



MINISTÈRE
DE L'ÉDUCATION
NATIONALE
ET DE LA JEUNESSE

*Liberté
Égalité
Fraternité*

Conseil scientifique
de l'éducation nationale

RATIONALISER L'ORTHOGRAPHE DU FRANÇAIS POUR MIEUX L'ENSEIGNER

Synthèse de la recherche et recommandations

Texte rédigé par
Liliane Sprenger-Charolles,
Anne Abeillé,
Bernard Cerquiglini



Ce texte a été rédigé par **Liliane Sprenger-Charolles**, **Anne Abeillé** et **Bernard Cerquiglini**¹.
Remerciements à **Stanislas Dehaene**, **Franck Ramus**, au **secrétariat général** du CSEN, et tout particulièrement à **Anne Lenglet**, pour leur relecture attentive de la présente note de synthèse.

¹ Rédigé par **Liliane Sprenger-Charolles**, directrice de recherche émérite à l'université d'Aix-Marseille, au Centre de Recherche en Psychologie et Neurosciences, **Anne Abeillé**, professeure à l'université Paris Cité, Laboratoire de linguistique formelle et **Bernard Cerquiglini**, professeur émérite à l'université Paris Cité, Laboratoire de linguistique formelle, vice-président de la Fondation des Alliances françaises et membre de l'Académie royale de Belgique. Avec la contribution de **Stanislas Dehaene**, professeur de sciences cognitives expérimentales au Collège de France, **Franck Ramus**, directeur de recherche au CNRS et membre du Département d'Études Cognitives de l'École normale supérieure, **Anne Lenglet**, chargée d'études au secrétariat général du CSEN.

Résumé

Un double paradoxe caractérise les débats autour de la question de l'orthographe en France. D'une part, si l'orthographe française a beaucoup évolué au fil des siècles, les débats à propos des réformes successives n'ont, quant à eux, que peu varié. D'autre part, si les performances des élèves français dans ce domaine sont loin d'être satisfaisantes, ce constat n'est toutefois pas récent : il est concomitant des lois qui ont rendu l'instruction obligatoire pour tous. Afin de permettre de comprendre ce double paradoxe, cette note de synthèse présente les grandes lignes de l'évolution de l'orthographe du français, les caractéristiques actuelles de cette orthographe et les résultats des évaluations du niveau en orthographe de nos élèves.

En français, comme dans toutes les écritures alphabétiques, c'est le degré de transparence des relations entre phonèmes (unités de base du code oral) et graphèmes (unités de base du code écrit comme <a>, <ou>, , <ch>) qui facilite l'apprentissage de la lecture-écriture. Ainsi l'orthographe du français est moins transparente que celle de l'espagnol mais plus transparente que celle de l'anglais. Il est donc plus facile d'apprendre à lire-écrire en espagnol qu'en français et en français qu'en anglais. Toutefois l'orthographe du français se caractérise par des spécificités lexicales et grammaticales qui rendent l'apprentissage de l'écriture bien plus difficile que celui de la lecture. Concernant l'orthographe grammaticale, de nombreuses marques morphologiques s'écrivent mais ne s'entendent pas à l'oral (par exemple : *elle est jolie, les mouches volent*). Quant à l'orthographe lexicale, l'Académie française a toujours été partagée entre une tendance conservatrice (cf. l'ajout des marques étymologiques <s> et <p> dans *écriture*) et une réformatrice (la moitié des mots ont changé d'orthographe depuis la première édition de son dictionnaire). Ces explications sont confortées par les statistiques indiquant le degré de régularité (la consistance) des mots français en tenant compte des correspondances graphème-phonème et phonème-graphème ainsi que de la morphologie. Elles sont également confortées par les études indiquant que, au début de l'apprentissage, les mots irréguliers (*pied, femme...*) sont moins bien lus, et surtout moins bien écrits, que les réguliers (*porte, samedi...*). Toutefois, les erreurs les plus nombreuses concernent l'orthographe grammaticale, le cas le plus notoire étant celui de l'accord du participe passé avec l'auxiliaire avoir et le complément antéposé (maîtrisé par moins de 20% des élèves en fin de primaire). La conclusion présente les actions nécessaires pour améliorer l'orthographe et son enseignement.

Sommaire

Introduction	6
A. Les différents systèmes d'écriture et leur coût	7
1. Les écritures alphabétiques.....	7
a. Incidence de la transparence de l'orthographe dans l'apprentissage typique de la lecture	7
b. Incidence de la transparence de l'orthographe dans la dyslexie	8
2. Quelques spécificités de l'orthographe du français	8
B. L'Histoire de l'orthographe française: une longue réforme	9
1. Avant la création de l'Académie française	9
2. Un choix malheureux au XVII ^e siècle	10
3. Une réformatrice opiniâtre : l'Académie	12
4. Évolution de l'orthographe depuis la mise en place de l'instruction obligatoire	13
a. Les arrêtés de 1900, 1901 et 1976 sur l'orthographe	13
b. Les Rectifications de 1990	15
C. Données quantitatives sur l'orthographe actuelle	16
1. Données sur la fréquence des mots et des unités infralexicales	16
a. Données sur la fréquence des mots	16
b. Données sur la fréquence et la consistance des unités infralexicales	17
c. Utilisation de ces bases de données.....	18
2. Données provenant d'études descriptives et expérimentales	18
2.1 L'apprentissage de l'orthographe lexicale	19
a. EOLE : Échelle d'acquisition de l'orthographe lexicale.....	19
b. EQOL : Échelle Québécoise d'Orthographe lexicale.....	19
c. Résultats d'une série d'observations longitudinales.....	20
d. De la régularité à la consistance.....	21
2.2. Le poids de la morphologie lexicale et grammaticale dans l'orthographe du français.....	21
a. Un problème spécifique au français : celui des marques de flexion et dérivation	21
b. Données sur l'accord du participe passé	22
3. Que faudrait-il faire pour mieux enseigner l'orthographe ?	25
Ce qu'il faut retenir	26
Annexes	27
Bibliographie	36

Introduction

Cette synthèse est centrée sur deux questions : Pourquoi et comment rationaliser l'orthographe du français ? Comment mieux l'enseigner ? Afin de pouvoir répondre à ces questions, nous nous sommes appuyés sur trois types de travaux :

- des études historiques sur l'évolution de l'orthographe du français depuis le XII^e siècle,
- des bases de données qui présentent les caractéristiques actuelles de cette orthographe,
- des études qui ont évalué les capacités orthographiques de nos élèves.

Ces travaux permettent de réfuter **différents mythes**, en particulier celui qui repose sur l'affirmation de l'existence d'un passé idyllique dans lequel tout le monde écrivait correctement le français. En fait, les inquiétudes sur le niveau d'orthographe des élèves français sont concomitantes de l'instauration de l'obligation de l'instruction. Dès 1900, un arrêté signé par Georges Leygues, ministre de l'Instruction publique, stipule qu'il ne faut plus compter comme fautes, pour les examens et concours de son ministère, un certain nombre de cas difficiles concernant l'orthographe, dont l'accord du participe passé (Chervel, 1977 et 1991). Un autre mythe est lié à l'affirmation que, si l'on modifiait l'orthographe, il ne serait plus possible de lire les textes anciens. Comme l'a écrit Brunot dans sa lettre ouverte au Ministre de l'Instruction publique en 1905 (Cerquiglini, 2004) : « Pareille illusion se comprend chez ceux qui n'ont jamais ouvert que des éditions scolaires, ou qui oublient que le texte de la Collection des Grands Écrivains est publié dans une orthographe uniformisée, rajeunie, truquée, où on a juste laissé <oi> en souvenir du passé. »

Dans la suite, sont d'abord présentés les différents systèmes d'écriture, plus particulièrement les écritures alphabétiques et, dans ces écritures, celle du français. Cette orthographe se caractérise par des spécificités dues à des raisons historiques et à des choix politiques qui sont développées dans les sections suivantes et accompagnées par les résultats d'études ayant évalué le niveau de maîtrise de nos élèves dans ce domaine. La conclusion insiste sur les réformes nécessaires pour améliorer l'efficacité de l'enseignement de l'orthographe.

A. Les différents systèmes d'écriture et leur coût

Pour comprendre les fondements et les enjeux du débat sur l'orthographe, il faut connaître les types de codage utilisés pour transcrire la langue orale. Certaines écritures transcrivent surtout des unités orales ayant un sens : ce sont les écritures logographiques comme le mandarin (Xu, Tan & Perfetti, 2019). D'autres transcrivent surtout les unités sonores de l'oral : des syllabes pour les écritures syllabiques (les caractères Hiragana et Katakana du japonais) ; des phonèmes pour les écritures alphabétiques. L'écriture de l'arabe qui, comme celle de l'hébreu, se situe entre les deux grands systèmes précédents, est dite morpho-syllabique. Apprendre à lire est plus long dans une écriture logographique que dans une écriture alphabétique en raison du plus grand nombre d'unités à manipuler et de leur plus grande complexité (Verhoeven & Perfetti, 2017).

1. Les écritures alphabétiques

Dans une écriture alphabétique, l'apprentissage de l'écriture, comme celui de la lecture, repose sur **la compréhension du principe alphabétique** selon lequel les voyelles et les consonnes sont transcrites par des graphèmes (lettres ou groupes de lettres correspondant à un phonème : <u>, <ou> dans *pur*, pour ; <f>, <ch> dans *fou*, *chou*). Pour lire dans un tel système, il faut donc **mettre en relation des graphèmes avec des phonèmes** : c'est le principe du B.A-BA ou décodage utilisé pour la lecture, et vice-versa pour l'écriture. Une écriture alphabétique est plus économique qu'une écriture logographique, la maîtrise du principe alphabétique permettant de comprendre et de produire une infinité de mots, de phrases et de textes à l'aide d'un nombre limité de correspondances graphème-phonème pour la lecture et phonème-graphème pour l'écriture.

a. Incidence de la transparence de l'orthographe dans l'apprentissage typique de la lecture

Grâce à l'apprentissage d'un petit nombre de règles de correspondances graphème-phonème, l'élève peut retrouver les milliers de mots oraux qu'il a stockés dans sa mémoire bien avant d'avoir appris à lire (Ziegler & Sprenger-Charolles, 2023). Chaque décodage réussi active la forme sonore et le sens des mots connus à l'oral, ce qui permet l'instauration d'une représentation orthographique, qui sera renforcée par les répétitions (Share, 1995 ; Ziegler, Perry & Zorzi, 2014). Si le décodage est bien le *sine-qua-non* de l'apprentissage typique de la lecture (Share, 1995 ; Ziegler et al., 2014), la réussite de cet apprentissage doit dépendre de la transparence des correspondances graphème-phonème dans la langue dans laquelle il s'effectue. De fait, les travaux de recherche ont montré qu'apprendre à lire dans une langue ayant une orthographe opaque (l'anglais) est plus difficile que dans une langue ayant une orthographe très transparente (l'espagnol) ou moyennement transparente (le français). C'est ce qu'indiquent les résultats présentés dans la figure 1 (Dehaene, 2007 d'après Seymour, Aro & Erskine, 2003 ; Ziegler, 2018).

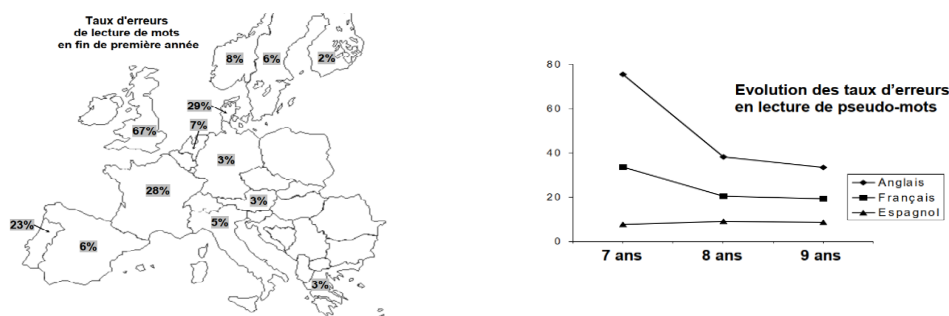


Figure 1. Erreurs (%) en lecture de mots isolés après un an d'apprentissage dans 15 pays d'Europe. On constate que dans les langues qui ont une orthographe très régulière (espagnol, finnois) les erreurs sont moins nombreuses que dans celles dont la transparence de l'orthographe est moyenne (français) ou faible (anglais). Cette observation est confirmée par l'évolution de 7 à 9 ans des erreurs en lecture de mots inventés en espagnol, en français et en anglais (Dehaene, 2007 d'après Seymour et al., 2003).

b. Incidence de la transparence de l'orthographe dans la dyslexie

La dyslexie du développement se caractérise par un déficit d'automatisation de l'identification des mots écrits qui se manifeste par des difficultés en lecture de mots isolés. Ce déficit, qui résulte généralement d'un trouble de la composante phonologique du langage (Ramus, in press), est souvent inattendu par rapport aux autres capacités cognitives du dyslexique et à l'enseignement qu'il a reçu. Les conséquences secondaires peuvent inclure des problèmes de compréhension en lecture (Lyon, Shaywitz & Shaywitz, 2003). Comme dans le développement typique (Verhoeven & Perfetti, 2017), l'opacité de l'orthographe rend plus difficile l'apprentissage de la lecture chez les futurs dyslexiques (Verhoeven, Perfetti & Pugh, 2019). Il a en effet été relevé que leur déficit d'identification des mots écrits s'observe dès la précision de la réponse dans des langues qui ont une orthographe opaque (en anglais) mais plutôt dans le temps de réponse dans des orthographe très transparentes (en espagnol ou en italien) ou moyennement transparente (en français, cf. la synthèse de Sprenger-Charolles, Siegel, Jimenez & Ziegler, 2011) y compris chez des dyslexiques adultes (cf. Paulesu et al., 2001, pour une comparaison entre des adultes anglais, français et italiens).

2. Quelques spécificités de l'orthographe du français

Dans un monde linguistique idéal, la graphie devrait être la transcription fidèle de l'oral, « la peinture de la voix », disait Voltaire. En français, pour diverses raisons, ce n'est pas vraiment le cas. Une de ces raisons vient du fait qu'il n'est pas possible de transcrire les 10 à 16 voyelles du français avec les 6 voyelles de l'alphabet latin. Il a donc été nécessaire d'utiliser soit des signes diacritiques (accent sur <e>), soit des suites de lettres (<ou>, <on>) pour différencier <e> de <é> et <è> (cf. élève), <u> de <ou> (russe versus rousse) ainsi que les voyelles orales (<a>, <o>, <i> et <u>) des voyelles nasales (<an>, <on>, <in> et <un>). A cet inventaire, s'ajoutent des allographes, c'est-à-dire des graphèmes différents pour un même phonème (<en> et <ant> dans *enfant*). Les différences dans le décompte des voyelles proviennent de variations régionales, certains francophones ne différenciant pas, par exemple, le <a> *patte* de celui de *pâte* alors que d'autres ne perçoivent pas bien la différence entre *brin* et *brun*.

De plus, notre orthographe fut constituée au Moyen Âge, et nous écrivons encore d'innombrables consonnes finales que l'on prononçait à cette période mais qui ont cessé de s'entendre à partir de la fin du XV^e siècle. La graphie médiévale des XII^e et XIII^e siècles était simple et assez fidèle à la prononciation. Les scribes ont essayé de transcrire certains phonèmes spécifiques en inventant des digraphes (<sch> dans *schavoir, schavant...*). Certains perdurent, bien que le temps les ait rendus inutiles. Par exemple, l'adjectif féminin *bone* s'est prononcé /bɔ-n/ ; l'apparition d'une nasalisation du /ɔ/ a entraîné la prononciation /bɔ̃-n/, ce qui a nécessité de doubler le <n>, et cette double consonne s'est maintenue alors que le son /ɔ̃/ a perdu sa nasalité. Il en va de même du verbe *sonner*, qui n'est plus prononcé /sɔ̃-ne/, et ne devrait donc avoir qu'un seul <n> comme dans *sonore, sonate*, mots apparus après la dénasalisation. Il n'est donc pas facile de décider quand une consonne doit être doublée, sauf quand elle suit la lettre <e> (*effet, elle, erreur, cette...*)¹. À partir du XII^e siècle, la diffusion de l'écrit a figé la graphie, tandis que la langue orale évoluait rapidement, en particulier dans le domaine de la phonologie.

Dans ce domaine, le français actuel, qui se distingue des autres langues romanes par un plus large inventaire des voyelles, se différencie de l'anglais par une plus grande stabilité des sons vocaliques. En effet, comme l'a souligné Delattre (1965), la majorité des voyelles de l'anglais se transforment en un <e> muet (/ə/) quand elles sont dans une syllabe non accentuée, ce qui a un fort impact négatif sur la lecture, mais aussi et surtout sur l'écriture.

