HAL
open science

Description des pratiques orthographiques de scripteurs ordinaires et de scripteurs en difficulté avec l'écrit à partir de l'analyse d'une dictée

Carole Blondel, Catherine Brissaud, Fanny Rinck

To cite this version:

Carole Blondel, Catherine Brissaud, Fanny Rinck. Description des pratiques orthographiques de scripteurs ordinaires et de scripteurs en difficulté avec l'écrit à partir de l'analyse d'une dictée. 5e congrès mondial de linguistique française (CMLF 2016), Jul 2016, Tours, France. 10.1051/shsconf/20162707003 . hal-01865976

HAL Id: hal-01865976

https://hal.univ-grenoble-alpes.fr/hal-01865976

Submitted on 2 Sep 2018

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.

SHS Web of Conferences **27**, 07003 (2016) DOI: 10.1051/shsconf/20162707003
Congrès Mondial de Linguistique Française - CMLF 2016

Description des pratiques orthographiques de scripteurs ordinaires et de scripteurs en difficulté avec l'écrit à partir de l'analyse d'une dictée

Carole Blondel, Catherine Brissaud et Fanny Rinck
Laboratoire de Linguistique et Didactique des Langues Etrangères et Maternelles – LIDILEM - Université Stendhal Grenoble
carole.blondel@u-grenoble3.fr, catherine.brissaud@ujf-grenoble3.fr, fanny.rinck@ujf-grenoble.fr

Résumé. Alors que la compétence orthographique des adultes est rarement investiguée dans les enquêtes internationales ayant trait à la littératie, l'enquête d'utilité publique IVQ (Information et vie quotidienne), conduite en France en 2011, a permis de disposer de plus de 2000 dictées réalisées par des adultes repérés en difficulté en réception de l'écrit. Ce corpus de textes, inédit dans les enquêtes de cette envergure, permet de préciser les compétences orthographiques et phonographiques (capacité à transcrire la chaine sonore) des enquêtés considérés en difficulté dans cette enquête. Pour mieux évaluer les compétences spécifiques de cette population, nous avons fait passer la dictée (une liste de courses composée de mots et de pseudo-mots, inventés pour l'exercice) à 87 enquêtés de 2011 qui ne présentaient pas de difficulté particulière avec l'écrit (et qui n'avaient donc pas été évalués en production d'écrit). Cet article présente les résultats de la comparaison d'une partie du corpus recueilli en 2011 (434 textes) et de celui recueilli dans le cadre de l'enquête complémentaire en 2014 (87 textes). Il propose une description des compétences des enquêtés : après avoir présenté la méthodologie utilisée, nous comparons les performances orthographiques et morphographiques des deux populations, puis leurs performances phonographiques. Nous observons enfin les choix graphiques concernant les finales des pseudo-mots et nous dégageons des comportements de scripteurs. Les conclusions s'inscrivent dans une perspective didactique.

Abstract. While adult spelling ability is rarely considered in international studies of literary competence, a public survey, IVQ (Information and daily life), was carried out in France in 2011, in which 2000 dictations were given to adults considered to have difficuly in written transcription. This body of texts, hitherto unparalled in a survey of this importance, made it possible to pinpoint the levels of spelling and phonography (ability to translate sounds into words) of those considered to be in difficulty. In order to evaluate more precisely the specific level of these people, we gave the same dictation (a shopping list composed of real and made-up words, invented for the purpose) to 87 people who did not present any particular difficulty in writing, and who therefore had not been tested for writing skills in 2011. This article presents the results of the comparison between part of the 2011 survey (434 texts) and those used in the subsequent survey in 2014 (87 texts). It offers a description of the competence of the group tested. After a description of the methodology used, we shall compare the spelling and morphographical performances of the two groups, and then their phonographical performances. Finally we shall study the graphic choices in the endings of the made-up words, and propose the writers' profiles. The conclusions are framed within a didactic perspective.

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

Le développement d'une société de l'information et l'accroissement des supports de communication dans la vie quotidienne s'accompagnent de la nécessité de disposer d'une bonne maitrise des compétences fondamentales que sont la lecture, l'écriture, la compréhension orale et le calcul pour ne pas rester en marge (Guérin-Pace, 2009). Ces compétences qui sont mobilisées dans les situations de la vie courante relèvent alors du concept de littératie et s'éloignent d'une approche plus scolaire de la lecture et de l'écriture notamment. Si la définition de la littératie par l'OCDE renvoie à une hétérogénéité des compétences visées : « Aptitude à comprendre et à utiliser l'information écrite, dans la vie courante en vue d'atteindre des buts personnels et d'étendre ses connaissances et ses capacités » (OCDE, 2000 : 12), la polyvalence de ce concept est soulignée par Jaffré (2004 : 31), qui, au terme de son tour d'horizon de différentes définitions, désigne la littératie comme « l'ensemble des activités humaines qui impliquent l'usage de l'écriture, en réception et en production ». Les différentes définitions abordent néanmoins la question des compétences de base dans leur dimension fonctionnelle au regard des situations de la vie quotidienne : il s'agit bien de lire pour comprendre, d'écrire pour s'exprimer (Moreau et al., 2013).

Les différentes enquêtes internationales[1] destinées à rendre compte du niveau de littératie des populations menées depuis les années 1990, ne mesurent pas la capacité des adultes à écrire (et donc à orthographier). Une des raisons provient très certainement de la complexité à comparer les performances évaluées dans des systèmes d'écriture différents et dont le niveau de difficulté peut être très variable. Par ailleurs, nous verrons dans l'analyse présentée dans cet article que même dans un système en particulier (par exemple en français), le choix du type de test ainsi que son contenu lorsqu'il s'agit d'évaluer l'écriture et l'orthographe, est loin d'être évident. En premier lieu, que cherche-t-on à évaluer exactement : la capacité à communiquer par écrit pour un usage courant ou la maitrise parfaite du système orthographique d'une langue ? De quelles compétences a-t-on besoin pour communiquer par écrit et quels critères prendre en considération pour évaluer ces compétences ? Enfin, où se situe la frontière entre la maitrise du système ou d'une partie du système et sa non maitrise ?

Pour mesurer le niveau de littératie de la population âgée de 16 à 65 ans, résidant en France métropolitaine, la France a mis au point une enquête à très grande échelle, menée par l'INSEE, dont le protocole conséquent[2] a permis d'interroger 13750 personnes au niveau national en 2011. Cette enquête, intitulée *Information et Vie quotidienne* (désormais IVQ ; INSEE, 2011) a également la spécificité de prendre en compte tous les niveaux de compétences, y compris les faibles niveaux (parfois écartés des enquêtes internationales) pour rendre compte du nombre de personnes en situation de grande difficulté face à l'écrit et pour déterminer le taux d'illettrisme. Dans IVQ, un module appelé module « bas » est ainsi destiné à cerner les compétences à l'écrit de personnes repérées en difficulté dans ce domaine. Il inclut un test de production d'écrit (une dictée), qui fait l'objet du présent article. Le protocole d'IVQ prédit donc de possibles difficultés en production écrite pour les personnes qui n'ont pas réussi les tests de lecture et/ou de compréhension (avec des scores < à 40% de réussite) dans la mesure où les autres enquêtés n'ont pas passé de test d'écriture. Mais avoir des faiblesses en lecture et/ou en compréhension signifie-t-il pour autant ne pas maitriser l'orthographe ? Et passer avec succès les deux premières épreuves est-il une garantie de maitrise de l'orthographe ?

L'enquête IVQ nous permet de disposer de plus de 2000 dictées réalisées par des adultes repérés en difficulté en réception de l'écrit. Ce corpus d'écrits est inédit dans les enquêtes de cette envergure qui visent à évaluer le niveau de maitrise de la littératie chez les adultes, la capacité d'écriture et la compétence orthographique n'étant jamais investiguées (Jeantheau, 2015).

