



**HAL**  
open science

**Les zones de fragilité orthographique dans les  
productions sous dictée contrainte d'un texte classique  
par des adultes diplômé×e×s**

Manuel Pérez, Claudine Garcia-Debanc

► **To cite this version:**

Manuel Pérez, Claudine Garcia-Debanc. Les zones de fragilité orthographique dans les productions sous dictée contrainte d'un texte classique par des adultes diplômé×e×s. Congrès Mondial de Linguistique Française 2016, Jul 2016, Tours, France. pp.7013 - 7013, 10.1051/shsconf/20162707013 . hal-01403700

**HAL Id: hal-01403700**

**<https://hal.science/hal-01403700>**

Submitted on 27 Nov 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Les zones de fragilité orthographique dans les productions sous dictée contrainte d'un texte classique par des adultes diplômé·e·s

Manuel Pérez<sup>1,1</sup>, et Claudine Garcia-Debanc<sup>2</sup>

<sup>1</sup>Membre associé de CLLE-ERSS, Université de Toulouse, CNRS (UMR5263), SFR AEF ESPE Académie de Toulouse, UT2J, France.

<sup>2</sup>CLLE-ERSS, Université de Toulouse, CNRS (UMR5263), SFR AEF ESPE Académie de Toulouse, UT2J, France.

**Résumé.** L'étude conduite s'appuie sur des dictées recueillies auprès d'étudiant·e·s de 1<sup>o</sup> et de 2<sup>o</sup> années du Master MEEF 1<sup>er</sup> degré en préparation au Concours de Recrutement des Professeurs des Écoles. Lors d'une séance de remise à niveau en orthographe, les étudiant·e·s ont produit un extrait tiré d'un texte écrit par Fénelon entre 1701 et 1712. Cet extrait a fait l'objet de trois enquêtes étalées de 1873 à 2005 destinées à établir un bilan comparatif des connaissances des élèves du Cours Moyen 2<sup>o</sup> année à la 3<sup>o</sup> de collège. L'objectif de formation de cette séance était une réflexion sur l'orthographe du français et son enseignement à partir de pratiques orthographiques étudiantes. Pour l'objectif de recherche, il s'agissait d'analyser les zones de fragilité orthographique du français. Selon les travaux coordonnés par Lucci et Millet (1994), celles-ci se manifestent par les variations orthographiques contenues dans les productions, sont inhérentes au système et ne relèvent ni des scripteurs ni des situations d'écriture. Si tel est le cas, elles doivent aussi apparaître dans les productions sous dictée d'adultes diplômé·e·s se destinant à l'enseignement primaire. Afin de favoriser leur émergence dans les productions de participant·e·s destiné·e·s à véhiculer la norme orthographique, la dictée a été contrainte. Les résultats indiquent un faible pourcentage de variations, concentrées, quoique dans un ordre différent dû à la spécificité de la tâche, sur les mêmes zones de fragilité orthographique que celles dégagées par Pach, Jacquemin, Millet et Billiez (1994) : les diacritiques et auxiliaires d'écriture, les morphogrammes non verbaux de nombre et les lettres dérivatives ainsi que les morphogrammes verbaux, particulièrement l'accord du participe passé avec *avoir* ou les finales en /E/.

---

<sup>1</sup> Auteur de correspondance : manuel.perez@univ-tlse2.fr

**Abstract. Spelling fragility areas in classic French texts written under constrained dictation by graduates.** This study is based on dictations collected from students of Master 1 and 2 of Education. During a training session in spelling, students had to write a sample extracted from a text written by Fénelon between 1701 and 1712. This sample led to three large scale inquiries carried out from 1873 to 2005 and meant to compare 4<sup>th</sup> to 9<sup>th</sup> graders spelling knowledge. With regard to formation, this session aimed at leading a reflection from students' productions about French spelling and its teaching. The research aim was to assess the areas of French system spelling fragility. According to studies coordinated by Lucci and Millet (1994), these areas are reflected in spelling variations shown by productions, are inherent to the system and don't depend neither on writers nor on writing situation. If so, they should come to light in productions under dictation of graduates who intend to teach at primary schools and who can be considered as vectors of standard spelling. In order to foster the advent of these spelling variations, the dictation traditional processes have been changed. Results state a small percentage of variations. Even if in a different order due to task specific aspects, they concentrate on the same spelling fragility areas than those identified by Pach, Jacquemin, Millet and Billiez (1994): diacritical marks and writing auxiliary, non verbal morphograms of number, derivative letters and verbal morphograms, especially agreement of the past participle with auxiliary verb *avoir* or finals in /E/.

Les productions sous dictée contrainte d'un texte classique par des adultes en passe de devenir des vecteurs dans la transmission de l'orthographe du français laissent-elles transparaître des zones de fragilité du système telles celles dégagées dans les travaux coordonnés par Lucci et Millet [1] ? C'est ce que nous avons cherché à savoir en proposant à des étudiant-e-s en préparation au Concours de Recrutement des Professeurs des Écoles (CRPE) la dictée *Les arbres*. Celle-ci a fait l'objet de trois enquêtes de grande ampleur étalées de 1873 à 2005 [2, 3] qui ont permis de dresser un bilan comparatif des connaissances orthographiques des élèves du Cours Moyen 2<sup>o</sup> année à la 3<sup>o</sup> de collège sur plus de cent trente ans.

L'étude conduite ne saurait se situer dans la lignée de ces enquêtes antérieures : non seulement elle n'a aucune visée historique mais elle ne prétend pas décrire l'orthographe de futur-e-s formateur-ice-s, description qui requerrait un protocole différent et surtout d'autres types de tâches, plus proches de celles auxquelles les participant-e-s sont habitué-e-s dans le contexte universitaire. Son intérêt va plutôt vers les zones de fragilité du système orthographique que Pach, Jacquemin, Millet et Billiez [4 : p. 45] ont identifiées à partir d'un corpus d'écrits produits par des adultes « *« ordinaires » n'ayant aucune relation particulière avec l'orthographe [et de] « futurs professionnels », étudiants en licence de lettres et élèves de lycée professionnel se préparant au métier de secrétaire* ». Selon ces travaux, ces zones de fragilité ne relèvent ni des usagers ou usagères ni des situations d'écriture mais sont inhérentes au système de l'orthographe du français. Elles se manifestent à la faveur des variations orthographiques contenues dans les productions, l'utilisation de ce concept en lieu et place de termes tels que « erreur » ou « faute » présentant l'intérêt d'éviter tout jugement ou toute connotation morale et « *permet[tant] de rattacher la problématique de la faute à celle, plus générale, de la variabilité qui accompagne la mise en œuvre des systèmes symboliques (comme on parle, par exemple, de variation phonétique)* » [5 : p. 33]. Si les zones de fragilité dépendent du système orthographique, elles devraient également transparaître dans les productions d'adultes se

destinant à l'enseignement primaire, en dépit de situations scripturales différentes : alors que Pach *et al.* [4] ont analysé un corpus de demandes d'emploi, correspondance libre, rendus d'examen et/ou prises de notes, nous avons utilisé une situation de dictée contrainte pour le présent travail.

Cette étude, inscrite dans un contexte de formation professionnelle, a pour objet le domaine peu exploré [6] de l'orthographe étudiante. Quelques recherches y ont pourtant été consacrées dans le monde francophone. Au Québec, Simard [7] cherche à identifier les écueils rencontrés en orthographe d'usage à partir de trois tâches (identification d'erreurs, correction de textes et rédaction) à effectuer en verbalisant l'activité intellectuelle. Lefrançois [8] évalue la compétence orthographique de futur·e·s enseignant·e·s par l'inventaire de leurs stratégies lorsque des problèmes d'orthographe doivent être résolus en correction ou en rédaction de textes. Pour ce faire, elle demande aux étudiant·e·s des commentaires et/ou des justifications sur leurs productions, ce qui la conduit à proposer un programme de remédiation. En Belgique, Legros [9] interroge la pertinence d'un socle de compétences orthographiques pour l'enseignement supérieur. À partir d'un corpus de productions écrites, il pointe les conditions nécessaires à l'amélioration de ces compétences. Enfin, dans une étude française plus récente, David [10] se penche sur les difficultés orthographiques d'étudiant·e·s de 1<sup>ère</sup> année de lettres modernes identifiées à partir d'un corpus de productions écrites et d'entretiens semi-directifs portant sur les pratiques d'écriture, sur lesdites difficultés et sur le vécu orthographique des participant·e·s. Il suggère trois axes qui permettraient d'avancer sur la question des écarts orthographiques : l'axe institutionnel lié à la survalorisation de la norme ; l'axe de la recherche, sa mise en relation avec la formation des enseignant·e·s et l'implication de la sociolinguistique ou de la psychologie sociale ; l'axe des apprentissages par un travail portant et sur les connaissances et sur les procédures de contrôle et de révision.