¹ Il reste toutefois des cas ambigus (cf. *ennemie* et *ennui*) ainsi que des prononciations atypiques (cf. *femme* et les adverbes en *emment* : *évidemment, fréquemment...*).

En résumé

Les travaux de recherche ont montré que l'opacité de l'orthographe, c'est-à-dire la faible transparence des correspondances entre graphèmes et phonèmes complique l'apprentissage de la lecture chez de futurs lecteurs typiques. Ainsi, apprendre à lire est plus difficile en anglais que dans des langues qui ont une orthographe moins opaque que celle de l'anglais, par exemple l'espagnol et l'allemand, et même le français. En outre, les difficultés des dyslexiques sont aggravées par l'opacité de l'orthographe. Par exemple, elles se notent dès la précision de la réponse en lecture de mots isolés chez des anglophones et surtout dans le temps de réponse chez des dyslexiques hispanophones, italophones et même francophones. Toutefois en français, les relations graphème-phonème sont plus régulières que les relations phonème-graphème. Il est donc plus facile d'apprendre à lire que d'apprendre à écrire dans cette langue. Ces différences entre lecture et écriture s'expliquent par des raisons historiques ainsi que par des choix politiques qui sont présentés dans les sections suivantes.

B. L'Histoire de l'orthographe française : une longue réforme

1. Avant la création de l'Académie française

Que le français, davantage que les autres langues romanes, possède des phonèmes spécifiques éloignant notablement cette langue du latin, vient de raisons qui tiennent à l'histoire : la **forte influence germanique** ; le français est la langue romane des « Francs ». Cela a eu une autre conséquence : l'érosion phonétique due à la germanisation qui a produit de très nombreux homophones. Le français parlé fourmille d'équivoques que le contexte, à l'oral, peut lever, mais pas à l'écrit. Par exemple /eldäse/ peut être du singulier ou du pluriel : en écrivant cet énoncé, il faudra choisir entre elle dansait et elles dansaient². D'une façon générale, l'évolution phonétique de la langue française a multiplié les ambiguïtés que la graphie s'est chargée de lever ; dans des domaines divers, citons la chute du /b/ final des infinitifs du 1^{er} groupe (le <er> de *laver* se prononçait de la même façon que celui de *mer*), ainsi que celle du /z/ final de la 5^e forme du présent (le <ez> de *lavez* se prononçait comme celui du *canal de Suez*). En raison de ces changements de prononciation, *laver* et *lavez* deviennent homophones de *lavé*. En outre, on ne prononce plus le /s/ final des noms et adjectifs, les pluriels devenant homophones des singuliers (*poires* et *poire*). Ce qui explique le phénomène général du surmarquage de la morphologie grammaticale auquel procède la graphie. Ainsi, entre les deux phrases « Ce jeune journaliste s'exprime bien » et « Ces jeunes journalistes s'expriment bien » il n'y a qu'une différence à l'oral (ce / ces) mais quatre à l'écrit : sur le déterminant, l'adjectif, le nom et le verbe. La distinction des homophones (*saint*, *sain*) est l'une des caractéristiques de l'orthographe française. On la pratique dès la fin du Moyen Âge, en puisant notamment dans les « consonnes étymologiques » non prononcées : par exemple, <s> et le <p> de *écriture* (qui vient du latin classique *scriptura*). C'est un aspect de la « relatinisation » de la langue à laquelle, à partir de la fin du XIV^e siècle, les scribes s'attachent : cette relatinisation concerne bien des aspects, dont le lexique et la graphie. L'orthographe française devient savante, hérissée de consonnes étymologiques (*écriture*, *poulmon*), lettres de noblesse qui tentent de renouer un lien patrimonial que la parole a perdu : graphie baroque qui abuse des <y> et des consonnes à grandes hampes. C'est de cette graphie qu'hérite le XVII^e siècle, époque où la question de la norme va se poser.

² Sauf dans certains cas dans lesquels le mot suivant commence par une voyelle (*Elles aimaient danser* se prononce /elzemedäse/). Sur les liaisons, voir Encrevé (1988). Voir aussi les études sur la maîtrise de certaines liaisons en compréhension orale par les jeunes enfants dès 30 mois, voire avant (Nazzi et al, 2011).

2. Un choix malheureux au XVII^e siècle

Au XVII^e siècle, face à cette graphie traditionnelle, élitiste et savante, le camp réformateur comptait notamment le bataillon des imprimeurs calvinistes. On sait que la révocation de l'édit de Nantes les fit s'installer aux Pays-Bas, d'où ils diffusaient en France des ouvrages commodes, au petit format, adoptant une orthographe simplifiée : à la fin du XVII^e siècle, cette orthographe « hollandaise » était de fait une concurrente de la graphie traditionnelle. Le camp du progrès orthographique bénéficiait de nouveaux membres : les femmes, auxquelles on n'enseignait pas le latin et qui se fiaient à leur oreille... les Précieuses n'étaient point si ridicules. Depuis la fondation de l'Académie française en 1635, on attendait la publication de son dictionnaire. Chargée, selon l'article 24 de ses statuts, de « travailler avec tout le soin et toute la diligence possible à donner des règles certaines à notre langue », la jeune Académie devait adopter une norme orthographique pour son dictionnaire, laquelle ferait foi. Le travail, commencé par Vaugelas qui mourut à la tâche, fut repris, longuement muri, retardé. Enfin le dictionnaire fondateur et prescripteur parut en 1694.

L'Académie eut largement le temps d'examiner la situation de l'orthographe, partagée entre l'usage par défaut (latinisant et baroque) et diverses innovations, notamment la graphie simple et régulière des imprimeurs hollandais³ et des Précieuses. Les choses ne pouvaient rester en état ; de tous côtés, y compris le pouvoir royal, on pressait l'Académie de trancher. Elle discuta longuement et choisit pour son dictionnaire la graphie la plus traditionnelle (Cerquiglini, 2004), « L'Académie s'est attachée à l'ancienne Orthographe receüe parmi tous les gens de lettres ». La préface du dictionnaire ne laisse aucun doute : l'orthographe française sera latinisante, étymologique « parce qu'elle ayde à faire connoistre l'Origine des mots ». L'Académie s'oppose au courant réformateur, ainsi qu'aux divers usages innovants, tant dans l'esprit (la graphie se règlera moins par la prononciation que selon l'étymologie) que dans sa pratique : maintien des consonnes étymologiques et redoublées, refus du circonflexe que beaucoup emploient désormais à la place d'une consonne (un <s>, généralement) ou d'une voyelle disparue. L'orthographe académique écrit donc, suivant ces choix : *obmette*⁴, *poulmon*, *raffraischir*, *apostre*, *descouvrir*. La route est longue jusqu'à la norme graphique actuelle qui est à peu près celle des réformateurs de l'époque.

Considérons le début de la fable de La Fontaine Le Corbeau et le Renard (1668), qu'apprennent tous les écoliers de France sans la lire dans son orthographe d'origine (Cerquiglini, 1989) :

Le Corbeau & le Renard
Majstre Corbeau, sur un arbre perché,
Tenoit en son bec un fromage.
Majstre Renard, par l'odeur alleché,
Luy tint à peu près ce langage :
Et bon-jour, Monsieur du Corbeau,
Que vous estes joly ! que vous me semblez beau !
Sans mentir, j'i vostre ramage
Se rapporte à vostre plumage,
Vous estes le Phoenix des hostes de ces Bois.

Jusqu'en 1990, quinze mots, soulignés ci-dessous, vont changer de graphie : le mot et était écrit &, le <i> final <y>⁵ et les accents modernes étaient absents.

³ Livre imprimé en Hollande par les imprimeurs Elsevier, ou par leurs imitateurs, entre la fin du XVI^e et le début du XVIII^e, se caractérisant par une orthographe simplifiée ainsi que par un petit format (généralement in-12).

⁴ Du latin *ob* pour devant et *mittere* pour envoyer, laisser aller (d'après le TLFi, <https://www.cnrtl.fr/etymologie/>)

⁵ En fait <y> pour /i/ a été introduit dans la 1^{ère} édition du dictionnaire de l'Académie, l'édition de 1740 (sous la direction de l'abbé d'Olivet) a supprimé de nombreux <y> (*ceci, ici, voici, asile, analyse, paroxisme...*) certains ayant été réintroduits ultérieurement (*analyse*).

Le Corbeau et le Renard
 Maître Corbeau, sur un arbre perché,
 Tenait en son bec un fromage.
 Maître Renard, par l'odeur alléché,
 Lui tint à peu près ce langage :
 Et bonjour, Monsieur du Corbeau,
 Que vous êtes joli ! que vous me semblez beau !
 Sans mentir, si votre ramage
 Se rapporte à votre plumage,
 Vous êtes le Phénix des hôtes de ces bois.

Autre exemple, les fameuses Stances qu'apprennent les collégiens ont cette forme dans l'édition originale (1637) du Cid de Pierre Corneille :

Percé iufques au fonds du cœur
 D'vne atteinte impreueuë auffi bien que mortelle,
 Miserable vangeur d'vne iufte querelle,
 Et malheureux objet d'vne iniufte rigueur,
 le demeure immobile, & mon ame abbatuë
 Cede au coup qui me tuë.
 Si prés de voir mon feu recompensé,
 O Dieu l'eftrange peine !
 En cét affront mon pere eft l'offensé,
 Et l'offenfeur le pere de Chimene.
 Qve ie fens de rvdes combats !
 Contre mon propre honneur mon amour s'intereffe
 Il faut vanger vn pere, & perdre une Maijtreffe,
 L'vn m'anime le cœur, l'autre retient mon bras.

L'édition qu'on en offre aux élèves opère des corrections typographiques (introduction de <s> court de Maître et de <u>⁶, et orthographiques (par exemple, la suppression du <s> devant le <t> de maîtresse et maître et l'ajout du tréma sur le <e> final dans tuë, abbatuë indiquant l'allongement de la voyelle précédente). **Ainsi, entre l'édition originale de 1637 du Cid de Corneille et celle de 1990, vingt-trois mots vont changer de graphie :**

Percé jusques au fond du cœur
 D'une atteinte imprévue aussi bien que mortelle,
 Misérable vengeur d'une juste querelle,
 Et malheureux objet d'une injuste rigueur,
Je demeure immobile et mon âme abattue
Cède au coup qui me tue.
 Si près de voir mon feu récompensé,
 O Dieu l'étrange peine !
 En cet affront mon père est l'offensé,
 et l'offenseur le père de Chimène.
Que je sens de rudes combats !
 Contre mon propre honneur mon amour s'intéresse
 Il faut venger un père, et perdre une Maîtresse,
 L'un m'anime le cœur, l'autre retient mon bras.

6 Avant l'ajout de la voyelle <u> pour le /y/ de lune, une lettre diacritique permettait de différencier le <v> consonne de la voyelle (par exemple, l'ajout du <h> qui permet de différencier hvitre de vitre).

3. Une réformatrice opiniâtre : l'Académie

L'histoire de la norme graphique, par la suite, est celle d'une lente et prudente, mais honnête et courageuse, dans l'ensemble efficace résipiscence. Sans se déjuger, refusant tout bouleversement, mais sensible aux critiques, s'adjoignant de bons grammairiens, l'Académie entreprend un travail pluriséculaire d'autocorrection. Dès la troisième édition (1740) de son dictionnaire, confiée à l'abbé d'Olivet, excellent phonéticien qu'elle a admis en son sein, l'Académie « des Lumières » opère des aménagements systématiques et spectaculaires : suppression de consonnes superflues ; réglementation du pluriel (les mots en <t> qui le perdaient au pluriel le gardent désormais : parents, enfants). En particulier l'Abbé d'Olivet met en place un vrai régime d'accentuation : emploi de l'accent grave formant couple avec l'aigu, introduction de l'accent circonflexe. L'accentuation française est établie, comme on le voit dans les deux textes précédents.

Il serait fastidieux d'énumérer les multiples « rectifications » que l'Académie française apporta, d'édition en édition du dictionnaire, à sa propre orthographe. **Au total, un mot sur deux a changé de graphie depuis la première édition en 1694** (cf. Catach 1989). Contentons-nous de noter que ces aménagements purent être de grande ampleur : ainsi, la 4^e édition (1762) régularisa les pluriels en <z> (*amitez, bontez* devinrent *amitiés, bontés*) et en <oix> (*loix, roix* deviennent *lois, rois*). En 1835 (6^e édition), l'Académie changea les terminaisons de l'imparfait et du conditionnel (*avois, aurois*), et les finales de certains mots (*françois, anglois*, etc.) écrites désormais <ais> comme on les prononçait⁷. On reconnaît désormais aisément un texte publié avant cette date (cf. l'édition originale de 1637 du *Cid* de Corneille et celle de 1990).

Ces changements furent parfois audacieux. Ainsi la 7^e édition (1878) fit perdre à *poème, poète* et *poésie* leur tréma traditionnel au profit d'un accent : le tollé fut immense (la *Poésie* était défigurée !), quelques académiciens menaçaient de claquer la porte. Un changement inverse déclencherait sans doute aujourd'hui la même tempête.

Certaines améliorations furent malheureusement incomplètes. Dès sa 3^e édition (1740), la Compagnie entreprit de simplifier les doubles consonnes des verbes en *app* : *appaiser* devint *apaiser*, mais *apparaitre*, par exemple, ne fut pas touché. L'édition de 1878 supprima les doubles <h> (*rhythme* devint *rythme*), mais laissa subsister les <th> et <rh>.

Un tel processus d'ajustement courut enfin sur plusieurs siècles. La plus récente édition complète du dictionnaire (8^e ; 1932-1935) corrigea environ 500 graphies, manifestant, notons-le au passage, une faveur pour la soudure des composés tels que *contrecoup, entracte, primesautier*, etc. Mais elle ajouta aussi quelques erreurs, comme *nénufar* devenu *nénuphar* par confusion avec *nymphéa*, alors qu'il vient non du grec mais du persan.

En résumé

Deux éléments cruciaux expliquent l'origine des différences entre l'oral et l'écrit en français. **En premier, la double filiation du français : romane et germanique.** Cela s'est traduit par des évolutions de la phonologie qui ont contribué à creuser l'écart avec l'écrit. Plus que les autres langues romanes, le français a des phonèmes spécifiques, la prononciation de certains ayant fortement évolué. Ainsi, des voyelles orales se sont transformées en voyelles nasales avant de retrouver leur oralité tout en gardant, pour certaines, un double <n> à l'écrit (cf. *sonner* versus *sonore*). Les différences avec les autres langues romanes tiennent aussi à l'influence germanique sur le marquage de la morphologie grammaticale : l'introduction du déterminant devant le nom et du pronom sujet devant les verbes a

⁷ Dès la fin du XII^e siècle, on écrivait <oi> que l'on prononçait /oj/ (comme dans *cowboy*). Par la suite ce groupe s'est prononcé /wɛ/ comme dans *rouet*, puis /ɛ/ pour les formes verbales (*avois*), et /wa/ pour les autres (*roi, moi*)

accompagné la disparition à l'oral, mais pas à l'écrit, de la plupart des marques de flexions (entre *ce chanteur, il chante à l'opéra* versus *ces chanteurs, ils chantent* à l'opéra, il n'y a qu'une seule différence à l'oral contre 4 à l'écrit).

En second, le rôle de l'Académie. Alors que le choix initial était celui d'une orthographe savante, environ un mot sur deux a changé d'orthographe depuis la première édition du dictionnaire de l'Académie. À l'aune de ces changements les *Rectifications* de 1990, ne paraissent ni démesurées, ni extravagantes.

4. Évolution de l'orthographe depuis la mise en place de l'instruction obligatoire

Comme le rappelle Chervel (1977, 1991), la III^e République, en rendant obligatoire l'instruction, a souhaité réformer l'orthographe pour la rendre moins opaque. Jules Ferry s'adressait ainsi aux instituteurs en 1880 « Aussi, Messieurs, ce que nous vous demandons à tous, c'est de nous faire des hommes avant de nous faire des grammairiens ! [...] Oui, vous avez compris qu'il faut réduire dans les programmes la part des matières qui y tiennent une part excessive ; vous avez compris qu'aux anciens procédés, qui consomment tant de temps en vain, à la vieille méthode grammaticale, à la dictée – à l'abus de la dictée – il faut substituer un enseignement plus libre, plus vivant et plus substantiel [...] épargnons ce temps si précieux qu'on dépense trop souvent dans les vétilles de l'orthographe, dans les règles de la dictée qui font de cet exercice une manière de tour de force et une espèce de casse-tête chinois » (Discours au Congrès pédagogique, 2 avril 1880). Des écrivains comme Sainte-Beuve et Anatole France se sont mobilisés, et de nombreux linguistes ont publié des propositions (cf. Darmesteter, 1888).

