.5 et le graphique IV.5 présentent le nombre d'erreurs produites et leur répartition selon le type d'erreurs pour chaque niveau de classe.

Les résultats montrent que :

- les erreurs phonologiques seules sont rares quel que soit le niveau scolaire (moins de 0.3%) ;
- le pourcentage d'erreurs phono-lexicales diminue entre le CM2 et la 5^e et augmente entre la 5^e et la 3^e. Ce type d'erreurs est alors plus fréquent en 3^e (21,25%) que dans les deux autres classes (14,22% pour les CM2 et 12,55% pour les 5^e) ;
- les erreurs lexicales seules (erreurs phonologiquement plausibles) sont fréquentes quel que soit le niveau scolaire mais leur pourcentage diminue d'une classe à l'autre (CM2 > 5^e > 3^e) ;
- le pourcentage d'erreurs phono-morphosyntaxiques reste bas mais constant d'une classe à l'autre (moins de 4,7%) ;
- les erreurs morphosyntaxiques seules sont les plus fréquentes quel que soit le niveau scolaire (entre 54,24% et 59,41%) ;
- la hiérarchie des proportions d'erreurs est similaire d'une classe à l'autre : phonologique seule < phono-morphosyntaxiques < phono-lexicale < lexicale seule < morphosyntaxiques seule.

Tableau IV.5 : Nombre et répartition des types d'erreurs selon le niveau scolaire.

| Nombre et % d'erreurs par type | | CM2 | 5 ^e | 3 ^e |
|-----------------------------------|---|---------------|----------------|----------------|
| Nombre total d'erreurs | | 1357 | 1761 | 1981 |
| Phonologique seule | N | 1 | 5 | 5 |
| | % | 0.01% | 0.28% | 0.25% |
| Phono-lexicale | N | 193 | 221 | 298 |
| | % | 14.22% | 12.55% | 21.25% |
| Lexicale seule | N | 367 | 418 | 421 |
| | % | 27.04% | 23.74% | 21.25% |
| Phono- morphosyntaxique | N | 58 | 82 | 80 |
| | % | 4.27% | 4.66% | 4.04% |
| Morphosyntaxique seule | N | 736 | 1035 | 1177 |
| | % | 54.24% | 58.77% | 59.41% |



Graphique IV.5 : Répartition des types d'erreurs selon le niveau scolaire.

2. Débit d'écriture

Le tableau IV.6 présente l'ensemble des données concernant les liens entre le débit d'écriture (en millisecondes par caractère) et les types d'erreurs orthographiques produites et ce, selon le niveau scolaire.

Tableau IV.6 : Débit d'écriture (en ms par caractère) selon le type d'erreurs et le niveau scolaire.

| Niveau scolaire | Types d'erreurs | N | Débit | Moyenne | Écart-type |
|-------------------|-----------------|-------------------|-------------------|---------------|--------------|
| CM2 | Lex seule | 135 | Débit | 6.65 | 0.31 |
| | | | Débit brut | 451.64 | 21.29 |
| | Morpho seule | 405 | Débit | 6.62 | 0.26 |
| | | | Débit brut | 417.83 | 16.53 |
| | Phono-lex | 62 | Débit | 6.69 | 0.16 |
| | | Débit brut | 490.95 | 11.81 | |
| 5 ^e | Phono-morpho | 22 | Débit | 6.53 | 0.21 |
| | | | Débit brut | 341.15 | 11.09 |
| | Total | 624 | Débit | 6.63 | 0.27 |
| | | | Débit brut | 428.17 | 17.18 |
| | 5 ^e | Lex seule | 177 | Débit | 6.55 |
| Débit brut | | | | 353.21 | 8.47 |
| Morpho seule | | 579 | Débit | 6.55 | 0.17 |
| | | | Débit brut | 353.39 | 8.91 |
| Phono-lex | | 68 | Débit | 6.56 | 0.12 |
| | | Débit brut | 360.84 | 6.61 | |
| 5 ^e | Phono-morpho | 41 | Débit | 6.49 | 0.16 |
| | | | Débit brut | 308.49 | 7.66 |
| | Total | 865 | Débit | 6.55 | 0.16 |
| | | Débit brut | 351.65 | 8.62 | |

| | | | | | |
|----------------|--------------|-------------------|-------------------|---------------|-------------|
| 3 ^e | Lex seule | 184 | Débit | 6.49 | 0.13 |
| | | | Débit brut | 312.03 | 6.25 |
| | Morpho seule | 659 | Débit | 6.48 | 0.15 |
| | | | Débit brut | 300.01 | 6.84 |
| | Phono-lex | 128 | Débit | 6.50 | 0.12 |
| | | Débit brut | 319.51 | 5.99 | |
| | Phono-morpho | 27 | Débit | 6.45 | 0.14 |
| | | | Débit brut | 282.70 | 6.10 |
| | Total | 998 | Débit | 6.48 | 0.14 |
| | | | Débit brut | 304.15 | 6.64 |

L'ANOVA a mis en évidence un effet significatif du facteur Type d'erreurs orthographiques sur le débit en CM2 ($F(3,620)=2.62, p=.05$). Le Tukey test montre un effet tendanciel entre les erreurs phono-morphosyntaxiques et les erreurs phono-lexicales ($p=.077$). Le débit est donc tendanciellement plus lent lorsque le mot comporte au moins une erreur phono-lexicale que lorsque l'erreur est de type phono-morphosyntaxique.

Cet effet n'est pas observé en 5^e et en 3^e, ($F(3,861)=1.88, p=.132$ en 5^e et $F(3,994)=2.20, p=.086$ en 3^e). La vitesse d'écriture ne varie donc pas significativement selon l'erreur commise par l'élève.

3. Pauses

Les liens entre les temps de pauses et les types d'erreurs orthographiques produites ont été analysés à la fois pour les pauses avant les mots et pour les pauses internes.

3.1. Pauses précédant les mots erronés

Le tableau IV.7 présente l'ensemble des données concernant les liens entre les pauses précédant les mots (en millisecondes) et les types d'erreurs orthographiques produites et ce, selon le niveau scolaire. Le graphique IV.6 résume ces données et met en évidence les résultats significatifs ($p<.05$).

Tableau IV.7 : Pauses (en ms) avant les mots selon le type d'erreurs et le niveau scolaire.

| Niveau scolaire | Types d'erreurs | N | Pauses avant | Moyenne | Écart-type |
|------------------|-----------------|------------------|------------------|---------------|--------------|
| CM2 | Lex seule | 152 | P_AVA | 6.83 | 0.42 |
| | | | P_AVAbrut | 673.54 | 40.98 |
| | Morpho seule | 468 | P_AVA | 6.82 | 0.39 |
| | | | P_AVAbrut | 659.32 | 37.53 |
| | Phono-lex | 67 | P_AVA | 6.82 | 0.37 |
| P_AVAbrut | | | 658.54 | 36.05 | |
| Phono-morpho | 29 | P_AVA | 6.80 | 0.34 | |
| | | P_AVAbrut | 631.18 | 31.35 | |
| Total | 716 | P_AVA | 6.82 | 0.39 | |
| | | P_AVAbrut | 661.07 | 37.82 | |

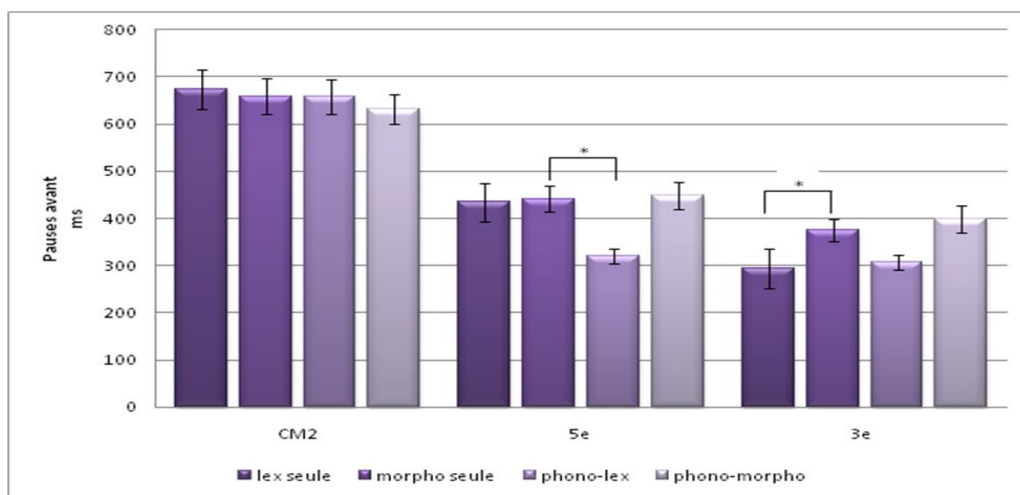
| | | | | | |
|----------------|--------------|--------------------|--------------------|---------------|--------------|
| 5 ^e | Lex seule | 196 | P_AVA | 6.64 | 0.40 |
| | | | P_AVAb Brut | 434.25 | 26.27 |
| | Morpho seule | 668 | P_AVA | 6.64 | 0.40 |
| | | | P_AVAb Brut | 441.38 | 26.85 |
| | Phono-lex | 77 | P_AVA | 6.51 | 0.33 |
| | | P_AVAb Brut | 320.36 | 16.10 | |
| | Phono-morpho | 45 | P_AVA | 6.65 | 0.43 |
| | | | P_AVAb Brut | 447.50 | 28.72 |
| | Total | 986 | P_AVA | 6.63 | 0.40 |
| | | | P_AVAb Brut | 429.35 | 25.92 |
| 3 ^e | Lex seule | 210 | P_AVA | 6.47 | 0.36 |
| | | | P_AVAb Brut | 294.01 | 16.35 |
| | Morpho seule | 760 | P_AVA | 6.57 | 0.41 |
| | | | P_AVAb Brut | 375.11 | 23.54 |
| | Phono-lex | 135 | P_AVA | 6.49 | 0.34 |
| | | P_AVAb Brut | 307.22 | 16.00 | |
| | Phono-morpho | 32 | P_AVA | 6.60 | 0.47 |
| | | | P_AVAb Brut | 398.27 | 28.48 |
| | Total | 1137 | P_AVA | 6.55 | 0.40 |
| | | | P_AVAb Brut | 350.80 | 21.38 |

Les résultats de l'ANOVA montrent un effet significatif du type d'erreurs orthographiques sur les pauses précédant les mots, en 5^e comme en 3^e. L'effet est très marqué chez les élèves de 3^e ($F(3,982)=2,85$, $p=.037$ pour les élèves de 5^e et $F(3,1133)=5,13$, $p=.002$ pour les élèves de 3^e). Il n'y a en revanche pas d'effet significatif observé dans les productions d'élèves de CM2 ($F(3,714)=0,05$, $t=ns$).

Le Tukey test a permis de mettre en évidence le fait que, au sein de la classe de 5^e, les pauses avant les mots comportant au moins une erreur « morpho. seule » sont significativement plus longues que celles précédant les mots avec au minimum une erreur « phono-lexicale » ($p=.020$). Ce résultat suggère que le traitement des mots comportant une ou plusieurs erreurs morphosyntaxiques est différent de ceux comprenant au moins une erreur lexicale et entraînant une altération phonologique.

En 3^e, les pauses sont significativement plus longues devant des mots comportant au moins une erreur « morpho. seule » que devant des mots avec erreur « lex. seule » ($p=0,004$). Le coût cognitif diffère selon que le mot comporte une erreur morphosyntaxique au minimum ou une (ou des) erreur(s) lexicale(s).

Il n'est pas observé d'autre effet du type d'erreurs orthographiques sur les pauses précédant les mots erronés.



Graphique IV.6 : Liens entre les pauses (en ms) avant les mots erronés et le type d'erreurs selon le niveau scolaire.

3.2. Pauses internes aux mots erronés

Le tableau IV.8 présente l'ensemble des données concernant les liens entre les pauses internes aux mots (en millisecondes) et les types d'erreurs orthographiques produites et ce, selon le niveau scolaire. Le graphique IV.7 résume ces données et met en évidence les résultats significatifs ($p < .05$).

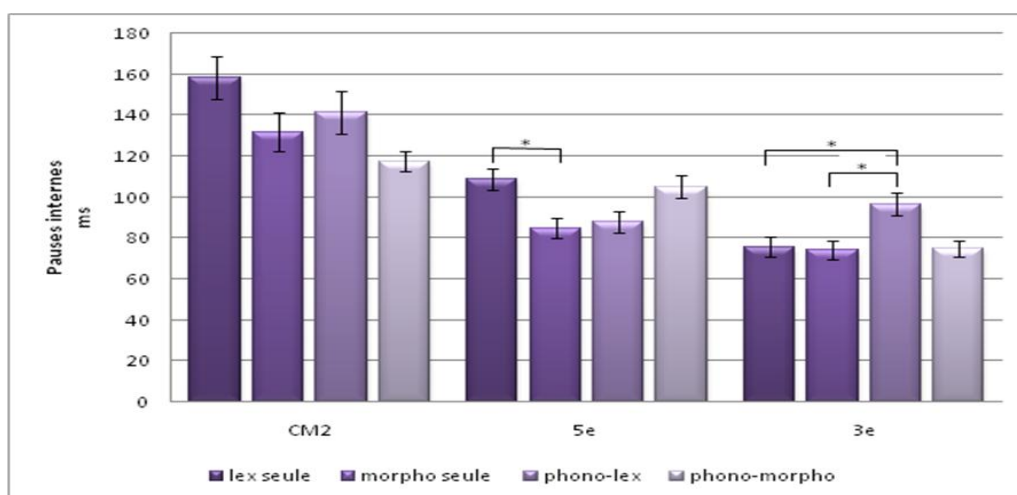
Tableau IV.8 : Pauses (en ms) internes aux mots selon le type d'erreurs et le niveau scolaire.