Dans cet article, nous utilisons les données recueillies dans le sud-est de la France, dans les régions Auvergne, Bourgogne, Provence-Alpes-Côte d'Azur et Rhône-Alpes, soit un quart du corpus national disponible. Dans ces quatre régions, 447 personnes (sur les 3237 interrogées) ont été dirigées vers le module bas après leurs résultats au test d'orientation, dont 13 enquêtés qui n'ont pas produit de dictée ou n'ont écrit que quelques items avant d'abandonner. Nous disposons ainsi de 434 textes produits par des enquêtés repérés en difficulté avec l'écrit dans l'enquête (Auvergne : 28 dictées ; Bourgogne : 21 dictées ; PACA : 279 dictées, Rhône-Alpes : 106 dictées)[3].

L'épreuve de production écrite dans l'enquête IVQ

L'épreuve de production écrite se présente sous la forme d'une liste de courses que les enquêtés ont reproduite de manière manuscrite et qui est composée de mots (courants et rares) ainsi que de pseudo-mots (ou mots inventés pour l'exercice) présentés comme des noms de marques ou des noms propres : *pharmacie – anti-rhume micatol ou sirape – alcool à 90° dux – épicerie – tomate(s) de pays – cerise(s) – confiture(s) – sel gobar – trois fromages qui ne sentent pas fort – librairie – le fiancé solennel de joc pévanor.* Les items ont été choisis par l'équipe du PsyEf pour rendre compte des connaissances orthographiques des enquêtés (à partir de critères comme le nombre de phonèmes à transcrire, leur caractère régulier ou irrégulier ainsi que les procédures de traitement[4] que les mots requièrent pour être transcrits) et pour tester leurs compétences en phonographie. La transcription des pseudo-mots demande en effet d'opérer une sélection parmi les correspondances phonie-graphie plausibles en français. Enfin, l'accord en nombre d'un nom (*fromages*) et d'un verbe (*sentent*) doit permettre de définir leurs compétences en morphographie.

Voici un exemple de dictée produite par un enquêté du module bas interrogé en 2011 :

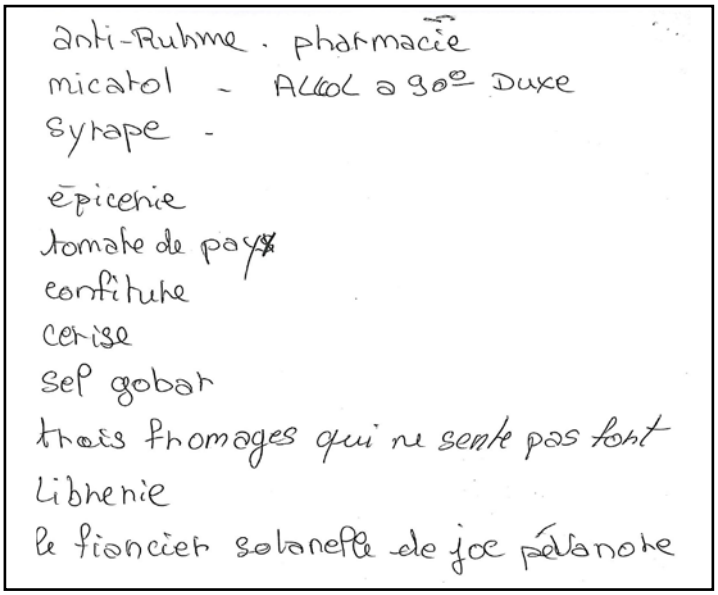
anti-Ruhme . pharmacie
micatol - Alcol a 90° Duxe
Syrape -
épicerie
tomate de pays
confiture
cerise
sel gobar
trois fromages qui ne sente pas fort
Librenie
le fiancier solanelle de joc pévanore

Figure 1. Exemple de dictée produite par un enquêté du module « bas »

Pour rendre compte du degré de difficulté de la dictée, nous présentons le classement des mots selon leur indice de consistance[5] (du plus consistant au moins consistant) et de fréquence[6] calculés dans la base Manulex-infra. Selon Lété (2008), la consistance peut être définie comme le « degré d'incertitude associé à l'écriture d'une association phonème-graphème ou à la lecture d'une association graphème-phonème. Ce niveau d'incertitude, plus communément appelé "consistance", permet de prédire la difficulté à écrire ou lire un mot ». Le degré de consistance met donc en relief le degré de variabilité des graphèmes pour un phonème en écriture, révélant ainsi le niveau de complexité orthographique ». Dans la base Manulex-Infra (Ortéga & Lété, 2010), cet indice est calculé selon la fréquence des associations phonème-graphème en fonction de la position des syllabes dans le mot.

Lemmes	Consistance	Rang « consistance »	Fréquence	Rang « fréquence »
solennel	86,64[7]	1	2,53	14
tomate	82,99	2	23,87	6
fromage	81,72	3	118,16	3
sentir	81,04	4	22,68	7
sel	80,52	5	78,43	4
pays	79,25	6	309,72	1
fort	78,31	7	302,95	2
librairie	77,63	8	11,13	11
confiture	76,01	9	42,33	5
cerise	72,03	10	21,41	8
pharmacie	67,17	11	16,84	10
épicerie	64,86	12	7,14	13
rhume	61,74	13	20,21	9
alcool	61,47	14	7,35	12
fiancé	59,19	15	0,46	15

Tableau 1. Classement des mots de la dictée IVQ selon leur degré de consistance phonologie-orthographe (consistance PO totale, du CP au CM2, calcul par Token) et leur indice de fréquence

Dans cet exercice, les enquêtés ont ainsi été confrontés à des mots plutôt réguliers (dont le degré de consistance est élevé) avec *tomate*, *fromage*, *sentir*, *sel* et à des mots plus irréguliers (dont le degré de consistance est plus faible) comme *rhume*, *alcool*, *fiancé*. La fréquence des mots peut également constituer une explication aux difficultés rencontrées par les scripteurs, notamment lorsque la graphie est plutôt irrégulière. S'opposent ainsi des mots plutôt fréquents comme *pays*, *fort*, *fromage* et des mots dont l'exposition semble bien moindre : *épicerie*, *solennel*, *fiancé*.

Le corpus de dictées disponible permet de comparer les compétences orthographiques et phonographiques des enquêtés considérés en difficulté dans l'enquête mais il ne permet pas de confronter ces résultats avec ceux qu'auraient obtenu les enquêtés dirigés vers le module « haut ».

Une enquête complémentaire à IVQ 2011 : le projet « Compétences orthographiques des adultes et contextes sociaux d'usage »

Pour mieux évaluer les compétences spécifiques dans le domaine de l'écrit des personnes jugées en difficulté, nous avons fait passer la dictée IVQ à des enquêtés de 2011 qui ne présentaient pas de difficulté particulière avec l'écrit (et qui n'avaient donc pas été évalués en production d'écrit). Cette opportunité de réinterroger des enquêtés ayant participé à l'enquête de 2011 et de disposer ainsi d'un corpus qui oppose deux catégories de personnes, au niveau de compétence à l'écrit différent (selon les résultats d'IVQ), n'a pas été une tâche aisée. Trois ans après l'enquête de l'INSEE, il a fallu retrouver ces enquêtés à partir du peu d'informations dont nous disposions (180 enquêtés retrouvés à partir de 300 fiches adresses) et convaincre ces derniers de participer à une extension de la précédente enquête, non obligatoire à la différence des enquêtes d'utilité publique. Dans la région Rhône-Alpes, 28% des enquêtés retrouvés ont refusé de participer à cette nouvelle enquête et 27% n'étaient pas disponibles lors du recueil, ce qui a abouti au recueil de 87 dictées.