C'est dans ce contexte que le ministre de l'Instruction publique, Georges Leygues, a publié en 1900 un arrêté de tolérances pour ne plus compter comme fautes un certain nombre de problèmes concernant l'orthographe lexicale (soudure des mots composés) et grammaticale (pluriel des mots composés et des mots étrangers, accord du participe passé). Compte tenu de l'opposition de l'Académie française, il publie un second arrêté plus modeste et toujours en vigueur (Leygues, 1901). Quelques années plus tard, l'Académie approuve le Rapport Faguet, qui reprend certaines propositions de Brunot (1905), dont la régularisation des pluriels en <ous>, et le remplacement de <y> voyelle par <i> (*stile, analyse*), mais sans l'appliquer dans son dictionnaire.

En 1965, le rapport Beslais relance les propositions de simplification de l'orthographe comme la suppression du <x> marque du pluriel pour les noms et les adjectifs (*des bijoux* comme *des bisous*), mais aussi marque de la personne pour les verbes (*je peux* comme *je meus*). Certaines de ces simplifications, qui avaient été proposées par Darmesteter (1888) et Brunot (1905) sont reprises dans l'arrêté de 1976 du ministre de l'Éducation nationale René Haby : ainsi ne seront plus comptées comme fautes, aux examens et concours nationaux, l'absence d'accent circonflexe sur les voyelles autres que <e> (*drole, maitre*), ni le <s> après vingt ou cent (*quatre-vingts-deux, deux cents deux*).

a. Les arrêtés de 1900, 1901 et 1976 sur l'orthographe

Ces arrêtés ont pour but d'améliorer l'enseignement de l'orthographe grammaticale, l'accord ne s'entendant ni au pluriel à l'exception des liaisons (*les [z]enfants, ils [z]arrivent*) ni au féminin sauf pour une trentaine de verbes (par exemple, *dire [dit-dite], faire [fait-faite], mettre [mis-mise]*). Ils portent principalement sur l'accord du participe passé, un des problèmes les plus épineux de la grammaire française (Abeillé et Godard, 2021)⁸

⁸ Jusqu'au XVII^e siècle, on pouvait prononcer le <s> final au pluriel et, jusqu'au début du XX^e siècle, et encore de nos jours dans certaines régions, le <e> final pouvait se faire entendre par un allongement de la voyelle qui précède (*je l'ai lu ≠ je l'ai lue*) mais cette prononciation est de plus en plus rare.

Employé aux temps composés, le participe passé suit la règle commune avec l'auxiliaire être, c'est-à-dire l'accord est avec le sujet (*Ils partent. Ils sont partis*), sauf pour les verbes pronominaux (*Elle s'est fait mal*). Il n'en va pas de même avec l'auxiliaire avoir. Dans ce cas, jusqu'au XIII^e (Marchello-Nizia 1999 ; Cerquiglini 2021), l'ordre des mots étant assez libre, l'accord se faisait quelle que soit la position du participe passé. Ainsi, dans *La Queste del saint Graal* (1225-1230), Marchello-Nizia (1999) observe un accord à 94% quand l'objet précède le participe (*J'ai les lettres écrites*) et à 71% quand il le suit. Au XVI^e siècle, l'ordre des mots est plus fixe (auxiliaire – participe – objet), l'accord est donc moins nécessaire. Toutefois, en 1538, Marot propose d'accorder le participe avec l'objet quand celui-ci « va devant » (*Dieu en ce monde nous a faits*), sur le modèle de l'italien (*Dio noi a fatti*), très à la mode à l'époque, mais cette règle n'est pas vraiment appliquée, comme en témoigne le poème de Ronsard (1550) qui accorde le participe avec l'objet qui le suit (sa robe) : *Mignonne, allons voir si la rose / Qui ce matin avait desclose / Sa robe de pourpre au soleil*.

Au milieu du XVII^e siècle, l'académicien Vaugelas (1647) note à propos de l'accord du participe passé « il n'y a rien de plus important ny de plus ignoré ». Selon lui, l'accord avec l'objet ne se fait que si le participe est en position finale, le participe restant invariable quand il est suivi d'un attribut (*les habitants nous ont rendu maîtres de la ville*), d'un complément prépositionnel (*les maux qu'il a fait à son fils*), d'un infinitif (*Je l'ai fait peindre [cette fille]*) ou d'un sujet inversé (*la peine que m'a donné cette affaire*). Cette règle est reprise par le grammairien Bouhours (1675) et utilisée par des écrivains tels que Molière et Racine⁹. A cette époque, l'accord peut se faire avec en, et le participe des verbes pronominaux peut s'accorder avec le sujet, même s'il y a un objet direct (cf. Wagner et Pinchon, 1962 : *Pendant ces derniers temps combien en a-t-on vus ? La Poule aux œufs d'or, La Fontaine ; Ils se sont donnés l'un à l'autre une promesse de mariage, l'Avare*, Molière, 1668). A la fin du XVII^e siècle (1679), l'Académie française va étendre l'accord du participe passé avec l'objet qui précède aux verbes pronominaux (*la chose qu'il s'est dite*).

Les grammaires scolaires du XIX^e siècle ne vont pas reprendre la règle de Bouhours (1675) et Vaugelas (1647), sauf pour faire + infinitif, mais multiplier les exceptions : l'accord se fait avec l'objet, mais pas s'il s'agit du pronom en (*Des pommes, j'en ai mangé*), ni d'un complément de mesure (*la somme que ça a couté*), ni d'un verbe impersonnel (*les efforts qu'il a fallu*). En ce qui concerne les verbes pronominaux, on aboutit à trois cas : invariabilité (*Il s'est dit des choses*), accord avec le sujet (*des choses se sont dites*), accord avec l'objet (*les choses qu'il s'est dites*).

Comme le notait Bescherelle dans la *Grammaire Nationale* (1852), l'accord du participe passé avec l'objet est « la règle la plus artificielle de la langue française » et « l'épouvantail des enfants ». Les autres grammairiens des XIX^e et XX^e siècles sont d'accord sur ce point. Ainsi, dans le *Bon Usage* (1936), Grevisse parle d'une « pente instinctive vers l'invariabilité du participe passé » avec avoir. De fait de nombreux écrivains ne l'accordent pas avec l'objet ou généralisent l'accord avec le pronom en¹⁰. De même, dans la *Grammaire méthodique du français*, Riegel et al. (1994) observent que « vu la multiplication des cas d'espèce due aux interférences avec d'autres règles, elle a toujours posé de redoutables problèmes d'application » et « la tendance actuelle est à l'invariabilité du participe non seulement à l'oral où une phrase comme *Quelle impression vous a-t-il produite ?* sonne bizarre mais aussi chez les meilleurs auteurs modernes ».

Plus récemment, Abeillé et Godard (2021) notent dans la *Grande Grammaire du français* : « La question de l'accord du participe passé se pose plus à l'écrit qu'à l'oral, où on l'entend rarement. Elle suscite de nombreux commentaires mais il est souvent difficile de savoir si le participe est accordé ou non ». De plus, compte tenu des exceptions, la règle d'accord empêche de bien analyser comme objet direct le pronom en avec les verbes transitifs directs (*j'en ai mangé, j'en ai acheté* ≠

9 De fait, dans *Les femmes savantes* (1672), Molière n'accorde pas le participe passé à l'acte IV (c'est une femme qui parle) : « jamais il ne m'a prié de lui rien dire » (v.1138), car il n'est pas final.

10 « As-tu vu la tête qu'il a fait quand il s'est aperçu qu'elle n'était pas là ? » (Proust, *A la recherche du temps perdu*, Un amour de Swann, 1913) ; « C'est celle [une doctrine] que nous ont transmis nos ancêtres » (Gide, *Les Faux Monnayeurs*, 1925) ; « Il vous vaudra plus de surprises, [...] plus de bonheur aussi que je ne vous en ai jamais donnés ». (d'Ormesson, *Histoire du juif errant*, 1990).

je les ai **mangés**, je les ai **achetés**), ainsi que certains compléments de mesure (les sommes que ça a **couté** ≠ les sommes que j'ai **payées**), le complément des verbes impersonnels (les efforts qu'il a **fallu**) et le complément de faire suivi d'un infinitif (Je les fais cuire versus les sardines que j'ai **fait** cuire).

C'est en raison de ces différents problèmes que, dès le début du XX^e siècle, le ministre de l'Instruction publique Georges Leygues a proposé dans un arrêté de 1900 de tolérer l'invariabilité du participe avec avoir et les pronominaux, car l'accord avec l'objet « tombe en désuétude. Il paraît inutile de s'obstiner à maintenir artificiellement une règle qui n'est qu'une cause d'embarras dans l'enseignement, qui ne sert à rien pour le développement de l'intelligence et qui rend très difficile l'étude du français aux étrangers. » Ces tolérances ont été reprises partiellement, en 1976 par le ministre Haby dans un nouvel arrêté, toujours en vigueur, qui demandait de tolérer, pour les examens et concours nationaux, l'invariabilité du participe devant un infinitif (la pomme que j'ai **vu** tomber), et son accord avec en et les compléments de mesure (j'en ai mangés, les sommes que ça a coutées), alors que les Rectifications de 1990 prônent seulement l'invariabilité de laissé devant infinitif, sur le modèle de faire (je les ai **laissé** cuire).

b. Les Rectifications de 1990

Ces rectifications (cf. tableau 1 ; voir aussi Béchenec & Sprenger-Charolles, 2014 ; Cerquiglini, 2018) concernent surtout l'orthographe lexicale. Elles reprennent en partie les simplifications et rationalisations demandées au XIX^e siècle, ainsi que les tolérances des arrêtés de 1900 et 1976.

Tableau 1. Principales Rectifications orthographiques de 1990

Orthographe	Règles	Exemples	Exceptions
Orthographe lexicale			
Soudure (Leygues 1900)	Soudure des mots composés commençant par un adjectif, un verbe, une préposition (contre, entre...), et ceux d'origine étrangère	rondpoint, portefeuille, contrattaque, weekend	grand-père, petit-fils, prie-Dieu, sèche-cheveu, et ceux avec des prépositions (après-midi, avant-garde, sans-papier, sous-préfet...)
Traits d'union	Systématisation dans l'écriture des numéraux complexes	cent-un, deux-mille, quatre-centième	
Accent circonflexe	Supprimé sur <i> et <u>	cout, gout, maitre, parait	dû, sûr ; imparfait du subjonctif (fît, sût)
Accent aigu	Le <è> remplace <é> en fonction de la prononciation	évènement, cèderai	
Conjugaison des verbes en -eler et -eter	Le <è> remplace <e> et est suivi par consonne simple dans ces verbes	elle ruissèle, j'époussète, j'étiquête...	appeler et jeter (plus ceux de la même famille) j'appelle, je jette
Tréma	Tréma ajouté sur le <u> prononcé devant voyelle	ambigüe, argüer	
Orthographe grammaticale			
Pluriel des noms composés ou empruntés	La marque du pluriel est sur le second nom	des après-midis, des gratte-ciels	des prie-Dieu (exemple qui signale la part de l'idéologie dans l'orthographe)

Ces modifications sont intégrées comme options dans les traitements de texte et figurent dans tous les dictionnaires. Supposées enseignées depuis 2008 en France et « prioritairement » en Belgique, elles figurent dans la plupart des manuels du primaire depuis 2016. Mais elles ne sont pas appliquées par tous les manuels du secondaire ni par les médias, peu par les auteurs et les éditeurs, et sont méconnues du grand public.

En résumé

Les rectifications de l'orthographe de 1990 concernent surtout l'orthographe lexicale alors que les arrêtés de 1900, 1901 et 1976 portaient surtout sur l'orthographe grammaticale (en particulier sur l'accord du participe passé). Les deux reprennent en partie les simplifications et rationalisations demandées par l'Académie depuis sa création. Certaines sont entrées dans l'usage, par exemple, pour les rectifications de 1990, les accents graves (cf. *événement*), d'autres moins. Nous verrons dans la suite que ces rectifications facilitent l'apprentissage de l'orthographe. C'est pourquoi il est important de les appliquer systématiquement, dans tous les manuels, mais aussi dans tous les textes mis à disposition des élèves.

C. Données quantitatives sur l'orthographe actuelle

1. Données sur la fréquence des mots et des unités infralexicales

Le développement d'outils informatiques a permis la mise en place de bases de données qui se situent dans le prolongement des travaux pionniers de Catach (1984), Véronis (1998), Cerquiglini (1989). Ces données permettent de connaître les caractéristiques majeures de l'orthographe du **français** : la fréquence des unités infralexicales (correspondances graphème-phonème et phonème graphème, CGPh et CPhG, entre autres) et leur consistance dans le lexique écrit adressé aux adultes (cf. *Lexique-infra* : Gimenes, Perret & New, 2020) ou aux enfants (cf. *Manulex-infra* : Peereman, Lété & Sprenger-Charolles, 2007 ; *Manulex-Morpho* : Peereman, Sprenger-Charolles & Messaoud-Galusi, 2013 ; *EQOL* : Stanké et al., 2019).

La présentation qui suit prend en compte les bases de données issues d'un corpus sur le lexique écrit adressé aux enfants : celles dérivées, d'une part, de *Manulex* (Lété et al., 2004) et, d'autre part, d'*EQOL* (Echelle Québécoise d'Orthographe lexicale, Stanké et al., 2019). *Manulex* présente la fréquence des mots du lexique écrit adressé à l'enfant d'après des manuels français du CP au CM2. Trois sous-bases sont dérivées de *Manulex* : *Manulex-Infra* (Peereman et al., 2007), *Manulex-Morpho* (Peereman et al., 2013; Peereman & Sprenger-Charolles, 2018) et *Morphiste* (Peereman et al., en préparation)¹¹. Les dernières tiennent compte de la morphologie dans une partie du corpus de *Manulex* (les 10 000 mots les plus fréquents). La base de données *EQOL* a été développée plus récemment, en suivant les mêmes principes, mais sur un corpus issu du lexique écrit adressé aux enfants du Québec.

a. Données sur la fréquence des mots

La représentativité du corpus utilisé est une question cruciale. Si on compare les deux bases issues d'un corpus du lexique écrit adressé aux enfants, *Manulex-infra* (base de données française) et *EQOL* (base de données québécoise), on constate que le recouvrement est important. En effet, 76% des 14 800 items de *EQOL* se retrouvent dans *Manulex* (soit 11 246 mots) et la corrélation entre la fréquence des mots communs à ces deux bases est élevée (0,97).

¹¹ Voir aussi pour la morphologie dérivationnelle en français la base *Démonette* (Hathout & Namer, 2023)

b. Données sur la fréquence et la consistance des unités infralexicales

La figure 2 indique la **consistance** des relations graphème-phonème dans Manulex-Morpho en tenant compte, ou non, de la morphologie. La consistance est un indice qui varie de 0 à 100. Ainsi, *onze* est un mot très consistant pour la lecture (100%) mais pas pour l'écriture (44-45%) ce qui provient du fait que <on> et <z> suivi par un <e> de fin de mot ne peuvent se lire que d'une seule façon alors que ces mêmes graphèmes peuvent s'écrire de différentes manières (avec, entre autres, le graphème <om>, <s> intervocalique et sans le <e> de fin de mot ; voir l'annexe 1 pour plus de détails sur les modes de calcul de la consistance). La figure 2 permet de constater que la consistance des CGPh est très élevée, surtout pour l'initiale des mots, sans changement notable en fonction de la prise en compte de la morphologie. En revanche, l'impact de la morphologie est très fort en écriture pour les finales des mots.

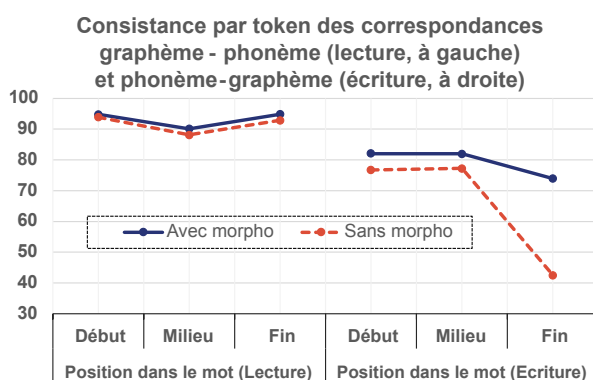


Figure 2. Consistance des correspondances graphème-phonème (CGPh, lecture) et phonème-graphème (CPhG, écriture) en tenant compte, ou non, de la morphologie (D'après Manulex-Morpho, Sprenger-Charolles & Peereboom, 2023)

La figure 3 indique la répartition des lettres muettes écrites à la fin des noms apparaissant dans Manulex, soit environ 15000 mots, dont 70% se terminant par une lettre muette (Peereboom, travail en cours). Le <s>, qui est la plus fréquente, code dans 96% des cas une flexion du nombre. Quelques noms fréquents ont un <s> final au singulier (*temps*, censé faire le lien avec le latin *tempus* - pourtant ce lien n'est pas marqué dans d'autres mots, comme *champ* issu de *campus*).

