

Des phonèmes aux graphèmes

From phonemes to graphemes

Anne DISTER – UCLouvain – Saint-Louis – Bruxelles

Marie-Louise MOREAU – Université de Mons

Résumé : Nous analysons les correspondances phonèmes-graphèmes dans un corpus de près de 12 000 dictées produites lors d'un championnat d'orthographe. Cette analyse montre qu'en fin de scolarité primaire, les élèves maîtrisent bien ces correspondances, le pourcentage des écarts, calculé sur le nombre de phonèmes, étant toujours inférieur à 1 %. Cela ne veut pas dire qu'ils écrivent selon la norme orthographique, mais qu'ils utilisent bien les graphèmes-cibles correspondant aux phonèmes-sources. Si le français était une langue transparente comme l'italien ou le finnois, les élèves seraient des champions en orthographe. Notre attention se porte également sur trois règles spécifiques de l'écrit qui concernent la combinaison des graphèmes : n qui devient m devant b ou p dans les voyelles nasales, q+u et la transcription du phonème /ɛ/ en syllabe écrite fermée. Si les deux premières sont bien maîtrisées, c'est moins le cas du phonème /ɛ/. Que ce soit pour les correspondances phonèmes-graphèmes ou les règles spécifiques qui gouvernent l'écrit, les achoppements se condensent autour de certains mots, plus rares, qui semblent ne pas faire partie du lexique mental des élèves.

Mots-Clés : graphie ; phonie ; orthographe ; dictée

Abstract: We analyze phoneme-grapheme correspondences in a corpus of nearly 12,000 dictations produced during a spelling championship. This analysis shows that by the end of primary school, pupils have a good command of these correspondences, the percentage of deviations, calculated on the number of phonemes, always being less than 1%. This does not mean that they write according to the orthographic norm, but that they use the target graphemes corresponding to the source phonemes. If French were a transparent language like Italian or Finnish, our students would be spelling champions. Our attention is also drawn to three specific rules of writing which concern the combination of graphemes: n which becomes m before b or p in nasal vowels, q+u and the transcription of the phoneme /ɛ/ into a closed written syllable. While the first two are well mastered, this is less the case for the phoneme /ɛ/.

Whether it's phoneme-grapheme correspondences or the specific rules governing writing, the stumbling blocks are condensed around certain, rarer words that don't seem to be part of the pupils' mental lexicon.

Keywords: grapheme; phoneme; orthograph; dictation

Introduction

La dictée du Balfroid est, en Belgique francophone, une véritable institution. Depuis sa première édition en 1988, des milliers d'enfants de 6^e année primaire ont participé à ce célèbre concours d'orthographe. Les élèves, s'ils n'ont pas redoublé ou n'ont pas sauté une ou plusieurs classes, ont donc, au moment de l'épreuve, 11 ou 12 ans. Or, malgré son succès et les discours récurrents sur la crise de l'orthographe, aucune étude n'a jamais analysé les résultats de ces championnats. C'est pour pallier ce manque que le projet de recherche Orthocor a vu le jour en 2015, grâce à un financement de la Fédération Wallonie-Bruxelles.

Chaque année, le concours est organisé en 6 demi-finales (une par province belge francophone) et une finale, à laquelle accèdent les élèves qui ont commis moins de 6 fautes lors de l'épreuve éliminatoire. Dans le cadre de notre recherche, nous avons rassemblé les dictées qui se sont déroulées de 2012 à 2016. En tout, cela fait donc 35¹ dictées différentes de 132 mots en moyenne, pour lesquelles nous comptabilisons 11 298 copies.

Les épreuves se déroulent dans une grande salle, capable d'accueillir tous les participants, dont le nombre varie, pour notre corpus, de 192 à 618 selon les sessions. La responsable du championnat, sur une estrade et équipée d'un micro, lit le texte, intégralement d'abord, segment de phrase par segment de phrase ensuite, puis de nouveau intégralement. Sa diction est normée. Elle ne réalise pas les doubles consonnes ou les liaisons de manière appuyée, à la Topaze (Mortamet et Conseil, 2019).

Les textes, qu'elle a elle-même composés, correspondent à peu près au niveau linguistique des enfants. À peu près : ils contiennent souvent des mots qui ne figurent pas dans le *Vocabulaire orthographique de base* (Ters *et al.*, 1968 ; dorénavant *VOB*), mais ne sont pas hérissés des difficultés orthographiques qu'affectionnent d'autres championnats de ce type (voir dans le tableau 1 les mots qui concentrent le plus de difficultés phonographémiques dans

¹ On compte 7 sessions par an (6 demi-finales et 1 finale) pendant 5 ans.

6 des dictées). On ne relève aucune systématique dans le choix des phonèmes ou des groupes de phonèmes.

Dans le présent article, nous portons notre attention sur la correspondance entre phonèmes et graphèmes. Dans un premier temps, nous présentons des résultats sur un sous-corpus composé des 6 dictées proposées aux demi-finales de 2016, soit 1584 copies. Nous considérerons ensuite, dans l'ensemble des 11 298 copies, trois règles spécifiques de l'écrit régissant la combinaison des graphèmes : *n* devant *b* ou *p* ; *q* suivi de *u* ; *e* en syllabe graphique fermée.