Ce second recueil de données a été rendu possible grâce à un projet complémentaire intitulé « Compétences orthographiques des adultes et contextes sociaux d'usage » qui a reçu le soutien financier de la DGLFLF (Délégation générale à la langue française et aux langues de France) et l'approbation du comité du label de la statistique publique. Le protocole, conçu au sein du laboratoire LIDILEM à Grenoble et du laboratoire DYSOLA à Rouen[8], est constitué de la dictée IVQ augmentée de quelques

items (un rayon bricolage a été ajouté) choisis pour présenter des difficultés, notamment morphographiques, en nombre plus important que celles de la dictée d'IVQ[9] : *bricolage – deux produits parfumés qui dégraissent bien – quatre paquets de vis plates – un ciseau à bois normal – une boite d'allumettes – une pince pour arracher les clous – un pot de peinture – quelques mètres de tuyau d'arrosage*.

Dans cet article, nous différencions les deux corpus en appelant les enquêtés ayant passé le module bas en 2011, enquêtés « bas », et ceux qui ont participé à l'enquête complémentaire en 2014, enquêtés « haut ». Nous disposons ainsi de 434 dictées réalisées par des enquêtés « bas » et de 87 dictées (qui contiennent en plus un rayon « bricolage ») réalisées par des enquêtés « haut ». Dans cette analyse, nous proposons de décrire les compétences orthographiques et phonographiques des enquêtés. Après avoir présenté la méthodologie utilisée, nous comparons les performances orthographiques des deux populations, puis leurs performances phonographiques. Nous observons enfin les choix graphiques concernant les finales des pseudo-mots et nous dégageons des comportements de scripteurs. Les conclusions s'inscrivent dans une perspective didactique.

1 Evaluer les compétences en littératie à partir d'une dictée : la nécessité d'une analyse particulière

L'exploitation de ce corpus pose les bases d'une réflexion méthodologique sur la manière de traiter des données produites par des scripteurs dont on sait qu'ils ont du mal à atteindre la norme orthographique. Ainsi, une évaluation classique qui consisterait à repérer les graphies incorrectes et à établir des scores de réussite ne nous permettrait pas de rendre compte des capacités des enquêtés à communiquer par écrit ni de définir précisément leur niveau de compétences : comment réagissent-ils lorsqu'ils ne connaissent pas l'orthographe d'un mot ? Parviennent-ils à en restituer la forme phonologique ? Les choix graphiques erronés tiennent-ils compte des règles de position et de distribution relatifs à l'écriture du français ?

Pour répondre à ces questions, nous avons tenu compte de la spécificité de ce corpus et avons adopté un traitement différent de celui appliqué par l'INSEE.

1.1 L'informatisation du corpus : deux points de divergence

Le traitement de l'accentuation

L'INSEE a fait le choix de neutraliser tous les accents produits par les enquêtés et de considérer comme correctes toutes les formes non accentuées produites (ex. *fiance* pour *fiancé*, *epicerie* pour *épicerie*).

Par souci de cohérence, nous n'avons pas pris en compte l'absence d'un accent sur le « e » initial d'*épicerie*, celui-ci ayant très souvent été transcrit avec une majuscule non accentuée. En effet, nous ne souhaitions pas imputer de « e » accentué (*Epicerie* = *épicerie*), tout en considérant incorrect du point de vue phonographique les « e » minuscules non accentués (*epicerie* = **epicerie*). L'accentuation du « e » initial d'*épicerie* a donc été neutralisée (*Epicerie, epicerie, épicerie* = *épicerie*). En revanche, nous avons considéré comme erronées les adjonctions d'accents (**épicèrie*, **cérise*, **rhûme*) ou de diacritiques (**çerises*). Nous avons également considéré comme erronée l'utilisation du graphème « e » pour noter le phonème /e/ à la fin de *fiancé*.

Les imputations et les graphies illisibles

L'INSEE a également procédé à un travail d'imputation de mots orthographiés correctement, pour les personnes qui n'ont pas répondu à tous les items : « Ces travaux d'imputation […] se sont appuyés sur les données individuelles collectées lors de la passation du module biographique » (Accord cadre INSEE – ANLCI, 2003). Nous avons corrigé ces imputations en indiquant la présence de non-réponse dans le fichier et avons attribué un code « illisible », pour les graphies dont nous avons considéré la reconnaissance des lettres trop aléatoire, et que nous ne souhaitions pas non plus indiquer comme une réponse manquante.

Un travail de vérification et de correction des graphies a donc été réalisé pour disposer d'un fichier qui corresponde aux graphies proposées par les enquêtés.

1.2 Critères et attentes de l'INSEE

L'épreuve de production écrite fait partie, dans l'enquête IVQ, du module « bas » (appelé aussi module « ANLCI » car conçu avec la collaboration de l'Agence Nationale de Lutte contre l'Illettrisme), centré sur l'évaluation des personnes qui présentent des difficultés dans les domaines fondamentaux de l'écrit. Rappelons que l'un des objectifs de cette enquête d'utilité publique était de mesurer le taux d'illettrisme en France. Les enquêtés ont ainsi été interrogés en lecture, en écriture et en compréhension écrite pour déterminer leur niveau de difficulté à l'écrit sur une échelle de 1 à 4 selon leur pourcentage de réussite dans les trois composantes : graves difficultés (moins de 40 % de réussite), fortes difficultés (entre 40 % à 60 % de réussite), difficultés partielles (entre 60 % à 80 % de réussite) ou absence de difficulté (plus de 80% de réussite). Les personnes enquêtées ont alors été considérées en situation d'illettrisme si elles se trouvaient en difficulté grave ou forte dans au moins une des compétences évaluées (lecture, écriture, compréhension écrite). Au terme d'IVQ 2011, l'INSEE a ainsi considéré que 7% des français se trouvaient en situation d'illettrisme[10].

Pour évaluer les performances en production écrite et établir des niveaux de compétences, l'INSEE a réalisé des scores de réussite des items de la dictée, à partir d'une analyse quantitative des données. C'est la conformité à la norme orthographique pour les mots et les attentes graphophonologiques des concepteurs du test pour les pseudo-mots qui ont été jugés. Un point a été accordé par réponse correcte, ce qui donne un score sur 21 (15 mots et 6 pseudo-mots).

La réussite en morphosyntaxe a été davantage récompensée puisque deux points ont été attribués à chaque marque de nombre réussie (*ent* pour l'accord du verbe *sentent* et *s* pour le nom *fromages*). La dictée a ainsi été notée sur 25 points, qui déterminent le niveau de l'enquêté sur l'épreuve de production écrite. Les scores de réussite ont permis de considérer en grande difficulté dans ce domaine, les enquêtés qui ont réussi moins de 9 mots sur les 21 qui ont été corrigés.

1.3 Le barème utilisé pour notre étude

Pour rendre compte des compétences orthographiques et phonographiques des enquêtés, nous avons procédé à un nouveau calcul du barème.

1.3.1 Analyse des mots

Les performances orthographiques des enquêtés sont observées d'après le respect de la forme orthographique de 16 mots de la dictée : *pharmacie – anti*[11] *– rhume – alcool – épicerie – tomate(s) – pays – confiture(s) – cerise(s) – sel – fromage(s) – sente(nt) – fort – librairie – fiancé – solennel*. Nous avons retiré *qui*, analysé par l'INSEE, car les divers mots de liaison n'ont pas tous été transcrits par les enquêtés et nous avons ajouté les items *fromage(s)* et *sente(nt)* dont seules les marques de nombre avaient été prises en compte.

La compétence en morphosyntaxe a également été analysée d'après les deux marques de nombre présentes sur les items *fromages* et *sentent*. En revanche, nous avons tenu compte de toutes les marques de nombre proposées, y compris celles qui diffèrent de la marque attendue sur le verbe *sentir* (*sentent*, *sentes*). Les items *tomate(s)*, *cerise(s)*, *confiture(s)* qui, sans déterminant, peuvent porter ou non la marque de pluriel, pouvaient être écrits soit au pluriel soit au singulier. Enfin, nous n'avons pas relevé la présence d'une marque de nombre sur l'adverbe *fort* car sur cet item, c'est l'absence de marque qui correspond à une réussite et nous ne pouvons déterminer si ceux qui ont réussi ont eu conscience de ne pas devoir marquer le nombre sur un adverbe. Nous considérons que l'invariabilité des adverbes est une connaissance complexe qu'il n'est pas opportun de juger dans une dictée destinée à évaluer les compétences orthographiques d'un public en difficulté.