Le <e>, arrive en seconde position. Dans la plupart des cas, il ne correspond pas à une marque morphologique (par exemple, *femme*, *arbre*, *lycée*). Certains noms en <e> sont des épécènes, c'est-à-dire qu'ils ont des deux genres (*un / une adulte*, *athlète*, *élève*, *guide*, *libraire* ; *artiste*, *journaliste*, *touriste*...). D'autres sont soit uniquement masculins (*un personnage*, *monarque*, *génie*, *parasite*, *fantôme*, *cadavre*, *squelette*), soit uniquement féminins (*une canaille*, *brute*, *personne*, *vedette*...). En outre, le sens d'un nom féminin peut différer de celui qu'il a au masculin. Ainsi, le féminin de *médecin* n'est pas *médecine*, celui de *mandarin* n'est pas *mandarine*, tout comme celui d'*ingénieur* n'est pas *ingéneuse*, ou encore celui de *gourmet* n'est pas *gourmette*.

Parmi les trois autres lettres finales muettes rencontrées dans plus de 100 noms qui apparaissent dans la figure 3, la plus fréquente est <t> suivie par <x> et <d> (respectivement 966, 212 et 112 occurrences). Dans près de 50% des cas, le <t> est une marque morphologique (*candidat-candidate*, *abricot-abricotier*), mais pas dans d'autres (*escargot*, *rempart* ...). Toutefois on relève aussi des noms sans <t> final qui ont pourtant un dérivé en <t> (*abri* – *abriter*).

Le <x> final est le plus souvent une marque du pluriel pour les noms en <(e)au> (*noyau*, *cadeau*) et <eu> (*feu*). Mais cette graphie est le résultat d'une erreur : les scribes médiévaux utilisaient le signe **ʒ** (dénommé *um*) pour abrégé certaines lettres de fin de mot, en particulier <us>. Ce signe a été confondu avec <x>, et un <u> a été ajouté, ce qui explique pourquoi *feus* est devenu *feʒ* puis *feux*. Toutefois, le pluriel de *pneu* (qui est un mot récent) est *pneus*. Dans le cas des noms en <ou>, le pluriel en <s> est devenu majoritaire (*sous*, *bisous*, *nounous*), mais il reste sept exceptions (*bijou*, *caillou*, *chou*, *hibou*, *genou*, *joujou*, *pou*).

De plus, le <x> final est parfois présent au singulier (*choix, voix*), sans valeur morphologique et, de plus, avec une valeur étymologique instable : *voix* vient de *vox* en latin, mais *roi* qui vient de *rex* n'a pas de <x>, et *choix* ne vient pas d'un mot en <x>. En outre, les <x> à la fin des chiffres deux, dix et six sont des anomalies. En effet, les deux derniers se prononcent /dis/ et /sis/ et leurs dérivés s'écrivent avec un <x> (*deuxième, dixième* et *sixième*) qui se prononce /z/.

Enfin, le <d> final peut être une marque de flexion (*bavard* vs *bavarde*) ou de dérivation (*bavardage*). Toutefois dans quelques mots, il n'est pas dans le lemme (*cauchemar* vs *cauchemarder*) alors qu'il est présent dans d'autres mots sans fonction morphologique (*foulard*). S'ajoutent à ces cas les nombreux mots ayant une rime en <ar>, sans <d> (*car, char, bar*).



Figure 3. Répartition des lettres finales dans les noms apparaissant dans Manulex, environ 15000 mots (Peereman, travail en cours).

Pour les mots d'autres catégories, en particulier les mots grammaticaux, qui sont très fréquents, les lettres finales muettes posent des problèmes de lecture et d'écriture, la morphologie n'étant le plus souvent d'aucun secours.

c. Utilisation de ces bases de données

Pour le moment, ces bases de données sont utilisées surtout par les chercheurs pour contrôler de façon systématique les variables qu'ils manipulent dans leurs études : par exemple, la fréquence et la régularité des mots et des unités infralexicales (phonèmes, graphèmes, marques de morphologie lexicale ou grammaticale). Les progressions pédagogiques, notamment celles utilisées dans les manuels scolaires, devraient elles aussi s'appuyer sur ces données qui permettent de guider cette progression : enseigner en premier les relations phonème-graphème les plus fréquentes et les plus régulières mais présenter rapidement le problème posé par les marques de morphologie lexicale (<d> de *accord*) et grammaticale (marque du genre comme le <e> dans *amie*, et marques du nombre comme <s> et <ent> dans *ils chantent*) qui doivent s'écrire alors qu'elles ne se prononcent généralement pas, sauf dans les liaisons (voir la note 2 ; voir aussi Dehaene et al., 2011).

2. Données provenant d'études descriptives et expérimentales

Depuis la fin des années 1990, de nombreuses études ont porté sur la maîtrise de l'orthographe lexicale et grammaticale, principalement, mais pas uniquement, par les enfants des classes élémentaires et du collège. Ces nouvelles données offrent des pistes de travail aux enseignants, leur permettant d'améliorer leurs pratiques afin de mieux accompagner les élèves dans ces apprentissages complexes.

La première partie de cette section porte sur l'orthographe lexicale avec deux séries de travaux : ceux provenant des *Échelles d'Orthographe Lexicale* EQOL et EOLE présentant des données prises dans les

cinq niveaux de l'école élémentaire et ceux issus d'études avec, entre autres, des données longitudinales (du milieu du CP et à la fin du CM1). La suite est centrée sur la morphologie lexicale et grammaticale et contiennent des données sur la maîtrise de l'accord du participe passé avec les auxiliaires être et avoir.

2.1 L'apprentissage de l'orthographe lexicale

a. EOLE : Échelle d'acquisition de l'orthographe lexicale

Ce corpus, publié en 2003¹², comporte 12 000 mots fréquents du français écrit relevés dans la presse nationale qui ont été dictés en contexte à environ 50 000 élèves scolarisés dans 2 000 classes des 5 niveaux de l'école élémentaire (cf. Pothier & Pothier, 2003).

L'examen des mots correctement écrits par au moins 75% des enfants de CP, soit un peu plus de 300 (cf. Annexe 2) permet de constater qu'ils ont presque tous des correspondances phonème-graphème consistantes (moyenne sur 100 d'après Manulex-infra : 79,4 ; écart-type : 11,98). Par exemple, un seul mot comprend une double consonne non motivée (ballon), 4 contiennent un graphème avec une prononciation exceptionnelle (le <p> de *sept*, le <x> de *six*, les deux <ll> de *fil* et le <ez> de *nez*) et 5 ont une consonne finale muette qui peut permettre de créer des mots dérivés (le <t> de *chat* et de *cent*, le <x> de *faux*) ou non (le <p> de *loup*, la *loupe* n'étant pas sa femelle).

Cette étude a également évalué l'impact de l'orthographe rectifiée. Les résultats (Pothier & Pothier, 2016 et Annexe 3) portent sur une partie des mots pour lesquels on trouve dans EOLE une comparaison entre l'ancienne et la nouvelle orthographe. La plupart des modifications considérées (95 sur 136) concernent la suppression de l'accent circonflexe sur <i> et <u> (68 occurrences), les autres portent sur le remplacement de l'accent aigu par un accent grave (comme dans *évènement*, 7 occurrences), le déplacement du tréma (comme dans *ambigüité*, 4 occurrences) ainsi que la « rationalisation » de 9 mots français (par exemple, *charriot*) et 7 étrangers (par exemple, *speakeur*). Pratiquement tous ces changements ont permis d'augmenter le pourcentage d'élèves maîtrisant mieux cette orthographe que la traditionnelle. EOLE a donc mis en relief l'impact positif des rectifications de l'orthographe en particulier celui de la suppression des accents sur <i> et <u>¹³.

Cette étude a aussi mis en relief le fait que, au début de l'apprentissage de l'écriture (fin de CP) les élèves s'avèrent majoritairement capables d'écrire correctement surtout des mots réguliers et ayant une structure syllabique simple consonne-voyelle, même longs (*rapidité*) et rares (*avidité*). Toutefois certains mots correctement orthographiés par la majorité des enfants de CP ne le sont plus ultérieurement : par exemple, *mécano*, *parmi* et *cou* (désignant une partie du corps). Il se peut que les élèves écrivent par la suite ces mots de façon plus complexe mais en suivant d'autres types de régularités, celles sur les fins de mots : *cou* avec un <p> en lien avec *couper*, *coupure* et *mécano* avec <ot> (comme dans *idiot*, *tricot*...) ou <eau> (comme dans *beau*, *bateau*, *drapeau*...) qui sont les finales les plus fréquentes pour les mots se terminant par le son /o/. Cette étude ne permet pas de vérifier cette hypothèse, qui a toutefois été évaluée dans des recherches présentées dans la suite de cette synthèse (cf. Lété, Peereman & Fayol, 2008).

b. EQOL : Échelle Québécoise d'Orthographe lexicale

Environ 1200 mots les plus fréquents des 15 000 mots de ce corpus ont été sélectionnés pour cette étude. Intégrés dans 66 exercices contenant chacun 20 phrases (un mot test par phrase), ces mots ont été dictés à près de 5 000 élèves des 5 premiers niveaux du cursus élémentaire, **scolarisés** principalement dans des écoles publiques. Les données ont été recueillies en trois temps distincts (mai et septembre 2015 et mai 2016), sur une période de trois semaines consécutives. A la différence d'EOLE, les élèves faibles en orthographe (10^{ème} centile) ont été éliminés (cf. Stanké et al., 2019).

12 L'édition 2022 comportait des erreurs de catégorisation des mots. Par exemple (p. 60) 'verbe' pour *chômage*, 'nom' pour *chômer*. Ce problème, qui provenait probablement d'un décalage de lignes lors de l'intégration des données avec les mots de l'orthographe rectifiée, a été corrigé dans la nouvelle édition de 2024, qui remplace celle de 2022.

13 Pour la Belgique, Dister et Moreau (2019) ont étudié la dictée du Balfroid, championnat d'orthographe annuel de fin de primaire, de 2013 à 2016, et observent que les élèves qui appliquent les Rectifications (sans circonflexe sur <i> et <u>) font moins d'erreurs d'accent que ceux qui utilisent l'orthographe ancienne.

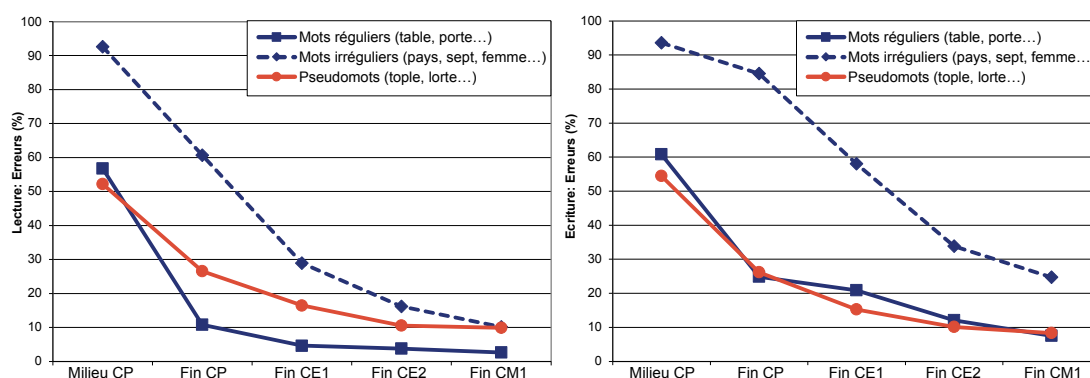
Pour la plupart des mots communs à cette échelle et à EOLE les écarts de scores sont minimes (autour de 15%). Toutefois certaines différences peuvent dépasser 70%. En outre, le plus souvent, l'orthographe des mots tend à être acquise plus précocement chez les élèves québécois que chez les français. Selon les auteurs, cette différence peut s'expliquer, entre autres, par le fait que les élèves ayant des difficultés orthographiques sévères ont été exclus de EQOL mais pas de EOLE. Au-delà de ces différences, ces deux études montrent qu'il faut environ cinq ans pour que les élèves soient en mesure d'orthographier les mots selon la norme.

c. Résultats d'une série d'observations longitudinales

Les publications précédentes (EOLE et EQOL) présentent des résultats d'études transversales : ce ne sont pas les mêmes enfants qui ont été vus à chaque niveau scolaire. Les différences observées pourraient donc s'expliquer par des différences de population. Les résultats qui suivent sont issus d'une série d'études utilisant des données longitudinales du milieu du CP à la fin du CM1.

Si comme le suggèrent les résultats précédents, les enfants utilisent essentiellement la procédure phonologique (c'est-à-dire CPhG en écriture) au début de l'apprentissage, les mots réguliers fréquents doivent être écrits – et lus – aussi bien que des pseudomots (des mots inventés de même niveau de complexité orthographique que les mots réguliers : *table* versus *tople*) et mieux que des mots irréguliers au niveau de CPhG ou des CGPh (*pays*, *sept* ou *femme*). Ces hypothèses ont été examinées dans une série d'études pionnières avec des enfants français qui ont été suivis du CP au CM1 (cf. Sprenger-Charolles, Siegel & Bonnet, 1998 et Sprenger-Charolles, Siegel, Béchenneq & Serniclaes, 2003 ; voir la synthèse de Deacon, Desrochers & Levesque, 2017). Cette étude s'est étalée sur 6 ans (une année de pré-tests, suivie par les premières observations en début de maternelle, suivies par celles de mi-CP, puis par les 4 sessions de fin CP, CE1, CE2 et CM1). Les données sur la fréquence et la consistance des unités infralexicales n'étaient alors pas disponibles, ce qui explique le choix de la variable binaire régularité (mots réguliers versus *mots irréguliers*).

Les résultats attendus ont été observés en milieu de CP en lecture, comme en écriture (cf. figures 4 et 5) : dans les deux modalités, les erreurs sur les mots réguliers ne sont alors pas plus nombreuses que celles sur les pseudomots. Les différences entre mots réguliers et pseudomots deviennent toutefois rapidement significatives en lecture (dès la fin du CP) mais pas en écriture. C'est entre ces deux périodes que se produisent les changements les plus notables en lecture : toutes les erreurs diminuent alors significativement, mais davantage celles sur les mots réguliers (moins 46%) que celles sur les mots irréguliers et les pseudomots (moins 25 à 33%).



Figures 4 et 5. Évolution des erreurs en lecture et en écriture du milieu du CP à la fin du CM1 (d'après Sprenger-Charolles et al., 2003)

Les résultats observés en lecture s'expliquent par le fait que les mots réguliers de cette étude sont à la fois réguliers et fréquents, ce qui n'est le cas ni pour les pseudomots (qui sont réguliers mais pas fréquents), ni pour les mots irréguliers (qui sont fréquents mais pas réguliers). Ces résultats permettent de comprendre pourquoi les progrès des enfants espagnols, qui n'ont à lire pratiquement que des mots réguliers, sont si rapides alors que l'apprentissage de la lecture est lent et laborieux pour les enfants anglais (Dehaene, 2007 ; Seymour et al., 2003 ; Ziegler, 2018).

Inversement, les mots réguliers ne sont pas mieux écrits que les pseudomots (Figure 5), ce qui peut s'expliquer par le fait que, pour écrire un mot, même régulier, il ne suffit pas de connaître les CPhG, il faut en plus maîtriser l'orthographe canonique : alors que les mots auto et beau ne peuvent se lire que d'une seule façon, ils peuvent s'écrire de plusieurs manières, une seule étant correcte. Cette contrainte n'existe pas pour l'écriture de pseudomots.

d. De la régularité à la consistance

Une autre étude (Lété et al. 2008) a utilisé les données de consistance des unités infralexicales issues de Manulex-Infra (Peereman et al., 2007) pour analyser les résultats des élèves examinés dans EOLE (Pothier & Pothier, 2003). L'effet de la consistance (comme celui de la régularité dans l'étude de Sprenger-Charolles et al., 2003) est le plus important, sans changement notable avec le niveau scolaire. En revanche, celui de la fréquence lexicale augmente surtout de la fin du CP à celle du CE1. Durant la même période, l'effet de la longueur des mots diminue.

La portée des résultats de cette étude est limitée par le fait que, à la différence de celle de Sprenger-Charolles et al. (1998 et 2003), celle de Lété et al. (2008) n'est pas longitudinale et n'a examiné que les données en écriture, sans comparaison avec la lecture. En outre, certains graphèmes sélectionnés (<emm> et <enn>) posent des problèmes. Ainsi d'après la version initiale de Manulex-Infra dans laquelle le mot femme a 3 graphèmes (<f>, <emm>, plus <e> final) sa consistance est très élevée, aussi bien pour la lecture que pour l'écriture (82,9 et 73,7%). En revanche, les nouveaux décomptes (<https://inframorph.github.io/index.html>), basés sur un découpage plus classique avec 4 graphèmes (<f>, <e>, <mm>, plus <e> final) sa consistance, qui est un peu moins élevée pour la lecture (75,9), est très faible pour l'écriture (37,7%).