1. Les élèves connaissent-ils les correspondances phonèmes-graphèmes ?

1.1. Méthode

Nous avons catégorisé les mots des 6 dictées proposées aux demi-finales de 2016 en différents ensembles suivant que l'oralisation des graphies adoptées par les élèves coïncidait ou non avec la cible orale :

- Nous avons considéré comme des écarts des productions telles que *pénètre*>*penaitre*, *renouveau*>*renouvant*, *valorisé*>*falorisé*, *insignifiant*>*insinifiant*. Les correspondances graphèmes-phonèmes ne permettent pas d'arriver, quand on les lit, à la même chaîne phonique que la cible. C'est sur cette catégorie que notre attention s'est focalisée.
- Nous avons en revanche traité comme conformes à la cible orale des graphies comme *accentuaient*> *axentuet* ; *certain*> *serteint*, *éphémère*> *éfémaire*, *insouciant*>*inssoussiant*, *foisonnant*>*foizonants*, *inconsciemment*> *inconciament* ; *pénètre*>*pénaitre* : elles sont certes déviantes par rapport à la norme, mais leur lecture à voix haute ne les distinguerait pas de la cible.

Le relevé ne prend pas en compte les mots absents et les mots inachevés, c'est-à-dire ceux dont l'élève n'a écrit au maximum que la moitié des phonèmes (*onéreux*>*on*), ainsi que pour les substitutions, lorsque la réalisation de l'élève correspond, phonétiquement, non pas à la cible, mais à un autre mot : *appliqué*>*impliqué*, *abdiqué* ; *cependant*>*pourtant* ; *condiment*>*continent* ; *parcelles*> *passerelle*, *partielles*.

1.2. Résultats globaux

Le pourcentage des écarts, calculé sur le nombre de phonèmes, est, dans les 6 dictées, toujours inférieur à 1 %, variant de 0,3 à 0,8 %. Le principe de la transcription alphabétique est donc très bien intégré. Les enfants connaissent bien les correspondances phonèmes-graphèmes, ils savent quelles sont les possibilités disponibles pour transcrire les différents phonèmes (Bosman et Van Orden, 1997 ; Fayol et Jaffré, 2008 ; Peereman *et al.* 2013 ; Sprenger-Charolles *et al.*, 2003). Si l'orthographe du français était comme celle du finnois, de l'italien, de l'espagnol, du néerlandais, de l'allemand, et de bien d'autres langues européennes ou non européennes, essentiellement phonologique, sa maîtrise serait assurée sans problème aucun, et certainement bien avant la 6^e primaire. C'est bien ce qu'indiquent déjà Seymour *et al.* (2003) dans une étude portant sur les capacités en lecture des enfants parlant dix des langues européennes.

Dans les écritures strictement phonologiques, un phonème et un seul est associé à un graphème et un seul, et inversement, comme dans l'alphabet phonétique international. On est loin de compte en français : à un même phonème correspondent en effet souvent plusieurs graphèmes et inversement², l'éventail des solutions étant cependant moindre dans le second cas. Il en résulte une asymétrie entre lecture et écriture : le signifiant oral d'un mot écrit est davantage prévisible que le signifiant écrit d'un mot oral. Ce que nos données indiquent, c'est que les enfants dont nous analysons les productions ne rencontrent vraisemblablement pas de difficultés dans la lecture³. Mais pour écrire correctement, ils doivent savoir quels graphèmes précisément, parmi les différentes possibilités théoriques, la norme orthographique retient pour les unités qu'ils ont à transcrire. Pour identifier la graphie qui convient, les enfants appliquent des stratégies graphotactiques : ils dégagent certaines régularités statistiques dans la manière dont les correspondances graphèmes-phonèmes se répartissent. Par exemple, alors que le système permet de transcrire /o/ par *o, ô, au, eau*, beaucoup d'enfants, dès la première primaire, repèrent que si *eau* est fréquent en finale de mots, il est rare au début (Fayol et Jaffré, 2008, p.191⁴). Il reste que la multiplicité des graphies théoriquement possibles pour transcrire

² Le phonème /s/, par exemple, est transcrit par *s* dans *soulier*, *ss* dans *passer*, *c* dans *ceci*, *ç* dans *lançait*, *t* dans *potion*, *sc* dans *scier*, *x* dans *dix* ; le graphème *s* se lit /s/ dans *soulier*, /z/ dans *poison*, ou il est muet (*pas*), le choix entre les possibilités étant généralement dicté par le contexte.

³ Nous envisageons ici le déchiffrement des mots, pas l'accès à leur sens. À fortiori, pas l'accès au sens des textes dans lesquels ils figurent.

⁴ Pour une synthèse sur les régularités graphotactiques, voir Pacton *et al.* (2019).

un même phonème ou une même suite de phonèmes est une des pierres d'achoppement du système orthographique du français.