| Niveau scolaire | Types d'erreurs | N | Pauses internes | Moyenne | Écart-type |
|-------------------|-----------------|-------------------|-------------------|---------------|--------------|
| CM2 | Lex seule | 132 | P_Int | 6.20 | 0.41 |
| | | | P_Int brut | 158.49 | 10.53 |
| | Morpho seule | 356 | P_Int | 6.12 | 0.45 |
| | | | P_Int brut | 131.83 | 9.59 |
| | Phono-lex | 60 | P_Int | 6.15 | 0.45 |
| P_Int brut | | | 141.25 | 10.27 | |
| Phono-morpho | 17 | P_Int | 6.07 | 0.26 | |
| | | | P_Int brut | 117.49 | 4.99 |
| 5 ^e | Lex seule | 172 | P_Int | 6.04 | 0.29 |
| | | | P_Int brut | 108.83 | 5.27 |
| | Morpho seule | 487 | P_Int | 5.93 | 0.35 |
| | | | P_Int brut | 84.99 | 5.03 |
| | Phono-lex | 63 | P_Int | 5.94 | 0.35 |
| P_Int brut | | | 88.07 | 5.20 | |
| Phono-morpho | 37 | P_Int | 6.02 | 0.33 | |
| | | | P_Int brut | 105.21 | 5.77 |
| Total | 759 | P_Int | 5.96 | 0.34 | |
| | | P_Int brut | 91.10 | 5.20 | |

| | | | | | |
|----------------|--------------|-------------------|-------------------|--------------|-------------|
| 3 ^e | Lex seule | 175 | P_Int | 5.88 | 0.38 |
| | | | P_Int brut | 75.84 | 4.90 |
| | Morpho seule | 554 | P_Int | 5.87 | 0.36 |
| | | | P_Int brut | 74.51 | 4.55 |
| | Phono-lex | 122 | P_Int | 5.98 | 0.34 |
| | | P_Int brut | 96.57 | 5.44 | |
| | Phono-morpho | 25 | P_Int | 5.87 | 0.31 |
| | | | P_Int brut | 74.75 | 3.95 |
| | Total | 876 | P_Int | 5.89 | 0.36 |
| | | | P_Int brut | 77.21 | 4.73 |

Les résultats de l'ANOVA mettent en évidence un effet du facteur Type d'erreurs orthographiques sur les pauses internes aux mots erronés dans les productions d'élèves de 5^e et 3^e ($F(3,755)=4,76$, $p=.003$ en 5^e et $F(3,872)= 3,46$, $p=.016$ en 3^e). L'effet est particulièrement marqué chez les élèves de 5^e. Cet effet n'est pas présent dans les productions des élèves de CM2 ($F(3,561)=1,27$, $p=.286$).

En 5^e, les pauses au sein de mots comportant au moins une erreur « lex. seule » sont significativement plus longues que lorsque le mot comporte une ou plusieurs erreurs « morpho. seules » ($p=.002$). En 3^e, les pauses avant les mots sont significativement plus longues lorsque les erreurs sont de type « phono-lex » que lorsqu'elles sont « morpho. seules » ($p=.010$) ou « lex. seules » ($p=.036$).



Graphique IV.7 : Liens entre les pauses (en ms) internes aux mots erronés et le type d'erreurs selon le niveau scolaire.

4. Révisions

Les révisions témoignent de la dynamique en temps réel de la production écrite. À ce titre, elles sont un indice des connaissances orthographiques puisqu'elles reflètent ce sur quoi le scripteur tente de se corriger.

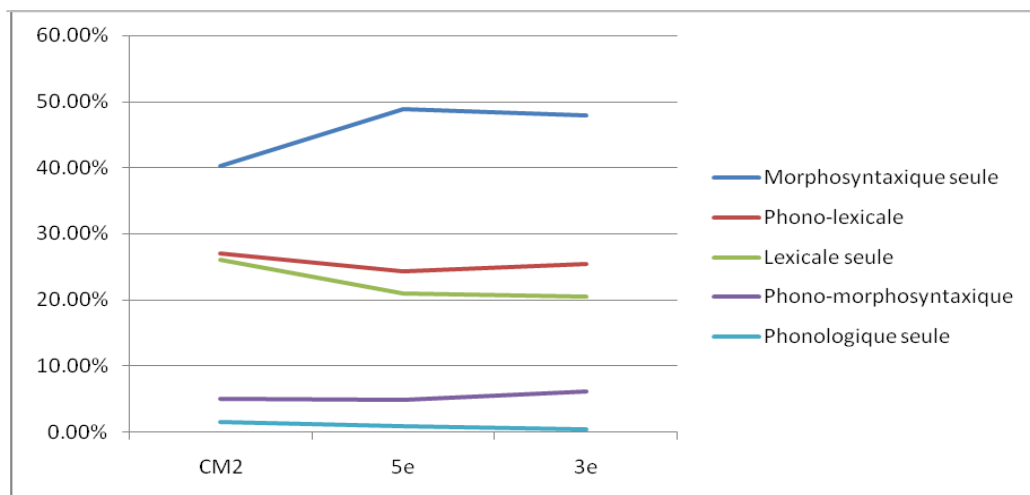
Le tableau IV.9 et le graphique IV.8 résument les principales données concernant les révisions produites dans ce corpus.

Tableau IV.9 : Révisions des mots erronés selon le niveau scolaire et le type d'erreurs produites.

| Niveau scolaire | Nombre de mots révisés | Pourcentage de mots révisés | Pourcentage de mots révisés selon le type d'erreurs | | | | |
|----------------------|------------------------|-----------------------------|---|-----------|-----------|--------------|-------------|
| | | | Morpho seule | Phono-lex | Lex seule | Phono-morpho | Phono seule |
| CM2 | 247 | 22.01% | 40.32% | 26.98% | 26.03% | 5.08% | 1.59% |
| 5^e | 330 | 21.98% | 48.94% | 24.24% | 20.94% | 4.94% | 0.94% |
| 3^e | 391 | 22.19% | 47.94% | 25.43% | 20.51% | 6.20% | 0.43% |
| Moyenne | 323 | 22.06% | 45.73% | 25.55% | 22.49% | 5.41% | 0.99% |

Les révisions mériteraient bien sûr un traitement et une analyse spécifique ainsi qu'un intérêt plus poussé. Ces calculs de pourcentage permettent d'observer que :

- le pourcentage de mots révisés reste relativement constant d'un niveau scolaire à l'autre (22.06% en moyenne) ;
- les erreurs morphosyntaxiques sont davantage révisées par les élèves de 5^e et 3^e (48.94% et 47.94%) que par les élèves de CM2 (40.32%) ;
- les erreurs lexicales seules (phonologiquement plausibles) sont davantage révisées par les élèves de CM2 (26.03%) que par les élèves de 5^e et de 3^e (20.94% et 20.51%) ;
- la hiérarchie des proportions de mots révisés par type d'erreurs est similaire d'une classe à l'autre (phonologique seule < phono-morphosyntaxique < lexicale seule < phono-lexicale < morphosyntaxique seule).



Graphique IV.8 : Révisions des mots erronés selon le niveau scolaire et le type d'erreurs produites.

Chapitre V
DISCUSSION DES RESULTATS

I. Rappels des données de l'étude

1. Objectif

L'objectif de cette étude était de replacer le processus orthographique dans le contexte global de la production écrite et de définir les liens pouvant exister entre erreurs orthographiques et paramètres en temps réel de la production écrite chez des élèves tout-venants scolarisés en ZEP et REP.

2. Hypothèses

Des variations des données de temps réel (débit d'écriture et pauses) étaient attendues, autant en comparant mots sans erreur orthographique et mots avec erreurs orthographiques qu'en confrontant ces paramètres aux différents types d'erreurs orthographiques produites.

3. Méthode

Pour cela, les temps de pause (avant et à l'intérieur des mots) ainsi que le débit d'écriture ont été mis en corrélation avec la présence ou l'absence d'erreurs orthographiques ainsi qu'avec le type d'erreurs orthographiques (phonologique seule, lexicale seule, morphosyntaxique seule, phono-lexicale et phono-morphosyntaxique). Chacune de ces mises en relation a été réalisée à trois niveaux scolaires (CM2, 5^e et 3^e).

II. Interprétation des résultats

Avant d'aborder les résultats en lien avec nos hypothèses, nous proposerons une analyse des données générales de notre corpus.

1. Données générales : fluence, pourcentages d'erreurs et pourcentages de types d'erreurs par niveau de classe

Ces données permettent la compréhension de la dynamique de la production orthographique sans pour autant traiter des paramètres temporels en tant que tels.

1.1. Comparaison de la fluence écrite par niveau scolaire

Les résultats montrent une augmentation de la quantité de mots produits entre le CM2, la 5^e et la 3^e.

L'automatisation complète du graphisme se fait en moyenne, comme le rappellent Alamargot, Plane, Lambert et Chesnet (2010), durant la classe de 3^e. Pour autant, il s'agit d'un processus d'acquisition relativement linéaire dont les performances ne cessent de s'améliorer avant même d'être automatisées. Graham et al. (1997) ont confirmé le poids de la transcription graphique sur la quantité de texte produit par des enfants. L'automatisation progressive du graphisme entre le CM2 et la 3^e peut donc expliquer en partie ce constat.

De plus, ce résultat va dans le sens des attentes pédagogiques exprimées par les programmes d'enseignement officiels (10 lignes en CM2, 30 lignes en 5^e et 40 lignes en 3^e ; Bulletins officiels 2008 a et b). L'augmentation de la fluence, compétence entraînée en milieu scolaire, s'observe donc au sein de notre étude.

1.2. Comparaison des pourcentages d'erreurs par niveau scolaire

Les résultats montrent une baisse progressive du pourcentage d'erreurs entre le CM2 et la 3^e. Le processus orthographique (de bas niveau), plus précisément lexical, n'est automatisé qu'à la fin du lycée mais est, comme le graphisme, en amélioration relativement linéaire des premiers apprentissages à l'expertise (Alamargot et al., 2010). Les capacités d'orthographe grammaticale, bien que jamais complètement automatisées, s'améliorent également avec le niveau scolaire.

Dans notre étude, cette amélioration des performances s'avère plus marquée entre le CM2 et la 5^e (8,57 points d'écart) qu'entre la 5^e et la 3^e (1,41 points). Même si l'automatisation de l'orthographe est progressive, l'évolution des exigences et des compétences entre le primaire et le collège pourrait expliquer cette nette différence.

Les élèves, à mesure qu'ils progressent en niveau scolaire, produisent à la fois moins d'erreurs orthographiques et des textes plus longs, ce qui démontre un double progrès.

2. Comparaison des pourcentages de types d'erreurs par niveau scolaire

2.1. Erreurs phonologiques seules

Ces erreurs sont très rares quel que soit le niveau scolaire. Elles sont en majorité liées à une méconnaissance de la valeur phonologique contextuelle de certaines graphies. Celles-ci pourraient être considérées, selon la définition de Jaffré et Fayol (1997), comme des « zones de fragilité » plus propices aux erreurs.

Ces résultats peuvent s'expliquer d'une part par la relative rareté de ces situations et d'autre part par le fait que le scripteur, face à une difficulté orthographique, est sensible à la forme phonologique du mot. Ainsi, faisant appel à ses compétences en langage oral et à sa compréhension du principe alphabétique, il peut corriger son erreur phonologique (par

l'ajout d'un accent par exemple). Le mot produit est donc phonologiquement plausible et l'erreur persistante est considérée selon notre typologie comme « lexicale seule ».

2.2. Erreurs phono-morphosyntaxiques

Les résultats montrent que ce type d'erreurs est rare et son pourcentage constant pour les trois classes étudiées. Nous pouvons supposer que les niveaux de langage oral et de transcodage des participants sont suffisants pour éviter ce type d'erreurs.

2.3. Erreurs phono-lexicales, lexicales seules

Le collectif de l'Inserm (2007) rapporte que les erreurs non phonologiquement plausibles se raréfient au cours de l'apprentissage alors que les erreurs phonologiquement plausibles restent fréquentes même dans les productions d'experts. Une même tendance pouvait être attendue pour les classes étudiées au sein de cette étude (CM2, 5^e et 3^e). Cependant, peu d'études rendent compte du pourcentage d'erreurs attendu à chaque niveau scolaire au sein d'une production écrite.

Nos résultats concernant les erreurs lexicales (phonologiquement plausibles) s'inscrivent dans la lignée du constat de l'Inserm (2007).

En revanche, le pourcentage d'erreurs phono-lexicales (non phonologiquement plausibles) diminue effectivement entre le CM2 et la 5^e mais augmente entre la 5^e et la 3^e, pour atteindre son plus haut niveau au sein de cette classe. Ce constat ne corrobore donc pas la raréfaction attendue de ce type d'erreurs. Nous développons à ce sujet deux interprétations :

- le fait que les élèves produisent des erreurs non phonologiquement plausibles suggère l'existence de difficultés d'application des conversions phonèmes-graphèmes, fortement liée aux capacités de langage oral et à la qualité des apprentissages premiers. Le fait que la population de notre étude soit issue d'un contexte socioculturel et socioprofessionnel défavorisé et soit scolarisée dans des établissements d'éducation prioritaire pourrait expliquer en partie ces résultats. À ce propos, Chevrie-Muller, Watier, Arabia, Arabia et Dellatolas (2005 ; cités par Dellatolas & Peralta, 2007) ont montré chez des enfants jeunes (3 ans et demi) une influence du niveau socio-culturel ainsi que du bilinguisme sur leur niveau de langage oral. Pour autant, concernant l'entrée dans le langage écrit, l'étude de Watier, Dellatolas et Chevrie-Muller (2006 ; cités par Dellatolas & Peralta, 2007) démontre, au CE1, une influence du niveau socio-culturel mais pas du bilinguisme ;
- les principales études portant sur l'orthographe analysent les performances des participants au sein de productions de mots isolés là où notre étude leur imposait la rédaction de textes libres. Au regard de la théorie capacitaire (McCutchen, 1994, 1996), la gestion des contraintes et des processus rédactionnels augmente nécessairement le coût cognitif alloué à la tâche et influe sur les performances orthographiques. Ici, cette augmentation du coût pourrait avoir eu un impact sur la quantité d'erreurs produites.

2.4. Erreurs morphosyntaxiques seules

Les erreurs morphosyntaxiques seules rassemblent la majorité des erreurs produites, quel que soit le niveau scolaire. Au vu de la typologie employée, ces erreurs correspondent à des difficultés portant sur la morphologie flexionnelle (déclinaison et conjugaison). Or, un apprentissage explicite et l'application consciente de règles sont nécessaires à la maîtrise de cette compétence. Cela explique, d'une part, que le coût attentionnel qui lui est associé soit supérieur à celui des autres traitements orthographiques et, d'autre part, que ce processus orthographique ne soit jamais complètement automatisé.