Notre étude n'a pas l'objectif des travaux mentionnés. Elle s'intéresse moins aux compétences des étudiant·e·s qu'aux zones de fragilité du système. Elle ne s'appuie ni sur une tâche de production écrite autonome ni sur la verbalisation de l'activité intellectuelle mais sur une unique tâche de dictée contrainte qui limite l'accès aux savoirs et aux stratégies des participant·e·s [7] et dont les conditions de production ont été conçues afin de favoriser l'émergence de variations. Le protocole que nous avons conçu répond donc à des objectifs de recherche plus que de formation.

On peut toutefois imaginer que le fait que de futur·e·s professeur·e·s des écoles se trouvent confronté·e·s à une tâche mettant en évidence les zones de fragilité du système contribue à une prise de conscience de leur part des spécificités de l'orthographe du français et des composantes de ce même système. Cette mise en situation nous paraît ainsi pertinente pour les aider à modifier leurs conceptions de l'orthographe comme objet de formation, dans la perspective de réfléchir à la mise en place d'un enseignement rénové cohérent dans leur future classe. En effet, vivre une situation de dictée contrainte les met en position de s'interroger sur cette activité et de dégager des indicateurs de complexité.

Au cours d'une séance de remise à niveau en orthographe en vue de la préparation au CRPE, les étudiant·e·s ont été invité·e·s à réfléchir à l'orthographe du français, à son enseignement et à leurs propres pratiques orthographiques en étant placé·e·s dans la situation d'élèves avec le texte *Les Arbres* à produire individuellement sous dictée contrainte. Les productions échangées ont été analysées à l'aide de la *Grille typologique des erreurs d'orthographe* [11 : pp. 281-283] et de son adaptation pour la classe [12], à la suite de quoi l'orthographe du texte a été collectivement examinée et ses obstacles potentiels identifiés.

Nous avons choisi la dictée au détriment d'une approche intégrative de l'enseignement de l'orthographe qui permet de travailler cette dimension de la langue lors de la production de textes [13]. La raison en est que, pour aussi scolaire et artificielle qu'elle soit et malgré

les réserves qu'elle suscite, la dictée reste emblématique de l'école [14] de l'orthographe à l'école [3] et peut-être même de l'orthographe tout court [15].

Dans un premier temps sera retracée la carrière du texte utilisé dans notre protocole dans l'étude du niveau orthographique. Seront ensuite présentées les contraintes auxquelles la production a été soumise ainsi que les caractéristiques linguistiques de l'extrait dicté. Puis, les aspects méthodologiques seront développés avant que ne soient présentés les résultats. Ces derniers seront discutés dans une dernière partie qui proposera également quelques perspectives de recherches futures et quelques pistes pour la formation des enseignant·e·s.

## 1 *Les Arbres*, un texte au brillant avenir

Le 9 mai 1873, à l'école des frères de la Doctrine chrétienne de Bayonne, était dicté pour la première fois, sans titre et sans mention de l'auteur, l'extrait d'un texte écrit par Fénelon entre 1701 et 1712, le *Traité de l'existence de Dieu*.

Les 59 copies recueillies étaient les premières d'un corpus promis à une longue carrière. Ce jour-là, Gaspard Beuvain d'Altenheim, Inspecteur Général de l'Instruction Primaire, ne pouvait se douter que cent dix et cent trente ans plus tard d'autres élèves produiraient le même texte dans les mêmes conditions ni que cent quarante ans plus tard, la même épreuve serait proposée à des adultes diplômé·e·s. Il savait toutefois que sa dictée deviendrait itinérante et que, par ses soins, des milliers d'élèves la produiraient aux quatre coins du pays, à la ville et à la campagne. De 1873 à 1877, des paquets de copies corrigées accompagnées de rapports d'inspection sur les écoles ayant accueilli l'inspecteur arrivèrent sur les bureaux des différents ministres.

Ce qu'il reste de ce double corpus, aujourd'hui conservé aux Archives Nationales, est ce fonds exploité par Chervel et Manesse pour en faire un « *témoignage global sur le niveau des connaissances en orthographe des élèves [...] dans les années qui précèdent Jules Ferry* » [2 : p. 10]. Plus encore, les deux chercheurs se sont appuyés sur ce corpus initial pour mener une comparaison historique à plus d'un siècle de distance en constituant un second corpus avec la collaboration de trois mille élèves de même âge que celles et ceux du siècle précédent. De décembre 1986 à juin 1987, les quatre phrases qui composent l'extrait de Fénelon (intitulé pour l'occasion *Les Arbres*) ont été produites dans les mêmes conditions qu'au temps de l'Inspecteur Beuvain. L'échantillon ainsi recueilli devait refléter le niveau en orthographe des élèves de la fin des années 80. En 2005, Manesse et ses collaboratrices repartaient sur les routes de France afin de constituer un deuxième corpus contemporain [3].

## 2 *Les Arbres* : caractéristiques linguistiques et conditions de production

### 2.1 Le texte

Les arbres s'enfoncent dans la terre par leurs racines // comme leurs branches s'élèvent vers le ciel. // Leurs racines les défendent contre les vents et vont chercher, // comme par de petits tuyaux souterrains, // tous les suc destinés à la nourriture de leur tige. // La tige elle-même se revêt d'une dure écorce // qui met le bois tendre à l'abri des injures de l'air. // Les branches distribuent en divers canaux la sève // que les racines avaient réunie dans le tronc.

Les doubles barres obliques correspondent aux pauses aménagées lors de la dictée du texte.

## 2.2 Analyse

L'extrait dicté est bref. Les 83 mots graphiques qui le constituent se répartissent en 36 mots de classe ouverte et 47 mots de classe fermée, classes qui correspondent approximativement aux mots lexicaux et aux mots grammaticaux [16]. Certains mots peuvent y apparaître plusieurs fois, qu'ils soient lexicaux (*e.g.* le substantif *racines*) ou, de façon plus prévisible, grammaticaux (*e.g.* le déterminant possessif *leurs*). Dans ce cas, chaque occurrence est suivie de son ordre d'apparition (*e.g.* *comme\_1* correspond à la première des deux occurrences de la conjonction *comme*).

Les écueils qu'a pu engendrer la production sous dictée de ce texte découlent en partie des impératifs auxquels ont été soumises les conditions de production, destinées à favoriser l'émergence de variations orthographiques et donc différentes d'une dictée traditionnelle. Dans le rituel que constitue celle-ci, on distingue 4 phases : 1) la réception auditive d'un stimulus oral, 2) la réponse consistant en une production écrite concomitante à la réception de parties du même stimulus dictées au moins deux fois, 3) la mise en correspondance de la production écrite et du stimulus oral qui assure une première révision et 4) la seconde révision de la production en l'absence du stimulus. Afin de recueillir le plus grand nombre possible de variations et d'éviter leur filtrage à posteriori, il a été décidé de complexifier la tâche : les participant·e·s ont dû écrire le texte au fur et à mesure de sa dictée, sans exposition préalable, sans répétition et sans possibilité de révision.

À ces obstacles liés aux conditions de production s'ajoutent les difficultés linguistiques inhérentes au texte et qui relèvent de la langue et de son orthographe : contraintes sémantiques d'une langue classique, entraves possibles à la compréhension y compris chez des adultes diplômé·e·s, zones de fragilité du système orthographique du français, effets de fréquence lexicale.