Si les correspondances phonographiques sont bien intégrées pour l'essentiel, il reste cependant une toute petite proportion de cas où les enfants recourent à un graphème qui n'a pas la valeur phonologique souhaitée. Quelles correspondances paraissent présenter le plus de difficultés ? La répartition des différents écarts en catégories (selon que, par exemple, le phonème /s/ de *poisson* est transcrit, erronément, par un seul *s*, ou par *c*, ou par *t...*, ou qu'un *u* prend la place d'un *ou*, etc.) indique rapidement que les chiffres sont étroitement tributaires des mots particuliers utilisés dans le texte dicté : on ne peut évidemment observer de difficulté sur les terminaisons en /aj/, si le texte ne comporte ni *caille* ni *épouvantail*, etc. Comme on peut le voir dans le tableau 1, les écarts s'agglomèrent en fait autour de certains mots : 10 des 100 à 140 mots que comptent ces dictées sont à l'origine de 57 à 89,1 % des écarts. Si l'on prend en compte les 15 formes les plus fréquemment fautives, on relève 60,3 à 93,4 % des écarts. Qu'on nous entende bien : ce ne sont pas les fautes d'orthographe que nous relevons ici, mais seulement, parmi celles-ci, les cas où la production des élèves n'est pas conforme aux correspondances prévues par le code entre phonèmes et graphèmes, des cas, autrement dit, où, quand on prononce ce qu'a écrit l'enfant, on n'aboutit pas au signifiant oral de la cible. Les mots concernés ne sont pas parmi les plus familiers, beaucoup d'entre eux (ci-dessous, ils sont signalés comme tels par un astérisque) ne figurent d'ailleurs pas dans le *VOB* (Ters *et al.*, 1968).

Tableau 1 : Mots pour lesquels les élèves respectent le moins les correspondances phonèmes-graphèmes

Dictée	Les mots qui provoquent le plus d'écarts	% pour les 15 premiers	% pour les 10 premiers
D16-Alimentation	quotidiennement*(18,2), alléchants* (17,0), terrasse* (16,8), réjouissant* (7,8), alternatives* (7,1), aménager* (6,9), souhaitée (3,9), désormais (3,0), pimpant* (1,6), adresser (1,6), lointains (1,6), sainement* (1,6), condiments* (1,1), trouverez (0,9), employés (0,9)	90,1	83,9
D16-Fenêtre	insignifiant* (17,6), effaçant (13, 0), insouciant* (7,9), gaieté (7,4), pénètre (7,0), aubépines* (5,0), jonquilles (4,4), témoins (4,0), bosquets (3,3), inaperçues* (3,3), dès (3,3), dépouillés (2,3), précoce* (2,9), immaculé* (1,6), bulbes* (1,5)	83,6	72,8
D16-Gelé	éphémère* (24,7), gazouillis* (13,3), déçus (5,4), effet (3,4), inconsciemment* (2,7), accentuait* (2,3), croyaient (1,8), clarté* (1,4), gelé (1,2),	60,3	57,2

	était (1,1), crocus* (0,8), gel* (0,8), été (0,7), frileux* (0, 5), brillait (0,4)		
D16-Ordinateur	intéressants (29,3), aisément* (17,7), semblaient (11,4), attellent* (5,5), soigneusement (5,2), éclaircissent (4,8), né (3,0), hebdomadaires* (2,0), désormais (2,0), est (2,0), complément (1,6), indispensable (1,6), bénéficiaire* (1,4), obscurs (1,4), révolue* (1,3)	90,4	81,0
D16-Perspective	impeccablement* (23,1), désespéraient* (21,6), ressortis (14,7), élançaient (11,7), réjouit* (6,5), perspective* (4,4), engins* (2,6), joues (2,1), représentait (1,3), loueurs* (1,3), espoirs (1,1), utilisés (0,9), attendaient (0,9), étaient (0,8), comblés (0,6)	93,4	89,1
D16-Tempête	déchaînés* (18,0), violemment* (10,9), annonçaient (9,6), flamboyant* (8, 5), empessant* (5,5), résistaient* (4,4), pressés (4,4), dépassait (4,1), tenaient (3,8), malgré (3,0), éléments* (2,7), sensations* (2,5), craindre (2,2), retournaient (1,9), incorrigibles* (1,4)	82,8	72,1