L'irrégularité de l'orthographe du français a un impact considérable sur les élèves. Selon les études passées en revue, il faut environ cinq ans pour que les élèves francophones soient en mesure d'orthographier les mots isolés selon la norme. Ces résultats corroborent ceux obtenus dans des comparaisons interlangues qui indiquent que, dans des systèmes d'écriture opaques comme l'est le français pour les relations phonème-graphème, l'apprentissage de l'orthographe lexicale est beaucoup plus lent que dans des systèmes d'écriture plus transparents (Caravolas 2004 ; Landerl & Wimmer 2008 ; Seymour et al. 2003). Ils indiquent que, au début de l'apprentissage de l'écriture (comme pour celui de la lecture) la régularité (ou la consistance) des mots a autant d'importance que leur fréquence (les mots réguliers sont aussi bien lus et écrits que des mots rares ou inventés de même niveau de difficulté orthographique). Toutefois certains résultats signalent aussi que des mots correctement orthographiés en CP ne le sont plus par la suite. C'est le cas d'un mot tel que le *mécano*, <ot> et <eau> étant des finales plus fréquentes que <o> (cf. *idiot*, *tricot* ; *bateau*, *troupeau*).

2.2. Le poids de la morphologie lexicale et grammaticale dans l'orthographe du français

a. Un problème spécifique au français : celui des marques de flexion et dérivation

De nombreuses marques de flexion qui s'écrivent ne se prononcent pas (par exemple, le <e> du féminin dans *ami*e, le <s> et le <ent> du pluriel dans ils chantent). Parmi celles qui se prononcent, on trouve des homophones (*chanté*, *chanter*, *chantez*...). Ces deux irrégularités permettent de comprendre pourquoi les enfants, tout comme les adultes, même lettrés, produisent des erreurs d'orthographe dans ce domaine (Fayol, 2020).

Plusieurs travaux de recherche ont porté sur les flexions du nombre (Fayol, 2020 ; Brissaud & Fayol, 2018), la marque <s> pour le pluriel des noms et des adjectifs, et celle du <ent> pour les verbes à la troisième personne. Trois résultats majeurs ressortent de ces recherches. D'une part, la compréhension de ces marques en lecture est plus précoce que leur écriture. D'autre part, la maîtrise de la flexion nominale <s> est plus précoce et mieux assurée que celle de la même flexion appliquée à l'adjectif. Enfin, la maîtrise de la flexion verbale <ent> est la plus tardive.

L'acquisition des marques du genre à l'écrit est plus problématique que celle des marques du nombre (Brissaud & Fayol, 2018), le <e>, marque du féminin, étant moins consistant que le <s>, marque du pluriel. Ainsi, en plus de ce qui a été présenté concernant <e> en fin de noms (cf commentaires de la figure 3), la part des adjectifs épïcènes (c'est-à-dire bi-genre) se terminant par <e> est très forte, principalement en raison du partage de certains suffixes: <able> (*admirable, confortable...*), <ible> (*pénible, sensible...*), <aire>, (*ordinaire, polaire...*), <ique> (*magique, magnifique...*) auxquels s'ajoutent de nombreux adjectifs non suffixés (*drôle, utile...*).

Les résultats signalent toutefois que des élèves du CE2 au CM2 oublient moins souvent d'écrire le <e> dans des mots féminins qui se terminent par cette marque morphologique (*fanfare*) que dans des mots masculins qui ont la même lettre finale (*cigare*). Dans le même sens, les élèves ajoutent significativement plus de <e> marque du féminin dans des mots comme *chaleur*, qui ne se terminent pas par cette marque morphologique, en comparaison avec des mots masculins comme *facteur*. En revanche, et comme attendu, ils omettent moins souvent le <d> de *bavard* que celui de *foulard* et ajoutent plus souvent un <d> à *cauchemar* en comparaison avec *racontar* (Pacton, Nys, Fayol & Peereman, 2022).

Ce sont toutefois les accords dans le groupe verbal qui posent le plus de problèmes : même les adultes commettent des erreurs de proximité (*Je vous disez*) et, même à l'oral, les enfants produisent des erreurs de généralisation (ils conjuguent le verbe *prendre* comme les nombreux autres verbes en <dre> tels que *rendre, vendre...* et ils disent *j'ai rendu*, comme *j'ai rendu* ou *j'ai vendu*, le participe passé en <u> étant la forme la plus fréquente pour ces verbes).

En outre, les études incluant des comparaisons entre différentes modalités d'enseignement ont mis en relief le fait que l'acquisition des marques du nombre par les enfants du CP au CE2 s'effectue de manière plus précoce et plus rapide quand les règles sont introduites explicitement et avec des exercices (Fayol, Grimaud & Jacquier, 2013). Par rapport à la simple exposition, ce type de pratique entraîne très rapidement des modifications profondes des performances : le marquage du pluriel apparaît alors dès la première année et la fréquence des accords corrects augmente pour toutes les catégories grammaticales.

b. Données sur l'accord du participe passé

De nombreuses études ont porté sur les élèves du CP à la classe de 6^{ème}. Dans plusieurs de ces études (Fayol, 2020 ; Brissaud & Fayol, 2018 ; Fayol & Pacton, 2006) des élèves (du CE1 à la 6^{ème}) devaient compléter des énoncés en sélectionnant une des quatre flexions possibles (-é, -ée, -és, -ées) dans des phrases dans lesquelles l'item test est soit épithète (*Les garçons coiffés par leur sœur sourient.*), soit attribut (*Les garçons sont coiffés par leur sœur et sourient.*), ou conjugué avec l'auxiliaire avoir (*Les garçons ont coiffé leur sœur et sourient.*). La proportion d'accords corrects ne diffère pas entre le CE1 et le CE2 (environ 50%), alors qu'elle est de plus en plus élevée du CM1 au CM2 et de ces deux classes à la 6^{ème}. Quand le participe est avec l'auxiliaire avoir, les performances restent très faibles. C'est ce qui a été relevé dans une étude de la DEPP (2022) avec des élèves de CM2 qui ont eu à écrire des phrases avec l'auxiliaire être (*Les gamins se sont certainement perdus*) et avoir (*les enfants..., elle les a peut-être vus*). Le pourcentage d'enfants répondant correctement, qui est faible dans le premier cas (36%), l'est encore plus dans le second (17%).

Qu'en est-il après le primaire ? Selon Chevrot et al. (2003, voir aussi Brissaud & Chevrot 2000) l'accord du participe passé avec le sujet représente la plupart (60%) des erreurs commises par les collégiens. Brissaud & Cogis (2008) ont relevé un accord réussi du participe avec l'objet dans seulement 17% des copies d'élèves de 3^e (*la sève que les racines avaient réunie dans le tronc*). D'après Fayol et Pacton (2006), à la différence de celui avec être, l'emploi du participe passé avec avoir déroge à la règle usuelle: il ne se réalise pas avec le sujet mais avec le complément, et seulement quand celui-ci est antéposé. Comme le signalent Brissaud & Cogis (2008) « cet accord pose des difficultés d'acquisition telles qu'il est absurde de prétendre l'exiger en fin de scolarité obligatoire¹⁴ ».

14 Bena (2004) a compté que près de 80 heures étaient consacrées à l'enseignement de l'accord du participe passé en Belgique (primaire et secondaire). Nous n'avons pas le chiffre pour la France, mais on peut suggérer que ces heures seront mieux utilisées à autre chose, notamment la compréhension ou la production de textes.

Qu'en est-il chez des adultes à l'oral ? Les résultats qui précèdent sont confortés par ceux issus d'études conduites avec des adultes à l'oral. Ainsi, dans le corpus de français parlé enregistré à Aix-en-Provence (Corpaix), Audibert-Gibier (1992) a mesuré le taux d'accord au féminin audible avec avoir et l'objet antéposé, pour les participes permettant de le faire entendre (300 cas). Avec le pronom personnel de 3^e personne, le taux d'accord correct était de 64% pour le participe en position finale (contre 33% en position non finale). En revanche, avec le relatif (que), ce taux baissait fortement (34% quand le participe était en position finale, contre 24% dans les autres cas). L'usage limiterait donc l'accord au pronom personnel de 3^e personne (comme en italien moderne), et suivrait la règle de Bouhours-Vaugelas (n'accorder que le participe final) plutôt que celles de l'enseignement scolaire.

Plus récemment, Gaucher (2015) a étudié l'accord du participe passé dans 4 grands corpus oraux (CFPP200 pour Paris, C-Oral-Rom et PFC pour la France et Valibel pour la Belgique) enregistrés entre 1980 et 2010, totalisant plus de 500 heures de parole et plus de 5 millions de mots. Il observe un taux d'accord au féminin allant de 93% pour le participe épithète à 90% pour le participe avec être, 80% pour les verbes pronominaux, contre 50% seulement pour le participe avec avoir, avec de nombreux cas de non-accord (*j'avais arrêté mes études et puis je les ai repris* après, issu du Corpus PFC ; *c'est une jupe que j'ai fait* euh l'année passée, Corpus Valibel, 1991). Comme Audibert-Gibier (1992), Gaucher (2015) observe aussi un meilleur taux d'accord quand le participe est en position finale que lorsqu'il ne l'est pas (60 % contre 36%).

Qu'en est-il chez des adultes à l'écrit ? Le Levier (2022) a examiné 178 dictées d'étudiants postbac (Section STS). Alors que l'accord du participe épithète est correct dans 75% des cas, le taux de réussite avec avoir n'est que de 57%. L'accord avec le sujet pour les verbes pronominaux ainsi que celui avec l'objet dans la subordonnée relative étant encore plus faible (30%). Ces résultats ont été obtenus alors que les étudiants étaient capables de réciter correctement les règles d'accord. L'autrice conclut que « la règle d'accord du participe passé avec un complément antéposé est devenue lettre morte pour une partie des élèves français » (en fait pour 70% d'entre eux).

Benzitoun & Flesch (2024) ont étudié l'accord du participe dans les relatives en que dans un corpus d'écrits spontanés (forums Reddit France) de septembre 2021 à septembre 2022. Sur un ensemble de 395 adultes (66 femmes et 329 hommes), d'âge moyen 30 ans, étudiants ou travaillant après des études supérieures (bac + 3 ou bac + 5), seulement 30 ne font aucune erreur. Le taux d'erreur global est de 39% pour 2972 participes passés. Les erreurs concernant l'accord du masculin singulier (accord avec le sujet) sont rares (exemple : *un navire de croisière qu'ils ont rachetés*) alors que celles concernant l'accord au féminin ou au pluriel (avec l'objet) sont fréquentes (exemples : *Des trucs qu'il a vu* sur YouTube ; *tu dois obéir aux lois que j'ai inventé*), en particulier quand masculin et féminin sont homophones, ce qui est le cas de 99% des verbes.

Afin de tester si le non-accord est encore perçu aujourd'hui comme erroné, Hemforth, Siqi-Chen & Lui (2023) ont fait passer en ligne une expérience de jugement d'acceptabilité sur une échelle de 1 (pas du tout acceptable) à 5 (parfaitement acceptable) à 47 adultes d'âge moyen 40 ans (âge médian 33 ans), monolingues, ayant toujours vécu en France. Vingt items (mêlés à des distracteurs) leur ont été proposés dans quatre conditions : avec ou sans accord et avec une reprise du complément soit par les, soit par que (par exemple : *Les bouteilles, Dilan les a perdu* ; *Les bouteilles, Dilan les a perdues* ; *Ce sont les bouteilles que Dilan a perdu* ; *Ce sont les bouteilles que Dilan a perdues*). Les participes étaient homophones au masculin et féminin, ce qui est le cas le plus fréquent. Comme l'indique la figure 6, le non-accord a fréquemment été jugé comme acceptable (note moyenne supérieure à 3/5), en particulier pour les relatives en que. Ce résultat est d'autant plus frappant que, vu le mode de recrutement (via la plateforme du Réseau d'information des sciences cognitives), le niveau de littéracie de ces participants est probablement supérieur à celui de la moyenne nationale.

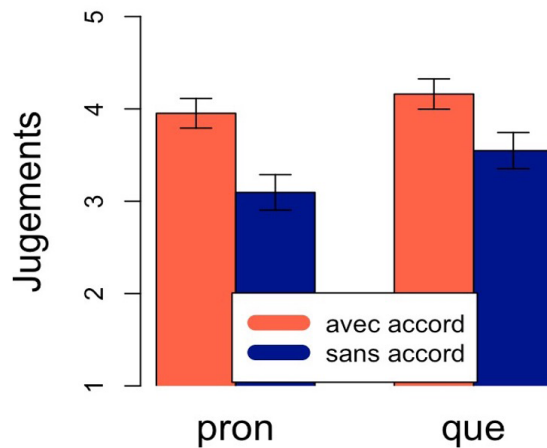


Figure 6. Jugement d’acceptabilité sur l’accord du participe passé (Hemfort et al., 2023)

Enfin, pour les verbes pronominaux, Barque et Candito (2019) ont montré que dans de grands corpus écrits, 99% des emplois sont des verbes pronominaux intrinsèques (*s’apercevoir*), médio-passifs¹⁵ (*se vendre*) ou réfléchis intransitifs (*se laver*), c’est-à-dire des cas d’accord avec le sujet ; les cas d’invariabilité (*Elle s’est dit des choses*) ou d’accord avec l’objet (*les choses qu’elle s’est dites*) représentant moins de 1% des cas.

Pour conclure sur l’accord du participe passé, on peut signaler que Béguelin (2022) parle d’une *simple servitude grammaticale moribonde* et *que*, selon Tseng (2021), en contexte informel, le non-accord n’est pas perçu comme une erreur ; en fait on ne le remarque pas du tout¹⁶.

En résumé

Concernant les marques du genre et du nombre, d’une part, leur compréhension est plus précoce que leur production. D’autre part, la maîtrise du pluriel des noms (<s>) est plus précoce et mieux assurée que celle de la même flexion appliquée à l’adjectif. De plus, la flexion du genre est plus problématique que celle du nombre, le <e>, étant moins consistant que le <s>. Ce sont toutefois les accords dans le groupe verbal qui posent le plus de problèmes. Concernant l’accord du participe passé avec *être* et *avoir*, les erreurs les plus fréquentes portent sur l’auxiliaire *avoir*. A l’école élémentaire, comme au collège, ou même postbac, les francophones sont peu sensibles, à l’oral comme à l’écrit, à cet accord qui va à l’encontre de la règle générale : accord sujet-verbe. Il serait donc bien de s’en tenir à deux règles simples : invariabilité avec *avoir*, accord systématique avec le sujet avec *être*. Le premier point demanderait un enseignement explicite, mais pas le second puisqu’il tombe sous la règle commune pour les accords verbaux¹⁷.

¹⁵ Forme verbale entre la voix passive et active (Les livres se sont bien vendus.)

¹⁶ In informal contexts, missing agreement with *avoir* is not perceived as an error; in fact it usually goes completely unnoticed

¹⁷ C’est ce que suggère le récent manuel de Poli et Monneret (2023): s’en tenir aux « cas les plus usuels » pour le participe passé (invariabilité et accord avec le sujet) jusqu’en 6^{ème}.

3. Que faudrait-il faire pour mieux enseigner l'orthographe ?

Les conclusions et les recommandations qui suivent s'appuient sur les résultats les plus notables des travaux de recherche, travaux qui ont mis en relief des pistes susceptibles d'améliorer l'enseignement de l'orthographe du français.

Comme pour l'apprentissage de la lecture, c'est principalement la régularité des mots au niveau des correspondances phonème-graphème qui explique la réussite de l'apprentissage de l'orthographe. Il faut donc utiliser surtout des mots réguliers à ce niveau au début de cet apprentissage.

Au-delà des correspondances phonème-graphème, **il faut tenir compte des différences entre code oral et code écrit s'expliquant principalement par la morphologie, tant grammaticale que lexicale.** Ainsi, de très **nombreuses marques de morphologie grammaticale** (en particulier les celles du genre et du nombre pour les noms et les adjectifs, et celles de la personne et du temps pour les verbes) **et lexicale** (le <t> d'*éclat*) **s'écrivent mais ne s'entendent pas** (par exemple : *les mouche*s *volent*, *elle est jolie*, *grand*-père). En outre, **parmi celles qui se prononcent, on trouve de nombreux homophones** (*chanté*, *chanter*, *chantez*...). **La morphologie du français devrait être l'objet d'un enseignement précoce et explicite focalisé sur les différences entre l'oral et l'écrit.**

Il faut également **savoir ce que les élèves des différents niveaux de l'école élémentaire s'avèrent capables d'orthographier correctement.** Il est en effet inutile de s'acharner à enseigner des orthographes maîtrisées par une minorité de francophones. C'est entre autres le cas **de l'accord du participe passé avec l'auxiliaire avoir et le complément antéposé. Cette règle**, qui va à l'encontre de la règle générale selon laquelle le verbe s'accorde avec le sujet, n'est maîtrisée que par peu d'enfants en fin de primaire (moins de 20%) et elle tombe en désuétude chez les adultes.