### 2.2.1 Contraintes sémantiques

Manesse *et al.* ont pointé la spécificité de la langue classique dont « *le propre [...] est qu'elle n'affirme pas ouvertement son étrangeté comme le fait, par exemple, l'ancien français* » [3 : p. 60].

Bien que les adultes diplômé·e·s de cette étude soient supposé·e·s connaître tous les mots du texte, les commentaires des difficultés relevées par les auteurs gardent toute leur pertinence. En effet, il n'est pas certain que la connaissance de ces mots soit celle du sens qu'ils avaient du temps de Fénelon, ce qui peut entraver la compréhension et, par ricochet, la performance orthographique. Nous reprenons dans leur intégralité les commentaires de Manesse *et al.* [3 : p. 61] qui reprennent en partie ceux de Chervel et Manesse [2 : pp.30-31] :

“comme leurs branches...” (première phrase) : aucun problème particulier d'orthographe ; mais ce type de propositions de comparaison est déroutant (quel sens donner à *comme* ?).

“leur tige” : qu'est-ce qu'une tige d'arbre ? se demandera peut-être encore l'élève [ou l'étudiant·e]. Entre le tronc, la branche, le rameau, il pourra hésiter. D'autant que l'usage a tranché depuis longtemps. On n'utilise plus le mot *tige* pour désigner le tronc comme le faisait encore, en 1694, le Dictionnaire de l'Académie : “la partie de l'arbre ou de la plante qui sort de la terre, & qui porte les branches et les feuilles.” Littré lui-même, malgré son goût prononcé pour l'archaïsme, avait rendu compte de ce changement.

“se revêt” : le verbe est rare et on en verra le témoignage dans la fuite massive vers un quasi-homophone...

“une dure écorce” : même incompréhension, en français, on dit “une écorce dure” et l'antéposition de l'adjectif a posé problème.

“des injures de l'air” : curieuse acception du mot pour les élèves...

“distribuer, réunir de la sève” : expressions pour le moins déconcertantes, ces deux verbes étant habituellement employés avec des objets ou des personnes qu’on peut compter, ce qui n’est pas le cas de la sève ! Même perplexité, chez les élèves, concernant les sujets : comment les branches pourraient-elles “distribuer”, et les racines “réunir” quoi que ce soit ?

“en divers canaux” : double difficulté. Canal est technique ou désuet dans ce sens. Quant à divers, il est ici impropre et ne peut que troubler les élèves, ceux du moins qui maîtrisent le mot : il signifie “plusieurs, pris sur un ensemble plus large”, et l’on voit mal pourquoi les branches ne diffuseraient pas la sève sur tous les canaux.

### 2.2.2 Les zones de fragilité du système graphémique du français

Utiliser l’appellation « système graphémique du français » permet de considérer le graphème comme unité minimale de la variation orthographique. Toutefois, cette appellation reste impropre puisqu’elle laisse entendre que les graphèmes formeraient un système unique. Or, comme le posent Catach *et al.* [11], ces derniers ont des fonctions qui les différencient et qui les renvoient à trois sous-systèmes dont la conjonction constitue non pas un système uniforme mais un plurisystème. Certains d’entre eux, les phonogrammes (*e.g.* *m*, *ai*, *s* et *on* dans *maison*), servent à transcrire les phonèmes et ont une valeur phonographique. D’autres, qui « fonctionnent “pour l’œil” » [5 : p. 36], ont une valeur visuographique et se départagent en deux sous-systèmes : 1) le sous-système sémiographique (ou idéographique) constitué de ceux qui véhiculent du sens, à savoir les morphogrammes qui transcrivent des morphèmes (*e.g.* le *d* de *grand*, le *s* marque du pluriel dans *poissons*) et les logogrammes qui servent à différencier des homophones (*e.g.* *sain*, *saint*, *cinq*, *seing*, *sein*, *ceint*), et 2) le sous-système mutographique regroupant les lettres historiques et étymologiques qui n’ont aucune fonction dans l’orthographe contemporaine, graphèmes muets sans valeur sémiographique (*e.g.* le *h* dans *huile*) parmi lesquels on trouve la plupart des consonnes géminées (*e.g.* le *pp* dans *appeler*).

Ce plurisystème se caractérise par des zones de fragilité orthographique qui concentrent les variations. Selon Pach *et al.* [4], ces dernières relèvent, par ordre décroissant, des diacritiques et des auxiliaires d’écriture, de la morphographie verbale (*e.g.* confusion de temps, de modes ou de personnes dans la flexion), de la morphographie non verbale (*e.g.* omission du pluriel dans le groupe nominal), des logogrammes, des consonnes géminées et des phonogrammes.

Toutes ces zones de fragilité sont présentes dans le texte *Les Arbres*.

Les diacritiques et les auxiliaires d’écriture y apparaissent inégalement : la cédille et le tréma en sont absents alors que les accents grave, aigu et circonflexe, l’apostrophe, les majuscules et le trait d’union y sont représentés (*e.g.* *élèvent*, *revêt*, *s’*, *Leurs*, *elle-même*). Ils relèvent des trois sous-systèmes : phonographique (*e.g.* l’accent aigu de *é* dans *réunie* qui renvoie au phonème [e]), sémiographique (*e.g.* l’accent grave de *à* qui différencie la préposition du verbe *avoir*) et mutographique (*e.g.* la majuscule du *l* dans *Les* qui marque le début d’une phrase).

La morphographie verbale renvoie souvent à la morphologie silencieuse du français et se manifeste par les flexions du verbe dans la relation qu’il entretient avec son sujet (*e.g.* *Les branches distribuent*). Elle apparaît également dans l’obstacle que représente l’accord au féminin du participe passé *réunie*, régi par le complément d’objet direct antéposé *la sève*. Enfin, l’infinitif *chercher* peut lui aussi être source de variations avec sa finale en /E/ dont l’homophonie avec d’autres formes du même verbe constitue une des plus grandes difficultés de l’orthographe du français [17, 18].

La morphologie silencieuse du français concerne également la morphographie non verbale qui relève des accords en genre et en nombre dans les syntagmes nominaux (*e.g.*

*une dure écorce, de petits tuyaux souterrains*) tout comme des lettres dérivatives muettes (e.g. *vents, petits, bois, tronc*).

Les logogrammes exclusivement orthographiques (i.e. à l'exception de ceux qui impliquent des diacritiques ou des auxiliaires d'écriture) sont au nombre de dix (*dans, terre, par, vers, vents, et, met, air, en* et *canaux*).

Le phénomène des consonnes géminées apparaît dans 5 mots du texte : *terre, comme* (deux occurrences), *souterrains* et *nourriture*.

Le dernier secteur à considérer est celui des phonogrammes. Majoritaire chez Catach *et al.* [11], il représente chez Pach *et al.* [4] la zone de fragilité qui concentre le moins de variations. Ce secteur est à associer avec la consistance phonographémique, notion qui renvoie à la relation que les phonèmes entretiennent avec les graphèmes et variable sous-lexicale reconnue pour avoir une influence considérable sur les productions orthographiques [19]. Plus le nombre de graphèmes candidats à la transcription d'un même phonème est réduit et plus la relation est consistante. Ainsi, selon Catach *et al.* [11], la correspondance phonographémique (CPG) [p] → *p* ou *pp* est plus consistante que [i] → *i, y, î, î, hi* ou *ee*. Par ailleurs, la fréquence de la CPG joue également un rôle important : par exemple [o] → *o* (75% d'utilisation) plus fréquente que [o] → *eau* (3% d'utilisation). Les graphèmes qui composent les 33 mots lexicaux de la dictée appartiennent en grande majorité aux CPG les plus fréquentes (ne sont ici considérés que les phonogrammes, qui ont un équivalent sonore) et seules six CPG minoritaires peuvent être recensées dans un ensemble de douze mots (Tableau 1).