Mais a-t-on vraiment affaire à une maîtrise incomplète des correspondances entre phonèmes et graphèmes ? En fait, on doit supposer que quand les enfants écrivent un texte sous dictée, ils ne transposent pas dans l'écrit la forme orale qu'ils entendent. Comme chez les adultes, celle-ci a pour fonction d'activer un signe intégré dans leur lexique mental, et c'est le signifiant de ce signe intériorisé qu'ils transcrivent. Ce n'est pas un hasard si les concentrations d'écarts se rencontrent dans des mots inconnus des enfants, ou peu familiers, sous leur forme écrite et parfois même orale. Comme leur lexique mental ne contient pas ces mots, les enfants tentent de recomposer ce qu'ils ont entendu, non sans distorsion quelquefois : notre corpus nous offre ainsi des cas tels que, par exemple, *aubépinés>aubutines*, *bosquets>paussetés*, *immaculé>inafumé*, *insouciantes>inssidence*, *perspective>perpsectise*, *perpextive*⁵. Si 20 enfants, dans la dictée 16D-Fenêtres, écrivent *aubépinés>opépinés*, *aupépinés*, *oppépinés*, etc., on ne peut en déduire qu'ils confondent sourde et sonore, ou qu'ils ignorent que /b/ se transcrit par le graphème *b*, et non par *p* ; de telles confusions ne sont d'ailleurs pas illustrées dans d'autres mots au sein de la même dictée. Ces erreurs indiquent surtout comment ils réagissent devant un mot non ou peu connu : en le rattachant à des unités voisines qu'ils connaissent (dans le

⁵ Le phénomène est très visible dans les graphies utilisées par les élèves lorsqu'ils écrivent certains des dix mots difficiles, qui doivent départager les ex-aequo, mots dont beaucoup ont sûrement le même statut pour les élèves que des logatomes, des mots fictifs : *okoumé>hocumé*, *acoumèe*, *ocouné*, *ocumer*, *ocomer*, *aucoubé*, *hopimmée*, *augoumé*, *agoumer*... autant de formes dont l'oralisation ne correspond pas à celle de la cible.

dernier exemple, *pépin* ?⁶), et en se fondant sur la connaissance intuitive qu'ils peuvent avoir des régularités dans la distribution des phonèmes. Une dictée teste donc parfois non seulement l'orthographe, mais aussi le répertoire lexical, dans sa composante orale autant qu'écrite.

Une autre hypothèse doit cependant être réservée pour expliquer ces distorsions : certains enfants pourraient souffrir de déficit auditif, ou à tout le moins ne pas bien entendre ce qui est dicté, d'autant que les conditions acoustiques où les épreuves se déroulent ne sont pas toujours optimales.

Les signifiants figurant dans le lexique interne des enfants présentent parfois une forme différente de la norme. Le corpus présente plusieurs cas de réduction de groupes consonantiques (*arbre*>*arbe*), ou, particularisme phonique répandu en Belgique francophone, de substitution d'une consonne finale sourde en lieu et place d'une sonore (*perspective*>*perspectif*). De même, le corpus contient plusieurs cas de *-ij/* final transcrit *ie* (*jonquille*>*jonquie*)⁷. Mais le phénomène est tout à fait marginal dans les dictées que nous avons étudiées.

Certaines correspondances se révèlent cependant problématiques. On les voit mises à mal dans des mots de toutes les catégories, et dans toutes les dictées. Les plus importantes concernent les points suivants.

1.3.1. Phonèmes /e/ et /ɛ/

Ces cas sont illustrés par les graphies *désormais*>*desormais*, *depuis*>*dépuis*, *éléments*>*élèments*, *est*>*et*, *hebdomadaires*>*hébdomadaires*, *trouverez*>*trouverais*, par exemple. Le plus généralement, les graphies erronées tiennent à une omission d'accent là où il en faudrait un, ou l'inverse, moins fréquemment, ou encore au recours à un accent aigu là où on attend un grave (525 cas), ou le contraire (145 cas) ; elles indiquent aussi parfois une certaine difficulté à distinguer les phonèmes /e/ et /ɛ/. Les écarts sur ces deux phonèmes représentent entre 31 et 73 % des cas où les enfants ne respectent pas les correspondances phonèmes-graphèmes. On a donc là une zone de fragilité (Lucci et Millet, 1994) dans les compétences orthographiques des enfants, sur laquelle l'enseignement aurait à se pencher. Cette fragilité tient à différentes raisons :

⁶ Chez d'autres élèves, on trouve *aubépines*>*dos d'épines*, *abondamment*>*à bonds d'amants*, *éphémères*>*effet-maire*, etc.

⁷ On a aussi, moins souvent, des cas tels que *carie*>*carille*.

- Qu'on écrive *e*, *é* ou *è*, voire *ê*, cela n'impacte guère la physionomie iconique des mots, ce qui conduit certains enfants à les percevoir comme accessoires. Cette perception est d'ailleurs celle de certains maîtres aussi, qui ne retranchent qu'un demi-point pour une faute d'accent, contre un pour les fautes d'usage et deux pour les fautes grammaticales (Chevrot *et al.*, 1994).
- Si les accents modifient généralement la valeur phonologique du *e*⁸, il n'en va pas ainsi dans d'autres cas : *ça* se prononce de la même manière que *çà*, *la* que *là*, *ou* que *où*, etc.
- Il existe depuis plusieurs décennies une tendance à neutraliser à l'oral l'opposition entre ces deux unités, en particulier dans la norme de France⁹, qui ne coïncide pas tout à fait sur ce point avec celle de la Belgique, ce qui ajoute à la confusion.
- Les règles qui gouvernent l'emploi des accents ne sont pas des plus complexes, mais elles sont nombreuses, et leur formulation ne peut éviter les exceptions, parfois en cascade (voir l'annexe).