En conséquence, **il faut avoir des données sur la régularité des relations phonème-graphème en tenant compte de la morphologie grammaticale et lexicale. Il faut également avoir des données sur ce que la majorité des enfants s'avèrent capables de maîtriser au moins au milieu et à la fin du CP, puis à la fin des autres niveaux du primaire.**

Une meilleure utilisation des statistiques devrait permettre de **construire une progression pédagogique efficace pour enseigner l'orthographe. Nous recommandons donc que les statistiques fournies par les bases de données récentes** (par exemple, *Lexique*, *Manulex* ou *EQOL*) soient intégrées dans la **formation des enseignants.**

Nous proposons en plus une mesure simple qui reprend, en les généralisant, les arrêtés ministériels de 1901 et 1976 (toujours en vigueur) à propos de l'accord du participe passé : avec l'auxiliaire *avoir*, invariabilité du participe passé ; avec l'auxiliaire *être*, accord systématique avec le sujet. Ces deux choix libéreraient énormément de temps en classe, qui pourrait être consacré à des enseignements portant plus généralement sur la production écrite.

Il conviendrait également de mettre à disposition des enseignants et des élèves des manuels et des livres qui suivent au mieux l'orthographe rectifiée de 1990 aussi bien au niveau élémentaire qu'au collège ou au lycée. **Il faudrait encourager les éditeurs à éditer leurs publications en orthographe rectifiée.**

L'orthographe du français comprend toujours de nombreuses incohérences qui n'ont pas été corrigées par l'Académie, notamment les doubles consonnes et les pluriels en <x>. Comme l'a rappelé Maurice Druon, secrétaire perpétuel de l'Académie en présentant les Rectifications de 1990 au premier Ministre d'alors, Michel Rocard, **la langue évolue** et, disait-il, « il faudra recommencer le travail, dans trente ans, sinon même avant ».

34 ans ont passé. Il serait temps que soit créée **une nouvelle commission regroupant toutes les autorités et compétences en matière d'orthographe**, en France et hors de France, afin d'avancer avec une sage fermeté dans **la poursuite des régularisations de l'orthographe française**, tant lexicale que grammaticale, en tenant compte des avancées de la recherche.

Ce qu'il faut retenir

— Principales origines des différences entre l'oral et l'écrit en français.

Plus que les autres langues romanes, le français possède des phonèmes spécifiques, en particulier entre 10 et 16 voyelles qui ne pouvaient pas être transcrites par les 6 voyelles de l'alphabet latin. Il a fallu créer de nouveaux graphèmes (des lettres avec un accent ou des suites de lettres) pour différencier <e> de <é> et <è> (cf. élève), <u> de <ou> (cf. cette russe est rousse) ainsi que les voyelles orales (<a>, <o>, <i> et <u>) des voyelles nasales (<an>, <on>, <in> et <un>). A cet inventaire, s'ajoutent les allographes (<en> et <ant> dans *enfant*). Quant aux différences dans le décompte des voyelles, elles proviennent de variations régionales : certains francophones ne distinguant pas le <a> *patte* de celui de *pâte* alors que d'autres ne perçoivent pas la différence entre *brin* et *brun*.

La prononciation de certains phonèmes spécifiques du français a fortement évolué. Un exemple marquant est celui des voyelles orales qui se sont transformées en voyelles nasales avant de retrouver leur oralité tout en gardant, pour certaines, un double <n> à l'écrit (cf. *sonner* versus *sonate*).

Le surmarquage de la morphologie grammaticale à l'écrit. La plupart des marques de flexion qui s'écrivent ne s'entendent plus (*les bébés pleurent*), sauf dans les liaisons (les [z]amis).

Les décisions de l'Académie. Le choix initial était celui d'une orthographe savante dans laquelle des lettres étymologiques ont été ajoutées (le <s> et le <p> dans *écriture*) ou mises à la place d'autres (<y> pour noter <i> dans *luy* et *joly*, cf. les extraits du *Corbeau* et du *Renard* reproduits ci-dessus) D'un autre côté, l'Académie a effectué de nombreux ajustements au cours des siècles : environ un mot sur deux a changé d'orthographe depuis la première édition de son dictionnaire (cf. *écriture*, *savoir*, *lui*, *joli*, *asile*). De même, après avoir bloqué la réforme prévue par le ministre Leygues en 1900, elle a approuvé les Rectifications de 1990, mais ne promeut pas vraiment leur application.

— Apprendre à écrire est bien plus difficile qu'apprendre à lire

Concernant l'orthographe lexicale, **l'impact de l'irrégularité est plus fort en écriture de mots isolés qu'en lecture.** Ainsi une étude dans laquelle les mêmes élèves ont été vus à la fin des quatre premiers grades de l'école élémentaire a mis en relief le fait que des mots irréguliers (comme *pays*, *femme*, *pied*) sont systématiquement moins bien écrits que lus (par exemple, 90 contre environ 60% d'erreurs en fin de CP, 60 versus 30% en fin de CE1 et environ 30% contre 10% en fin de CM1).

L'existence de la norme orthographique : par exemple, le <o> final d'un mot tel que *mécano* peut s'écrire de plusieurs manières (les alternatives les plus fréquentes étant <eau> et <ot>, comme dans *beau* et *mot*) une seule étant correcte.

Le poids de la morphologie en écriture. D'une part, de **nombreuses marques morphologiques qui ne s'entendent pas doivent s'écrire** (les *chats miaulent*). D'autre part, la présence d'**homophones non homographes, en particulier pour les flexions verbales** (chant*er*, chant*é*, chant*ez*, chant*ait*, chant*aient*...).

— Incidence de l'opacité de l'orthographe dans le développement typique et la dyslexie

Dans le **développement typique**, c'est **l'automatisation progressive de la reconnaissance des mots écrits** (qui dépend de la transparence de l'orthographe) qui permet de comprendre ce qu'on lit.

La dyslexie se caractérise par des **difficultés de reconnaissance précise et/ou rapide des mots écrits** et par de **faibles capacités d'orthographe**. Ces difficultés qui sont aggravées par l'opacité de l'orthographe, résultent généralement d'un déficit phonologique, et sont à **l'origine des difficultés de compréhension écrite** observées chez la plupart des dyslexiques.

— L'utilisation des rectifications de l'orthographe de 1990 (tout au moins de certaines d'entre elles, cf. Annexe 3) permet une amélioration des performances des élèves.

Annexes

Annexe 1. Manulex-MorphO (Peereman et al., 2013)

Cette base de données fournit les valeurs de fréquence et de consistance des relations Graphème-Phonème (G-Ph) et Phonème-Graphème (Ph-G) d'environ 10000 mots issus de *Manulex*, base qui répertorie les mots issus de livres scolaires du CP au CM2 (Lété et al., 2004). Cette sélection correspond à environ 20% des entrées lexicales de *Manulex*, et à plus de 90% des entrées textuelles. Les associations G-Ph et Ph-G sont analysées distinctement selon qu'elles correspondent ou non à des marques morphologiques souvent silencieuses à l'oral : flexions nominales (<e>, <s>, <x>) et verbales (<ez>, <ent>...); dérivations (<d> de *bond*...). L'analyse de ces marques a été effectuée au niveau des relations G-Ph et Ph-G dans 4 cas : flexions nominales, flexions verbales (<er> mais pas 'ir' ou 'oir' qui ont plus d'un phonème), consonnes finales qui peuvent être silencieuses dans le mot racine (le <d> de *grand* versus *grande* et *grandeur*) et la finale <ent> des adverbes de manière (<ment> : *vraiment*...).

Pour chaque mot, *Manulex-morpho* fournit entre autres sa fréquence ainsi que la fréquence et la consistance des associations G-Ph et Ph-G, dans l'ensemble et en fonction de leur position (initiale, finale et intermédiaire). Ces estimations sont réalisées par type (fréquence lexicale) et par token (fréquence textuelle) pour les associations G-Ph et Ph-G. On peut illustrer la différence entre ces deux décomptes par l'exemple du graphème <ch> qui apparaît dans 562 mots français, 9 avec <ch> prononcé /k/ (*chorale*) et 553 avec <ch> prononcé // (*chat*). La consistance G-Ph, exprimée en pourcentage, est le rapport entre la fréquence propre de la correspondance et la fréquence globale du graphème. Ainsi, pour l'association <ch> - /k/, la consistance lexicale est de 1,60% ($9 / [553 + 9] \times 100 = 1,60\%$). La consistance textuelle est obtenue par le même calcul mais à partir de la fréquence textuelle : la fréquence totale des 9 mots avec <ch> prononcé /k/ étant de 184 et celle des 553 mots avec <ch> prononcé //, de 22.388, la consistance textuelle de l'association <ch> - /k/ est de 0,82% ($184 / [22.388 + 184] \times 100 = 0,82\%$).

Le tableau présenté dans la page suivante indique les principales rubriques correspondant aux colonnes A à AI de la base Manulex-Morpho. Il permet de constater que la fréquence des mots (colonne I) varie fortement, certains étant très fréquents (> 500 / 1 million) d'autres peu fréquents. Ce tableau permet surtout de constater les similarités et/ou de différences entre les relations G-Ph et Ph-G. Ainsi l'adjectif bon est très consistant aussi bien pour la lecture que pour l'écriture (100% par type et par token). En revanche, le mot *cinq* est plus facile à lire qu'à écrire (consistance par type et token : 70 et 71 pour les correspondances G-Ph et 21 et 22 pour celles Ph-G). C'est le cas pour de nombreux autres items. Une autre information d'intérêt est dans les dernières colonnes (AB à AI), qui indiquent, pour chaque mot, la relation G-Ph et Ph-G la moins fréquente et la moins consistante : par exemple, le <e> de *femme* et le <x> *dix* et *six*. Ces informations permettent également de comprendre pourquoi *chat* est plus facile à lire et à écrire que *loup*.

Téléchargements : <https://inframorph.github.io/index.html>.

Pour les codes phonétiques et morphologiques :

<https://inframorph.github.io/morpho/phonetique.html>

et <https://inframorph.github.io/morpho/morphologie.html>

Annexe 1 (suite) : Exemples d'entrées de Manulex MorphO, <https://inframorph.github.io/index.html>

			Fréq /	Segmentation G-Ph			Segmentation Ph-G		
Ortho	Phono	Cat	1 million	Graphème	Phonème	Match	Phonème	Graphème	Match
abri	abRi	NC	65	.a.b.r.i	.a.b.R.i	(a-a.b-b.r-R.i-i)	.a.b.R.i	.a.b.r.i	(a-a.b-b.r-R.i-i)
bon	bʃ	ADJ	492	.b.6on	.b.6ʃ	(b-b.6on-6ʃ)	.b.6ʃ	.b.6on	(b-b.6on-6ʃ)
bons	bʃ	ADJ	53	.b.+6on.3s	.b.6ʃ.3#		.b.6ʃ.3#	.b.+6on.3s	(
chat	Sa	NC	847	.ch.a.6t	.S.a.6#	(ch-S.a-a.6t-6#)	.S.a.6#	.ch.a.6t	(ch-S.a-a.6t-6#)
chou	Su	NC	37	.ch.ou	.S.u	(ch-S.ou-u)	.S.u	.ch.ou	(ch-S.ou-u)
choux	Su	NC	25	.ch.+ou.3x	.S.u.3#		.S.u.3#	.ch.+ou.3x	
cinq	s5k	ADJ	319	.c.in.6q	.s.5.6k	(c-s.in-5.6q-6k)	.s.5.6k	.c.in.6q	(c-s.in-5.6q-6k)
dix	dis	ADJ	226	.d.i.6x	.d.i.6s	(d-d.i-i.6x-6s)	.d.i.6s	.d.i.6x	(d-d.i-i.6x-6s)
femme	fa	NC	261	.f.e[CC]. mm.e	.f.a.m.°		.f.a.m	.f.e.mme	(f-f.e-a.mme-m)
feu	f2	NC	464	.f.eu	.f.2	(f-f.eu-2)	.f.2	.f.eu	(f-f.eu-2)
feux	f2	NC	42	.f.+eu.3x	.f.2.3#	(f-f.eu-2).3x-3#)	.f.2.3#	.f.+eu.3x	(f-f.eu-2).3x-3#)
loup	lu	NC	544	.l.ou.p	.l.u.#	(l-l.ou-u.p-#)	.l.u	.l.oup	(l-l.oup-u)
lune	lyn	NC	240	.l.u.n.e	.l.y.n.°	(l-l.u-y.n-n.e-°)	.l.y.n	.l.u.ne	(l-l.u-y.ne-n)
mardi	maRdi	NC	70	.m.a.r.d.i	.m.a.R.d.i		.m.a.R.d.i	.m.a.r.d.i	
midi	midi	NC	91	.m.i.d.i	.m.i.d.i	(m-m.i-i.d-d-i-i)	.m.i.d.i	.m.i.d.i	(m-m.i-i.d-d-i-i)
mot	mo	NC	1340	.m.o.t	.m.o.#	(m-m.o-o.t-#)	.m.o	.m.ot	(m-m.ot-o)
mots	mo	NC	2319	.m.o.+t.3s	.m.o.#.3#		.m.o.3#	.m.+ot.3s	
mur	myR	NC	199	.m.u.6r	.m.y.6R	(m-m.u-y.6r-6R)	.m.y.6R	.m.u.6r	(m-m.u-y.6r-6R)
nage	naZ	NC	22	.n.a.g.e	.n.a.Z.°	(n-n.a-a.g-Z.e-°)	.n.a.Z	.n.a.ge	(n-n.a-a.ge-Z)
nez	ne	NC	321	.n.ez	.n.e	(n-n.ez-e)	.n.e	.n.ez	(n-n.ez-e)
niche	niS	NC	36	.n.i.ch.e	.n.i.S.°	(n-n.i-i.ch-S.e-°)	.n.i.S	.n.i.che	(n-n.i-i.che-S)
onze	ʃz	ADJ	51	.on.z.e	.ʃ.z.°	(on-ʃ.z-z.e-°)	.ʃ.z	.on.ze	(on-ʃ.z-z.e-z)
oui	wi	ADV	631	.ou.i	.w.i	(ou-w.i-i)	.w.i	.ou.i	(ou-w.i-i)
rire	RiR	NC	73	.r.i.r.e	.R.i.R.°	(r-R.i-i.r-R.e-°)	.R.i.R	.r.i.re	(r-R.i-i.re-R)
rire	RiR	VER	180	.r.i.r.e	.R.i.R.°	(r-R.i-i.r-R.e-°)	.R.i.R	.r.i.re	(r-R.i-i.re-R)
sept	sEt	ADJ	157	.s.e.p.6t	.s.E.#.6t		.s.E.#.6t	.s.e.p.6t	(
six	sis	ADJ	190	.s.i.6x	.s.i.6s	(s-s.i-i.6x-6s)	.s.i.6s	.s.i.6x	(s-s.i-i.6x-6s)
trois	tRwa	ADJ	984	.t.r.oi.6s	.t.R.wa.6#			.t.r.oi.6s	
un	1	DET	18065	.6un	.61	(6un-61)	.61	.6un	(6un-61)