Une augmentation du pourcentage d'erreurs morphosyntaxiques seules est observée entre le CM2 et la 3^e et plus particulièrement entre le CM2 et la 5^e (4.53 points d'écart entre le CM2 et la 5^e et 0.64 points entre la 5^e et la 3^e).

Ceci peut être expliqué par le poids des contraintes rédactionnelles qui pèse davantage sur les élèves de collège que sur les élèves de primaire. Les programmes d'enseignement officiels l'attestent. Alors que les élèves de CM2 découvrent les premières exigences formelles de la rédaction, il est demandé aux élèves de 5^e de « maîtriser la narration et la description [...] de se faire comprendre, de prendre en compte son destinataire et de s'adapter à la situation de communication définie par les consignes » (BO 2008 b, p.8). Dans l'optique de l'entrée au lycée, il est attendu des élèves de 3^e « de développer et d'affirmer [leur] point de vue dans l'argumentation, de mettre l'accent sur l'implication et l'engagement (opinion, conviction, émotion), ou au contraire, la mise à distance et le détachement (objectivité, distance critique, humour) » (BO 2008 b, p.12).

Les collégiens, entraînés et évalués sur des aspects plus complexes de l'activité rédactionnelle, sans pour autant avoir complètement automatisé les processus « mécaniques » (orthographe et graphisme), connaissent une compétition des processus rédactionnels que l'on peut supposer supérieure à celle pesant sur les élèves de primaire. Au sein des traitements orthographiques, il est envisageable que l'application de la morphologie flexionnelle soit d'autant plus problématique.

Finalement, toutes les interprétations des résultats concernant le pourcentage d'erreurs par type sont compatibles avec la théorie de la triple détermination des erreurs verbales de Fayol et Miret (2005). Ainsi les erreurs orthographiques relevées dans ce corpus s'expliquent par une surcharge de coût cognitif, la fragilité d'une composante et/ou par l'impact de l'objectif poursuivi par le scripteur sur son énoncé.

3. Paramètres en temps réel et erreur orthographique

Étudier les paramètres en temps réel en lien avec les erreurs orthographiques au sein d'une production écrite de texte s'avère complexe. Les données correspondant aux différentes hypothèses de notre étude en sont le reflet et excluent en général une simple confirmation ou infirmation de celles-ci.

3.1. Retour sur les hypothèses

Nos résultats ne confirment qu'en partie l'hypothèse générale selon laquelle la présence d'erreurs orthographiques influe sur la valeur des paramètres en temps réel (débit d'écriture, pauses avant les mots, pauses internes). Ces influences, complexes puisque variant avec la nature du paramètre temporel étudié et le niveau scolaire considéré, nécessitent une description détaillée (cf. Analyse de résultats et confrontation à la littérature).

La première hypothèse opérationnelle selon laquelle les valeurs des paramètres d'analyse en temps réel diffèrent selon que les mots contiennent des erreurs orthographiques ou non est confirmée. Cependant, les résultats ne suivent pas toujours la tendance attendue :

- le débit d'écriture n'est pas plus lent pour les mots avec au moins une erreur orthographique que pour les mots sans erreur ;
- les pauses avant les mots erronés ne sont pas plus longues que pour les mots sans erreur ;
- les pauses internes sont effectivement plus longues lorsque le mot est erroné.

En somme, il n'existe pas de relation simple entre la production d'erreurs et la variation des paramètres temporels.

En ce qui concerne la deuxième hypothèse opérationnelle, les résultats confirment en partie l'influence du type d'erreurs sur les paramètres en temps réel. Ils sont en effet hétérogènes et dépendants du niveau scolaire.

Pour tous ces résultats, une analyse plus précise est proposée.

3.2. Analyse des résultats et confrontation à la littérature

Les paramètres en temps réel (débit d'écriture et pauses) ainsi que les révisions sont autant d'éléments sur lesquels s'appuie la recherche concernant la dynamique de la production écrite.

3.2.1. Comparaison des mots avec ou sans erreur(s) orthographique(s)

a. Pauses internes aux mots

Les pauses internes sont significativement plus longues (et ce dans les trois classes) lorsque le mot est erroné que lorsqu'il est correctement orthographié. Cela laisse supposer qu'elles seraient liées au traitement de l'erreur orthographique.

Ces observations peuvent être analysées à la lumière du modèle hiérarchique d'interprétation des pauses d'écriture rapporté par Chesnet et Alamargot (2005). Celui-ci distingue plusieurs types de pauses qui, selon leur localisation et leur durée dans l'énoncé,

seraient le reflet de traitements différents et de niveaux de complexité variables. Ainsi, le scripteur peut marquer une pause avant un paragraphe, une phrase, une proposition grammaticale, un syntagme, un mot ou à l'intérieur de celui-ci.

Au regard de ce modèle, les pauses internes de notre corpus concernent l'unité mot. Il semble donc plus aisé d'en fournir une interprétation. En effet, Chesnet et Alamargot (2005) rappellent que les pauses entre les mots sont dévolues à l'exécution motrice, à la sélection lexicale et au traitement orthographique. Les pauses internes, encore davantage centrées sur l'unité mot devraient d'autant plus être allouées à des traitements de bas niveau (orthographe et graphisme).

Cela confirme notre hypothèse : le coût cognitif étant supérieur en cas d'erreur orthographique, il est cohérent que les pauses internes aux mots soient plus longues lorsque le mot traité est erroné.

b. Pauses précédant les mots

Dans notre corpus, les pauses avant les mots ne sont pas distinguées selon la taille de l'unité qu'elles précèdent. Ceci en complique donc l'interprétation. Elles peuvent être en effet tout autant le reflet de traitements de haut niveau (planification, génération de texte et révision) que de bas niveau (exécution motrice et traitement orthographique). Déterminer précisément quels sont les processus effectivement associés à ces pauses est impossible, du moins sans autre indicateur de l'activité mentale (protocoles verbaux ou enregistrement de l'activité oculaire) (Chesnet & Alamargot, 2005). Ainsi, nous ne pouvons inférer que les pauses avant les mots de notre corpus soient strictement allouées au traitement orthographique.

Au sein de notre corpus, les pauses avant les mots sont, quel que soit le niveau scolaire, plus longues pour les mots ne comportant pas d'erreurs que pour les mots erronés. En revanche, cette différence n'est significative que pour les classes de 5^e et 3^e. Une interprétation rapide de ces résultats laisse supposer que ces pauses n'interviennent pas dans le traitement de l'erreur orthographique. Il s'agit d'approfondir cette analyse.

Foulin (1998 ; cité par Chesnet & Alamargot, 2005) postule que la durée d'une pause reflète le nombre et/ou la complexité des traitements. Ainsi, plus une pause est courte, moins le traitement serait coûteux. Dès lors, le fait que les pauses précédant les mots correctement orthographiés soient significativement plus longues pourrait être la conséquence de leur allocation à des traitements plus complexes ou plus nombreux. Les pauses avant les mots erronés, plus courtes, pourraient :

- traduire le fait que les sujets n'ont pas le sentiment de commettre ou de risquer de commettre une erreur ;
- être le reflet d'un traitement moins complexe, peut-être de bas niveau. Il n'est donc pas exclu que les pauses précédant les mots erronés soient en partie dévolues au traitement de l'erreur orthographique.

c. Débit d'écriture

Alors qu'aucune différence n'est observée au sein de la classe de CM2, le débit d'écriture est plus lent pour les mots sans erreur dans les classes de 5^e et de 3^e. La présence d'une difficulté orthographique ayant pour conséquence une augmentation du coût cognitif et la baisse du débit d'écriture étant un moyen de répondre à cela, nous pouvons nous attendre à une diminution du débit d'écriture pour les mots erronés.

Nous formulons à ce sujet deux interprétations :

- l'augmentation du coût cognitif associé à la production d'une erreur pourrait être gérée par la présence et l'allongement d'une pause interne plutôt que par la diminution du débit. Les résultats concernant les pauses internes corroborent cette idée ;
- la diminution du poids des processus de bas niveau sur la production écrite permet d'allouer les ressources cognitives disponibles aux traitements de haut niveau. Le ralentissement du débit d'écriture pour les mots correctement orthographiés laisse supposer que les ressources disponibles sont utilisées pour des traitements de plus haut niveau (planification, génération de texte, analyse du texte déjà produit) ou de bas niveau mais ne portant pas sur le mot lui-même.

Nous observons également une augmentation globale du débit d'écriture d'une classe à l'autre, que le mot soit erroné ou non. Ceci témoigne de l'automatisation progressive du graphisme.

3.2.2. Comparaison des types d'erreurs orthographiques

a. Pauses précédant les mots erronés

La durée des pauses précédant les mots erronés relevés dans le corpus de la classe de CM2 est homogène. Quel que soit le type d'erreurs, le temps de latence se situe entre 631 et 674 millisecondes. Ces résultats laissent supposer qu'il n'existe pas de différence de traitement de l'erreur orthographique selon son type. Tous les types d'erreurs orthographiques (i.e. tous les sous-processus orthographiques) semblent donc avoir, en CM2, le même coût cognitif.

En 5^e, les pauses précédant les mots contenant des erreurs lexicales seules, morphosyntaxiques seules ou phono-morphosyntaxiques sont comprises, de façon relativement homogène, entre 429 et 448 millisecondes. En revanche, celles précédant les mots contenant des erreurs phono-lexicales sont plus courtes (320 ms) et ce significativement en comparaison avec les erreurs morphosyntaxiques seules. Cela pourrait laisser supposer que les pauses précédant les erreurs phono-lexicales ne sont pas dévolues à leur traitement. Le coût cognitif associé au traitement des trois autres types d'erreurs est de plus supérieur (les pauses étant plus longues).

Pour la classe de 3^e, la pause précédant les mots comportant une erreur morphosyntaxique seule est significativement plus longue que dans le cas d'une erreur

lexicale seule. Néanmoins, la durée des pauses précédant les erreurs phono-lexicales et lexicales seules est relativement homogène (respectivement 307 et 294 ms). De la même façon, les pauses précédant les mots contenant des erreurs morphosyntaxiques seules et phono-morphosyntaxiques présentent des durées relativement proches (375 et 398 ms). Ces résultats laissent supposer qu'il existe une différence de coût associé au traitement des erreurs de type lexical ou morphosyntaxique (interprétation à nuancer de par l'absence de significativité de certains écarts).

La comparaison des trois niveaux scolaires montre une baisse progressive de la durée des pauses quel que soit le type d'erreurs. Ceci rejoint les modèles développementaux concernant l'automatisation progressive de l'orthographe et du graphisme (Alamargot et al., 2010). La dissociation de traitement selon le type d'erreurs observée en 3^e et en 5^e ne se retrouve pas en CM2. L'analyse des pauses précédant les mots erronés suggère que le traitement orthographique prend des formes différentes selon le niveau scolaire.

Cependant, ces interprétations sont là encore à nuancer au regard des critiques déjà exposées concernant l'analyse de ce type de pauses. Leur rôle et leur signification sont difficilement analysables au vu des données disponibles. Le postulat de cette analyse est qu'à une pause précédant un mot correspond le coût cognitif associé au traitement orthographique de CE mot. Pourtant, aucune étude jusqu'à présent n'a prouvé cela. Le traitement orthographique d'un nom peut en effet se situer avant son déterminant. De même, il s'agit de ne pas oublier qu'une « pause avant », dans un énoncé, est également « la pause après » du mot précédent...

b. Pauses internes aux mots erronés

Les pauses internes aux mots erronés sont en moyenne (toutes classes confondues) plus courtes que les « pauses avant » (entre 74 et 159 ms contre 307 à 674 ms pour les « pauses avant »). Ce constat va dans le sens du modèle hiérarchique des pauses déjà présenté (Chesnet & Alamargot, 2005). Les « pauses internes », contrairement aux « pauses avant », seraient plutôt dévolues à un traitement de bas niveau (orthographe et graphisme).

La comparaison des « pauses internes » selon le type d'erreurs n'a montré, pour la classe de CM2, aucune différence significative. Qualitativement, les pauses associées aux erreurs lexicales seules sont les plus longues.

Les pauses internes aux mots produits par les élèves de 5^e sont significativement plus longues pour les erreurs lexicales seules que pour les erreurs phono-lexicales. Ce résultat laisse supposer que le traitement de l'erreur lexicale, phonologiquement plausible, est plus coûteux que celui de l'erreur phono-lexicale qui ne l'est pas. Cette différence de coût pourrait être imputée à l'effort de mise en adéquation phonologique entre les formes orale et écrite du mot.

À l'inverse, dans la classe de 3^e, les pauses internes aux mots contenant une erreur phono-lexicale sont significativement plus longues que celles concernant les erreurs morphosyntaxiques seules et lexicales seules. Ce résultat est à mettre en lien avec le pourcentage particulièrement élevé d'erreurs phono-lexicales observé pour la classe de 3^e.

Les erreurs non phonologiquement plausibles étant généralement rares à ce niveau d'âge, il n'est pas impossible que les élèves, produisant l'une d'elles, y soient particulièrement sensibles et marquent une pause plus longue pour la traiter.

En 5^e comme en 3^e, les pauses associées aux erreurs morphosyntaxiques seules sont toujours les plus courtes. La durée de ces pauses suggère que le traitement de ce type d'erreurs ne s'effectue pas préférentiellement à l'intérieur des mots. Or, le traitement morphosyntaxique étant, des traitements orthographiques, le moins automatisable, il était attendu que son coût, plus important que pour les autres types d'erreurs, soit perceptible au travers de ces pauses (particulièrement dévolues aux traitements de bas niveau). Ce coût de traitement pourrait alors être retrouvé avant (par anticipation) ou après le mot (au cours d'une révision). Les résultats concernant les « pauses avant » sont, comme présenté auparavant, difficilement analysables. En revanche, les mots comportant au moins une erreur morphosyntaxique sont les plus révisés du corpus.