À la difficulté d'utiliser la bonne correspondance graphème-phonème pour /e/ et /ɛ/, s'ajoutent, chez beaucoup d'élèves, des écarts par rapport aux règles spécifiques de l'écrit, qui dictent comment les graphèmes se combinent entre eux : on observe ainsi de nombreux cas où un *e* est pourvu d'un accent, alors qu'il est en syllabe écrite fermée (*effaçant*>*èffaçant*, p.ex.). On y reviendra au point 3.3.

1.3.2. Les phonèmes /s/ et /z/

Dans 2,3 % à 33,4 % des écarts, c'est la transcription de /s/ et /z/, à l'aide de *s*, *ss*, *c*, *ç*, *z*... qui semble difficile à gérer : *consultée*>*conculcée*, *insouciant*>*insousiant*, *bise*>*bisse*, p. ex. Ici aussi, la composition des textes impacte fortement les résultats. Le pourcentage d'enfants qui achoppent sur les correspondances en cause, pour chacune des 6 dictées, s'établit à 4,6 %, 15,4 %, 19,6 %, 24,6 %, 37,6 % et 61,5 %. Même s'il apparaît clairement qu'une dictée n'est pas l'autre¹⁰, le problème semble bien ne pas être marginal. Et encore, pour ces graphèmes dont la valeur est sensible au contexte, les pourcentages fournis ici ne tiennent pas compte, rappelons-

⁸ Encore que *foret* se lise comme *forêt*.

⁹ La tendance est relevée déjà par Martinet (1971), et confirmée par des études plus récentes (Fagyal, Hassa et Ngom, 2002, p.ex.).

¹⁰ Nous le montrions déjà à propos de la difficulté relative de l'orthographe grammaticale et de l'orthographe lexicale (Dister et Moreau, 2019).

le, de graphies comme *merçi*, qui sont certes erronées, mais ne peuvent se lire que comme la cible.

1.3.3. Le phonème /ɲ/

Dans l'ensemble du corpus, on relève parfois une accumulation de fautes sur des mots contenant /ɲ/, pourtant associé univoquement au groupe *gn*. Mais pas sur tous les mots : /ɲ/ dans *accompagnant, campagne, campagnard, montagne, enseignes* et *ignore*, ne pose de problème qu'à moins de 5 % des élèves ; *soigneusement, ognon et mignon* à entre 6 % et 9,5 %. Ce n'est pas le cas de *dignité* (25,5 %), *insignifiant* (50,5 %) et *égratigné* (59,9 %). Comme ces trois mots ne figurent pas dans le *VOB*, on pourrait en inférer que beaucoup des élèves ont développé une approche de l'orthographe plus globale qu'analytique. Toutefois, le *VOB*, qui compte au total 42 mots contenant le groupe *gn*, propose *signe* et *signifier*, et il est curieux que bien des élèves n'aient pas utilisé ici les règles de dérivation qu'ils ont bien appliquées dans *campagnard*, pourtant hors *VOB* lui aussi. Une autre hypothèse interprétative serait, comme plus haut, que ces mots ne figurent pas dans le répertoire lexical des enfants, et que ceux-ci ne maîtrisent déjà pas bien leur signifiant oral.

1.3.4. Les groupes vocaliques avec /j/

La difficulté à traiter ces groupes mériterait également une attention particulière de la part des enseignants. Beaucoup d'enfants échouent à écrire correctement des mots comme *gazouillis, groseilliers...* Le nombre limité d'unités de cette catégorie dans notre corpus rend cependant délicat de généraliser les observations.

Les autres écarts par rapport aux correspondances phonèmes-graphèmes s'éparpillent sur une multitude de catégories dont aucune ne se révèle problématique de façon systématique.

2. Les élèves connaissent-ils les règles spécifiques de l'écrit gouvernant la combinaison des graphèmes ?

La conversion de phonèmes en graphèmes doit répondre aux correspondances phonographiques, mais aussi à certaines règles propres à l'écrit, qui fixent la manière dont les

graphèmes se combinent, indépendamment des phonèmes transcrits. Pour étudier la manière dont les enfants respectent ou non ces règles, notre travail porte ici sur l'ensemble des 35 dictées.

2.1. *N devant b ou p*

Parmi les règles spécifiques de l'écrit, une des mieux identifiées veut que les voyelles nasales soient transcrites à l'aide d'un *n*, mais avec un *m* devant les graphèmes *p* et *b*. Peu de mots échappent à cette règle : *bonbon* et ses dérivés, *bonbonne*, *bonbonnière*¹¹...