Pour les mots du rayon bricolage, produits uniquement par les enquêtés « haut » dans le cadre de l'enquête complémentaire, nous avons appliqué un barème sur 22 points qui prend en compte l'orthographe lexicale et la morphographie[12], avec un point attribué par item réussi : *bricolage – produits – parfumés – dégraissent – paquets – vis – plates – ciseau – à – bois – normal – boite – allumettes – pince – arracher – clous – pot – peinture – quelques – mètres – tuyau – arrosage.*

1.3.2 Analyse des pseudo-mots

Les pseudo-mots ont été comptabilisés selon leur conformité au principe phonographique. Ne disposant pas de données concernant la manière dont les items ont été dictés et notamment sur une éventuelle particularité de prononciation des enquêteurs, nous avons comptabilisé comme correctes des formes graphiques qui peuvent différer de la forme sonore attendue et qui peuvent provenir de la prononciation inhérente à l'enquêteur ou à l'enquêté. C'est ainsi que nous avons pris en considération toutes les graphies des voyelles antérieures /œ/ et /ə/ dont le point d'articulation est proche et de la voyelle /ɔ/, plus postérieure, ces trois voyelles formant une série au sein d'un système vocalique particulièrement riche. Par exemple, les graphies *micateul* pour *micatol* ou *jek*, *jeuq* pour *joc* ont été comptabilisées comme phonographiquement correctes. Par ailleurs, même si une forme graphique est plus probable qu'une autre en français (ex. *joque* plutôt que *jok*), les deux ont été comptabilisées comme phonographiquement correctes.

En tenant compte des biais éventuels liés à la prononciation des enquêteurs sur les pseudo-mots et en rectifiant le barème appliqué sur les deux items porteurs d'une marque de nombre (un point pour l'orthographe de l'item et un point pour la marque du nombre, même si cette dernière ne correspond pas à celle attendue), nous avons établi un barème qui sanctionne moins durement les enquêtés qui ne maitrisent pas la morphographie. Ce barème que nous appelons « orthographique », consiste à noter la dictée sur 24 points, avec un point attribué par item réussi (sur les 16 mots et six pseudo-mots) et deux points pour la morphographie (un point pour le *s* de *fromages* et un point pour le *nt* ou *s* de *sentent*).

Pour mettre en relief les compétences d'écriture de scripteurs qui, bien qu'éloignés de la norme orthographique, produisent des formes graphiques parfaitement lisibles et qui permettent d'établir une communication minimale, nous avons également établi un score de réussite phonographique sur les mots de la dictée. Ce second barème, que nous appelons « phonographique », permet de rendre compte des compétences phonographiques des enquêtés « bas » sur les 16 mots de la dictée (ex. *frommage* pour *fromages* ; *solanel* pour *solennel*, etc.). Nous n'avons pas procédé à cette analyse pour les enquêtés « haut » car nous n'avons pas observé de difficultés d'ordre phonographiques dans les graphies erronées de ces derniers.

Les scores de réussite sont présentés sous forme de pourcentages et ont été calculés en fonction du nombre de réponses produites. Dans ce calcul, nous avons choisi de ne pas prendre en compte les réponses qui diffèrent de l'item attendu (ex. *formation* pour *pharmacie*), les graphies illisibles et les non réponses. L'absence de graphie ne signifie en effet pas nécessairement que l'enquêté n'était pas capable de produire une forme correcte. Certains enquêtés ont en effet arrêté l'exercice après avoir orthographié correctement une partie des items, d'autres ont reproduit tous les produits de la liste de courses sans reproduire le nom des rayons (*pharmacie*, *épicerie*, *librairie*), ce qui semble davantage provenir d'un choix ou d'une incompréhension de la consigne.

2 Les performances orthographiques et morphographiques dans la dictée : comparaison des enquêtés « bas » et « haut »

Nous présentons dans un premier temps, les performances orthographiques des enquêtés « bas » et « haut » sur la partie commune de la dictée. Après avoir comparé leur score, nous présentons le résultat obtenu par les enquêtés « haut » sur le rayon bricolage. Nous procédons ensuite à une analyse plus fine des résultats en présentant les performances lexicales des enquêtés sur chaque mot de la dictée et leurs compétences en morphographie observées d'après l'accord en nombre de *fromages* et *sentent* ainsi que

sur l'absence d'accord de *fort*. Dans chaque analyse, nous comparons les résultats des 434 enquêtés « bas » et des 87 enquêtés « haut ».

2.1 Les compétences orthographiques, phonographiques et morphographiques : répartition du score obtenu sur chaque dictée

2.1.1 Comparaison des scores des enquêtés « bas » et « haut » sur la dictée IVQ

Nous comparons dans la figure 2, le score obtenu par les enquêtés « bas » et les enquêtés « haut » sur la dictée IVQ. Nous rappelons que les items ont été codés selon le respect de la forme orthographique pour les mots, le respect de la forme phonographique pour les pseudo-mots et deux marques de nombre (le *s* de *fromages* et le *nt* ou *s* de *sentent*) pour un score de total de 24 points.

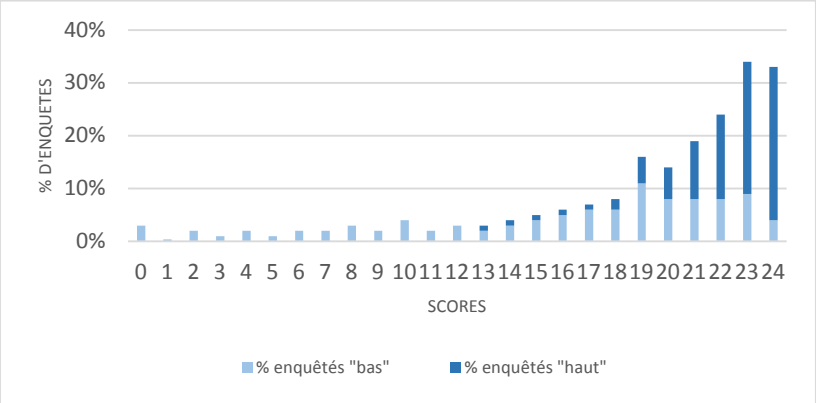

Figure 2. Répartition du score obtenu à la dictée IVQ par les enquêtés « bas » et « haut »

La figure 2 montre une dispersion maximale des scores chez les enquêtés « bas » (de 0 à 24 sur 24) que nous n'observons pas chez les enquêtés « haut ». En revanche, 88% des enquêtés « haut » et 33% des enquêtés « bas » ont un profil orthographique similaire, avec une bonne réussite à la dictée (plus de 80%). Ce résultat montre qu'un faible niveau de compétences en lecture ou en compréhension tels que révélés par le test d'orientation pour les enquêtés « bas », n'affecte pas nécessairement les compétences orthographiques évaluées dans l'exercice.

La plus faible dispersion des résultats observée chez les enquêtés « haut » montre que si le test d'orientation a effectivement permis de repérer un niveau minimum de compétences orthographiques chez ces derniers (aucun enquêté n'a obtenu de score inférieur à 13 et la moyenne de 22 sur 24 ainsi que la médiane de 23 sur 24 sont particulièrement élevées), il se révèle en revanche insuffisant pour affirmer qu'il existe un lien entre les compétences orthographiques et le reste des compétences évaluées (lecture et compréhension écrite). La dispersion des scores parmi les enquêtés « bas » révèle en effet une réalité multiforme des compétences orthographiques, y compris chez un public repéré en difficulté avec l'écrit.