**Tableau 1.** Correspondances phonographémiques minoritaires dans les mots de l'extrait dicté.

Mots du texte	C.P.G. minoritaire
<i>air</i>	[ɛ] → <i>ai</i>
<i>revêt</i>	[ɛ] → <i>ê</i>
<i>canaux – tuyaux</i>	[o] → <i>aux</i>
<i>ciel – enfoncent – racines – écorce</i>	[s] → <i>c</i>
<i>nourriture – souterrains – terre</i>	[R] → <i>rr</i>
<i>souterrains</i>	[ɛ̃] → <i>ain</i>

### 2.2.3 Les mots du texte et leur fréquence

La fréquence lexicale pouvant avoir un impact sur la production orthographique, y compris en dictée adulte [20], nous avons cherché celle des 33 mots de classe ouverte. Selon la base de données *Lexique 3.8* [21] et plus particulièrement le corpus de livres (par million d'occurrences), tous ces mots affichent des fréquences très variées qui s'étalent de 3,18 pour *suc* à 2854,93 pour *vont*. Ces fréquences sont celles des lemmes pour 32 mots, le lemme étant l'unité linguistique non soumise à des phénomènes de flexion et pouvant constituer une entrée de dictionnaire (e.g. l'infinitif pour les verbes : *distribuer* pour *distribuent* ; le singulier pour les noms : *vent* pour *vents* ; le masculin singulier pour les adjectifs : *dur* pour *dure*). La forme orthographique *vont* étant trop éloignée de son lemme *aller*, c'est sa fréquence et non pas celle de son lemme qui a été prise en compte.

### 3 Méthode

#### 3.1 Participant-e-s et recueil des données

Les productions ont été recueillies par des formateurs et formatrices en didactique du français de l'ESPE que nous remercions, auprès de 220 étudiant-e-s inscrit-e-s en 1<sup>ère</sup> ou 2<sup>o</sup> année du Master MEEF 1<sup>er</sup> Degré (Métiers de l'Enseignement, de l'Éducation et de la Formation). Nous ne sommes pas en mesure de fournir d'indications relatives ni à leur âge ni à leur genre d'appartenance.

#### 3.2 Procédure

Avant l'activité proprement dite, les étudiant-e-s ont été informé-e-s des conditions de passation. L'extrait a été dicté en continu à débit moyen, sans répétition (nécessité éventuelle de sauter des mots) mais avec aménagement de pauses aux virgules et aux points (mentionnés lors de la dictée) et après certains syntagmes (pauses symbolisées, dans le texte présenté ci-dessus, par les doubles barres obliques).

Afin d'éviter tout biais relatif à l'attention, le projet de constitution et d'analyse du corpus n'a été présenté qu'après la réalisation de la tâche. La dictée a alors été contextualisée en référence aux trois enquêtes préalables, sans qu'il soit fait mention des zones de fragilité du système orthographique. L'analyse des performances sous dictée contrainte d'adultes diplômé-e-s et l'intention de recueillir des variations spontanées et non soumises à révision dans les productions d'un texte à priori exempt de difficultés orthographiques ont été les objectifs présentés. Enfin, après avoir garanti l'anonymat des productions, il a été précisé que le corpus serait constitué sur la base du volontariat.

#### 3.3 Données analysées

Sur les 220 productions initialement prévues, 11 ont été écartées pour non respect du protocole. Au final, ont donc été analysées 209 dictées.

Des 17347 formes attendues, 1573 ne correspondent pas aux mots du texte. Parmi ces dernières, 252 n'ont pas été produites, effet probable des conditions de production ou conséquence d'une arrivée tardive à la séance. Elles n'ont pas été prises en compte tout comme les 30 formes tronquées à la reproduction après photocopie, les 7 formes abrégées (*e.g.* *ds* pour *dans*) et les 136 formes différentes (*e.g.* *plusieurs* pour *divers*). L'analyse conduite s'appuie ainsi sur 16922 mots, dont 1148 formes contenant une ou plusieurs variations relevant soit des graphèmes soit des diacritiques et des signes d'écriture.

#### 3.4 Typologie des variations

La typologie des variations établie par Lucci et Millet [1] dans leur ouvrage préfacé par Nina Catach distingue les variations relatives aux diacritiques et aux auxiliaires d'écriture (Figure 1) des variations orthographiques (Figure 2) qui ne concernent que les graphèmes. Une telle distinction est certes discutable dans la mesure où diacritiques et auxiliaires ne peuvent s'envisager indépendamment des lettres avec lesquelles ils constituent le graphème. Certes, elle n'apparaît pas dans la typologie de Catach *et al.* [11], référence privilégiée par les didacticien-ne-s et dans laquelle les variations sur les diacritiques seraient réparties dans différentes rubriques. Toutefois, elle a l'avantage de mettre en lumière ce secteur traditionnellement considéré comme peu important ainsi que le rappelle Catach dans sa préface. Il nous a semblé nécessaire de sensibiliser les futur-e-s

formateur-riche-s à tous les secteurs de la variation dont l'importance est certes inégale en fonction des situations de communication mais dont aucun ne doit être négligé dans l'exercice de leur futur métier. La typologie de Lucci et Millet [1] nous paraît plus à même de répondre à notre objectif qui n'est pas d'analyser les difficultés orthographiques des étudiant-e-s mais de vérifier que les zones de fragilité du système sont indépendantes des utilisateur-riche-s et des situations d'écriture, ainsi que l'avancent Pach *et al.*[4]. C'est pourquoi les deux grilles ont été reprises dans la présente étude avec quelques adaptations mineures requises par les variations obtenues.

La grille des variations diacritiques (VD) affiche l'omission (*e.g.* \**grace* pour *grâce*), l'adjonction (*e.g.* \**merçi* pour *merci*), la substitution (*e.g.* \**èventuel* pour *éventuel*) et la sélection (*e.g.* *chaîne* ou *chaîne*) qui tient compte des *Rectifications de l'orthographe française* de 1990 et du choix qu'elles laissent aux usagers et aux usagères de l'orthographe de certains mots. A ces quatre modalités s'ajoute la neutralisation, phénomène qui implique un « *accent tracé à la verticale ou à l'horizontale, ou comme un point* » [5 : p. 37]. La sélection n'a pas été retenue puisqu'aucun mot du texte ne laisse le choix d'une forme alternative. Par contre, ont été rajoutées l'adjonction d'un blanc graphique (*e.g.* \**sous terrain* pour *souterrains*) et la neutralisation de l'accent circonflexe (*e.g.* \**même* pour *même*).