Enfin, d'une classe scolaire à l'autre et comme nous l'avions observé pour les pauses précédant les mots, la durée des « pauses internes » diminue et ce, quel que soit le type d'erreurs.

c. Débit d'écriture

Quel que soit le niveau scolaire, aucun effet significatif n'est observable lorsque l'on met en relation débit d'écriture et types d'erreurs orthographiques. Cela conduit à penser que tous les types d'erreurs orthographiques influent relativement de la même façon sur le débit d'écriture. Considérant les résultats déjà obtenus lors de la comparaison du débit d'écriture des mots avec ou sans erreurs, ce constat confirme qu'à ces âges, le traitement de l'erreur orthographique ne serait pas reflété globalement par des variations du débit d'écriture.

d. Révisions

Un premier résultat montre une constance du pourcentage de mots révisés d'une classe à l'autre (1/5 des mots erronés produits). Cela suggère que, en toute logique, le scripteur, lorsqu'il écrit un texte libre, produit ce qu'il est capable de produire et révisé ce qu'il est capable de réviser, quel que soit son âge.

Les mots comportant au moins une erreur morphosyntaxique sont les plus révisés, et ce, quel que soit le niveau scolaire. Cette observation montre que les élèves semblent davantage attachés à la correction des erreurs morphosyntaxiques lorsqu'ils révisent leur production. Nous pouvons supposer que les élèves, plus sensibles à ce type d'erreurs, les détectent plus facilement. Les expériences et les apprentissages peuvent en être à l'origine. Même s'il s'agit, pour la classe de CM2, du type d'erreurs le plus révisé, il est à noter qu'elles le sont moins que pour les classes de 5^e et de 3^e (8 points d'écart entre le CM2 et la 5^e et seulement 1 entre la 5^e et la 3^e). Ceci peut être expliqué par des capacités méta-morphosyntaxiques et une connaissance des règles d'accords plus développées au collège qu'en primaire, ces compétences étant liées à la maturation cognitive et à l'expérience des scripteurs. Toutefois, Fayol (2008) déplore le peu de recherches visant à montrer l'efficacité de l'apprentissage explicite des règles morphosyntaxiques.

De plus, la possibilité de se détacher progressivement du coût alloué au graphisme et à l'orthographe d'usage devrait permettre d'aborder les problèmes relatifs à la phrase et aux traitements de plus haut niveau (tel la révision), comme le suggère la théorie de la capacité (Just et Carpenter, 1992 ; McCutchen, 1994, 1996).

Les erreurs lexicales seules (phonologiquement plausibles) sont davantage révisées par les élèves de CM2 que par les élèves de 5^e et de 3^e (5 points d'écart entre le CM2 et la 5^e et 0,5 points entre la 5^e et la 3^e). Les apprentissages implicites et explicites de la fin de l'école primaire étant axés principalement sur l'orthographe d'usage, cela pourrait expliquer cette différence.

III. Biais et limites de l'étude

1. Population

Pour être représentatif d'une population plus globale et permettre une généralisation des résultats, un échantillon se doit d'être minutieusement sélectionné. Le corpus analysé lors de cette étude a été recueilli auprès d'un échantillon exclusivement constitué d'élèves scolarisés en zone d'éducation prioritaire (ZEP) et réseau d'éducation prioritaire (REP).

De la même façon, le nombre de participants par niveau scolaire n'est pas strictement égal (46 ; 47 ; 46 élèves), ni pondéré selon le sexe (22 filles/24 garçons ; 23/24 ; 22/24).

Cette constitution n'enlevant en rien sa valeur aux résultats, elle nécessite cependant de la prudence quant à leur généralisation.

2. Création d'une typologie

La création d'une typologie d'erreurs orthographiques est dépendante de l'objectif visé et du champ de compétences de son concepteur. Ainsi :

- Catach et al. (1996) proposent une typologie à visée pédagogique, afin d'offrir aux enseignants la possibilité d'analyser les erreurs de leurs élèves en les classant ;
- Moats (1995) utilise sa typologie comme un outil de description linguistique (analyse de la langue) ;
- Estienne (1999) offre aux rééducateurs une typologie des erreurs à visée clinique ;
- Simon (1956 ; cité par Bellone, 2003) décrit ses catégories d'erreurs orthographiques dans le but d'évaluer les compétences orthographiques lors d'une épreuve de dictée.

Notre typologie se devait d'allier des connaissances linguistiques (connaissances développementales, repères d'acquisition), cognitivistes (temporalité de l'écriture), ainsi qu'un point de vue clinique à visée rééducative (causes des erreurs, mécanismes en jeu). Elle a su s'adapter au gré des exceptions rencontrées. Certains choix de classification

d'erreurs ont été faits au terme de discussions d'ordre clinique et d'échanges plutôt qu'en regard d'argumentations purement théoriques.

Relativement à notre champ d'étude et à nos compétences (davantage cliniques et rééducatives que conceptuelles et descriptives), notre souci a été de construire une typologie applicable statistiquement au domaine peu connu du temps réel.

3. Analyse

3.1. Données inhérentes au corpus

La plupart des données disponibles concernant le corpus analysé proviennent d'autres études. Ainsi, c'est lors d'analyses antérieures qu'ont été distingués les mots erronés des mots ne contenant pas d'erreur. Ce type d'analyse impliquant de nombreux correcteurs génère toujours un pourcentage d'erreurs. Certains items considérés à tort comme erronés ont ainsi été repérés mais notre analyse n'exclut pas la possibilité de quelques oublis et erreurs de jugement de notre part. La taille du corpus autorise heureusement cette marge d'erreurs.

3.2. Analyse des erreurs

L'analyse du corpus d'erreurs a été faite item par item, plusieurs fois, afin de pouvoir constituer une typologie efficace et éviter les oublis et erreurs de jugement inhérents à une analyse de cet ordre. L'analyse de l'orthographe au sein de la production écrite, par essence dynamique, ne pouvait se faire de la même façon que lors d'analyses de mots isolés a posteriori.

Pourtant, même si l'objectif était ici d'analyser la production orthographique au sein de l'ensemble dynamique qu'est la production écrite de texte, l'analyse des erreurs s'est faite nécessairement a posteriori. D'où ce paradoxe : l'analyse en temps réel de l'orthographe ne peut réellement se faire qu'a posteriori... Ainsi, même si certaines erreurs ont été aisément classées, d'autres auraient mérité une analyse plus fine des intentions de l'enfant et des processus en jeu. Ces exceptions restent rares.

IV. Perspectives, apports professionnels et personnels

1. Perspectives de recherche

1.1. Nature des mots

L'analyse de ce corpus ne distingue pas les mots selon leur nature. Or, le traitement orthographique pourrait différer selon ce critère. Étudier cette hypothèse à travers une analyse de paramètres en temps réel pourrait permettre d'affiner les connaissances sur le processus de transcription orthographique.

1.2. Modélisation des pauses en quartiles

Une analyse plus approfondie des pauses nécessite de prendre en compte, en plus de leur localisation, leur durée. Alamargot et al. (2010) proposent, dans la lignée du modèle hiérarchique d'interprétation des pauses d'écriture, une analyse des durées de pauses et un classement de celles-ci en quartiles afin de les différencier selon qu'elles sont allouées à des processus de haut ou de bas niveau. Selon cette méthodologie, les deux premiers quartiles rassembleraient les pauses dévolues au traitement graphique, le troisième correspondrait au traitement orthographique et le quatrième aux processus de haut niveau (planification, génération de texte et révision).

L'application de cette méthodologie permettrait de clarifier en partie la constitution des pauses, notamment celles précédant les mots, difficilement analysables.

1.3. Apport de l'analyse de l'input visuel

L'utilisation de tablettes graphiques et du logiciel Eye and Pen a permis le recueil de textes entiers et de paramètres temporels leur étant associés. Pourtant, la seule utilisation de la modalité graphique d'Eye and Pen prive l'analyse de nombreux résultats.

Ainsi, Chesnet et Alamargot (2005) critiquent le modèle hiérarchique d'interprétation des pauses d'écriture qui ne rend pas précisément compte de la dynamique de la production écrite. Celui-ci ne permet pas, en effet :

- de distinguer clairement les processus en jeu pendant les temps de pause, principalement pour les pauses les plus longues ;
- de décrire les traitements parallèles à la production graphique ;
- de mettre en évidence les retours sur le texte produit en amont, tout particulièrement lorsqu'aucune trace graphique n'en découle (recherche d'informations, vérifications).

Ces auteurs proposent de coupler les modalités graphique et oculomotrice (fixations et saccades). Le dispositif complet Eye and Pen est prévu pour pallier ces biais. C'est alors l'étude de cas, avant celle de grands effectifs, qui pourrait permettre une analyse plus fine de la production écrite et notamment du traitement orthographique dans ce contexte.

2. Apports pour la profession d'orthophoniste

C'est par la connaissance et la compréhension du fonctionnement cognitif, chez le tout-venant comme chez le patient porteur d'un trouble, que l'orthophoniste peut élaborer des techniques de rééducation et les appliquer dans sa pratique clinique. Il en est de même pour le pédagogue soucieux de proposer des méthodes d'enseignement et d'entraînement adaptées et le cas échéant des outils de remédiation efficaces. Ainsi, cette étude entend rendre compte du fonctionnement cognitif associé au traitement orthographique et plus particulièrement au traitement de l'erreur orthographique au sein de la production écrite de texte.

L'activité de rédaction libre de texte constitue à la fois un objectif évolutif de l'enseignement et la tâche majeure du scripteur expert. Elle est en ce sens davantage écologique que des tâches de dictée ou de production d'énoncés plus courts. Il paraît donc d'autant plus pertinent d'en connaître les modalités et les compétences qui les sous-tendent.

3. Apports personnels

Les recherches concernant le fonctionnement cognitif normal comme pathologique bousculent ou confirment nos idées et nos pratiques professionnelles. Ce travail nous aura confortées dans notre conviction quant à la nécessité de rester attentives à ces publications. D'un point de vue technique, l'expérience de la manipulation, de la compréhension et de l'interprétation de données statistiques autorisent une meilleure appréhension des articles scientifiques. Ce premier travail nous aura de plus permis d'entrouvrir la grande porte de la recherche.

Les difficultés d'orthographe constituent l'une des principales plaintes à l'origine de la consultation chez l'orthophoniste. En ce sens, des connaissances théoriques approfondies dans le domaine ne peuvent qu'asseoir notre future pratique clinique.

Connaître la nature des interactions existant entre les différents processus rédactionnels au sein de la production écrite constitue un savoir qui, bien que peu enseigné, semble indispensable à la pratique orthophonique. Les liens entre processus de bas niveau (orthographe et graphisme), automatisables et processus de haut niveau, qui ne peuvent s'élaborer que sur la base de cette automatisation, justifient nombre d'actions de rééducations et d'aménagements éducatifs.

Travailler à la consolidation du stock lexical orthographique paraît d'autant plus justifié que cela permettra l'accession aux autres compétences rédactionnelles. L'apprentissage de formes orthographiques d'abord isolées ne répond donc pas seulement à la plainte concernant l'orthographe mais est un préalable indispensable (comme l'automatisation du graphisme) aux compétences de plus haut niveau. De même, décharger l'enfant porteur de trouble du poids du graphisme et de l'orthographe par l'utilisation d'un ordinateur et d'un correcteur orthographique (par exemple) paraît indispensable afin de permettre la pleine expression de ses capacités rédactionnelles.

La gestion des contraintes est, pour l'orthophoniste, un enjeu majeur de la rééducation. Les activités proposées doivent en effet être ciblées afin de ne pas faire intervenir simultanément trop de compétences. Proposer une graduation des objectifs selon leur difficulté constitue en ce sens l'une des clés de l'automatisation des compétences travaillées. Ainsi, une meilleure conscience des coûts attentionnels associés aux différents processus de la production écrite permet de mieux penser les objectifs et micro-objectifs thérapeutiques.

CONCLUSION

L'erreur tout comme les données temporelles (débit d'écriture et pauses) pourraient constituer des traces des coûts cognitifs des opérations impliquées dans la production verbale. En ce sens, notre étude se proposait d'analyser, au sein de productions de textes, les erreurs orthographiques commises en lien avec les paramètres en temps réel leur étant associés et donc d'évaluer dans quelle mesure ces deux dimensions de la production écrite sont liées. Il s'agissait de décrire les variations des paramètres en temps réel de l'écriture lors de la production d'erreurs orthographiques. Les données temporelles des mots avec et sans erreurs orthographiques ainsi que celles des mots selon le type d'erreurs qu'ils contiennent ont été comparées.

Les résultats montrent essentiellement que les paramètres en temps réel de la production écrite ne sont pas associés de manière univoque à l'apparition et au traitement des erreurs orthographiques. Ainsi, les pauses internes aux mots sont clairement corrélées avec le traitement des erreurs orthographiques. Les pauses avant les mots, quant à elles, constituent des données temporelles difficilement analysables. Nos résultats montrent leur nature multifactorielle et font apparaître l'importance de les analyser sur la base d'une méthodologie spécifique capable de prendre en considération les variables qui influent sur leurs variations. Enfin, pour le débit d'écriture, les résultats rapportent que son ralentissement n'intervient pas systématiquement en lien avec le traitement d'erreurs orthographiques.

Ces résultats globaux méritent d'être nuancés. Suivant la nature de l'erreur considérée et le niveau scolaire du scripteur, les résultats présentent des variations dont la présentation détaillée rend compte. L'automatisation progressive des compétences orthographique et graphique ainsi que les apprentissages implicites et explicites des scripteurs peuvent expliquer les variations inter-classes. Les traitements orthographiques (phonologique, lexical et morphosyntaxique) entraînent des coûts cognitifs distincts qui pourraient expliquer les variations observées selon le type d'erreurs.

Ce champ d'étude, encore récent, demande donc le développement d'une méthodologie de recueil et d'analyse précise et adaptée. Les paramètres en temps réel de la production écrite s'inscrivent par nature dans une dynamique globale. Isoler un traitement (orthographique par exemple) pour l'analyser exige donc :

- de prendre en considération tous les traitements engagés dans une rédaction libre (haut et bas niveau) ;
- de proposer des méthodes de recueil adaptées (alliant les données de l'output graphique avec par exemple des données de l'input visuel par l'enregistrement de l'oculomotricité) ;
- d'analyser les résultats selon une méthodologie qui tienne compte de la dynamique de la production écrite (par l'analyse des zones oculaires, des durées et des localisations des pauses).

Notre travail de recherche constitue une étape dans la compréhension de la place de l'orthographe au sein de la rédaction de texte, performance centrale du scripteur expert. Il

s'inscrit dans une description dynamique du fonctionnement cognitif rédactionnel qui saura être utile à l'orthophoniste dans la définition et la gradation de ses objectifs rééducatifs et au pédagogue dans l'élaboration de méthodes d'enseignement, d'entraînement et de remédiation adaptées.