Nous observons ensuite les compétences orthographiques des enquêtés « haut » sur le rayon bricolage, dont les difficultés lexicales et morphographiques sont plus significatives que dans la dictée IVQ, avec du lexique et des structures syntaxiques plus complexes.

2.1.2 Répartition des scores des enquêtés « haut » sur le rayon bricolage

La figure 3 présente la répartition des scores obtenus par les enquêtés « haut » sur le rayon bricolage. Nous rappelons que le barème prend en compte l'orthographe et la morphographie pour un score total de 22 points.

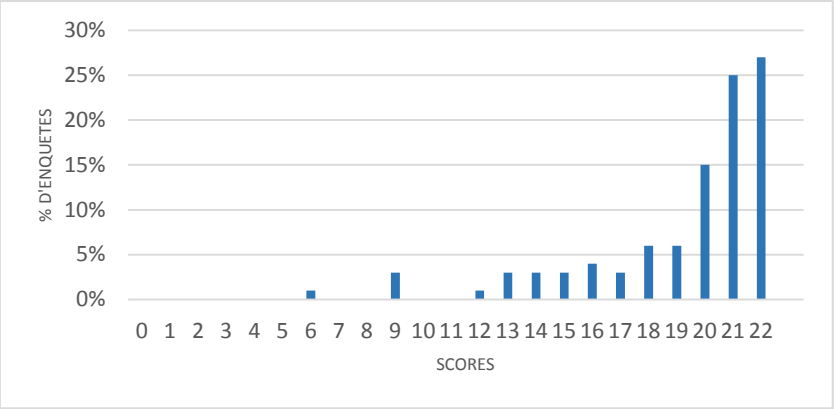

Figure 3. Répartition des scores obtenus par les enquêtés « haut » sur le rayon bricolage

Les écarts qui subsistent entre les enquêtés « haut » sont plus importants sur cette partie de la dictée, avec des scores plus dispersés que sur la partie IVQ, compris entre 6 et 22 malgré une médiane qui reste élevée (21 sur 22). La proportion de très bons scores est ainsi moins élevée avec 79% des enquêtés « haut » qui ont obtenu plus de 80% de réussite, contre 88% pour la partie IVQ. Quelques enquêtés « haut » ont éprouvé des difficultés parfois importantes (trois scores inférieurs à 10 sur 24), qui n'apparaissaient pas sur la partie IVQ de la dictée.

Nous détaillons ensuite les scores de réussite obtenus sur chaque item, pour chaque dictée et pour chaque population. Nous proposons une analyse différente des résultats obtenus sur le lexique et sur la morphographie pour identifier ce qui a posé problème aux enquêtés « bas » et à certains enquêtés « haut » sur la dictée IVQ et aux enquêtés « haut » sur le rayon bricolage.

2.2 Les performances lexicales des enquêtés : classement des mots selon leur réussite orthographique

Nous analysons les compétences lexicales[13] des enquêtés sur la dictée et classons les mots (hors pseudo-mots) selon leur pourcentage de réussite orthographique. Les marques de nombre et de genre ont été neutralisées (ex. *forts, fort = fort* ; *fiancé, fiancée = fiancé*).

2.2.1 Classement des mots de la dictée IVQ et comparaison des enquêtés « bas » et « haut »

Nous observons dans la figure 4, les compétences lexicales des enquêtés « bas » et « haut » sur la dictée IVQ et classons les mots (hors pseudo-mots) selon leur pourcentage de réussite. Pour faciliter la comparaison, les mots sont classés du mieux réussi au moins bien réussi par les enquêtés « bas ».

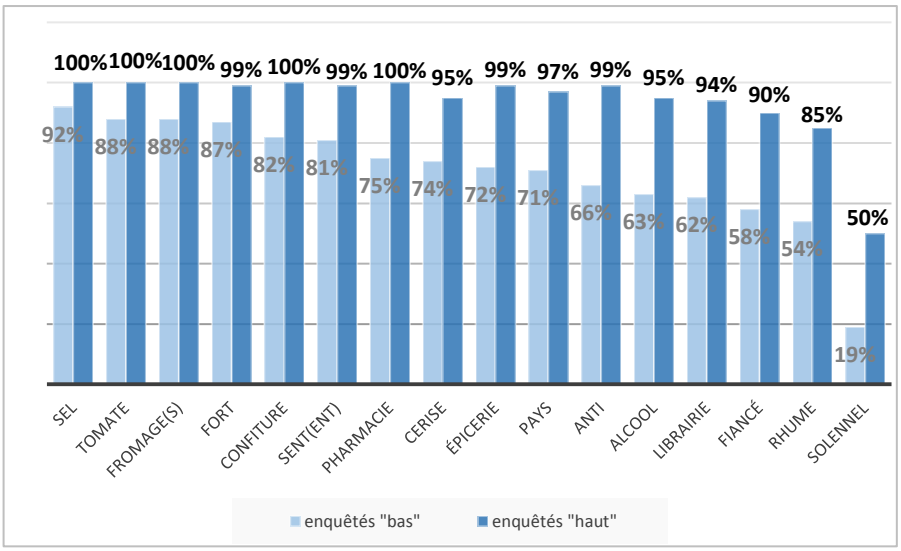

Figure 4. Compétences lexicales des enquêtés « bas » et « haut » pour chaque mot de la dictée IVQ

Dans les deux corpus, les mots particulièrement consistants (indice > à 80), qui sont aussi les plus fréquents tels que *sel, tomates, fromages* ont été particulièrement bien réussis, contrairement aux mots dont le degré de consistance est moindre (*cerise*, *épicerie*, *pays*) ou dont l'indice de fréquence est particulièrement faible (*fiancé*, *solennel*). L'effet de fréquence impacte particulièrement la réussite chez les enquêtés « haut » lorsqu'il s'agit de mots rares, ce que nous observons dans l'écart de réussite orthographique entre les mots de la dictée IVQ (réussis en moyenne à 97%) et *solennel* que seule la moitié des enquêtés « haut » a correctement restitué. L'effet de fréquence s'observe également dans les graphies des enquêtés « bas ». L'item *pays*, qui apparait comme le plus fréquent du corpus tout en possédant une structure rare, qui demande au scripteur de l'avoir intégré en mémoire (le transcodage des relations phonèmes-graphèmes ne pouvant aboutir à la graphie correcte), a été plutôt bien réussi par les enquêtés « bas » (71% de réussite). Parmi les graphies erronées, nous observons par ailleurs des traces de la structure complexe de ce mot : **pay*, **payi*, **pais.*

La réussite sur les items semble donc davantage provenir de l'effet de consistance et donc de la régularité des combinaisons de phonèmes et de graphèmes. La présence de correspondances phonie-graphie irrégulières fait ainsi baisser les résultats des enquêtés « bas » y compris sur les mots fréquents (ex. *tomate* réussi à 88% vs *pharmacie* réussi à 75%) et les enquêtés « bas » et « haut » sur les mots plus rares (ex. *fiancé*, *rhume*, *solennel*).

2.2.2 Classement des mots du rayon bricolage

Dans la figure 5, nous observons les compétences lexicales des enquêtés « haut » sur 22 mots du rayon bricolage. Comme pour la dictée IVQ, les marques de nombre et de genre ont été neutralisées.