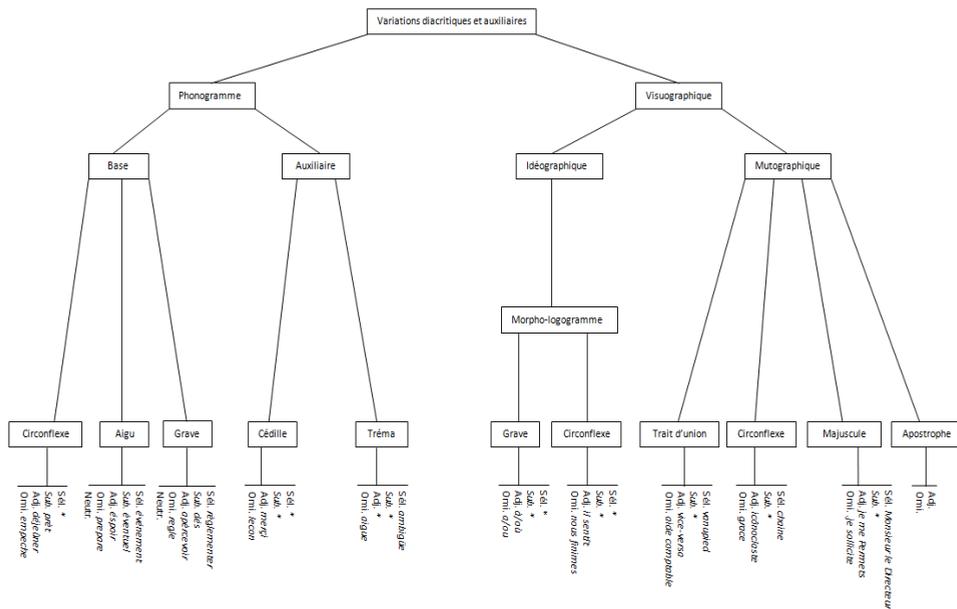
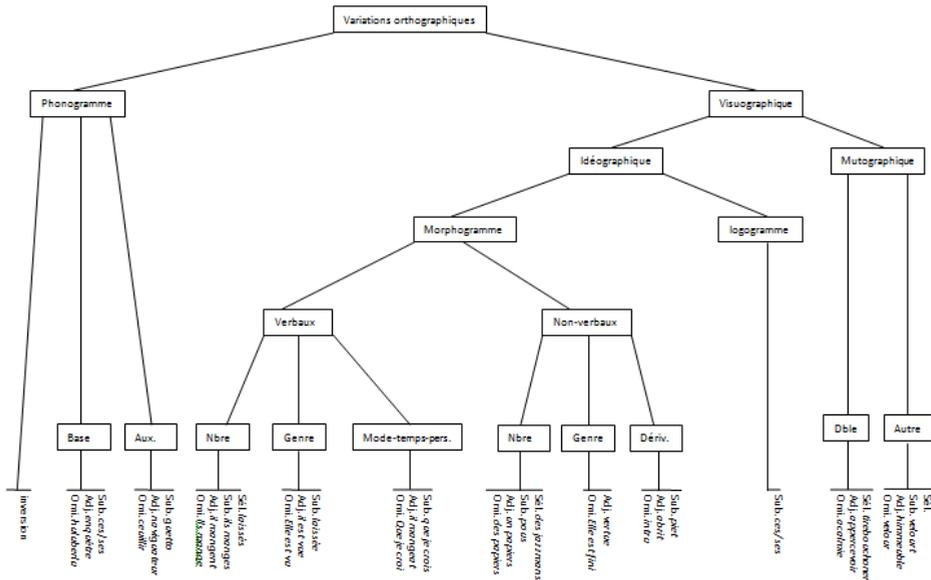


Fig. 1. Variations diacritiques et auxiliaires d'écriture [1 : p. 248].

La grille consacrée aux variations orthographiques (VO) recueille l'omission (*e.g.* \**velour* pour *velours*), l'adjonction (*e.g.* *il est* \**vue* pour *il est vu*), la substitution (*e.g.* \**navigateur* pour *navicateur*) et la sélection (*e.g.* *des jazzmans* ou *des jazzmen*) qui n'a pas été retenue. La substitution du genre par le nombre (\**réunis* pour *réuni*) a été rajoutée pour les morphogrammes verbaux.



**Fig. 2.** Variations orthographiques [1 : p. 247].

A cause de leur poids important, nous avons complété ces deux types de variations par une catégorie que nous avons appelée variations sémantiques (VS). Elle regroupe les variations qui ne peuvent véritablement être considérées comme des VO. Ainsi, dans le syntagme *\*Leurs racine*, le substantif contient une VO par omission du morphogramme non verbal de nombre. *A contrario*, les syntagmes *\*Leur racine* (pour *Leurs racines*), *\*le vent* (pour *les vents*) ou *\*leurs tiges* (pour *leur tige*) dans lesquels les chaînes d'accord sont respectées ne contiennent aucune variation purement orthographique : ils ne font que s'éloigner de la norme représentée par l'extrait de Fénelon, conséquence probable de la mauvaise compréhension par les participant·e·s de la langue classique utilisée et/ou des conditions de production.

## 4 Résultats

### 4.1 Données quantitatives

Tous les mots de l'extrait ont subi une transformation pour l'ensemble du corpus, qu'ils n'aient pas été produits, qu'ils aient été abrégés ou remplacés par d'autres ou encore qu'ils contiennent des variations.

Ces derniers atteignent un taux de 6,78% qui correspond au nombre de mots contenant au moins une variation sur le nombre total de mots pris en compte. Il est supérieur aux 5,35% obtenus par Pach *et al.* [4] dans le corpus de prises de notes et/ou de rendus d'examen des futur·e·s enseignant·e·s ou secrétaires et cette supériorité numérique pourrait s'expliquer par l'impossibilité pour les participant·e·s de la présente étude de réviser leurs productions.

Les variations, pour leur part, s'élèvent à 1270 et se répartissent en 34,96% de variations diacritiques (VD), 37,87% de variations orthographiques (VO) et 27,17% de variations sémantiques (VS).

## 4.2 Aspects qualitatifs de la variation

Nous reprenons l'ordre des variations tel qu'il apparait dans la section précédente. Ce choix ne préjuge pas de l'importance de telle ou telle catégorie, l'analyse du corpus ne pouvant s'appuyer que sur les données qu'il délivre. Ainsi, l'équivalence des pourcentages de VD (34,96%) et de VO (37,87%) ne peut être perçue comme le reflet des compétences orthographiques de futur·e·s formateur·rice·s : elle correspond à des données recueillies au moyen d'une tâche spécifique, dans des conditions qui ne le sont pas moins et ne saurait constituer une description exhaustive de ces compétences.

### 4.2.1 Variations diacritiques (VD)

Ainsi que l'indique le tableau 2, les VD à valeur phonographique (81,53%) concernent principalement les phonogrammes de base (78,15%) : celles qui relèvent de l'accent circonflexe (9,01%) apparaissent sur trois mots, *revêt*, *même* et *dure*, par omission (\**revet*), substitution (\**revêt*, \**même*), neutralisation (\**revêt*, \**même*) et adjonction (\**rêvet*, \**dûre*) ; l'accent aigu affiche le plus grand pourcentage (58,78%) avec, dans l'ordre décroissant d'importance numérique, des neutralisations (e.g. \**ēcorce*) et des substitutions (e.g. \**dēfendent*), quelques omissions (e.g. \**reunie*, \**destines*) et de façon plus périphérique des adjonctions (e.g. \**dēstinés*) ; celles liées aux accents graves phonographiques (10,36%) correspondent à des neutralisations (e.g. \**sēve*), des omissions (e.g. \**élevet*), des substitutions (e.g. \**sēve*) ou des adjonctions (e.g. \**dēstinés*). Les phonogrammes auxiliaires présentent aussi des variations qui ne concernent que la cédille (3,38%) ajoutée au mot *enfoncent* (\**enfonçent*).

Les VD à valeur visuographique (18,47%) représentantes du sous-système sémiographique (0,45%) ont trait à l'omission à deux reprises de l'accent grave idéographique dans le morpho-logogramme à.

**Tableau 2.** Répartition des variations diacritiques selon leur sous-système d'appartenance.

Phonogrammes de base	Accent circonflexe	9,01%	78,15%	81,53%
	Accent aigu	58,78%		
	Accent grave	10,36%		
Phonogrammes auxiliaires	Cédille		3,38%	
Morpho-logogrammes	Accent grave	0,45%		
Mutogrammes	Blanc graphique	0,68%	18,02%	18,47%
	Trait d'union	8,11%		
	Majuscule	9,23%		

Celles qui appartiennent au sous-système mutographique (18,02%) se concentrent sur le mot *souterrain*, le seul à avoir été segmenté par un blanc graphique (0,68%) ou un trait d'union (8,11%). Les 27 occurrences de \**elle même* dans lesquelles le trait d'union a été omis n'ont pas été prises en compte dans le calcul des pourcentages dans la mesure où nous avons choisi le mot pour unité d'analyse. Les autres mutogrammes concernent les majuscules (9,23%) qui apparaissent principalement dans le mot *comme* (\**Comme*), c'est-à-dire au début d'une circonstancielle comparative dépendant d'une proposition principale [22] à l'issue de laquelle une pause a été aménagée lors de la dictée. Plus rarement, les majuscules peuvent aussi ne pas être présentes en début de phrase (e.g. \**les*, \**leurs*, \**la*).

#### 4.2.2 Variations orthographiques (VO)

Comme le montre le tableau 3, les VO à valeur phonographique sont largement minoritaires (18,71%) comparativement à celles dont la valeur est visuographique (81,29%).