Dans l'ensemble, les élèves ont bien intégré cette règle. Dans 30 des 35 dictées, le pourcentage d'élèves qui y dérogent est inférieur à 2 %. Mais qu'ils aient à écrire un mot peu connu, et ils s'en remettent à la règle qui transcrit les voyelles nasales par une voyelle suivie de *n* : c'est le cas, par exemple, dans des dictées différentes¹², de *désemplassaient*, pour 20,8 % de l'effectif ; de *pimpant*, pour 20,1 % ; de *intempéries* pour 8,6 % ; de *empressant* et *flamboyant*, pour 9 % et 6,2 %. Aucun de ces mots ne figure dans le *VOB*.

2.2. *Q suivi de u*

Le graphème *q* est toujours suivi d'un *u*, sauf dans *cinq* et *coq*, et une poignée de mots utilisés seulement par les amateurs de scrabble (*maqam*, *qat*...). Ce *u* a parfois une valeur fonctionnelle, et transcrit /w/, comme dans *quartz*. Mais, le plus souvent, il n'en a aucune (*qui*, *quand*...). Cette règle-là aussi est bien acquise par les élèves, même si on note toujours quelques accrocs, de quelques unités (*consacré*>*consaqré*, *queue*>*queue*, *technologies*>*tecqnologies*, *syntaxiques*>*s'attaqsiquent*, p.ex.). Un cas est plus important : 7,7 % des enfants éprouvent des difficultés à écrire *cinquième* avec un *u*, sans doute influencés par la perte du *u* dans les abréviations *cinq^e*, *cinq^{ème}*, *cinq^{ième}*, qu'on rencontre d'ailleurs dans les copies, malgré la consigne explicite d'écrire les chiffres en toutes lettres.

¹¹ *Bonhomme* et ses dérivés ne figurent pas parmi ces exceptions : la voyelle qui précède le *b* ne correspond pas à une nasale ; le /n/ est d'ailleurs bien perceptible.

¹² Il s'agit respectivement de 13D-Retour, 16D-Alimentation, 13D-Bonjour, 16D-Tempête (pour les deux derniers mots).

2.3. /ɛ/ en syllabe écrite fermée

Moins connue sans doute est la règle¹³ qui exclut qu'on mette un accent sur un *e* dans une syllabe écrite fermée : *e* suivi de deux consonnes (eCC), écrites (*terre, escalier*) ou orales (*examen*), sauf si la seconde est une liquide (*espiègle, pétrifier*), ou si les deux consonnes sont *ch* (*sèche*) ou *gn* (*duègne*) ; de même, on ne met pas d'accent sur *e* devant une consonne finale d'unité lexicale (*bec, secs, bel, hivers, met, etc.*) ; dans certains mots, le *e* peut cependant être surmonté d'un accent circonflexe (*forêt*). Bien que la règle portant sur les finales soit beaucoup mieux observée que la première, nos décomptes ne distingueront pas les écarts à l'une et à l'autre, pas plus que les cas impliquant des accents aigus, graves ou circonflexes.

Contrairement à celles qu'on a examinées en 3.1 et 3.2, ces régularités-ci sont mal intégrées par beaucoup d'enfants : le corpus fourmille de cas tels que *ardennais>ardénnaï*, *effacer>èffacer*, *électroniques>élèctroniques*. La proportion d'enfants concernés varie sensiblement selon les dictées : il oscille

- entre 4,2 et 9,6 % dans 9 des 35 dictées
- entre 10 et 19,4 % dans 13 autres
- entre 20,3 et 29,1 % dans 5 autres encore
- entre 30 et 45 % dans 6 cas
- entre 50,7 et 64,9 % dans les deux derniers cas.

Le problème ne peut donc pas être considéré comme mineur. Les différences de proportions tiennent à la composition des textes : le nombre d'erreurs de ce type est fonction du nombre de mots présentant des séquences eCC, et de la plus ou moins grande difficulté de ces mots pour les enfants. Certaines unités sont en effet plus exposées : 50,7 % des élèves achoppent sur au moins une des deux occurrences de *électroniques* dans une même dictée, et ailleurs, 27,6 % sur *électrique*, 25,7 % sur *efficace*, 17,5 % sur *prestement*, 15,3 % sur *pressante...* Sans doute désemparés par le caractère peu familier de ces mots¹⁴, ou par leur longueur, ou par leur structure syllabique, ou quelque autre difficulté, les enfants recourent à une stratégie de base, qui consiste à apparier chaque phonème avec un graphème : « /ɛ/, je l'écris è ; /s /, je l'écris ss. » Contre cette hypothèse, on pourrait faire valoir que pour transcrire /ɛ/, les élèves emploient bien davantage un *e* avec accent aigu (58,2 à 69,1 % des totaux de chacune des années), qu'un grave (27 à 39,4 %) ou un circonflexe (dont la proportion ne dépasse pas 3,9 %). Mais cette

¹³ On peut poser l'hypothèse que ce n'est pas enseigné comme une règle en tant que telle, à la différence des exemples qui précèdent.

¹⁴ Parmi ces exemples, seul *prestement* ne figure cependant pas dans le *VOB*.

tendance est cohérente avec celle qu'on observait précédemment. Elle indique soit une lacune dans les correspondances phonographémiques, soit la connaissance diffuse d'une régularité statistique (l'accent aigu sur le *e* est en effet bien plus fréquent que le grave¹⁵), soit encore le recours à l'accent aigu comme s'il s'agissait d'un accent universel, en lui assignant le rôle que jouerait l'accent plat qu'appellent de leurs vœux certains réformistes (Catach, 1992).