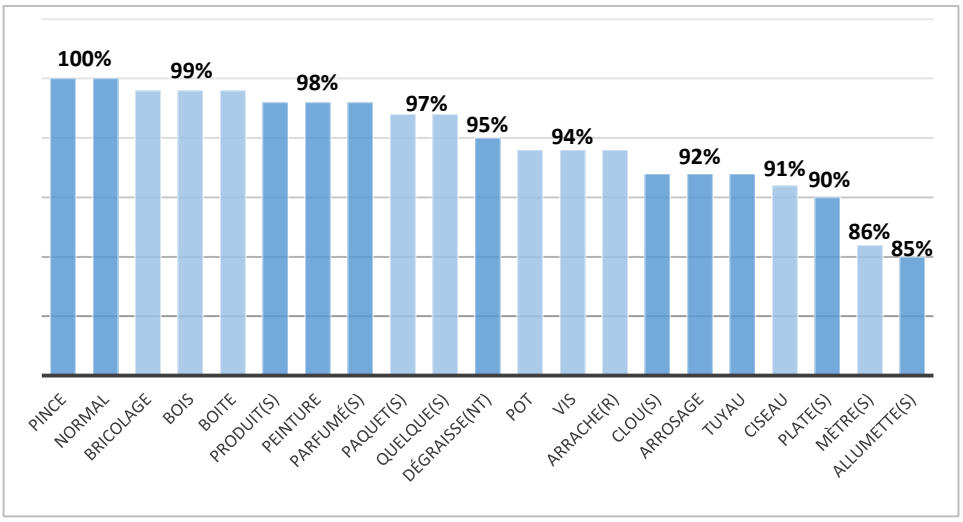

Figure 5. Compétences lexicales des enquêtés « haut » pour chaque mot du rayon bricolage

Sur le rayon bricolage, c'est également sur les mots les plus réguliers (*pince*, *bricolage*) et les plus fréquents (*bois*, *deux*) du corpus, que les taux de réussite sont les plus élevés. Alors que les items de la dictée complémentaire contiennent des difficultés lexicales un peu plus importantes, on observe effectivement une proportion plus importante de mots qui obtiennent moins de 95% de réussite. En revanche, aucun mot n'obtient moins de 85% de réussite.

Nous avons observé le type d'erreurs produites sur les quatre mots les moins bien réussis du rayon bricolage : *plate(s)*, *ciseau(x)*, *mètre(s)*, *allumette(s)*. Les erreurs relevées sur *plate(s)*, *mètre(s)* et *allumette(s)*, concernent principalement le redoublement de consonnes (**platte*, **mettre*, **alumette*) et l'ajout d'un accent circonflexe (**mêtre*, **plâte*). Sur *ciseau*, les erreurs concernent le choix du graphème pour les phonèmes /s/ (**siseau*, **sciseau*), /z/ (**cizeau*, **sicau*) et pour un seul enquêté, le phonème /o/ (**sicau*).

2.3 Comparaison des performances en morphographie des enquêtés

Nous évaluons ensuite la compétence en morphographie des enquêtés « haut » à partir des onze marques de nombre présentes dans l'intégralité de la dictée, avec l'accord de six noms : *fromages – produits – paquets – allumettes – clous – mètres*, de trois verbes : *sentent – dégraissent – arracher*, du déterminant *quelques* et de deux adjectifs : *parfumés – plates*. Nous avons également regardé si les enquêtés ont respecté l'absence d'accord pour trois mots (l'adverbe *fort* et les noms : *ciseau – tuyau*). Les marques de nombre erronées (ex. *les *cloux*, trois occurrences) sont également comptabilisées, seule la présence d'un accord nous intéressant.
Nous comparons le pourcentage de réussite sur la marque de nombre de *fromages*, *sentent* et l'absence d'accord de *fort* (présents dans la partie commune de la dictée pour les deux corpus) avec celui obtenu par les enquêtés du module bas.

Dans la figure 6, les taux de réussite (accord en nombre ou respect du singulier) ne concernent que la morphographie.

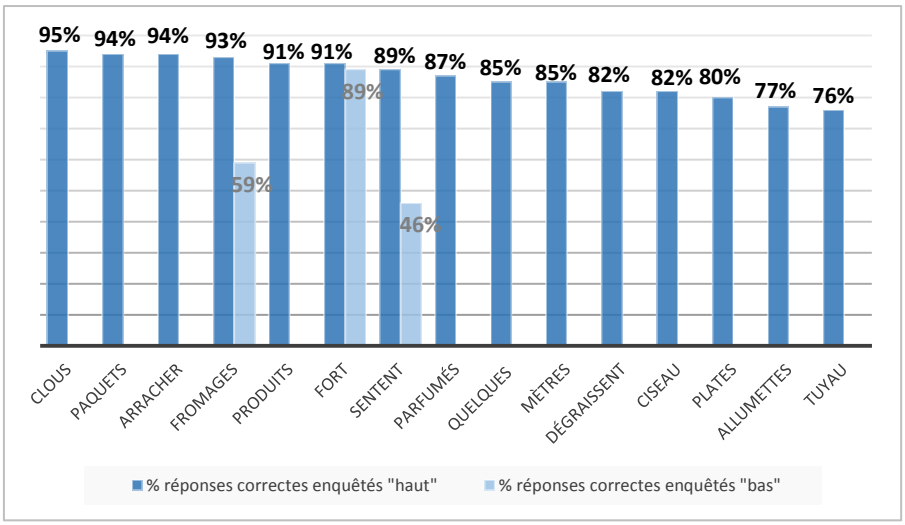

Figure 6. Proportion de réponses correctes obtenues sur la morphographie pour les enquêtés « haut » et « bas »

Chez les enquêtés « bas », il semble plus facile de produire une forme non marquée, la marque de nombre sur le nom et le verbe ayant été plutôt mal restituée alors que *fort* obtient 89% de réussite, score très proche de celui des enquêtés « haut ». En revanche sur l'accord en nombre du nom *fromages* et du verbe *sentent*, la différence de performances entre les deux populations est très marquée, notamment sur l'accord du verbe, avec 43 points d'écart entre les enquêtés « haut » (89% de réussite) et les enquêtés « bas » (46% de réussite) mais aussi sur celui du nom *fromages* (34 points de différence : 93% contre 59%).

Il aurait cependant été intéressant de davantage tester les compétences en morphographie des 33% enquêtés « bas » qui ont obtenu, comme les enquêtés « haut », de très bons résultats à la dictée IVQ (plus de 80% de réussite). Ceux-ci ont plutôt bien réussi la morphographie de la dictée, en obtenant 89% de réussite sur *fromages*, 95% de réussite sur *sentent* (score supérieur à la moyenne des scores obtenus par les enquêtés « haut » sur cet accord).

L'accord en nombre du nom *fromages* et celui du verbe *sentent* semble donc constituer un bon indicateur pour distinguer les personnes qui ont de grandes difficultés en morphographie mais ils ne suffisent pas, en revanche, pour cerner les compétences dans ce domaine des personnes qui ont de meilleures connaissances orthographiques. Par contraste, la pertinence de l'item *fort* est discutable dans la mesure où c'est l'absence de marque qui correspond à une réussite et qu'il n'est pas possible, sans interroger le scripteur sur ses choix, de déterminer s'il a conscience de ne pas marquer le nombre sur un adverbe.

Chez les enquêtés « haut », on observe une bonne restitution des marques de nombre dans les constructions simples (avec présence d'un déterminant : *les clous*, *quatre paquets*, etc.), mais les résultats baissent lorsqu'il s'agit de tournures complexes où aucun signal ne vient aider au marquage du pluriel : par exemple, le *s* de *allumettes* dans *une boite d'allumettes* a été produit par seulement 77% des enquêtés « haut », difficulté que l'on retrouve en fin de scolarité au collège (Brissaud, 2015). En revanche, l'absence d'accord pour *tuyau* dans *quelques mètres de tuyau* a posé problème à 24% des enquêtés « haut » qui semblent avoir réalisé l'accord entre le déterminant *quelques* et le nom *tuyau*.

3 Les performances phonographiques des enquêtés dans la dictée IVQ : analyse des mots et des pseudo-mots

Dans cette partie, nous présentons les compétences phonographiques des enquêtés « bas » et « haut » à partir des pseudo-mots et des graphies erronées des mots (d'un point de vue orthographique). Nous valorisons ainsi la capacité des scripteurs à coder les phonèmes en proposant une forme graphique du

pseudo-mot ou du mot, conforme au système phonographique du français. Cette analyse nous permet de distinguer la proportion d'enquêtés capables de restituer la forme phonologique des mots dictés, de ceux pour qui les connaissances phonographiques sont insuffisantes pour passer du canal oral au canal écrit. Les compétences phonographiques représentent en effet une étape intermédiaire et essentielle avant l'acquisition des conventions d'écriture (Fayol, 2013).