BIBLIOGRAPHIE

Ouvrages

Astolfi, J.-P., 1997, *Un outil pour enseigner*. Paris : ESF.

Babin, N. (2003). *Programmes et pratiques pédagogiques pour l'école élémentaire*. Paris : Hachette Livre.

Bellone, C. (2003). *Dyslexies et dysorthographies ; connaissances de base théoriques et pratiques : d'hier à aujourd'hui et demain*. Isbergues : Ortho Édition.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, NJ : Lawrence Erlbaum Associates.

Bonin, P. (2003). *Production verbales de mots. Approche cognitive*. Bruxelles : De Boeck Université.

Catach, N., (1989). *Les délires de l'orthographe : en forme de dictionnaire*. Paris : Plon.

Catach, N., Gruaz, C., & Duprez, D. (1996). *L'orthographe française : traité théorique et pratique avec des travaux d'application et leurs corrigés*. Paris : Nathan.

Ecalte, J. & Magnan, A. (2002). *L'apprentissage de la lecture : fonctionnement et développements cognitifs*. Paris : Armand Colin.

Estienne, F. (1999). *Méthode d'initiation à l'écrit pour les dyslexiques et les dysorthographiques*. Issy-les-Moulineaux : Masson.

Gombert, J.E. (1990). *Le développement métalinguistique*. Paris : PUF.

Institut National de la santé et de la recherche médicale (INSERM). (2007). *Dyslexie, dysorthographe, dyscalculie. Bilan des données scientifiques*. Paris : Les éditions Inserm.

Jaffré, J.-P., & Fayol, M. (1997). *Orthographe. Des systèmes aux usages*. Paris : Flammarion.

Levelt, W. J. M. (1989). *Speaking : from intention to articulation*. Cambridge, MA : the MIT press.

Ministère de l'éducation nationale, de la jeunesse et de la vie associative (2008a). *Cycle des approfondissements – Progressions pour le cours élémentaire deuxième année et le cours moyen*. (Bulletin officiel (BO) Hors-série n°3 du 19 juin 2008). Repéré à <http://www.education.gouv.fr/bo/2008/hs3/approfondissements.html>

Ministère de l'éducation nationale, de la jeunesse et de la vie associative (2008b). *Programmes du collège : Programmes de l'enseignement de français*. (Bulletin officiel (BO) spécial n°6 du 28 août 2008). Repéré à http://media.education.gouv.fr/file/special_6/21/8/programme_francais_general_33218.pdf

Moats, L. C. (1995). *Spelling : development, disabilities, and instruction*. Timonium, MA : York Press.

Pothier, B. (1994). *Comment les enfants apprennent l'orthographe*. Paris : Retz.

Reboul, O. (1980). *Qu'est-ce qu'apprendre ?* Paris : PUF.

Sprenger-Charolles, L., & Casalis, S. (1996). *Lire. Lecture/écriture : acquisition et troubles du développement*. Paris : PUF.

Chapitres d'ouvrages

Alamargot, D., & Fayol, M. (2009). Modelling the development of written composition. In R. Beard, D. Myhill, M. Nystrand, & J. Riley (Eds), *Handbook of Writing Development*. (pp. 23-47). Sage : United Kingdom.

Berninger, V.W., & Swanson, H.L. (1994). Modification of the Hayes and Flower model to explain beginning and developing writing. In E. Butterfield (Ed.), *Advances in cognition and educational practice. Vol. 2 : Children's writing : toward a process theory of the development of skilled writing*. (pp 57-82). Greenwich, CT : JAI Press.

Bock, K., & Levelt, W. J. M. (1994). Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics*. (pp. 945-984). New York : Academic Press.

Carlisle, J.F. (1995). Morphological awareness and early reading achievement. In L. Feldman (Ed.), *Morphological aspects of language processing*. (pp.189-209). Hillsdale, NJ : Lawrence Erlbaum Associates.

Content, A., & Zesiger, P. (1999). L'acquisition du langage écrit. In J.A. Rondal & X. Seron (Eds.), *Troubles du langage : bases théoriques, diagnostic et rééducation*. (pp. 179-209). Sprimont : Mardaga.

Dellatolas, G. & Peralta, F. (2007). Épidémiologie des troubles du langage oral et écrit de l'enfant. In C. Chevrie-Muller & J. Narbona (Eds.), *Le langage de l'enfant : Aspects normaux et pathologiques*. (pp. 283-297). Issy-les-Moulineaux : Masson.

Fayol, M. (2008). Les erreurs : manière d'appréhender la production orthographique. In M. Fayol, & J-P. Jaffré (Eds), *Orthographier* (pp. 123-136). Paris : Presses Universitaires de France.

Fayol, M. (2008). Enseigner pour faire apprendre. In M. Fayol, & J-P. Jaffré (Eds), *Orthographier* (pp. 211-227). Paris : Presses Universitaires de France.

Fayol, M. (2010). Apprendre l'orthographe. In M. Crahay et al. (Eds.), *Manuel de psychologie des apprentissages scolaires*. (pp. 257-272). Bruxelles : De Boeck.

Fayol, M., Foulin, J-N., Maggio, S., & Lété, B. (2010 in press). Towards a dynamic approach of how children and adults manage text production. In Grigorenko, E., Mambrino, E. & Preiss, D.D. (Eds.), *Handbook of writing : a mosaic of perspectives*. New York: Psychology Press.

Fayol, M., Thévenin, M.G., Jarousse, J.P., & Totereau, C. (1999). From learning to teaching to learning french written morphology. In T. Nunes (Ed.), *Learning to read : An integrated view from research and practice*. (pp. 43-64). Dordrecht : Kluwer.

Flower, L. S., & Hayes, J. R. (1980). The dynamics of composing : making plans and juggling constraints. In L. W. Gregg & E. R., Steinberg (Eds.), *Cognitive processes in writing*. (pp 31-50). Hillsdale : Lawrence Erlbaum Associates.

Foulin, J.N. (1997). Bons et faibles orthographes. In S. Guillard (Ed.), *Lecture et dyslexies*. (pp. 181-196). Grenoble : Les Pluriels de Psyché

Frith, U. (1980). Unexpected spelling problems. In. U. Frith (Ed.), *Cognitive processes in spelling*. (pp. 495-515). London : Academic Press.

Frith, U. (1985). Beneath the surface of developmental dyslexia. In K.E. Patterson, J.C. Marshall & M. Coltheart (Eds), *Surface dyslexia. Neuropsychological and cognitive studies of phonological reading*. (pp. 301-330). London : Erlbaum.

Jaffré, J.-P. (2006). L'orthographe du français est-elle une bonne orthographe ? In Les journées de l'observatoire national de la lecture (ONL) (Ed.), *Enseigner la langue : orthographe et grammaire*. Repéré à <http://onl.inrp.fr/ONL/publications/publi2006/enseigner>

Jaffré, J.-P. (2008). La mixité des orthographe. In M. Fayol, & J-P. Jaffré (Eds), *Orthographier* (pp. 85-101). Paris : Presses Universitaires de France.

Lété, B. (2003). Building the mental lexicon by exposure to print : A corpus-based analysis of French reading books. In P. Bonin (Ed.), *Mental lexicon. "Some words to talk about words"*. (pp. 187-214). Hauppauge, NY : Nova Science Publisher.

Lété, B. (2004). MANULEX : Le lexique des manuels scolaires de lecture. Implications pour l'estimation du vocabulaire des enfants de 6 à 11 ans. In E. Calaque & J. David (Eds.), *Didactique du lexique : Contextes, démarches, supports*. (pp. 241-257). Bruxelles : De Boeck.

McCutchen, D. (1994). The magical number three, plus or minus two: working memory in writing. In E. C. Butterfield (Ed.), *Advances in cognition and educational practice. Vol. 2: Children's writing : toward a process theory of the development of skilled writing*. (pp 1-30). Greenwich, CT : JAI Press.

Scardamalia, M., & Bereiter, C. (1983). The development of evaluative, diagnostic and remedial capabilities in children's composing. In M. Martlew (Ed.), *The psychology of written language: a developmental approach*. (pp. 67-95). London : Wiley.

Articles

Alamargot, D., & Lebrave, J. L. (2010). The study of professional writing : a joint contribution from cognitive psychology and genetic criticism. *European Psychologist*, 15, 12-22.

-
- Alamargot, D., Plane, S., Lambert, E., & Chesnet, D. (2010). Using eye and pen movements to trace the development of writing expertise : Case studies of a seventh, ninth and twelfth grader, graduate student, and professional writer. *Reading and writing*. Online first. doi : 10.1007/s11145-009-9191-9
- Berninger, V.W., Fuller, F., & Whitaker, D. (1996). A process model of writing development across the life span. *Educational Psychology Review*, 8(3), 193-218.
- Chanquoy, L., & Alamargot, D. (2002). Mémoire de travail et rédaction de textes : évolution des modèles et bilan des premiers travaux. *L'Année Psychologique*, 102, 363-398. Repéré à <http://http://cogprints.org/3425/1/WMandtext.pdf>
- Chesnet, D., & Alamargot, D. (2005). Analyses en temps réel des activités oculaires et graphomotrices du scripteur: intérêt du dispositif 'Eye and Pen'. *L'Année Psychologique*, 105(3), 477-520. doi : 10.3406/psy.2005.29706
- Chesnet D., Guillabert F., & Espéret E. (1994). G-STUDIO : un logiciel pour l'étude en temps réel des paramètres temporels de la production écrite. *L'année psychologique*, 94(2), pp. 283-293. doi : 10.3406/psy.1994.28757
- David J. (2003). La dimension orthographique dans les écrits des jeunes enfants. *Les dossiers des sciences de l'éducation*, 9, 29-39.
- Favart, M., & Olive, T. (2005). Modèles et méthodes d'étude de la production écrite. *Psychologie Française*, 50(3), 273-285.
- Fayol, M. (en préparation). Dynamique de la production écrite.
- Fayol, M., Miret, A. (2005). Écrire, orthographier et rédiger des textes. *Psychologie française*, 50(3), 391-402.
- Flower, L.S., & Hayes, J.R. (1981). A cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-87.
- Ganier, F. (2006). La révision de textes procéduraux. *Langages*, 164, 71-85.
- Gentry, J.R. (1978). Early spelling strategies. *The elementary school journal*, 79(2), 88-92.
- Gentry, J.R. (1982). An analysis of developmental spelling in GNYS AT WRK. *The reading teacher*. 36, 192-200.
- Graham, S., Berninger, V.W., Abbott, R.D., Abbott, S.P. & Whitaker, D. (1997). Role of mechanics in composing of elementary school students : A new methodological approach. *Journal of Educational Psychology* 89, 170-182.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension : individual differences in working memory. *Psychological Review*, 99(1), 122-149.
-

Largy, P. (2001). La révision des accords nominal et verbal chez l'enfant. *L'année psychologique*, 101(2), 221-245. doi : 10.3406/psy.2001.29554.

Lété, B., Sprenger-Charolles, L. & Colé, P. (2004). MANULEX : a grade-level lexical database from French elementary school readers. *Behavior Research Methods, Instruments & Computers*, 36(1), 156-166.

Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95(4), 492-527.

Martinet, C., & Valdois, S. (1999). L'apprentissage de l'orthographe d'usage et ses troubles dans la dyslexie développementale de surface. *L'Année Psychologique*, 99(4), 577-622.

McCutchen, D. (1996). A capacity theory of writing: working memory in composition. *Educational Psychology Review*, 8(3), 299-325.

Pacton, S., & Casalis, S. (2006). L'utilisation d'informations morphologiques en orthographe rend-elle le cauchemar(d?) des lettres muettes moins cauchemardesque ? *Rééducation Orthophonique*, 225, 129-144.

Pacton, S., Fayol, M., & Perruchet, P. (2005). Children's implicit learning of graphotactic and morphological regularities in French. *Child Development*. 76(2), 324-339.

Plane, S., Alamargot, D., & Lebrave, J.-L. (2010). Temporalité de l'écriture et rôle du texte déjà produit dans l'activité rédactionnelle. *Langages*, 177, 7-28. doi : 10.3917/lang.177.0007

Rey, A., Pacton, S., & Perruchet, P. (2005). L'erreur dans l'acquisition de l'orthographe. *Rééducation Orthophonique*. 222, 101-120.

Véronis, J. (1988). From sound to spelling in French: simulation on a computer. *Cahiers de Psychologie Cognitive*. 8(4), 315-334.

Dictionnaires et Encyclopédies

Rey, A. (1995). « Erreur ». Dans Dictionnaire historique de la langue française (Tome 1). Paris: Le Robert.

Rey, A. (1995). « Faute ». Dans Dictionnaire historique de la langue française (Tome 1). Paris: Le Robert.

GLOSSAIRE

Cordial

Logiciel de correction lexicale et grammaticale et d'aide à la rédaction de langue française (cf. Synapse Développement).

Eye and Pen

« Le logiciel Eye and Pen permet d'enregistrer et d'analyser en temps réel les mouvements du crayon des rédacteurs pendant qu'ils écrivent ou effectuent des pauses de réflexion. Si l'on dispose d'un oculomètre, Eye and Pen permet aussi d'enregistrer (de manière synchrone) les mouvements oculaires du rédacteur : identifier précisément la nature des informations regardées pendant l'écriture fournit des indications précieuses sur le déroulement des mécanismes impliqués dans la production de documents écrits. » (Description extraite du site internet officiel <http://www.eyeandpen.net>)

Manulex

Manulex (pour lexique des manuels) est une base de données actualisée rendant compte, au travers de l'analyse de 54 manuels scolaires, du lexique écrit adressé à l'enfant (du CP au CM2, soit 1,9 millions de mots au total) (Lété, 2003, 2004 ; Lété, Sprenger-Charolles & Colé, 2004). Deux lexiques ont été distingués : un lexique des formes orthographiques et un lexique des lemmes (unités autonomes du lexique d'une langue), chacun comprenant des informations sur la fréquence d'apparition des mots.