Conclusions

« Pratique sociale emblématique » (pour reprendre le titre d'un numéro de la revue *Glottopol* qui lui est consacré), la dictée est un exercice particulier, qui souvent – et c'est sans doute encore plus vrai lors de concours – condense les difficultés : les participes passés foisonnent, à côté des pluriels irréguliers et autres embûches.

L'analyse de nos données montre qu'en fin de scolarité primaire, les élèves maîtrisent bien les correspondances phonèmes-graphèmes. Dans nos dictées, le pourcentage des écarts, calculé sur l'ensemble total de phonèmes à transcrire dans ces dictées, est toujours inférieur à 1 %. Même s'ils n'écrivent pas nécessairement comme la norme orthographique l'exigerait, les enfants utilisent un graphème plausible pour transcrire le phonème-source. Si le système de transcription du français était bi-univoque, les enfants en fin de scolarité primaire maîtriseraient assurément bien l'orthographe du français, et les francophones auraient un rapport totalement dépassionné avec la dictée.

Les écarts se concentrent autour de certains mots : 10 des 100 à 140 mots que comptent ces dictées sont à l'origine de 57 à 89 % des écarts. On peut faire l'hypothèse que ces mots, moins courants et absents du *Vocabulaire orthographique de base*, n'appartiennent pas au lexique mental des élèves.

Certaines correspondances sont toutefois moins bien maîtrisées par les élèves : c'est le cas pour /e/ et /ɛ/, /s/ et /z/, et le phonème /ɲ/, qui mériteraient une attention particulière de la part des enseignants.

En français, des règles spécifiques de l'écrit gouvernent la combinaison des graphèmes. Si la règle du *n* qui devient *m* devant *b* ou *p* dans les voyelles nasales ou celle du *q+u* semblent bien maîtrisées, il n'en va pas de même pour le phonème /ɛ/ devant deux consonnes ou une consonne finale. Il semble d'ailleurs que cette règle ne fasse pas l'objet d'un apprentissage

¹⁵ Lexique.org fournit ainsi 44.109 mots contenant le premier, contre 4.195 pour le second.

spécifique. Comme pour les correspondances graphèmes-phonèmes, les difficultés sont concentrées autour de certaines unités, moins familières pour les enfants.

D'autres correspondances mériteraient d'être explorées, comme les graphèmes dont la valeur est sensible au contexte : /s/ transcrit *s* ou *ss*, /g/ qui correspond à *g* ou *gu*, etc. (Dister et Moreau, à par.). Néanmoins, même si des questions subsistent, notre analyse suggère qu'il existe une hiérarchie dans la disponibilité des règles : celles qui dictent les correspondances graphèmes-phonèmes auraient le pas sur les règles spécifiques de l'écrit.

Bibliographie

- Aro, M. et Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology*, *94*, 143–174.
- Bosman, A. M. T., et Van Orden, G. C. (1997). Why spelling is more difficult than reading. Dans C. A. Perfetti, L. Rieben & M. Fayol (dir.), *Learning to spell: Research, theory, and practice across languages*, 173-194. Mahwah, NJ, Lawrence Erlbaum.
- Brissaud, C. et Cogis, D. (dir.)(2011). *Comment enseigner l'orthographe aujourd'hui*. Hatier.
- Bousquet, S., Cogis, D., Ducard, D., Massonnet, J. et Jaffré, J.-P. (1999). Acquisition de l'orthographe et modes cognitifs. *Revue française de pédagogie*, *126*, 23-37.
- Catach, N. (1992). À propos de l'accent plat. *Recherche et réforme, Liaisons HESO*, *19-20*, 36-37.
- (1995). *L'orthographe française*. Nathan (3^e édition).
- Chevrot J.-P., Fontana, L., Freychet, L., et Simon, J.-P. (1994). Comment les enseignants corrigent et évaluent les variations orthographiques. Dans V. Lucci et A. Agnès (dir.) (1994), *L'orthographe de tous les jours. Enquête sur les pratiques orthographiques des Français*. Champion, 137-190.
- Csecsy, M. (1981a). Les trois accents : aigu, grave et circonflexe. *Le français dans le monde*, *162*, 20-29.
- (1981b). Les trois accents : aigu, grave et circonflexe. *Le français dans le monde*, *164*, 57-63.
- Dister, A. et Moreau, M.-L. (2019). Le poids de l'orthographe grammaticale et de l'orthographe lexicale dans les dictées. Dans Cl. Mortamet (dir), *L'orthographe : pratiques d'élèves, pratiques d'enseignants, représentations*. Presses universitaires de Rouen et du Havre, 21-42.
- (à paraître, 2024). What can a corpus of dictations tell us about spelling variation and mastery of context-sensitive spelling? *Nota Bene*.