Une seule inversion de phonogrammes dans le corpus représente 0,21% des VO : *\*nourrutire*. Les variations à valeur phonographique ont trait, dans leur grande majorité, aux phonogrammes de base (18,50%) : il s'agit principalement de substitutions de phonogrammes (e.g. *ch* ou *que* pour *c* dans *sucs*), mais on trouve aussi des omissions et des adjonctions altérant la valeur phonique du mot (e.g. *\*checher*, *\*tyaux*, *\*sousterrains*).

**Tableau 3.** Répartition des variations orthographiques selon leur sous-système d'appartenance.

Phonogrammes		0,21%	18,71%			
	Base	18,50%				
Morphogrammes verbaux	Genre/Nbre	5,41%	33,26%	70,48%	71,52%	81,29%
	Nbre	10,19%				
	Genre	14,76%				
	Mode-temps-pers.	2,91%				
Morphogrammes non verbaux	Nbre	16,01%	37,21%	1,04%	9,77%	
	Genre	0,83%				
	Dériv.	20,37%				
Logogrammes						
Mutogrammes	Dble					

Parmi les variations à valeur visuographique, celles qui relèvent du sous-système sémiographique (71,52%) l'emportent amplement sur celles qui appartiennent au sous-système mutographique (9,77%). La plupart du temps, ce sont des morphogrammes (70,48%), les logogrammes affichant la valeur périphérique de 1,04% atteinte par une occurrence de *\*mais* pour *met*, une occurrence de *\*aire* pour *air* et trois occurrences de *\*canots* pour *canaux*. En ce qui concerne les morphogrammes verbaux (33,26%), l'accord du participe passé *réunie* avec le complément d'objet direct antéposé concentre un nombre considérable de variations. La forme *\*réunis* a été considérée comme consécutive à la substitution de la marque du genre par celle du nombre et représente, avec ses 26 occurrences, 5,41% des VO. De son côté, la forme concurrente *\*réunies*, obtenue par l'adjonction d'un *s*, a été classée dans la catégorie « Nombre » (10,19%) au même titre que *\*enfonce*, *\*défende*, *\*distribue* et *\*avait* (omission de la marque du pluriel). On peut penser que, pour cette forme, la variation est imputable à une méconnaissance des règles d'accord du participe passé employé avec l'auxiliaire *avoir*, accordé avec le sujet. La forme *\*distribus* a été classée dans la même catégorie même s'il s'agit moins ici d'une variation relevant de la flexion – dans la mesure où la forme n'existe pas dans le paradigme verbal – que du marquage catégoriel avec une confusion entre deux classes grammaticales (substitution d'une marque du pluriel verbal par celle du pluriel nominal). Les pourcentages les plus élevés s'affichent sur le genre (14,76%) avec l'omission de la marque du féminin dans *\*réuni*, due soit à une réelle omission de la marque soit à une erreur sur l'accord du participe passé avec le complément d'objet direct *que*, lui-même non marqué en genre. Enfin, les 2,91% de la catégorie « Mode-Temps-Personne » complètent les

morphogrammes verbaux : il s'agit de substitutions (e.g. \**défendre*), d'adjonctions (e.g. \**revêts*) ou d'omissions (e.g. \**distribu*). Les variations sur les morphogrammes non verbaux (37,21%) concernent les marques du nombre (16,01%) qui ont été omises (e.g. \**arbre*), ajoutées (e.g. \**terres*) ou substituées par confusion de marquage catégoriel (les 3 occurrences de \**branchent*), et de façon marginale, les marques du genre (0,83%) par omission (e.g. \**dur*) ou adjonction (e.g. *petite*). Le pourcentage le plus élevé de cette catégorie correspond aux lettres dérivatives (20,37%) et est presque entièrement lié au mot *abri* auquel une telle lettre a été ajoutée ce qui est compréhensible, étant donné le dérivé *abriter*. Seule une occurrence de \**tron* pour *tronc* (omission) vient compléter la série.

Pour terminer, les variations visuographiques relevant du sous-système mutographique atteignent 9,77% des VO. Une seule catégorie y est représentée, celle des consonnes doubles qui ont été ajoutées deux fois plus souvent qu'elles n'ont été omises (e.g. \**canneaux*, \**souterains*).

#### 4.2.3 Variations sémantiques (VS)

Comme l'indique le tableau 4, les VS pèsent d'un poids lourd sur l'ensemble de la variation. Elles sont toutefois très localisées puisqu'elles se concentrent sur 8 syntagmes nominaux et un accord sujet-verbe. Elles apparaissent majoritairement sur les morphogrammes non verbaux par adjonction (75,94%) ou omission (23,77%) du pluriel. L'omission du morphogramme verbal de nombre (0,29%) correspond au maintien de la chaîne d'accords dans la relation sujet-verbe.

**Tableau 4.** Répartition des variations sémantiques selon leur modalité.

Adjonction d'un morphogramme non verbal de nombre	* <i>leurs tiges</i> pour <i>leur tige</i>	75,36%	75,94%	99,71%
	* <i>les bois tendres</i> pour <i>le bois tendre</i>	0,58%		
Omission d'un morphogramme non verbal de nombre	* <i>leur racine</i> pour <i>leurs racines</i>	3,48%	23,77%	
	* <i>leur branche</i> pour <i>leurs branches</i>	0,58%		
	* <i>Leur racine</i> pour <i>Leurs racines</i>	1,16%		
	* <i>le vent</i> pour <i>les vents</i>	17,39%		
	* <i>de l'injure</i> pour <i>des injures</i>	1,16%		
Omission d'un morphogramme verbal de nombre	* <i>leur branche s'élève</i> pour <i>leurs branches s'élèvent</i>	0,29%		

#### 4.2.4 Mots et variations

Ne sont inventoriées dans cette section que les unités concernées par plus de 1% de variations, exception faite des variations sémantiques.

Le mot qui a donné lieu au pourcentage le plus élevé de variations est, sans véritable surprise, *réunie* (15,59%) avec des VO (11,10%) qui l'emportent largement sur les VD (4,49%) et 198 variations réparties dans 17 formes différentes. Parmi elles, 71 sont des omissions du morphogramme de genre, 41 des adjonctions du morphogramme de nombre et 23 des substitutions du morphogramme de genre par le morphogramme de nombre, les

formes les plus présentes étant les 49 occurrences de \*réuni, les 32 de \*réunies et les 18 de \*réunis. Ces résultats reflètent la complexité de l'accord du participe passé avec le complément d'objet direct antéposé, une des zones les plus problématiques de la morphographie verbale déjà signalée par Chervel et Manesse [2].

Le deuxième mot à avoir recueilli le plus haut pourcentage de variations est *souterrains*, avec une proportion plus grande de VO (5,04%) que de VD (3,07%) et 12 variantes. Les variations les plus fréquentes, dans une quasi égalité numérique, sont les 36 adjonctions du trait d'union et les 35 omissions du morphogramme de nombre, les variantes les plus représentées étant les 18 occurrences de \**sous-terrains* et les 18 de \**sous-terrain*. L'omission de la marque écrite du pluriel pourrait s'expliquer, compte tenu des contraintes de production, par l'absence à l'oral de cette même marque dans l'ensemble du syntagme nominal prépositionnel, absence qui renvoie à la morphologie silencieuse du français. Quant au trait d'union, son adjonction est plus difficile à interpréter : peut-être l'adjectif \**sous-terrain* subit-il une sur-segmentation en morphogrammes consécutive à une prégnance trop forte du substantif *terrain* et de la préposition *sous*, l'adjonction du trait d'union correspondant à une contraction du syntagme nominal prépositionnel *sous le terrain*.