Librement téléchargeable à l'adresse <http://unpc.univ-lyon2.fr/~lete/manulex/index.htm> et interrogeable sur le site de Lexique à l'adresse <http://www.lexique.org/moteur/>

Paradigme expérimental de double tâche

Méthode d'analyse en temps réel classique en psychologie cognitive « [reposant] sur le postulat selon lequel le système cognitif dispose d'une quantité limitée de ressources utilisées par les opérations de traitement de l'information. Le fait de demander aux individus d'accomplir simultanément deux tâches (principale et secondaire) les oblige à répartir leurs ressources entre ces deux tâches. [...] Ainsi, lorsqu'une interférence est observée, cette méthode permet de statuer sur l'utilisation d'une ressource commune aux deux tâches ou sur la présence d'un processus commun, et donc d'identifier les traitements réalisés pour produire un texte. » (Favart & Olive, 2005)

Tablette graphique

Périphérique raccordable à un ordinateur via une connexion filaire ou bluetooth et accompagné d'un stylet, permettant la saisie et l'analyse de données graphiques (production écrite, dessin, infographisme, etc.)

ANNEXES

Liste des annexes

Annexe I : Informations disponibles au sein des tableaux d'analyse

Annexe II : Productions et analyse des erreurs orthographiques d'un élève de CM2

Annexe III : Productions et analyse des erreurs orthographiques d'un élève de 5^e

Annexe IV : Productions et analyse des erreurs orthographiques d'un élève de 3^e

Annexe I : Informations disponibles au sein des tableaux d'analyse

1. Données concernant le mot N

- **Clé** : clé de classement des items de 1 à 5254
- **N° de ligne**
- **Level** : G5 = CM2 ; G7 = 5ème ; G9 = 3ème
- **Subject_index** : numéro du participant
- **Sex** : f = female ; m = male
- **Texttype** : e = expositif ; n = narratif
- **Order** : ECRIT1 = écrit puis oral ; ECRIT2 = oral puis écrit
- **Expression** : d = avec difficultés ; w = sans difficulté
- **Page** : numéro de page
- **Position** : rang de l'item dans le texte
- **SeqBegin** : temps du début de la séquence
- **SeqEnd** : temps de la fin de la séquence
- **LabelSeq** : texte écrit codé pour Clan
- **Erreur phonologique seule**
- **Erreur phono-lexicale**
- **Erreur lexicale seule**
- **Erreur morphosyntaxique seule**
- **Erreur phono-morphosyntaxique**
- **Total Erreurs**
- **Combinaisons d'erreurs**
- **Nfic** : nom du fichier concerné
- **Mots seuls** : forme orthographique correcte
- **Dispo** : RL = retour ligne ; RP = retour paragraphe
- **Révisions** : indique s'il s'agit de l'origine, du site ou du contenu de la révision
- **Ortho** : ERR = mot comportant une erreur ; ERR-E = mot erroné révisé mais pas corrigé ; ERR-C = mot erroné révisé et corrigé
- **nbmotsInSeq** : nombre de mots dans la séquence
- **isynt** : catégorie morphologique selon Manulex-infra (ABSMANLEX = item absent de la base de données Manulex)
- **icateg** : catégorie morphologique selon Cordial Analyseur
- **tcateg** : catégorie calculée pour faire correspondre Manulex-infra et Cordial Analyseur
- **IndexCordial** : rang du mot dans la phrase selon Cordial Analyseur
- **Cordial_corres** : forme orthographique (Cordial)
- **Cordial_lemme** : lemme (Cordial)

-
- **Cordial_codegram** : code grammatical (Cordial)
 - **Cordial_syntagme** : codage de la structure de la phrase (Cordial)
 - **Cordial_fonction** : fonction syntaxique (Cordial)
 - **Cordial_NumProp** : rang de la proposition (Cordial)
 - **Cordial_TypProp** : type de proposition (Cordial)
 - **Manulex_FRFO** : fréquence de la forme orthographique (Manulex)
 - **Freq_FO** : log de la fréquence de la forme orthographique
 - **Manulex_SYLL** : nombre de syllabes (Manulex)
 - **Manulex_LET** : nombre de lettres (Manulex)
 - **Manulex_PHON** : nombre de phonèmes (Manulex)
 - **Manulex_GRAPH** : nombre de graphèmes (Manulex)
 - **VPG_S** : nombre de voisins phonographiques
 - **VPG_F** : fréquence du voisinage phonographique
 - **Freq_VPG** : log de la fréquence du voisinage phonographique
 - **Manulex_FRBIG_T** : fréquence des bigrammes
 - **Manulex_FRBIP_T** : fréquence des biphones
 - **Manulex_SYL_I** : fréquence de la syllabe initiale
 - **Manulex_SYL_M** : fréquence de la syllabe médiane
 - **Manulex_SYL_F** : fréquence de la syllabe finale
 - **Manulex_FRSYL_moy** : fréquence moyenne de la syllabe
 - **Freq_SYLmoy** : log de la fréquence moyenne de la syllabe
 - **Manulex_CO_T** : consistance phonème-graphème totale
 - **Manulex_CO_I** : consistance phonème-graphème initiale
 - **Manulex_CO_M** : consistance phonème-graphème médiane
 - **Manulex_CO_F** : consistance phonème-graphème finale
 - **Manulex_freq_lemme** : fréquence du lemme
 - **Lexique_freq_lemme**
 - **Dur_T** : durée de la pause
 - **Code_T** : 0 = intra ; 1 = devant le mot ; 2 = devant retour ligne ; 3 = devant ponctuation ; 4 = devant ponctuation finale
 - **P_AVAbrut** : durée de la pause précédant le mot (en ms)
 - **P_AVA** : log de la durée de la pause précédant le mot
 - **nb_pause_intra** : nombre de pauses à l'intérieur du mot
 - **somme_pause_intra** : somme des pauses à l'intérieur du mot
 - **P_INTbrut** : rapport somme des pauses intra/ nombre de lettres
 - **P_INT** : log du rapport somme des pauses intra/ nombre de lettres
 - **index_pause_intra**
 - **durée_pause_intra_la+longue**
 - **localisation_pause_intra_la+longue** : 0 = au début du mot ; 1 = à la fin du mot
 - **débit_nbcarpsec** : débit en nombre de caractères par seconde
 - **débit_msecpcar** : débit en millisecondes par caractère (sans les pauses intra)
-

-
- **DEBIT** : log du débit en millisecondes par caractère
 - **débitbrut_msecpar** : débit en millisecondes par caractère (avec les pauses intra)

2. Données concernant le mot N+1

- **Position_n+1** : rang de l'item n+1 dans le texte
- **LabelSeq_n+1** : texte écrit codé pour Clan du mot suivant
- **MotsSeuls_n+1**
- Informations générales (linguistiques et temporelles)

3. Données concernant le mot N-1

- **Position_n-1** : rang de l'item n-1 dans le texte
- **LabelSeq_n-1** : texte écrit codé pour Clan du mot précédent
- **MotsSeuls_n-1**
- Informations générales (linguistiques et temporelles)

Annexe II : Productions et analyse des erreurs orthographiques d'un élève de CM2

Élève de sexe féminin, sans difficulté scolaire.

1. Brouillon du texte expositif

Il y a souvent des problèmes lorsqu'on ne connaît pas la personne et qu'on la juge sans savoir ce qu'elle est. Ces problèmes là sont souvent lorsqu'une nouvelle personne vient dans l'endroit et que vous ne connaissez pas cette personne vous vous permettez de juger mais ceci peut faire beaucoup de mal à la personne car elle ne connaît personne avec qui jouer, discuter, je pense qu'on a pas le droit de faire du mal à une personne lorsqu'on aime pas que cela nous arrive.

2. Brouillon du texte narratif

J'étais avec deux copines on s'amusaient ensemble et puis à un moment une de mes copines ma dit de rester avec elle et de laisser tomber l'autre et puis je l'ai laissée tomber toute seule et je suis allée avec mon autre ami et je me suis rendu compte que j'avais fait une erreur.

3. Analyse des erreurs

Ce tableau résume les informations principales ayant été nécessaires à notre analyse.

| Position | LabelSeq | phono seule | phono- lex | lex seule | phono- morpho | morpho seule | Combinaison | P_AVA brut | P_INT brut | débitbrut_ msecpcar |
|----------|--|----------------|---------------|--------------|------------------|-----------------|-----------------|---------------|---------------|------------------------|
| 6 | problème@err [: problèmes] | | | | | 1 | morpho seule | 474 | 97.88889 | 493.3333 |
| 10 | conner@err [: connaît] | | | | | 1 | morpho seule | 248 | 413.2857 | 1333.714 |
| 13 | personnes@err [: personne] | | | | | 1 | morpho seule | 492 | 131.375 | 557 |
| 27 | <ses@err [: ces]> [\$ rev:s:ref:w:2] | | | | | 1 | morpho seule | 259 | 91.66667 | 636.3333 |
| 29 | là@err> [: problèmes+là] | | | | | 1 | morpho seule | 804 | | nb lettres inconnu |
| 32 | [C] lorsqu^e@err_c [: lorsqu'] [\$ srev:3] [///] 0 [\$ orev:3] | | 1 | | | | phono-lex | 344 | 631.2857 | 1171.286 |
| 36 | [RL] personm@err_c [: personne] [\$ srev:4] [///] 0 [\$ orev:4] | | 1 | | | | phono-lex | 999 | 33 | 415 |
| 48 | [RL] connaissez@err [: connaissez] | | | 1 | | | lex seule | 912 | 120.9 | 490 |
| 55 | <permétez@err [: permettez]> [\$ rev:s:ref:w:7] | | | | | 1 | morpho seule | 916 | 368.1111 | 750.5556 |
| 73 | conner@err [: connaît] | | | | | 1 | morpho seule | 447 | 25.57143 | 787 |
| 74 | [RL] personnes@err [: personne] | | | | | 1 | morpho seule | 1090 | 99.875 | 532 |
| 85 | as@err [: a] | | | | | 1 | morpho seule | 571 | 1636 | 2784 |
| 99 | aimerez@err [: aimerait] | | | | | 1 | morpho seule | 336 | 172.5 | 558.625 |
| 104 | [RL] arrivent@err [: arrive] | | | | | 1 | morpho seule | 1045 | 301 | 1119 |
| 2 | était@err [: étails] | | | | | 1 | morpho seule | 227 | 177.6 | 704.8 |
| 8 | amusaient@err [: amusait] | | | | | 1 | morpho seule | 292 | 77.71429 | 814.2857 |
| 18 | [RL] copine@err [: copines] | | | | | 1 | morpho seule | 892 | 60.85714 | 403.4286 |
| 19 | ma@err [: m'a] | | | | | 1 | morpho seule | 1755 | | nb lettres inconnu |
| 36 | laisser@err [: laissée] | | | | | 1 | morpho seule | 1020 | 122.5714 | 544.7143 |
| 43 | aller@err [: allée] | | | | | 1 | morpho seule | 336 | 75.2 | 471.4 |
| 47 | ami@err [: amie] | | | | | 1 | morpho seule | 450 | 45.75 | 453 |

Annexe III : Productions et analyse des erreurs orthographiques d'un élève de 5^e

Élève de sexe féminin, avec difficultés scolaires.

1. Brouillon du texte expositif

il y a certaine chose a ne pas faire.
Comme recopier notre camarade de classe en plein contrôle. Car un contrôle est fait pour voir si nous avons compris le sujet. l'école sert a apprendre. Quand nous faisons un travail en équipe on peut mettre sous nos idées en commun. Comme en sport nous faisons un jeu avec 2 équipes dans un jeu d'équipes on doit faire jouer sous le monde. Si a la fin de la partie une équipe a perdu elle ne doit pas taper ses adversaire car c'est un mauvais joueur. Aussi si un jour on se promene dans la rue est qu'une personne fait tomber un objet nous devons lui rendre et non le garder. Si une nouvel gille arrive dans la classe est veut faire connaissance il faut nous aussi se faire connaître est non la laisser toute seul. Je pense que chaque personne sont pareil nous avons tous le même droit

2. Brouillon du texte narratif

C'était avec une copine on était pas dans la même classe donc on avait pas tous le temps les mêmes horaires. Comme certain jour nous finissons a la même heures elle ma demander de l'attendre a la sortie. Donc moi je l'attends jusqu'a se que les grilles du collège se ferma alors je partie. Le lendemain je lui demande pourquoi elle ne m'avait pas attendu est elle chercha des excuses. Comme c'était la première fois nous ne nous sommes pas fait la tête. Le jour d'après je l'attendais et malheureusement elle était partie, mais elle me disait que c'était sa mère qui lui avait dit de rendre vite. Un jour j'en ai eu marre donc je lui est dit que je l'attendrais plus. Donc maintenant nous nous attendons plus. Si a la sortie on se voit on rendre ensemble sinon chacune par de son côté

3. Analyse des erreurs

Ce tableau résume les informations principales ayant été nécessaires à notre analyse.