- Fagyal, Z., Hassa, S. et Ngom, F. (2002). L'opposition [e]-[ɛ] en syllabes ouvertes de fin de mot en français parisien : étude acoustique préliminaire. *Actes des XXIVèmes Journées d'étude sur la parole*, Nancy, 24-27 juin 2002, 165-168.
- Fayol, M. et Jaffré, J.-P. (1999). L'acquisition/apprentissage de l'orthographe, *Revue française de pédagogie*, 126, 143-170.
- (2008). *Orthographier*. Presses universitaires de France.
- Gak, V. G. (1976). L'orthographe du français. Essai de description théorique et pratique. SELAF.
- Jaffré, J.-P. (2003). La linguistique et la lecture-écriture : de la conscience phonologique à la variable « orthographe ». *Revue des sciences de l'éducation*, 29 (1), 37-49.
- Lucci, V. et Millet, A. (dir.) (1994). *L'orthographe de tous les jours. Enquête sur les pratiques orthographiques des Français*. Champion.
- Martinet, A. (1971). *La prononciation du français contemporain*. Droz.
- Mortamet, C. et Conseil, J. (2019). Où en sont les *moutonss* de Topaze? Les indices d'orthographe à l'oral quand maitres et maitresses dictent un texte à leurs élèves. Dans Cl. Mortamet (dir), *L'orthographe : pratiques d'élèves, pratiques d'enseignants, représentations*. Presses universitaires de Rouen et du Havre, 75-96.
- Pacton, S., Fayol, M., Nys, M. et Peereman, R. (2019). Implicit statistical learning of graphotactic knowledge and lexical orthographic acquisition. Dans C. Perret et Th. Olive, *Spelling and Writing Words*, 39, BRILL, 41-66.
- Peereman, R., Sprenger-Charolles, L., Messaoudi-Galusi (2013). The contribution of morphology to the consistency of spelling-to-sound relations ; a quantitative analysis based on French elementary school readers. *L'année psychologique*, vol.113/1 ; 3-33. Seymour, P. H. K.
- Sprenger-Charolles, L., Siegel, L. S., Béchennec, D. et 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, 194-217.
- Ters, Fr., Mayer, G. et Reichenbach, D. (1968). *Vocabulaire orthographique de base*. OCDL.

Annexe

Comment transcrire avec un *e* les phonèmes /ə/, /e/ et /ɛ/?¹⁶

Les règles qui suivent valent seulement lorsqu'on utilise le graphème *e* pour transcrire les phonèmes /ə/, /e/ et /ɛ/. Elles ne prennent pas en compte les graphies en *ai*, *ay*, *ei*, *eu*...

- /ə/, le phonème généralement caduc, s'écrit avec un *e*, sans accent (*demain*, *le*) ; la lettre *e* est employée aussi dans certains mots, sans valeur phonologique : *gaieté*, *gréement*, *rudolement*...
- /e/ s'écrit le plus souvent avec un accent aigu, sauf...
 - ... dans les terminaisons verbales *-er*, *-ez* (*aller*, *allez*)
 - ... dans les terminaisons nominales ou adjectivales *-ier* (*fermier*, *journalier*)
 - ... dans les déterminants *ces*, *des*, *les*, *mes*, *tes*, *ses*
 - ... dans la conjonction *et* ainsi que dans les noms *pied*, *chez*, *nez*, *rez*
 - ... dans *canoë*.
- /ɛ/ est écrit avec accent grave, sauf en syllabe écrite fermée, c'est-à-dire...
 - ... dans les noms en *-ec*, *-ect*, *-ef*, *-el*, *-et* (*bec*, *respect*, *fief*, *fiel*, *fouet*)
 - ... dans certains mots en *-er* (*fer*, *mer*...) ou en *-es*...
 - ... quand il est suivi de deux consonnes orales (examen) ou écrites (*vienne*, *aspect*), y compris lorsqu'il s'agit de liquides (*terre*, *elle*), sauf...
 - ... si ces deux consonnes sont *ch* (*lèche*, *sèche*) ou *gn* (*duègne*, *imprègne*, *règne*...)
 - ... si la deuxième consonne seulement est une liquide (*lèvre*, *siècle*)
- *ê* suit les mêmes règles que *è*, sauf...
 - ... on peut avoir un *ê* devant *-t* final (*arrêt*, *forêt*).
- *ë* est employé dans *Noël* et *boëtte*.

Si une réforme pouvait intervenir sur ce point, elle pourrait, minimaliste, légitimer le *e* avec accent plat (*ē*) en lieu et place de *é*, *è*, *ê* (Catach, 1992), ce qui résoudrait déjà bien des problèmes, sans perte de fonctionnalité (seraient homographes : *lés* VS *lès*, *prés* VS *près*, p. ex.). Plus ambitieuse, elle devrait en outre s'attacher à formuler des règles plus générales, en évitant ou en limitant les exceptions.

¹⁶ Cet inventaire des règles s'appuie sur ceux de Catach (1995), Csecsy (1981a et b) et de Gak (1976).