Ainsi, nous considérons des graphies comme *rume* ou *ruhme* acceptables du point de vue de la phonographie bien qu'elles soient éloignées de la norme orthographique. Nous avons accepté les formes contenant des segmentations non conformes (*en-ti*, *et pisseri*, etc.) avec ou sans traits d'union, ces segmentations ne modifiant pas la forme sonore du mot ainsi que les réalisations qui témoignent de marques d'oralité comme l'élision du « e » muet dans *épicerie* (*épissri*) ou l'assimilation du /i/ dans *fiancé* (*fillancé*).
Nous avons également accepté les réalisations graphiques plausibles de chaque phonème, en tenant compte des règles de position ou de distribution spécifiques au français. Ainsi, la graphie **épisserie* a été acceptée du point de vue de la phonographie, contrairement à la graphie **épiseri* dont la réalisation phonique /epizeri/ ne restitue pas la forme phonologique de l'item *épicerie*.

3.1 Classement des mots selon leur réussite phonographique chez les enquêtés « bas »

Nous présentons dans un premier temps, les compétences phonographiques des enquêtés « bas » pour chaque mot, en comparant la proportion d'enquêtés qui proposent une graphie qui restitue la forme phonologique des mots (graphies phonographiquement correctes additionnées aux graphies orthographiquement correctes) à ceux qui ont restitué la forme orthographique attendue. Nous n'avons pas procédé à cette analyse pour les enquêtés « haut », ces derniers n'ayant pas de difficultés d'ordre phonographique.

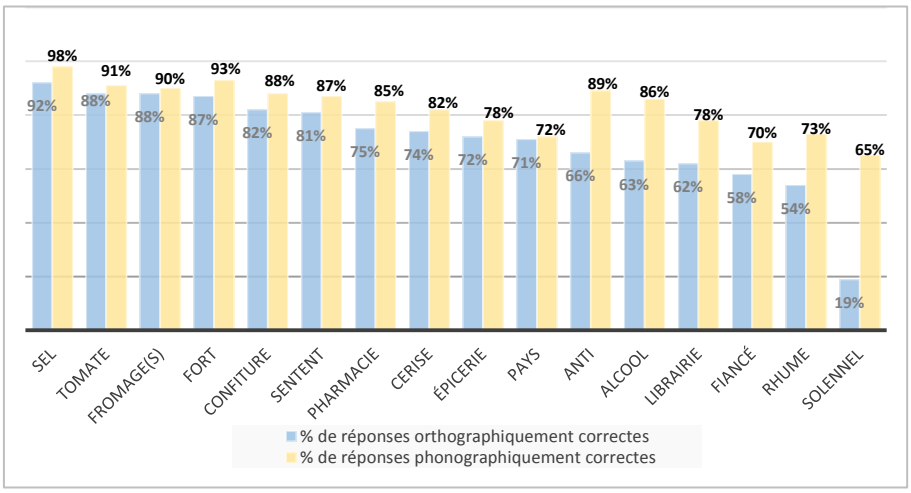

Figure 7. Proportion de réponses orthographiquement correctes et phonographiquement correctes proposées par les enquêtés « bas »

La prise en compte du critère phonographique fait augmenter les résultats sur chaque mot, de manière parfois très significative notamment sur les mots les moins bien restitués du point de vue orthographique comme *alcool*, *librairie*, *rhume* et surtout *solennel*. Cet item, qui a particulièrement posé problème aux enquêtés « bas » (19% de graphies correctes) passe à 65% (plus 46 points) avec la prise en compte du critère phonographique. Ainsi, des graphies comme *saulanel* ou *solannel* offrent une forme phonologique lisible du mot au-delà du respect de l'orthographe conventionnelle. Pour rendre compte de leurs connaissances des règles de position, nous observons une des difficultés de l'item *épicerie* : la restitution du phonème /s/. Parmi les enquêtés qui n'ont pas réussi à restituer la forme orthographique de

ce mot, 67% ont proposé un graphème plausible *ç*, *c* ou *ss*. Les autres enquêtés ont majoritairement utilisé le graphème *s* pour transcrire phonétiquement le phonème /s/ sans prendre en compte la règle de position de ce graphème, qui entouré de deux voyelles, prend la valeur phonologique /z/. Quelques enquêtés ont omis le *e* muet après le graphème *c* (« ce » devient alors « cr »), transformant ainsi la valeur phonologique de /s/ en /k/ : **epicri* que nous lisons /epiKri/.

3.2 Classement des pseudo-mots selon leur réussite phonographique chez les enquêtés « bas » et « haut »

Selon le même principe que pour les mots, nous avons classé les pseudo-mots selon leur pourcentage de réussite phonographique pour les enquêtés « bas » et « haut ».

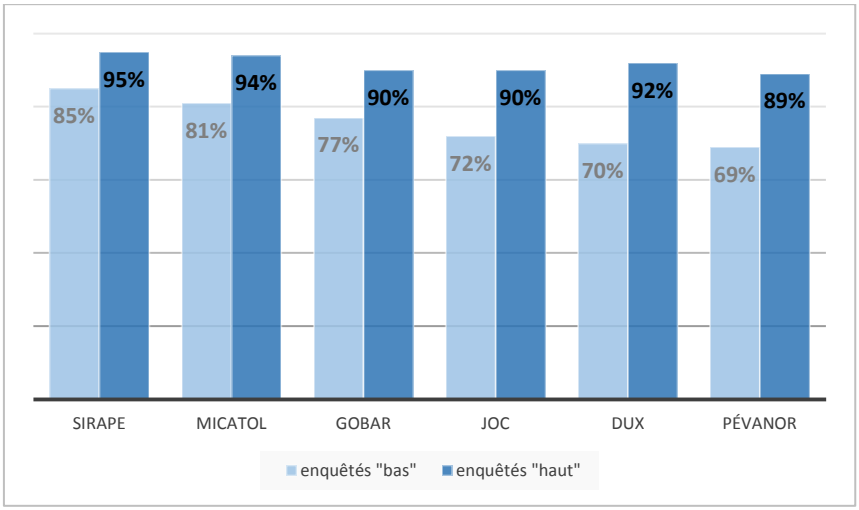

Figure 8. Proportion de réponses correctes pour les enquêtés « bas » et « haut » et pour chaque pseudo-mot selon le critère phonographique

Chez les enquêtés « haut », les taux de réussite sur les pseudo-mots sont compris entre 89% et 95%, aucun item n'atteignant les 100% de réussite observés sur certains mots. Les erreurs produites sur les mots sont majoritairement d'ordre orthographique, très peu d'erreurs d'origine phonographique ayant été relevées chez ces scripteurs. Sur *pévanor*, 11% des graphies sont phonographiquement incorrectes : **pavanord*, **bevanor*, **pevarnor*, ce qui nous interpelle sur un possible biais du protocole de l'enquête (prononciation des enquêteurs, bruits pouvant perturber l'audition, etc.). On peut également évoquer une difficulté liée à la surcharge cognitive, avec notamment en cause la place des pseudo-mots dans la dictée. La dernière phrase (qui est le titre d'un roman et son auteur) cumule les difficultés : « *Le fiancé solennel* de *Joc Pévanor* », avec deux mots difficiles et deux pseudo-mots à la suite dont l'un ressemble au prénom « Jacques » que nous retrouvons dans les réponses des enquêtés « bas » (22 occurrences) et « haut » (six occurrences) : *jack*, *jacques* et même *georges.*

Dans les deux populations, le classement des pseudo-mots par ordre de réussite est presque identique (sauf *dux* que les enquêtés « haut » ont mieux restitué). *Sirape* et *micatol*, les deux pseudo-mots les mieux réussis, présentent une structure syllabique différente (respectivement CV-CVC et CV-CV-CVC) et *pévanor*, le pseudo-mot le moins bien réussi par les enquêtés « bas » et « haut » présente néanmoins une structure identique à celle de *micatol* (CV-CV-CVC). La différence de réussite proviendrait donc plutôt de la composition phonémique de ces pseudo-mots, avec des correspondances phonie-graphie plus faciles à restituer que d'autres, peut-être parce qu'elles sont plus faciles à saisir à l'oral ou parce qu'elles relèvent de schémas plus fréquents à l'écrit.