| position | LabelSeq | phono seule | phono-lex | lex seule | phono-morpho | morpho seule | Combinaisons | P_AVA brut | P_INT brut | débitbrut_msecpcar |
|----------|---------------------------------|-------------|-----------|-----------|--------------|--------------|-----------------------|------------|------------|--------------------|
| 4 | certaine@err [: certaines] | | | | | 1 | morpho seule | 308 | 115.5556 | 423.8889 |
| 5 | chose@err [: choses] | | | | | 1 | morpho seule | 304 | 79 | 376 |
| 6 | [C] a@err [: à] | | | | | 1 | morpho seule | 3048 | 0 | 2043 |
| 25 | et@err [: est] | | | | | 1 | morpho seule | 404 | 72 | 323 |
| 39 | école@err [: école] | | | 1 | | | lex seule | 136 | 116.4 | 428 |
| 41 | [C] a@err [: à] | | | | | 1 | morpho seule | 219 | 0 | 432 |
| 46 | fesons@err [: faisons] | | | 1 | | | lex seule | 244 | 78.71429 | 435.2857 |
| 52 | [RL] peu@err [: peut] | | | | | 1 | morpho seule | 876 | 50 | 339 |
| 53 | metre@err [: mettre] | | | 1 | | | lex seule | 515 | 101.6667 | 458.6667 |
| 54 | tous@err [: toutes] | | | | 1 | | phono-morpho | 292 | 73 | 307.3333 |
| 56 | idée@err [: idées] | | | | | 1 | morpho seule | 820 | 137.2 | 449.4 |
| 58 | commen@err [: commun] | | 1 | | | | phono-lex | 316 | 123.8333 | 659.3333 |
| 64 | fesons@err [: faisons] | | | 1 | | | lex seule | 256 | 83.28571 | 429.7143 |
| 69 | équipes@err [: équipes] | | | 1 | | | lex seule | 475 | 122.4286 | 521.7143 |
| 74 | équipes@err [: équipe] | | | | | 1 | morpho seule | 111 | 180 | 604.6667 |
| 81 | +monde@err> [: tout+le+monde] | | | 1 | | | lex seule | 280 | | nb lettres inconnu |
| 84 | a@err [: à] | | | | | 1 | morpho seule | 220 | 0 | 460 |
| 100 | adversaire@err [: adversaires] | | | | | 1 | morpho seule | 164 | 109.7273 | 409.0909 |
| 114 | promene@err [: promène] | | 1 | | | | phono-lex | 223 | 131.1429 | 480.4286 |
| 118 | [C] [RL] est@err [: et] | | | | | 1 | morpho seule | 795 | 165.5 | 726 |
| 135 | [C] sy@cok@err [: si] | | | 1 | | | lex seule | 632 | 194 | 914 |
| 137 | nouvel@err [: nouvelle] | | | | | 1 | morpho seule | 280 | 69.375 | 291.5 |
| 143 | [C] est@err [: et] | | | | | 1 | morpho seule | 252 | 183 | 690.5 |
| 146 | connesence@err [: connaissance] | | 2 | 1 | | | phono-lex + lex seule | 263 | 87.58333 | 344.3333 |
| 153 | connetre@err [: connaître] | | | 1 | | | lex seule | 228 | 115.6667 | 365.2222 |
| 154 | [C] est@err [: et] | | | | | 1 | morpho seule | 5356 | 156.5 | 658 |
| 159 | seul@err [: seule] | | | | | 1 | morpho seule | 234 | 59.8 | 286.4 |
| 166 | sont@err [: est] | | | | 1 | | phono-morpho | 295 | 154.3333 | 600 |
| 17 | tous@err [: tout] | | | | 1 | | phono-morpho | 268 | 105 | 426 |
| 25 | certain@err [: certains] | | | | | 1 | morpho seule | 276 | 116.375 | 387.375 |
| 26 | jour@err [: jours] | | | | | 1 | morpho seule | 340 | 84.6 | 325.6 |
| 28 | [RL] finisons@err [: finissons] | | 1 | | | | phono-lex | 1040 | 107.2222 | 419 |
| 29 | a@err [: à] | | | | | 1 | morpho seule | 236 | 0 | 400 |
| 32 | heures@err [: heure] | | | | | 1 | morpho seule | 643 | 144.4 | 553 |

| | | | | | | | | | | |
|-----|--|--|---|---|--|---|-----------------------------|------|----------|-----------------------|
| 34 | ma@err [: m' a] | | | | | 1 | morpho seule | 336 | | nb lettres inconnu |
| 35 | demander@err [: demandé] | | | | | 1 | morpho seule | 7870 | 140.7143 | 552.4286 |
| 39 | a@err [: à] | | | | | 1 | morpho seule | 275 | 0 | 397 |
| 47 | atten@err [: attends] | | | 1 | | 1 | lex seule + morpho seule | 103 | 91.71429 | 333 |
| 49 | +a@err> [: jusqu'+à] | | | 1 | | | lex seule | 112 | | nb lettres inconnu |
| 50 | se@err [: ce] | | | | | 1 | morpho seule | 396 | | nb lettres inconnu |
| 57 | ferma@err [: fermèrent] | | | | | 1 | phono- morpho | 198 | 68.33333 | 310.6667 |
| 60 | parie@err [: partis] | | 1 | | | 1 | phono-lex + morpho seule | 240 | 127 | 501.3333 |
| 75 | [RL] attendu@err [: attendue] | | | | | 1 | morpho seule | 1092 | 133.875 | 456.375 |
| 76 | [C] est@err [: et] | | | | | 1 | morpho seule | 400 | 179.5 | 698 |
| 84 | +étais@err [: c'+était] | | | | | 1 | morpho seule | 124 | | nb lettres inconnu |
| 86 | premiere@err [: première] | | 1 | | | | phono-lex | 171 | 118.625 | 470 |
| 91 | somme@err [: sommés] | | | | | 1 | morpho seule | 260 | 75.16667 | 384.8333 |
| 100 | aprais@err [: après] | | | 1 | | | lex seule | 248 | 168.8 | 732 |
| 107 | etait@err [: était] | | 1 | | | | phono-lex | 340 | 150.6 | 496 |
| 117 | +étais@err [: c'+était] | | | | | 1 | morpho seule | 99 | | nb lettres inconnu |
| 125 | rentrez@err [: rentrer] | | | | | 1 | morpho seule | 284 | 113 | 449.8571 |
| 133 | <et@err [: ai]> [\$ rev:s:ref:w:w:4] | | | | | 1 | morpho seule | 888 | 316 | 1114 |
| 139 | est@err [: ai] | | | | | 1 | morpho seule | 284 | 161.5 | 624 |
| 148 | maintenat@err [: maintenant] | | 1 | | | | phono-lex | 336 | 241.2 | 644.3 |
| 155 | a@err [: à] | | | | | 1 | morpho seule | 136 | 0 | 392 |
| 160 | voie@err [: voit] | | | | | 1 | morpho seule | 188 | 114 | 442 |
| 164 | <ensemble@err [: ensemble]> [\$ rev:s:ref:l:l:5] | | | 1 | | | lex seule | 408 | 70 | 236.5 |
| 167 | par@err [: part] | | | | | 1 | morpho seule | 396 | 47 | 321 |
| 170 | côte@err [: côté] | | 1 | | | | phono-lex | 140 | 114.75 | 462 |

Annexe IV : Productions et analyse des erreurs orthographiques d'un élève de 3^e

Élève de sexe masculin, avec difficultés scolaires.

1. Brouillon du texte expositif

Les problèmes ~~raciales~~ raciales
les problèmes ^{raciales} ~~use~~ sont les plus important, on en voit
~~tous~~ les jours. tout ceux qui sont raciste vivent
entre eux. Y en a qui vivent avec d'autre qui ne sont
~~pas~~ raciste et ils s'énervent pour rien. Être raciste
c'est se tromper de colère. les gens sont
raciste ^{juste} parce que l'autre n'a pas la même
couleur que lui, il parle différemment que lui ou ils ne
ont pas les mêmes origines.

les problèmes familiales
les problèmes ^{familiales} ~~use~~ se portent souvent entres amis,
Par exemple une personnes insultes les parents
de l'autre personne mais ces parents sont mort
alors l'autre personnes va s'énervé et sa va
se disputé, se bagarqué et a la fin ils vont tous
les deux se retrouver à la clinique. Ses problèmes
sont souvent à l'adolescence, mais il faut toujours
en parlé à une grande personne.

2. Brouillon du texte narratif

Histoire 1 :

U

Un jour je marchais dans la rue avec mon copain, quand on a trouvé un portable on leva la tête il était au monsieur de devant je voulais le rendre mais mon copain ne voulait pas il ma dit on le garde et on le revend et on partage l'argent. Je lui ai dit: non laisse tomber viens on va le rendre sa vaudra mieux après j'ai couru pour rattrapper le monsieur et je lui ai dit: tenez monsieur vous avez fait tomber votre portable.

Histoire 2 :

Après avoir fait un foot j'allais rentrer chez moi et j'ai trouvé un billet de 20 euros, je regarde autour de moi pour voir s'il était à quelqu'un j'ai vu personne alors je les garde et je suis rentré chez moi.

3. Analyse des erreurs

Ce tableau résume les informations principales ayant été nécessaires à notre analyse.

| Position | LabelSeq | phono seule | phono-lex | lex seule | phono-morpho | morpho seule | Combinaisons | P_AVA brut | P_INT brut | débitbrut_msecpcar |
|----------|--|-------------|-----------|-----------|--------------|--------------|--------------|------------|----------------|--------------------|
| 3 | <familiales@err [: familiaux]> [\$ srev:2] [///] 0 [\$ orev:1] | | | | 1 | | phono-morpho | 916 | | nb lettres inconnu |
| 5 | <raciales@err [: raciaux]> [\$ rev:s:ref:w:2] | | | | 1 | | phono-morpho | 9195 | 2523.857 14 | 3130.57143 |
| 7 | problèmes@err [: problèmes] [\$ srev:4] | | | 1 | | | lex seule | 386 | 104.7777 78 | 408 |
| 8 | se@err [: ce] | | | | | 1 | morpho seule | 1035 | 91.5 | 426 |
| 12 | important@err [: importants] [\$ srev:5] | | | | | 1 | morpho seule | 175 | 89.3 | 342.3 |
| 21 | <raciales@err [: raciaux] | | | | 1 | | phono-morpho | 27651 | 72.42857 14 | 369.428571 |
| 24 | [C] tout@cab@err [: tous] | | | | 1 | | morpho seule | 6680 | 118.75 | 472 |
| 28 | raciste@err [: racistes] | | | | | 1 | morpho seule | 372 | | nb lettres inconnu |
| 34 | [C] y@cok@err [: il+y] | | | | | 1 | morpho seule | 6555 | | nb lettres inconnu |
| 41 | autre@err [: autres] | | | | | 1 | morpho seule | 124 | | nb lettres inconnu |
| 47 | raciste@err [: racistes] | | | | | 1 | morpho seule | 340 | | nb lettres inconnu |
| 88 | différemment@err [: différemment] | | 1 | | | | phono-lex | 276 | 119.0833 33 | 408 |
| 93 | ne@err [: n'] | | | | | 1 | morpho seule | 171 | 0 | 308 |
| 102 | familiales@err [: familiaux] | | | | 1 | | phono-morpho | 140 | | nb lettres inconnu |
| 107 | <familiales@err [: familiaux] | | | | 1 | | phono-morpho | 924 | | nb lettres inconnu |
| 111 | entres@err [: entre] | | | 1 | | | lex seule | 559 | 197 | 572.6 |
| 118 | <personnes@err_c [: personne]> [\$ srev:12] [///] 0 [\$ orev:12] | | | | | 1 | morpho seule | 208 | 103.875 | 452.125 |
| 120 | insultes@err [: insulte] | | | | | 1 | morpho seule | 368 | 150.5714 29 | 424.571429 |
| 128 | ces@err [: ses] | | | | | 1 | morpho seule | 140 | 108 | 417.333333 |
| 131 | mort@err [: morts] | | | | | 1 | morpho seule | 680 | 84.6 | 312.8 |
| 135 | personnes@err [: personne] | | | | | 1 | morpho seule | 658 | 154 | 483.875 |
| 138 | énervé@err [: énerver] | | | | | 1 | morpho seule | 138 | 129.1428 57 | 386 |
| 140 | sa@err [: ça] | | | | | 1 | morpho seule | 496 | 34 | 326.5 |
| 143 | disputé@err [: disputer] | | | | | 1 | morpho seule | 163 | 110.125 | 353.5 |
| 146 | bagarré@err [: bagarrer] | | | | | 1 | morpho seule | 274 | 80.25 | 347.625 |
| 162 | [C] ses@cok@err [: ces] | | | | | 1 | morpho seule | 24756 | 123.3333 33 | 1126.33333 |
| 173 | toujour@err [: toujours] | | | 1 | | | lex seule | 391 | 103.375 | 303.625 |
| 175 | parlé@err [: parler] | | | | | 1 | morpho seule | 144 | 43.66666 | 275.5 |

| | | | | | | | | | | |
|-----|--|--|--|--|---|---|------------------|------|----------------|-----------------------|
| | | | | | | | | | 67 | |
| 15 | [C] <quanb@err_c [: quand]> [\$ srev:1 [///] 0 [\$ orev:1] | | | | 1 | | lex seule | 4537 | 82 | 367.6 |
| 19 | trouve@err [: trouv ] | | | | | | phono- morpho | 199 | 145.6666 67 | 535.333333 |
| 43 | ma@err [: m' a] | | | | | 1 | morpho seule | 180 | | nb lettres inconnu |
| 51 | revent@err [: revend] | | | | | 1 | morpho seule | 449 | | nb lettres inconnu |
| 61 | dis@err [: dit] [\$ srev:2] | | | | | 1 | morpho seule | 323 | 82.33333 33 | 349.666667 |
| 72 | [C] sa@err [:  a] | | | | | 1 | morpho seule | 2320 | 39.5 | 329.5 |
| 80 | rattrapper@err [: rattraper] | | | | 1 | | lex seule | 480 | 114.8888 89 | 400.666667 |
| 87 | dis@err [: dit] [\$ srev:3] | | | | | 1 | morpho seule | 111 | 60.33333 33 | 329.333333 |
| 105 | fai@err [: fait] | | | | | 1 | morpho seule | 175 | 51.75 | 275 |
| 110 | rentr @err [: rentrer] | | | | | 1 | morpho seule | 368 | 81.28571 43 | 307.428571 |
| 116 | trouver@err [: trouv ] | | | | | 1 | morpho seule | 592 | 160.3333 33 | 446.666667 |
| 129 | <vois@err_c [: voir]> [\$ srev:5 [///] 0 [\$ orev:5] | | | | 1 | | phono- morpho | 409 | 84.75 | 322 |
| 131 | [C] si@err [: s'] | | | | | 1 | morpho seule | 570 | 81.5 | 348.5 |
| 144 | les@err [: l' ai] | | | | | 1 | morpho seule | 172 | | nb lettres inconnu |
| 145 | gard s@err [: gard ] | | | | | 1 | morpho seule | 387 | 103.2 | 500.4 |
| 149 | rentrer@err [: rentr e] | | | | | 1 | morpho seule | 884 | 98.42857 14 | 347.428571 |