Le mot *abri* arrive en troisième position du classement avec 7,64% de VO, des adjonctions d'une lettre dérivative réparties sur deux seules formes : les 91 occurrences de \**abris* et les 6 de \**abrit*. Ce résultat est à première vue assez surprenant. En effet, nous sommes en présence d'un nom au singulier dont les CPG sont non seulement les plus consistantes mais aussi les plus fréquentes. La variante minoritaire \**abrit* peut être vue comme liée au dérivé *abriter* et serait le résultat d'une construction logique telle que *cout/couter* ou *chant/chanter*. Quant à la forme majoritaire \**abris*, elle aurait pu être considérée comme consécutive à l'adjonction d'une marque de pluriel. Toutefois, cette interprétation ne nous satisfait pas dans la mesure où le *l'* impose un singulier, y compris à l'oral. La fréquence lexicale du lemme est de 56,76 par million selon la base de données *Lexique 3.8* [21] mais celle de *racine* est inférieure (31,96) et pourtant cette dernière unité n'apparaît pas dans le classement. La surreprésentation de \**abris* n'est peut-être donc pas due à des phénomènes de fréquence lexicale. Selon la base de données, parmi tous les mots se terminant par le phonème /i/, 125 finissent par le bigramme *ri* et 151 par le trigramme *ris*. Les premiers affichent une fréquence moyenne dans le corpus de livres de 3,29 occurrences par million alors qu'elle est de 7,57 pour les seconds et de 6,74 pour l'ensemble du corpus. Cette surreprésentation (43,54% des 209 formes produites) pourrait-elle alors s'expliquer par l'impact de la sensibilité aux régularités graphotactiques de la langue tel qu'il a été dégagé dans diverses études [23] ? Plus encore, la prédominance de la forme \**abris* sur la forme \**abrit* n'irait-elle pas dans le même sens que l'étude de Pacton, Fayol et Perruchet [24] selon laquelle les productions orthographiques sont sous la coupe de contraintes graphotactiques même lorsque des règles morphologiques peuvent être appliquées ?

Les positions suivantes dans le classement sont occupées par cinq mots qui contiennent plus de VD que de VO, ce qui confirme l'instabilité de la zone des diacritiques et des auxiliaires d'écriture :

- *écorce* (VD = 4,02% ; VO = 2,05%), *dépendent* (VD = 5,59% ; VO = 0,31%) et *élèvent* (VD = 5,51% ; VO = 0,08%) présentent principalement des VD qui relèvent soit de la neutralisation de l'accent soit de sa substitution. Si les VO de *dépendent* et de *élèvent* sont marginales, les 23 substitutions du phonogramme *ce* dans *écorce* méritent d'être signalées. Lorsqu'elles ont été opérées au profit de *se*, elles pourraient être dues à la prégnance de la CPG la plus fréquente (*i.e.* /s/ → *s*), alors qu'une substitution par *sse* pourrait signer un manque de prise en compte du contexte dans la mesure où les régularités graphotactiques du français ne permettent pas ce type de séquence après le phonogramme *r* ;

- parmi les 8 variantes de *revêt* (VD = 2,36% ; VO = 1,34%), la forme *\*revet* apparaît 16 fois et la forme *\*revait* 14 fois. Cette dernière fait écho aux propos de Manesse *et al.* [3] selon lesquels le verbe est rare et peut être confondu avec le quasi-homophone *rêvait* ;
- *destinés* (VD = 2,83% ; VO = 0,79%) affiche des neutralisations ou des substitutions de l'accent mais aussi des adjonctions d'accents principalement aigus sur la première syllabe (*\*désinés*) qui reflètent aussi un manque de prise en compte des lois de position. Les 10 VO apparaissent en grande majorité dans des formes telles que *\*destinaient*, *\*destiné*, *\*destinées*, illustration supplémentaire de la difficulté à transcrire les finales en /E/.

En queue de peloton arrivent les mots *canaux* (VD = 0% ; VO = 2,83%), *sève* (VD = 1,81% ; VO = 0,63%), *enfoncent* (VD = 1,18% ; VO = 0,47%), *comme\_1* (VD = 1,57% ; VO = 0%) et *tous* (VD = 0,24% ; VO = 1,10%). Les 5 variantes de *canaux* ne contiennent que des VO au nombre de 36, parmi lesquelles 23 doublements de la consonne *n*, 10 substitutions du phonogramme final *-aux* et les 3 substitutions du mot par son logogramme *\*canots*. Avec leurs 23 VD dues principalement à des neutralisations ou des substitutions de l'accent, les 10 variantes de *sève* sont à ranger aux côtés de *écorce*, *dépendent* et *élèvevent* ; les variations les plus fréquentes dans *enfoncent* sont les 15 adjonctions de cédille qui révèlent encore un manque de prise en compte du contexte ; les 20 majuscules ajoutées à *comme\_1* peuvent s'expliquer par la pause qui précède le mot lors de la dictée ; enfin, *tous* apparaît sous deux variantes soit à l'occasion de trois adjonctions d'une majuscule explicable par sa position après une virgule et donc un groupe de souffle lors de la dictée soit suite à 14 omissions du pluriel dans un syntagme nominal où ce dernier ne s'entend que dans le déterminant *les*.

## 5 Discussion

Cette étude avait pour objectif de vérifier que des zones de fragilité orthographique caractérisent l'écriture du français « *indépendamment des scripteurs et des situations scripturales* » [4 : p. 45]. Elle s'appuie sur les variations orthographiques recueillies dans un corpus de dictées réalisées par des étudiant·e·s en préparation au Concours de Recrutement des Professeurs des Écoles. En utilisant le même extrait en français classique de Fénelon, elle emprunte aux travaux de Chervel et Manesse [2] et de Manesse *et al.* [3] la description linguistique du texte dicté et de ses difficultés potentielles. Elle repose également sur l'enquête coordonnée par Lucci et Millet [1] dont elle s'inspire de la méthodologie pour l'analyse des variations. Ainsi reprend-elle la même typologie qui distingue variations diacritiques et orthographiques, distinction plus favorable à l'identification des secteurs fragiles du système qu'à l'analyse des difficultés de ses utilisateur·rice·s.

Afin de favoriser l'émergence de ces zones de fragilité dans des productions d'adultes en passe de devenir des vecteurs de la norme orthographique, nous avons transformé les conditions dans lesquelles se déroulent les dictées dites magistrales en supprimant ou en modifiant les phases de la procédure traditionnelle. Le texte qui n'a pas été lu au préalable a été produit au fur et à mesure de son unique dictée et les productions n'ont pu bénéficier d'aucune révision. Ces différentes contraintes ont sans nul doute eu un impact sur les performances en créant des états de surcharge cognitive [25] chez des participant·e·s qui ont dû à la fois écouter, comprendre, écrire et orthographier sans aucune possibilité de correction consécutive à la tâche. Si la procédure habituelle avait été maintenue, il est fort possible que la variation aurait été moindre. Pour le vérifier, il serait nécessaire de mener une autre étude qui utiliserait le même matériel mais proposerait des conditions différentes : une première condition contrôle qui comprendrait les quatre phases d'une dictée traditionnelle, une seconde condition dans laquelle seule l'exposition préalable à

l'intégralité du texte serait supprimée et une troisième condition qui empêcherait toute révision postérieure à la production. Ainsi seraient analysés les effets de l'exposition préalable et de la révision sur les textes orthographiés. Un tel travail rendrait certainement possible l'identification des erreurs de performance et des erreurs de compétence, à l'instar de l'étude conduite avec des adultes par Bonin, Peereman et Fayol [26] en dénomination écrite d'images. Dans celle-ci, la suppression, dans une des conditions, des contraintes temporelles et la demande faite aux participant·e·s de vérifier et de corriger leurs réponses permettait la délimitation des deux catégories d'erreurs. Son objectif semblable à celui du travail envisagé est toutefois différent de celui qui sous-tend la présente étude sans pour autant que l'un soit exclusif de l'autre. En effet, dans ce travail-ci, nous nous intéressons moins aux compétences et aux performances des individus qu'aux zones de fragilité orthographique inhérentes au système.