Consistance				Fréquence de l'unité				Consistance de l'unité			
G-Ph		Ph-G		la moins fréquente				la moins consistante			
Type/Tok		Typ /Tok		G-Ph		Ph-G		G-Ph		Ph-G	
99	99	84	85	71	i-i	71	i-i	96	b-b	43	i-i
100	100	100	100	227	6on-6§	227	6on-6§	100	b-b	100	b-b
100	100	100	100	227	6on-6§	227	6on-6§	100	b-b	100	b-b
98	98	99	99	211	ch-S	211	ch-S	96	6t-6#	97	a-a
99	99	79	74	34	ou-u	34	ou-u	99	ch-S	59	ou-u
100	100	86	83	34	ou-u	34	ou-u	99	ch-S	59	ou-u
70	71	21	21	2	6q-6k	2	6q-6k	11	c-s	8	6q-6k
58	59	68	70	2	6x-6s	2	6x-6s	3	6x-6s	8	6x-6s
76	76	38	39	10	e[CC]-a	10	e-a	2	e[CC]-a	0	e-a
100	100	81	83	25	eu-2	25	eu-2	100	f-f	66	eu-2
100	100	87	88	25	eu-2	25	eu-2	100	f-f	66	eu-2
88	91	53	54	5	p-#	3	oup-u	71	p-#	5	oup-u
94	94	94	94	270	l-l	131	ne-n	76	u-y	86	ne-n
100	100	87	88	71	i-i	71	i-i	100	d-d	43	i-i
93	93	85	86	71	i-i	71	i-i	73	i-i	43	i-i
68	69	54	54	130	t-#	15	ot-o	10	o-o	9	ot-o
76	77	70	70	130	t-#	15	ot-o	10	o-o	9	ot-o
92	92	99	99	57	6r-6R	104	6r-6R	76	u-y	96	u-y
91	92	99	99	162	n-n	148	ge-Z	66	g-Z	97	a-a
100	100	51	51	3	ez-e	3	ez-e	100	n-n	1	ez-e
93	93	99	99	162	n-n	85	che-S	73	i-i	98	i-i
100	100	44	45	6	on-§	6	on-§	100	on-§	13	ze-z
56	56	63	69	5	ou-w	5	ou-w	13	ou-w	43	i-i
93	93	80	80	779	r-R	264	re-R	73	i-i	43	re-R
93	93	80	80	779	r-R	264	re-R	73	i-i	43	re-R
36	36	58	58	13	6t-6t	15	p-#	2	p-#	0	p-#
58	59	63	65	2	6x-6s	2	6x-6s	3	6x-6s	8	6x-6s
97	97	81	82	62	6s-6#	30	6s-6#z	90	6s-6#	33	6s-6#z
100	100	100	100	4	6un-6l	4	6un-6l	100	6un-6l	100	6un-6l

Annexe 2. Mots correctement écrits fin CP par >75% des élèves (EOLE, 2003)

A - B	Cent**	Débuter	Foudre	Mirage	Pape	Redire	Suite
Acheminer	Chacune	Déchu	Franche	Mire	Parade	Refouler	Tabou
Adulte	Chantre	Délire	Franco	Miser	Parasite	Relative	Tacher
Alarme	Charbon	Démener	Friche	Mistral	Pardon	Relier	Taliban
Animal	Charlatan	Démonter	Fuite	Mode	Parfumer	Relire	Talon
Anormal	Chat**	Dénouer	G - N	Monde	Patron	Retenu	Tare
Armature	Chemin	Dictée**	Galopin	Mordre	Paver	Revenu	Tarte
Arme	Chicaner	Diluer	Garçon	Moto	Peluche	Riche	Tartine
Avaler	Chou	Dindon	Garer	Mou	Pépé	Ride	Tatami
Avec	Chouchou	Dodo	Gastéropode	Mouche	Piano	Rider	Timide
Avertir	Chute	Dominer	Gronder	Moucher	Pie**	Riposter	Tipi
Avidité	Chuter	Douche	Guitare	Mouler	Pile	Rire	Tirer
Aviron	Cinq	Duo	Ici	Mouton	Piler	Rite	Titre
Baba	Coca	Eau	Isoler	Mural	Pipe	Rive	Tolérable
Balise	Cocher	Ecrire	Joli	Muse	Pire	Robe	Tourbe
Ballon*	Cochon	Egal	Jupe	Mutualiste	Pitre	Roche	Trapu
Bar	Coco	Elu	Juste	Nage	Place	Ronde	Trente
Baron	Colmater	Emaner	Lame	Nature	Platane	Rouble	Tribu
Beau	Colonie**	Etabli	Louche	Navire	Plate	Rouge	Triple
Bébé	Conforme	Eté	Loulou	Néfasté	Poil	Rouler	Trois**
Biberon	Contraste	Eviter	Loup**	Néon	Pointure	Route	Troupe
Biche	Copine	Facture	Luire	Nez*	Poison	Routine	Tube
Bilan	Coq	Farine	Lune	Niche	Poivre	Ruche	Tulipe
Bio	Corde	Fatal	Mage	Noire	Ponte	Rude	Ultra
Bobo	Cordon	Fatalité	Maison	Nomade	Poste	Ruse	Uni
Bon	Corne	Faux**	Malade	Noter	Pou	S - Z	Unifier
Bonbon	Cote	Fétiche	Malin	Nuage	Poule	Sacoche	Vache
Boule	Cou	Feu	Maman	Nage	Profane	Sacre	Valable
Bousculer	Coucou	Fiable	Manipuler	O - P	Pur	Saluer	Vénéérable
Bouton	Coudre	Ficher	Marche	Octobre	Q - R	Samedi	Vent**
Boxer	Couler	Figure	Marcher	Onde	Quinze	Sapin	Verdure
Bravo	Couture	Filer	Mardi	Onduler	Raconter	Satin	Vidéo
Brutal	Cravate	Fille*	Marin	Onze	Rafle	Sept*	Vider
C	Cru	Filtrer	Mars	Orange	Rame	Sinon	Viril
Calme	Cuisine	Flacon	Martre	Orbite	Ranimer	Sirène	Vite
Camion	Cumul	Flore	Mécano	Ordre	Rapide	Six*	Vitrine
Cancan	D - F	Flou	Midi	Orme	Rapidité	Sonate	Vive
Canine	Darne	Foire	Mimer	Oui	Rare	Songe	Voile
Carton	Date	Fois**	Mine	Pacte	Rater	Sorte	Voiture
Cartouche	Dater	Fondé	Minute	Panache	Réalisme	Soutenir	Voter
Cave	Débouler	Fondre	Mirador	Papa	Rédiger	Star	Zéro

* Mots contenant une relation phonème graphème exceptionnelle (fille, nez, sept, six) ou rare (ballon)

** Mots avec une lettre finale muette, consonne (chat, cent, faux, fois, loup, trois, vent) ou voyelle (colonie, dictée, pie)

Annexe 3. Élèves (%) écrivant correctement les mots de l'orthographe rectifiée (EOLE, 2016)

Fond gris foncé, supérieur ou égal à 75%			Fond gris clair, supérieur ou égal à 50%				
			CP	CE1	CE2	CM1	CM2
OR 5		a fortiori	0%	2%	3%	2%	2%
		à fortiori	9%	1%	12%	4%	1%
OR 5		a postérieur	0%	0%	5%	0%	4%
		à posteriori	9%	4%	1%	0%	1%
OR 5		a priori (adverbe)	1%	7%	3%	8%	3%
		à priori	0%	0%	0%	14%	14%
OR 5	+	apriori (nom)	21%	31%	29%	42%	24%
		a priori	0%	0%	0%	14%	14%
OR 1	+	abime	76%	76%	66%	74%	69%
		abîme	12%	0%	7%	15%	36%
OR 1	+	abimer	73%	87%	83%	79%	77%
		abîmer	6%	18%	29%	15%	42%
OR 1		accroitre	0%	3%	2%	8%	17%
		accroître	0%	0%	2%	13%	17%
OR 1	+	affut	5%	7%	15%	27%	51%
		affût	0%	7%	7%	33%	14%
OR 2		aigüe	0%	0%	3%	13%	4%
		aiguë	0%	0%	0%	0%	0%
OR 1	+	ainé	7%	11%	34%	59%	53%
		aîné	0%	0%	18%	21%	22%
OR 2		ambigüe	0%	0%	0%	0%	3%
		ambiguë	0%	0%	0%	0%	21%
OR 2		ambiguïté	0%	0%	0%	0%	3%
		ambigüité	0%	0%	0%	0%	3%
OR 1		aout	18%	45%	34%	76%	81%
		août	8%	38%	73%	73%	83%
OR 1	+	apparaitre	10%	4%	27%	46%	63%
		apparaître	0%	0%	21%	22%	51%
OR 2	+	argüer	39%	54%	54%	76%	78%
		arguer	19%	40%	52%	71%	73%
OR 4	+	assoïr	9%	24%	44%	52%	53%
		asseoir	0%	8%	15%	27%	48%
OR 1	+	boite	58%	67%	81%	92%	86%
		boîte	59%	51%	50%	76%	82%
OR 1	+	boitier	19%	59%	72%	80%	91%
		boîtier	3%	8%	29%	69%	66%

Fond gris foncé, supérieur ou égal à 75%			Fond gris clair, supérieur ou égal à 50%				
			CP	CE1	CE2	CM1	CM2
OR 1	+	brulant	7%	28%	33%	53%	60%
		brûlant	0%	11%	18%	51%	48%
OR 1	+	bruler	47%	68%	80%	94%	93%
		brûler	0%	7%	42%	42%	96%
OR 1	+	brulure	42%	69%	68%	71%	73%
		brûlure	0%	15%	49%	31%	50%
OR 1	+	bucher	31%	52%	74%	76%	72%
		bûcher	10%	6%	34%	28%	45%
OR 1	+	bucheron	41%	62%	75%	79%	90%
		bûcheron	20%	13%	31%	26%	58%
OR 3		cèleri	3%	1%	6%	6%	6%
		céliéri	0%	11%	7%	7%	9%
OR 1	+	chaîne	12%	31%	54%	72%	72%
		chaïne	12%	11%	21%	70%	51%
OR 1		chainon	15%	12%	39%	33%	31%
		chaïnnon	0%	5%	3%	22%	13%
OR 4		charriot	4%	20%	31%	59%	53%
		chariot	9%	13%	46%	53%	78%
OR 1	+	comparaitre	6%	7%	21%	53%	52%
		comparaître	0%	0%	19%	12%	38%
OR 1	+	connaitre	1%	13%	49%	62%	81%
		connaître	0%	13%	47%	48%	71%
OR 4	+	corole	62%	59%	56%	58%	51%
		corolle	0%	3%	6%	37%	27%
OR 1	+	cout	19%	35%	38%	59%	65%
		coût	2%	0%	21%	24%	65%
OR 1	+	couter	53%	74%	95%	85%	96%
		coûter	7%	25%	20%	52%	74%
OR 1	+	couteuse	24%	47%	62%	76%	80%
		coûteuse	0%	0%	42%	37%	16%
OR 1	+	couteux	1%	7%	32%	58%	60%
		coûteux	0%	0%	9%	10%	15%
OR 1	+	croitre	36%	56%	47%	69%	62%
		croître	0%	0%	0%	13%	41%
OR 1	+	croute	59%	67%	66%	73%	78%
		croûte	2%	14%	28%	36%	57%
OR 1	+	crouton	42%	79%	89%	80%	89%
		croûton	0%	2%	0%	5%	11%

Fond gris foncé, supérieur ou égal à 75%			Fond gris clair, supérieur ou égal à 50%				
			CP	CE1	CE2	CM1	CM2
OR 1	+	dîner	65%	89%	91%	93%	94%
		dîner	6%	33%	51%	65%	80%
OR 1	+	disparaitre	11%	19%	50%	74%	84%
		disparaître	0%	35%	20%	28%	64%
OR 1	+	dument	5%	24%	39%	46%	45%
		dûment	0%	0%	8%	0%	1%
OR 1	+	enchainer	5%	26%	22%	66%	58%
		enchaîner	0%	06%	11%	32%	21%
OR 1	+	entrainement	4%	14%	43%	67%	68%
		entraînement	0%	0%	15%	38%	28%
OR 1	+	entraîner	13%	26%	57%	51%	73%
		entraîner	0%	0%	25%	23%	23%
OR 1	+	entraîneur	8%	27%	36%	65%	70%
		entraîneur	0%	8%	18%	30%	21%
OR 3	+	évènement	11%	34%	37%	65%	55%
		événement	19%	38%	47%	58%	50%
OR 4	+	exéma	4%	14%	20%	36%	37%
		eczéma	0%	0%	0%	0%	6%
OR 1	+	flute	47%	77%	85%	96%	98%
		flûte	0%	37%	52%	50%	53%
OR 1	+	flutiste	49%	72%	60%	68%	75%
		flûtiste	27%	0%	68%	12%	75%
OR 1	+	fraiche	17%	34%	58%	78%	76%
		fraîche	0%	0%	8%	31%	53%
OR 1	+	fraichement	4%	33%	68%	78%	93%
		fraîchement	4%	5%	47%	38%	57%
OR 1	+	fraicheur	8%	29%	49%	65%	74%
		fraîcheur	0%	42%	15%	26%	75%
OR 1	+	fut	21%	38%	44%	55%	48%
		fût	0%	2%	0%	39%	5%
OR 1	+	gite	23%	39%	50%	71%	69%
		gîte	0%	9%	16%	18%	46%
OR 1	+	gout	24%	59%	71%	71%	85%
		goût	5%	23%	52%	61%	75%
OR 1	+	gouter	67%	78%	87%	95%	86%
		goûter	9%	59%	65%	63%	75%
OR 4	+	greloter	46%	56%	71%	69%	74%
		grelotter	0%	5%	9%	3%	36%

Fond gris foncé, supérieur ou égal à 75%				Fond gris clair, supérieur ou égal à 50%			
			CP	CE1	CE2	CM1	CM2
OR 1	+	huitre	12%	47%	66%	83%	83%
		hôte	0%	0%	19%	20%	6%
OR 1	+	île	36%	67%	82%	87%	100%
		île	11%	77%	88%	83%	91%
OR 1		îlot	0%	12%	13%	34%	34%
		îlot	0%	7%	15%	32%	31%
OR 4	+	interpeler	31%	43%	63%	70%	62%
		interpeller	0%	0%	7%	23%	36%
OR 1	+	jeuner	8%	25%	44%	65%	72%
		jeûner	6%	9%	2%	11%	22%
OR 1		maitre	31%	66%	64%	84%	83%
		maître	12%	57%	89%	88%	96%
OR 1	+	maitresse	31%	43%	78%	80%	94%
		maîtresse	0%	50%	51%	72%	75%
OR 1	+	maitrise	6%	28%	36%	44%	44%
		maîtrise	0%	0%	2%	38%	42%
OR 1	+	maitriser	14%	19%	32%	51%	47%
		maîtriser	2%	1%	7%	16%	17%
OR 1	+	maraicher	4%	12%	15%	24%	17%
		maraîcher	0%	0%	0%	5%	
OR 1	+	mur	53%	75%	78%	88%	88%
		mûr	8%	27%	15%	61%	65%
OR 1		murir	33%	28%	37%	39%	29%
		mûrir	0%	2%	8%	12%	37%
OR 1	+	naitre	18%	26%	54%	80%	94%
		naître	5%	4%	28%	43%	90%
OR 4		ognon	14%	19%	34%	48%	54%
		oignon	5%	12%	29%	46%	84%
OR 1		paitre	9%	7%	18%	17%	22%
		paître	0%	0%	3%	2%	13%
OR 1	+	paraitre	8%	21%	39%	55%	77%
		paraître	2%	13%	8%	69%	30%
OR 1	+	piqûre	20%	23%	35%	48%	40%
		piqure	0%	0%	20%	24%	33%
OR 1		presqu'île	0%	0%	3%	22%	17%
		presqu'île	0%	2%	5%	6%	26%
OR 1		rafraichissant	0%	2%	21%	25%	60%
		rafraîchissant	0%	2%	9%	14%	41%

Fond gris foncé, supérieur ou égal à 75%			Fond gris clair, supérieur ou égal à 50%				
			CP	CE1	CE2	CM1	CM2
OR 4	+	rassoir	1%	12%	12%	47%	38%
		rasseoir	5%	0%	70%	14%	5%
OR 3		recèler	17%	19%	29%	31%	43%
		receler	11%	24%	24%	33%	53%
OR 1	+	reconnaitre	1%	13%	39%	40%	77%
		reconnaître	0%	4%	17%	38%	45%
OR 3	+	règlementer	14%	40%	60%	59%	78%
		réglementer	23%	44%	80%	68%	84%
OR3%	+	règlementaire	4%	14%	31%	35%	49%
		réglementaire	3%	13%	31%	40%	61%
OR 4	+	relai	23%	44%	30%	53%	53%
		relais	7%	12%	21%	13%	43%
OR 5	+	révolver	17%	30%	33%	59%	41%
		revolver	0%	4%	26%	13%	49%
OR 5	+	roqueur	5%	10%	43%	39%	82%
		rocker	2%	1%	27%	26%	31%
OR 3		sècheresse	0%	4%	14%	31%	22%
		sécheresse	0%	9%	13%	10%	38%
OR 5		speakeur	0%	0%	1%	3%	4%
		speaker	0%	0%	0%	0%	0%
OR 1		surcroît	1%	3%	9%	7%	10%
		surcroït	0%	1%	4%	18%	25%
OR 1	+	surement	8%	30%	43%	65%	41%
		sûrement	0%	0%	9%	5%	19%
OR 1	+	traîne	18%	42%	65%	75%	86%
		traîne	0%	11%	19%	14%	22%
OR 1	+	traineau	1%	32%	40%	63%	66%
		traîneau	0%	4%	17%	17%	29%
OR 1	+	trainer	11%	18%	44%	68%	72%
		traîner	0%	0%	6%	9%	17%
OR 1	+	traître	10%	27%	55%	71%	86%
		traître	10%	0%	21%	9%	42%
OR 1	+	transparaitre	0%	7%	16%	48%	47%
		transparaître	0%	0%	5%	14%	29%
OR 5	+	véto	69%	60%	62%	50%	50%
		veto	7%	0%	62%	20%	50%