TABLE DES ILLUSTRATIONS

Liste des graphiques et figures

| | |
|--|----|
| Figure I.1 : Les composants cognitifs de la production écrite selon Flower et Hayes (1980), schéma extrait de Favart et Olive (2005)..... | 17 |
| Figure III.1: Dispositif Eye and Pen (www.eyelandpen.net) | 27 |
| Graphique IV.I : Rapport des mots erronés au nombre total de mots produits par chaque participant selon le niveau scolaire (classement croissant)..... | 33 |
| Graphique IV.2 : Débit d'écriture (en ms par caractère) pour les mots avec et sans erreurs selon le niveau scolaire..... | 35 |
| Graphique IV.3 : Pauses (en ms) avant les mots avec et sans erreurs selon le niveau scolaire..... | 36 |
| Graphique IV.4 : Pauses (en ms) internes aux mots avec et sans erreurs selon le niveau scolaire..... | 37 |
| Graphique IV.5 : Répartition des types d'erreurs selon le niveau scolaire..... | 40 |
| Graphique IV.6 : Liens entre les pauses (en ms) avant les mots erronés et le type d'erreurs selon le niveau scolaire..... | 43 |
| Graphique IV.7 : Liens entre les pauses (en ms) internes aux mots erronés et le type d'erreurs selon le niveau scolaire..... | 44 |
| Graphique IV.8 : Révisions des mots erronés selon le niveau scolaire et le type d'erreurs produites..... | 45 |

Liste des tableaux

| | |
|--|----|
| Tableau III.1 : Exemples d'entrées du tableau rassemblant les données du corpus..... | 28 |
| Tableau IV.1 : Nombre de mots (total et erronés) et pourcentage de mots erronés selon le niveau scolaire..... | 33 |
| Tableau IV.2 : Débit d'écriture (en ms par caractère) pour les mots avec et sans erreurs selon le niveau scolaire..... | 34 |
| Tableau IV. 3: Pauses (en ms) avant les mots avec et sans erreurs selon le niveau scolaire..... | 36 |
| Tableau IV.4 : Pauses (en ms) internes aux mots avec et sans erreurs selon le niveau scolaire..... | 37 |
| Tableau IV.5 : Nombre et répartition des types d'erreurs selon le niveau scolaire..... | 39 |
| Tableau IV.6 : Débit d'écriture (en ms par caractère) selon le type d'erreurs et le niveau scolaire..... | 40 |
| Tableau IV.7 : Pauses (en ms) avant les mots selon le type d'erreurs et le niveau scolaire..... | 41 |
| Tableau IV.8 : Pauses (en ms) internes aux mots selon le type d'erreurs et le niveau scolaire..... | 43 |
| Tableau IV.9 : Révisions des mots erronés selon le niveau scolaire et le type d'erreurs produites..... | 45 |

TABLE DES MATIERES

| | |
|--|-----------|
| ORGANIGRAMMES | 2 |
| 1. <i>Université Claude Bernard Lyon1</i> | 2 |
| 1.1 Secteur Santé : | 2 |
| 1.2 Secteur Sciences et Technologies : | 2 |
| 2. <i>Institut Sciences et Techniques de Réadaptation FORMATION ORTHOPHONIE</i> | 3 |
| REMERCIEMENTS..... | 4 |
| SOMMAIRE..... | 5 |
| INTRODUCTION..... | 7 |
| PARTIE THEORIQUE | 8 |
| I. LES ERREURS ORTHOGRAPHIQUES | 9 |
| 1. <i>Acquisition de l'orthographe</i> | 9 |
| 1.1. Apprentissage du principe alphabétique..... | 9 |
| 1.2. Apprentissages orthographiques | 10 |
| 1.2.1. Acquisition des connaissances orthographiques générales : les régularités graphotactiques | 10 |
| 1.2.2. Apprentissages des connaissances lexicales spécifiques : constitution du lexique orthographique | 11 |
| 11 | |
| 1.3. Apprentissages morphologiques | 11 |
| 1.3.1. Acquisition de la morphologie dérivationnelle..... | 11 |
| 1.3.2. Acquisition de la morphologie flexionnelle..... | 12 |
| 2. <i>Statut de l'erreur</i> | 12 |
| 2.1. Statut social et pédagogique de l'erreur | 12 |
| 2.2. Intérêt et analyse de l'erreur orthographique dans la recherche scientifique | 13 |
| 3. <i>Étude des erreurs orthographiques</i> | 14 |
| 3.1. Origines des erreurs orthographiques | 14 |
| 3.2. Distribution des erreurs orthographiques au cours de l'apprentissage | 15 |
| II. LA PRODUCTION ECRITE DE TEXTE..... | 16 |
| 1. <i>Description des processus en jeu dans la production écrite</i> | 16 |
| 1.1. Modèle princeps d'Hayes et Flower | 16 |
| 1.2. Processus de formulation | 17 |
| 2. <i>Interactions entre les processus rédactionnels</i> | 18 |
| 2.1. Temporalité de l'écriture : partage attentionnel et théorie de la capacité | 18 |
| 2.2. Implications développementales..... | 18 |
| 2.2.1. Modèles développementaux | 18 |
| 2.2.2. Facteurs développementaux de la production écrite..... | 19 |
| 3. <i>Analyse en temps réel de la production écrite</i> | 20 |
| 3.1. Méthodes d'analyse de la production écrite | 20 |
| 3.2. Paramètres en temps réel et processus rédactionnels..... | 21 |
| 3.2.1. Débit d'écriture et temps de pause | 21 |
| 3.2.2. Révisions | 21 |
| PROBLEMATIQUE ET HYPOTHESES..... | 23 |
| I. PROBLEMATIQUE | 24 |
| II. HYPOTHESES..... | 25 |
| 1. <i>Hypothèse générale</i> | 25 |
| 2. <i>Hypothèses opérationnelles</i> | 25 |
| 2.1. Première hypothèse | 25 |
| 2.2. Deuxième hypothèse | 25 |
| PARTIE EXPERIMENTALE | 26 |
| I. METHODE DE RECUEIL DES DONNEES..... | 27 |
| 1. <i>Participants</i> | 27 |
| 2. <i>Matériel</i> | 27 |
| 3. <i>Procédure expérimentale</i> | 28 |
| II. CONSTITUTION DU CORPUS..... | 29 |
| 1. <i>Corpus général</i> | 29 |
| 2. <i>Corpus des erreurs orthographiques : constitution des critères d'analyse</i> | 30 |
| 2.1. « Erreurs phono-lexicales » | 30 |

| | | |
|---|--|-----------|
| 2.2. | « Erreurs lexicales seules » | 30 |
| 2.3. | « Erreurs morphosyntaxiques seules » | 31 |
| 2.4. | « Erreurs phono-morphosyntaxiques» | 31 |
| 2.5. | « Erreurs phonologiques seules » | 31 |
| 2.6. | Items problématiques | 32 |
| PRESENTATION DES RESULTATS | | 33 |
| I. | DONNEES DESCRIPTIVES DU CORPUS | 34 |
| II. | COMPARAISON DES MOTS AVEC OU SANS ERREUR(S) ORTHOGRAPHIQUE(S) | 35 |
| 1. | <i>Débit d'écriture</i> | 35 |
| 2. | <i>Pauses</i> | 36 |
| 2.1. | Pauses précédant les mots | 36 |
| 2.2. | Pauses internes aux mots | 38 |
| 3. | <i>Paramètres en temps réel et nature des mots</i> | 39 |
| III. | COMPARAISON DES DIFFERENTS TYPES D'ERREURS ORTHOGRAPHIQUES | 39 |
| 1. | <i>Données descriptives</i> | 39 |
| 2. | <i>Débit d'écriture</i> | 41 |
| 3. | <i>Pauses</i> | 42 |
| 3.1. | Pauses précédant les mots erronés | 42 |
| 3.2. | Pauses internes aux mots erronés | 44 |
| 4. | <i>Révisions</i> | 45 |
| DISCUSSION DES RESULTATS | | 47 |
| I. | RAPPELS DES DONNEES DE L'ETUDE | 48 |
| 1. | <i>Objectif</i> | 48 |
| 2. | <i>Hypothèses</i> | 48 |
| 3. | <i>Méthode</i> | 48 |
| II. | INTERPRETATION DES RESULTATS | 48 |
| 1. | <i>Données générales : fluence, pourcentages d'erreurs et pourcentages de types d'erreurs par niveau de classe</i> | 48 |
| 1.1. | Comparaison de la fluence écrite par niveau scolaire | 48 |
| 1.2. | Comparaison des pourcentages d'erreurs par niveau scolaire | 49 |
| 2. | <i>Comparaison des pourcentages de types d'erreurs par niveau scolaire</i> | 49 |
| 2.1. | Erreurs phonologiques seules | 49 |
| 2.2. | Erreurs phono-morphosyntaxiques | 50 |
| 2.3. | Erreurs phono-lexicales, lexicales seules | 50 |
| 2.4. | Erreurs morphosyntaxiques seules | 51 |
| 3. | <i>Paramètres en temps réel et erreur orthographique</i> | 51 |
| 3.1. | Retour sur les hypothèses | 52 |
| 3.2. | Analyse des résultats et confrontation à la littérature | 52 |
| 3.2.1. | Comparaison des mots avec ou sans erreur(s) orthographique(s) | 52 |
| a. | Pauses internes aux mots | 52 |
| b. | Pauses précédant les mots | 53 |
| c. | Débit d'écriture | 54 |
| 3.2.2. | Comparaison des types d'erreurs orthographiques | 54 |
| a. | Pauses précédant les mots erronés | 54 |
| b. | Pauses internes aux mots erronés | 55 |
| c. | Débit d'écriture | 56 |
| d. | Révisions | 56 |
| III. | BIAIS ET LIMITES DE L'ETUDE | 57 |
| 1. | <i>Population</i> | 57 |
| 2. | <i>Création d'une typologie</i> | 57 |
| 3. | <i>Analyse</i> | 58 |
| 3.1. | Données inhérentes au corpus | 58 |
| 3.2. | Analyse des erreurs | 58 |
| IV. | PERSPECTIVES, APPORTS PROFESSIONNELS ET PERSONNELS | 58 |
| 1. | <i>Perspectives de recherche</i> | 58 |
| 1.1. | Nature des mots | 58 |
| 1.2. | Modélisation des pauses en quartiles | 59 |
| 1.3. | Apport de l'analyse de l'input visuel | 59 |
| 2. | <i>Apports pour la profession d'orthophoniste</i> | 59 |
| 3. | <i>Apports personnels</i> | 60 |
| CONCLUSION | | 61 |

| | |
|--|-----------|
| BIBLIOGRAPHIE | 63 |
| GLOSSAIRE | 68 |
| ANNEXES | 69 |
| LISTE DES ANNEXES | 70 |
| ANNEXE I : INFORMATIONS DISPONIBLES AU SEIN DES TABLEAUX D'ANALYSE..... | 71 |
| 1. <i>Données concernant le mot N</i> | 71 |
| 2. <i>Données concernant le mot N+1</i> | 73 |
| 3. <i>Données concernant le mot N-1</i> | 73 |
| ANNEXE II : PRODUCTIONS ET ANALYSE DES ERREURS ORTHOGRAPHIQUES D'UN ELEVE DE CM2 | 74 |
| 1. <i>Brouillon du texte expositif</i> | 74 |
| 2. <i>Brouillon du texte narratif</i> | 74 |
| 3. <i>Analyse des erreurs</i> | 75 |
| ANNEXE III : PRODUCTIONS ET ANALYSE DES ERREURS ORTHOGRAPHIQUES D'UN ELEVE DE 5 ^E | 76 |
| 1. <i>Brouillon du texte expositif</i> | 76 |
| 2. <i>Brouillon du texte narratif</i> | 76 |
| 3. <i>Analyse des erreurs</i> | 77 |
| ANNEXE IV : PRODUCTIONS ET ANALYSE DES ERREURS ORTHOGRAPHIQUES D'UN ELEVE DE 3 ^E | 79 |
| 1. <i>Brouillon du texte expositif</i> | 79 |
| 2. <i>Brouillon du texte narratif</i> | 80 |
| 3. <i>Analyse des erreurs</i> | 81 |
| TABLE DES ILLUSTRATIONS..... | 83 |
| TABLE DES MATIERES..... | 85 |

Anaïs Deleuze

Élodie Murian

ERREURS ORTHOGRAPHIQUES ET PARAMÈTRES EN TEMPS RÉEL: Analyse de textes d'élèves de CM2, 5^e et 3^e.

87 Pages

Mémoire d'orthophonie -UCBL-ISTR- Lyon 2010

RESUME

La question développée est celle des liens pouvant exister entre les performances orthographiques et les paramètres en temps réel de la production écrite (débit d'écriture et pauses). Ces paramètres constituent en effet le reflet de la dynamique rédactionnelle, initialement décrite par Hayes et Flower (1981), où interagissent des processus de bas niveau (orthographe et graphisme) et de haut niveau (planification, génération de texte et révision), aux coûts attentionnels distincts. L'apprenti-scripteur apprend à gérer conjointement ces processus. L'automatisation progressive de son orthographe et de son graphisme autorise une allocation de son attention pour les processus plus complexes (McCutchen, 1994, 1996). Les brouillons de textes de 139 enfants de CM2, 5^e et 3^e, collectés en 2006 par le laboratoire Dynamique Du Langage (Université Lyon 2), ont été analysés. L'utilisation couplée de tablettes graphiques et du dispositif Eye and Pen a permis de recueillir la production écrite des participants assortie de paramètres en temps réel (débit d'écriture, pauses précédant les mots et pauses internes). Les données temporelles des mots contenant des erreurs orthographiques ont été comparées à celles des mots sans erreur. Les erreurs orthographiques ont de plus été classées par nos soins et recoupées statistiquement avec les données temporelles leur étant rattachées. Ce procédé non expérimenté jusqu'alors a pour but d'initier une compréhension des coûts cognitifs associés aux erreurs orthographiques au cours de la production écrite (et ceci selon le niveau scolaire du scripteur). Les résultats démontrent que les paramètres en temps réel ne rendent pas compte de manière univoque du traitement de l'erreur orthographique. Les pauses internes semblent davantage refléter le traitement orthographique que les pauses précédant les mots et le débit d'écriture. Une méthodologie adaptée permettrait de préciser l'analyse de ces paramètres. Coupler l'analyse graphique à l'analyse oculomotrice permettrait notamment de faire une part plus fine des processus en jeu.

MOTS-CLES

Erreur orthographique ; Production écrite ; Paramètres en temps réel ; Débit d'écriture ; Pause ; Pause précédant le mot ; Pause interne au mot.

MEMBRES DU JURY

Baldy Mouliner Florence, Chosson Christelle, Guillon-Invernizzi Fanny

MAITRE DE MEMOIRE

Michel Fayol

DATE DE SOUTENANCE

30 Juin 2011