Si nous avons suivi à la lettre la méthodologie proposée dans les travaux coordonnés par Lucci et Millet [1] en ne distinguant que VD et VO, les résultats auraient indiqué 34,96% pour les premières et 65,04% pour les secondes. Or, la présence, dans le corpus, d'occurrences telles que *\*leur tiges* et *\*leurs tiges* pour le syntagme *leur tige*, nous a conduit à scinder les VO en deux catégories : les VO relevant spécifiquement de l'orthographe (37,87%) et les VS davantage liées au sens (27,17%). L'intérêt de ces dernières réside dans leur concentration pour plus de 75% d'entre elles sur le syntagme *leur tige* dans la production duquel les variations potentielles avaient déjà été signalées par Chervel et Manesse [2] et Manesse *et al.* [3]. Le poids qu'elles font lourdement peser sur l'ensemble de la variation (20,48%) reflète le décalage entre la langue classique de Fénelon et le français contemporain – décalage renforcé par une découverte du texte concomitante à sa production sous dictée sans exposition préalable – mais pourraient aussi être le résultat de la suppression du processus de révision postérieur à la dictée en tant que telle. On peut en effet raisonnablement penser que lors de cette phase, la reprise du mot *tige* précédé du déterminant *la* dans laquelle la marque du singulier s'entend aurait conduit les participant·e·s à rectifier leur production. La présence de variations de même ordre sur d'autres syntagmes nominaux dans lesquels la marque du nombre s'entend (*le bois tendre*, *les vents*, *de l'injure*) ou ne s'entend pas (*leurs racines*, *leurs branches*) est plus difficile à comprendre sauf à la relier à l'état de surcharge cognitive engendré par les conditions de production.

En ce qui concerne les VD et les VO, leurs proportions inversées selon qu'elles relèvent de la zone phonographique (VD = 81,53% ; VO = 18,71%) ou visuographique (VD = 18,47% ; VO = 81,29%) renvoient au patron similaire de l'enquête de Pach *et al.* [4] : la phonographie pourrait bénéficier d'une vigilance plus grande lorsqu'elle s'exprime par les graphèmes alors que sa prise en charge par les diacritiques et les auxiliaires d'écriture tend à être négligée.

Pach *et al.* [4] avaient également classé par ordre d'importance décroissante les zones de fragilité du système avec les diacritiques et auxiliaires d'écriture, la morphographie verbale, la morphographie non verbale, les logogrammes, les consonnes géminées et les phonogrammes. Si toutes ces zones se manifestent dans la présente étude, elles occupent des proportions différentes. Cette différence partielle de résultats peut s'expliquer par le fait que la dictée donne lieu à des productions contraintes dans lesquelles le texte n'est pas conçu par le scripteur. Aussi sera-t-il intéressant de chercher ces zones de fragilité dans un corpus constitué de productions adultes autonomes telles que des devoirs ou des rendus d'examen. Ce corpus gagnerait à être enrichi des verbalisations relatives à la relation des participant·e·s à l'écriture, selon une méthodologie similaire à celles adoptées par les études précédemment citées [7, 8, 10]. Une telle démarche permettrait d'analyser tant les zones de fragilité du système auxquelles les étudiant·e·s doivent être sensibilisé·e·s que ses difficultés d'utilisation. Elle pourrait ainsi constituer un socle à partir duquel des dispositifs

de remédiation pourraient être envisagés, autres que ceux qui ne conduisent qu'à des impasses [9].

Enfin, la sélection du texte *Les Arbres* répond à deux critères : d'une part, en relevant toutes les difficultés potentielles contenues dans l'extrait, Chervel et Manesse [2] ont mené une analyse linguistique aussi rigoureuse que minutieuse, à même de sensibiliser les étudiant-e-s au caractère indispensable d'un tel travail de préparation au moment de choisir le texte à dicter en fonction des objectifs poursuivis et des notions visées ; d'autre part, la présentation d'une telle analyse a vocation à attirer l'attention des futur-e-s enseignant-e-s sur la nécessité de connaître solidement le système graphémique du français et son histoire, connaissances sans lesquelles nul examen des productions écrites des apprenant-e-s ne peut être raisonnablement envisagé, quelles que soient les tâches qui les actualisent (*i.e.* production de textes, dictée, copie, dénomination écrite d'images, etc.).

Cette situation de dictée telle qu'elle a été proposée dans des séances de formation initiale a conduit les étudiant-e-s à affronter les difficultés et à prendre conscience des zones de fragilité inhérentes au système orthographique du français. Le relevé et l'analyse de leurs propres erreurs leur a permis ensuite de s'interroger sur les raisons pour lesquelles ces erreurs sont produites. Ils peuvent aussi être utilisés pour dresser un portrait des zones de fragilité orthographique, définir un programme de remédiations pour chaque étudiant-e et développer leur vigilance sélective dans la phase de relecture. Cette activité contribue ainsi à une remise à niveau orthographique personnelle. Elle constitue également un préambule à un volet plus directement professionnel portant sur la typologie des erreurs orthographiques d'élèves de différents niveaux scolaires, les indicateurs de complexité dans la maîtrise des savoirs orthographiques tout au long du cursus et une réflexion relative aux types d'activités en orthographe. Enfin, la dictée contrainte conduit à s'interroger sur les finalités de cette tâche et à recenser les différentes modalités de mise en œuvre possibles, telles que proposées dans les recherches didactiques récentes [13, 14]. L'intérêt que peut représenter un tel recueil de données peut ainsi résider non seulement dans ses retombées en formation initiale mais également, pour les futur-e-s enseignant-e-s, dans l'influence qu'il pourra exercer sur leurs pratiques à venir.

## Références

1. V. Lucci, A. Millet, *L'orthographe de tous les jours : enquête sur les pratiques orthographiques des Français* (1994)
2. A. Chervel, D. Manesse, *La dictée : les Français et l'orthographe, 1873-1987* (1989)
3. D. Manesse, D. Cogis, M. Morgans, C. Tallet, *Orthographe : à qui la faute ?* (2007)
4. C. Pach, D. Jacquemin, A. Millet, J. Billiez, *L'orthographe de tous les jours : enquête sur les pratiques orthographiques des Français*, 43-136 (1994)
5. V. Lucci, *L'orthographe de tous les jours : enquête sur les pratiques orthographiques des Français*, 19-41 (1994)
6. C. Brissaud, *Pratiques*, **149-150**, 207-226 (2011)
7. C. Simard, *Revue des sciences de l'éducation*, **21** (1), 145-165 (1995)
8. P. Lefrançois, *La Littéracie. Conceptions théoriques et pratiques d'enseignement de la lecture-écriture*, 233-244 (2004)
9. G. Legros, « *Le français m'a tuer* ». *Actes du colloque « L'orthographe française à l'épreuve du supérieur »*, 69-79 (2006)
10. J. David, *Le français aujourd'hui*, **185** (2), 95-106 (2014)
11. N. Catach, C. Gruaz, D. Duprez, *L'orthographe française* (1995)
12. CRDP de Bretagne, *Lettres ouvertes aux enseignants de français, 14-15* (2001)
13. C. Brissaud, D. Cogis, *Comment enseigner l'orthographe aujourd'hui ?* (2011)
14. D. Cogis, C. Fischer, M. Nadeau, *Glottopol*, **26**, 69-91 (2015)
15. C. Brissaud, C. Mortamet, *Glottopol*, **26**, 2-10 (2015)

16. A. Polguère, *Notions de base en lexicologie* (2002)
17. C. Brissaud, J.-P. Chevrot, P. Lefrançois, *Langue Française*, **151(3)**, 74-93 (2006)
18. J. Gonac'h, *Glottopol*, **26**, 167-177 (2015)
19. P. Bonin, P. Collay, M. Fayol, *L'Année Psychologique*, **108**, 517-546 (2008)
20. B. Rapp, C. Epstein, M.-J. Tainturier, *Cognitive Neuropsychology*, **19 (1)**, 1-29 (2002)
21. B. New, C. Pallier, L. Ferrand, R. Matos, *L'Année Psychologique*, **101**, 447-462 (2001)
22. M. Riegel, J.-C. Pellat, R. Rioul, *Grammaire méthodique du français* (2009)
23. J. Danjon, S. Pacton, *Entretiens de Bichat - Orthophonie*, 35-41 (2009)
24. S. Pacton, M. Fayol, P. Perruchet, *Child Development*, **76**, 324-339 (2005)
25. L. Chanquoy, A. Tricot, J. Sweller, *La charge cognitive* (2007)
26. P. Bonin, R. Peereman, M. Fayol, *Journal of Memory and Language*, **45**, 688-720 (2001)