Bibliographie

- Abeillé A. & Godard D. (eds), (2021). La Grande Grammaire du Français (GGF). Arles : Actes Sud. Imprimerie Nationale Editions.
- Audibert-Gibier M. (1992). Etude de l'accord du participe passé sur des corpus de français parlé. *Langage et Société* 61:7-30.
- Barque L. & Candito M. (2019). La classification des verbes réfléchis à l'épreuve d'une annotation en corpus. *Langages*, 216, 121-137.
- Béchenec D. & Sprenger-Charolles L. (2014). Guide pratique de l'orthographe rectifiée, *Pratiques* 163-164. En ligne : <https://journals.openedition.org/pratiques/2227>.
- Béguelin M.-J. (2002). Faut-il simplifier les règles d'accord du participe passé ? Approches interlinguistiques de la complémentation verbale : quels savoirs pour l'enseignant ? Quels savoirs pour l'élève ? *Travaux neuchâtelois de linguistique (TRANEL)* 37, 163-189.
- Bena J.-M. (2004). Les déficits en matière de français langue maternelle : diagnostic et base de remédiation, (Thèse de doctorat, ULB) L'Harmattan, 2004.
- Benzitoun C. & Flesch M. (2024), La disparition qu'on a mis au jour : étude sur corpus de l'accord du participe passé avec avoir chez des adultes francophones, Actes du 9^e Congrès Mondial de Linguistique Française, Lausanne. <https://www.linguistiquefrancaise.org/>
- Bescherelle L-N. (1852). Grammaire nationale ou grammaire de Voltaire, de Racine, de Bossuet, Hachette, Paris.
- Beslais P. (1965). Rapport général sur les modalités d'une simplification éventuelle de l'orthographe (dit «Rapport Beslais»), Paris, Didier.
- Bouhours D. (1675/1687). Remarques nouvelles sur la langue française. Paris, S. Mabre-Cramoisy, 1975.
- Brissaud C. & Cogis D. 2008. L'accord du participe passé : Reconsidération d'un problème ancien à la lumière de données récentes sur l'acquisition. Durand, Habert, Laks (eds), Actes du CMLF, Paris. En ligne <https://www.linguistiquefrancaise.org/articles/cmlf/pdf/2008/01/cmlf08105.pdf>
- Brissaud C. & Fayol M. (2018). Écrire et rédiger : Etude de la langue et production d'écrits: Rapport scientifique de la conférence de consensus du CNET. En ligne : <https://www.cnesco.fr/wp-content/uploads/2018/04/VDEF-rapport-Brissaud-Fayol.pdf>
- Brunot F. (1905). La réforme de l'orthographe : Lettre ouverte au ministre de l'Instruction publique. Paris : A. Colin.
- Caravolas M. (2004). Spelling development in alphabetic writing systems: A cross-linguistic perspective, *European Psychologist*, 9(1): 3-14.
- Catach N. (1984). Les listes orthographiques de base (LOB). Nathan
- Catach N. (1989). Les délires de l'orthographe, un psychodrame français. Plon
- Catach N. (1999). La bataille de l'orthographe aux alentours de 1900, in Antoine & Martin (dir.). Histoire de la langue française 1880-1914. Editions du CNRS
- Cerquiglini B. (1989). Eloge de la variante : histoire critique de la philologie. Le Seuil
- Cerquiglini B. (2004). La Genèse de l'orthographe française (XII^e - XVII^e siècles). Paris, Champion
- Cerquiglini B. (2018). L'orthographe rectifiée : le guide pour tout comprendre. Paris, Libro.
- Cerquiglini B. (2021), Un participe qui ne passe pas, Paris, Points.
- Chervel A. (1977). Et il fallut apprendre à écrire à tous les petits français, Histoire de la grammaire scolaire, Paris, Payot.
- Chervel A. (1991). L'école républicaine et la réforme de l'orthographe (1879-1891). *Mots*, n°28, 35-55.
- Clédat L. (1903). Causerie sur l'Orthographe, Bulletin de la Société suisse de réforme orthographique, n°6, Lausanne.
- Conseil de la Langue française, des Langues régionales endogènes et des Politiques linguistiques de la FWB (2023), Avis d'initiative n° 3 (.pdf) (26 juin 2023) relatif à l'invariabilité du participe passé (PP) employé avec avoir.
- Darmesteter A. (1888). La question de la réforme orthographique. *La Revue Pédagogique*, 12, 581-604. En ligne https://education.persee.fr/doc/revpe_2021-4111_1888_num_12_1_2824
- Deacon H., Desrochers A. & Levesque K. (2017), Learning to Read French. The crosslinguistic study of reading and spelling acquisition: The case of French, in L. Verhoeven & C. Perfetti (eds.), Learning to Read across Languages and Writing Systems, Cambridge, Cambridge University Press, 243-269.
- Dehaene S. (2007). Les neurones de la lecture. Paris, Odile Jacob
- Dehaene S., Dehaene-Lambertz G., Gentaz E., Huron C. & Sprenger-Charolles L. (2011). Apprendre à lire : des sciences cognitives à la salle de classe. Paris, Odile Jacob.
- Delattre P. (1965). Comparing the phonetic features of English, German, Spanish and French. John Benjamins
- DEPP (2022). Les performances en orthographe des élèves-de-CM2. Note d'Information, n°22.37. En ligne : <https://www.education.gouv.fr/les-performances-en-orthographe-deseleves-de-CM2-toujours-en-baisse-mais-de-maniere-moins-marquee-343675>.
- Dister A. & Moreau M.-L. (2019). Les tribulations d'une réforme dans un championnat d'orthographe, Comment des élèves en fin de scolarité primaire ont intégré les Rectifications de 1990, *Le Français Moderne : Revue de linguistique française*, no.2, 281-295.
- Encrevé P. (1988). La liaison avec et sans enchaînements : phonologie tridimensionnelle et usages du français. Paris, Editions du Seuil
- Faguet E. (1905). Simplification simple de l'orthographe, Société française d'imprimerie et de librairie ancienne ; Ancienne librairie Lecène, Oudin et Cie, Paris.
- Fayol M. (2020). L'acquisition de l'écrit. Paris : PUF (Que sais-je?)

- Fayol M., Grimaud F. & Jacquier M. (2013). Une expérience d'enseignement explicite de l'orthographe lexicale. ANAE, n° 123, 156-163.
- Fayol M. & Pacton S. (2006). L'accord du participe passé : entre compétition de procédures et récupération en mémoire, *Langue Française*, 151, p. 59-73. En ligne : <https://www.cairn.info/revue-langue-francaise-2006-3-page-59.html>
- Gaucher D. (2015). Sémantique temporelle et accord du participe passé en français: une approche variationniste, *Journal of French Language Studies* 25 , 65–83. En ligne
- Gimenes M., Perret C. & New B. (2020). Lexique-Infra: Grapheme-phoneme, phoneme-grapheme regularity, consistency, and other sublexical statistics for 137,717 polysyllabic French words. *Behavior Research Methods*, 52, p. 2480-2488.
- Grevisse M. (1936). *Le Bon Usage*, 1^{re} édition, Duculot. [avec André Goosse, 16^e édition, 2016].
- Haby R. (1976). Arrêté du 28-12 1976 relatif aux tolérances grammaticales et orthographiques https://www.weblettres.net/guidetice/complements/arrete_1976.pdf
- Hathout N. & Namer F. (2023). Démonette : une base de données dérivationnelle du français, *Lexique* [En ligne], 33 | 2023, mis en ligne le 01 décembre 2023, consulté le 30 avril 2024. URL : <http://www.peren-revues.fr/lexique/1266>
- Hemforth B., Siqi Chen S. & Liu M. L. (2023), Acceptability judgement of French participle agreement: Chinese learners of French and Native speakers, *LLF*, Université Paris Cité.
- Kanas C. (2019). *L'orthographe pour tous*. Paris : Hatier (Collection le Bescherelle).
- Landerl K. & Wimmer H. (2008). Development of word reading fluency and spelling in a consistent orthography: An 8-year follow-up, *Journal of educational psychology*, 100(1) : 150-161.
- Le Levier H. (2022). Traces du discours grammatical scolaire chez des scripteurs experts: ce que révèle le traitement morphosyntaxique du participe passé chez des étudiants de l'enseignement supérieur technique. 8^e CMLF, Orléans. En ligne <https://doi.org/10.1051/shsconf/202213806021>
- Lété B., Peereman R. & Fayol M. (2008). Consistency and word-frequency effects on spelling among first-to fifth-grade French children. *Journal of Memory and Language* 58(4):952-977
- 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, 156-166.
- Leygues G. (1900). Arrêté relatif à la simplification de l'enseignement de la syntaxe française. In: *L'enseignement du Français à l'école primaire - Textes officiels*. Tome 2 : 1880-1939. Paris : Institut national de recherche pédagogique, 1995. pp. 191-198. (Bibliothèque de l'Histoire de l'Éducation, 5); https://www.persee.fr/doc/inrp_0000-0000_1995_ant_5_2_2019
- Leygues G. (1901). Arrêté relatif à la simplification de la syntaxe française. In: *L'enseignement du Français à l'école primaire - Textes officiels*. Tome 2 : 1880-1939. Paris : Institut national de recherche pédagogique, 1995. pp. 198- 203. (Bibliothèque de l'Histoire de l'Éducation, 5); https://www.persee.fr/doc/inrp_0000-0000_1995_ant_5_2_2020
- Mailhot H. et al. (2020). MorphoLex-FR: A derivational morphological database for 38,840 French words, *Behavioural Research*, 52, 1008-1025.
- Marchello-Nizia C. (1999). L'accord du participe passé avec l'objet direct en ancien français. *Verbum*, 21(3): 323-338.
- Nazzi T., Barrière I., Goyet L., Kresh S., Culbertson J. & Legendre G. (2011). Tracking irregular morpho-phonological dependencies in natural language: Evidence from the acquisition of subject-verb agreement in French. *Cognition*, 120, 119-135.
- Pacton S., Nys, Fayol M. & Peereman R. (2022). L'influence des connaissances graphotactiques dans l'acquisition de l'orthographe lexicale. In *Recommandations professionnelles concernant le langage écrit*. (pp. 23 2-237). CFO (Collège Français d'Orthophonie). <https://www.college-francais-orthophonie.fr/les-recommandations-de-bonne-pratique/>
- Paulesu E., Démonet J.F., Fazio F. et al. (2001). Dyslexia: cultural diversity and biological unity. *Science*, 16;291(5511):2165-7. doi: 10.1126/science.1057179. PMID: 11251124.
- Peereman R. (Work in progress). Principales spécificités de l'orthographe du français (lettres muettes de fin de mots, consonnes doublées...)
- Peereman R. et al. (Work in progress). Morphiste : a derivational morphological database for 10 000 French words.
- Peereman R., Lété B. & Sprenger-Charolles L. (2007). Manulex-Infra: Distributional characteristics of grapheme-phoneme mappings, infra-lexical and lexical units in child-directed written material. *Behavior Research Methods*, 39, 579-589. doi.org/10.3758/BF03193029 <https://inframorph.github.io/index.html>
- Peereman R. & Sprenger-Charolles L. (2018). Manulex-MorphO, une base de données sur l'orthographe du français intégrant les morpho-phonogrammes, *Langue Française*, 199, 99-109. <https://www.cairn.info/revue-langue-francaise-2018-3-page-99.htm>.
- Peereman R., Sprenger-Charolles L. & Messaoud-Galusi S. (2013). The contribution of morphology to the consistency of spelling-to-sound relations: A quantitative analysis based on French elementary school readers, *Année Psychologique*, 113, p. 3-33. doi.org/10.4074/S0003503313001012. <https://inframorph.github.io/index.html>
- Poli F., Monneret P. (2023). *Grammaire du CP à la 6^e*, Eduscol. En ligne.
- Pothier B. & Pothier P. (2003). ÉOLE - Échelle d'acquisition en Orthographe Lexicale Cycles 1, 2 et 3. Paris : Editions Retz.
- Pothier B. & Pothier P. (2016). *L'orthographe rectifiée à l'école. Quels impacts sur les apprentissages ?* Retz, ISBN : 978-2-7256-3515.
- Pothier B. & Pothier P. (2024). ÉOLE - Échelle d'acquisition en Orthographe Lexicale, Editions Retz, Paris (Version corrigée de l'édition 2022, voir la note 11).

- Ramus, F. (in press). Dyslexia, Phonological processing in. In H. Pashler (Ed.), *Encyclopedia of the Mind*. Thousand Oaks, CA: Sage http://www.lscop.net/persons/ramus/docs/Dyslexia_phonological_processing2.pdf
- Riegel M., Pellat J.-C. & Rioul R. (1994). *Grammaire méthodique du français*, Paris : PUF. (7^e édition, 2018)
- Seymour P. H. K., Aro M. & Erskine J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, 94, 143-174.
- Share D.L. (1995). Phonological recoding and self-teaching: sine qua non of reading acquisition. *Cognition*. 55(2),151-218 doi: 10.1016/0010-0277(94)00645-2. PMID: 7789090
- Sprenger-Charolles, Siegel L., Jimenez J. & Ziegler J. (2011). Prevalence and reliability of phonological, surface, and mixed profiles in dyslexia: A review of studies conducted in languages varying in orthographic depth. *Scientific Study of Reading*, 15 (6), pp.498-521. fahal-00733553f
- Sprenger-Charolles L. & Peereman R. (2023). Nouveaux arguments en faveur d'une rationalisation de l'orthographe du français, ANAE, 187.
- Sprenger-Charolles L., Siegel L.S., Béchenec D. & Serniclaes W. (2003). Development of phonological and orthographic processing in reading aloud, in silent reading and in spelling: a four-year longitudinal study. *Journal of Experimental Child Psychology*, 84, p. 194-217.
- Sprenger-Charolles L., Siegel L. S. & Bonnet P. (1998). Reading and spelling acquisition in French: the role of phonological mediation and orthographic factors. *Journal of Experimental Child Psychology*, 1998, 68 (2), .134-65. ff10.1006/jecp.1997.2422ff.
- Stanké B. et al. (2019). ÉQOL : Une nouvelle base de données québécoise du lexique scolaire du primaire comportant une échelle d'acquisition de l'orthographe lexicale. *Corpus [En ligne]*, 19 | 2019; <https://doi.org/10.4000/corpus.3818>
- TLFi : Trésor de la langue Française informatisé, <http://www.atilf.fr/tlfi>, ATILF - CNRS & Université de Lorraine
- Tseng J. (2021). Non-canonical arguments and past participle agreement, in Abeillé A. & Bonami O. (eds), *Constraint-based Syntax and Semantics*, Stanford: CSLI Publications.
- Vaugelas C. F. de (1647). *Remarques sur la langue française, utiles à ceux qui veulent bien parler et bien écrire*, Paris.
- Vaughn S. & Fletcher J. (2021). Explicit Instruction as the Essential Tool for Executing the Science of Reading, *Read Leag J.* May-Jun, 2(2), p. 4-11.
- Verhoeven L. & Perfetti C. (Eds, 2017). *Learning to Read across Languages and Writing Systems*, Cambridge, Cambridge University Press
- Verhoeven L., Perfetti C. & Pugh K., (Eds., 2019). *Developmental Dyslexia across Languages and Writing Systems*. Cambridge University Press
- Véronis J. (1988). From sound to spelling in French: simulation on a computer. *Cahiers de Psychologie Cognitive*, 8, 315-334.
- Wagner R.-L. & Pinchon J. (1962) *Grammaire du français classique et moderne*, Paris : Hachette.
- Xu M., Tan L.H. & Perfetti C. (2019). Developmental Dyslexia in Chinese. In: Verhoeven L., Perfetti C. & Pugh K., eds. *Developmental Dyslexia across Languages and Writing Systems*. Cambridge University Press; 2019:200-226.
- Ziegler J. (2018). Différences inter-linguistiques dans l'apprentissage de la lecture. *Langue Française*, 199, 35-49. En ligne : <https://www.cairn.info/revue-langue-francaise-2018-3-page-35.htm>.
- Ziegler J. C., Perry C. & Zorzi M. (2014), Modelling reading development through phonological decoding and self-teaching: Implications for dyslexia, *Philosophical Transactions of the Royal Society of London. B: Biological Sciences* 369 (1634), 20120397(17)
- Ziegler J. & Sprenger-Charolles, L. (2023). Apprendre à lire: du décodage à la compréhension écrite. ANAE, 182, 45-54.

Contact presse

01 55 55 30 10

spresse@education.gouv.fr

Contact Conseil scientifique de l'éducation nationale

cсен@education.gouv.fr

reseau-canope.fr/conseil-scientifique-de-leducation-nationale