Néanmoins, nous nous interrogeons sur la nécessité d'intégrer ces pseudo-mots dans une dictée pour repérer des difficultés d'ordre phonographique chez les scripteurs. Nous l'avons vu, les compétences en

phonographie des enquêtés « bas » peuvent être saisies dans l'analyse des graphies erronées des mots de la dictée et nous ne pouvons pas être certains que les erreurs phonographiques des enquêtés « haut » sur les pseudo-mots ne proviennent pas d'un biais du protocole. Nous émettons enfin une réserve sur l'utilisation du pseudo-mot *joc*, trop proche d'un prénom et de *dux*, proche de *luxe* (5 occurrences dans le corpus) qui ne permet par ailleurs d'observer que des compétences phonographiques restreintes : 92% des enquêtés « bas » ont correctement restitué la forme phonologique du début du mot /dy/, les erreurs portant sur la transcription de la suite /ks/.

L'utilisation de pseudo-mots, en revanche, peut permettre d'observer les comportements de scripteurs vis-à-vis des régularités du français, notamment l'utilisation des lettres finales muettes, analyse que nous proposons dans la partie suivante.

4 La graphie des pseudo-mots pour observer les comportements de scripteurs

Bien qu'ils aient été sélectionnés pour leurs compétences à l'écrit, les enquêtés au sein des deux populations ne forment pas un groupe de scripteurs homogène. Cette hétérogénéité observée dans les résultats nous incite à proposer une catégorisation pour tenter de mettre en évidence des comportements de scripteurs.

Les compétences en morphographie sont plutôt discriminantes dans les deux corpus. Pour reconstituer des groupes de scripteurs, nous opérons un nouveau classement (que celui proposé par l'INSEE) en quatre catégories, axé sur les compétences dans ce domaine : les enquêtés « haut » sont divisés en deux catégories, selon leur score de réussite en morphographie mesuré sur l'intégralité de la dictée (groupe 1 : ceux qui ont obtenu 100% de réussite dans ce domaine et groupe 2 pour ceux qui n'ont pas obtenu 100% de réussite). Nous ajoutons au groupe 2, les enquêtés « bas » qui ont réussi les deux marques de nombre de la dictée IVQ et partageons les autres enquêtés « bas » en deux groupes en fonction de leur score orthographique (groupe 3 : ceux qui ont obtenu un score de 13 à 22 et groupe 4 : ceux qui ont obtenu un score de 0 à 12). Nous utilisons le score de 12 comme frontière, parce qu'aucun enquêté haut n'a obtenu de score < à 13 sur 24 sur la dictée IVQ.

Nous observons dans la partie suivante, le comportement des groupes de scripteurs vis-à-vis de la graphie des finales des pseudo-mots (choix d'ajouter ou non une finale muette et choix de la finale muette). Le comportement des scripteurs face aux finales constitue-t-il un marqueur de réussite ou l'enregistrement de régularités ? Nous comparons alors les finales choisies avec la fréquence des finales sur les mots du français, pour voir si les choix des scripteurs seraient la marque d'une sensibilité à une régularité.

Avant de présenter cette analyse, nous soulignons que dans les deux corpus, nous observons une très bonne restitution des finales muettes sur les mots qui possèdent une marque flexionnelle (*pharmacie*, *épicerie*, *librairie*, *fort*) et ceux dont la présence d'un *e* muet final permet la prononciation de la consonne finale (*tomate*, *cerise*, *fromage*), avec plus de 90% de restitution pour chaque mot dans le corpus « bas » et très peu d'erreurs dans le corpus « haut ». Chez les enquêtés « bas », la restitution du *e* muet final est particulièrement élevée sur *tomate*, *confiture* et *fromage.* Sur *cerise* en revanche, nous constatons davantage d'erreurs, avec des graphies comme **ceriz ou *seris.* Face aux mots qui n'ont pas besoin de finale muette (*sel*, *alcool*), nous constatons que peu d'enquêtés « bas » ont ajouté une finale muette. Sur *sel* nous relevons l'ajout d'un *e* muet final chez 2% des enquêtés, et sur *alcool* chez 9% des enquêtés (**alecoole,*alcole).*

Dans les deux corpus, les enquêtés (y compris ceux qui ont le plus de difficultés), possèdent une bonne maitrise des régularités du français sur les finales muettes que l'on observe notamment dans la restitution des marques flexionnelles.

4.1 *Micatol* et *sirape* : *e* muet ou absence de finale muette ?

Nous classons dans le tableau 2, le choix des scripteurs dans chaque groupe concernant l'ajout ou non d'une finale muette (ici *e* muet final) pour les pseudo-mots *micatol* et *sirape*. Pour le calcul des

pourcentages, nous avons pris en compte toutes les graphies de *micatol* et *sirape* dont la dernière syllabe restitue la structure syllabique phonologique (CVC pour ces deux pseudo-mots) et le dernier phonème (/l/ pour *micatol* et /p/ pour *sirap*).

	MICATOL		**SIRAPE**	
	-le : *e* muet final	-l : pas de finale muette	-pe : *e* muet final	-p : pas de finale muette
groupe 1	12%	88%	31%	69%
groupe 2	9%	91%	40%	60%
groupe 3	33%	67%	59%	41%
groupe 4	56%	44%	74%	26%

Tableau 2. Proportion de finales muettes sur *micatol* et *sirape* pour chaque groupe de scripteurs

Pour *micatol*, l'ajout d'une finale muette caractérise le groupe le moins compétent (ajoute d'un *e* final plus d'une fois sur deux). Par contraste, les groupes 1 et 2 privilégient très clairement une graphie minimale, sans ajout de lettre muette.

Pour *sirape*, on observe toujours la même prégnance concernant la présence du *e* muet final, davantage présent dans les graphies des groupes 3 et 4. Le phénomène est marqué pour le groupe 4 (74% des graphies) alors que la consonne finale est davantage pratiquée par les groupes 1 et 2.

Ces résultats sont à interpréter à la lumière de la fréquence des finales sur les mots[14]. Quelques mots se terminent par -ap en français et la tendance est plutôt à la prononciation de la consonne : *rap*, *cap*, *gap* vs *drap*, *sparadrap*. Ainsi, tout se passe comme si les enquêtés les plus compétents enregistraient davantage que le *e* muet n'est pas obligatoire pour que le « p » en finale se prononce. Pour *micatol*, dont le « l » ne représente jamais une finale muette à la différence du « p » en finale, il existe davantage de mots se terminant par -ole (182 mots) que par -ol (141 mots) mais ici encore, le *e* muet n'est pas indispensable.

4.2 *Gobar* et *pévanor* : différentes finales

Dans le tableau 3, nous observons le comportement de chaque groupe de scripteurs concernant les finales sur les pseudo-mots *gobar* et *pévanor* : nous regardons si les scripteurs ont opté pour une lettre muette ou non et nous détaillons quelle finale est privilégiée (*d*, *e*, *t*, *s*). Pour le calcul des pourcentages, nous avons pris en compte les graphies dont la dernière syllabe restitue la structure syllabique phonologique CVC et le dernier phonème /r/.

	GOBAR				**PEVANOR**			
	-rd	-re	-rt	-r	-rd	-re	-rt	-r
groupe 1	58%	3%	21%	18%	39%	13%	3%	45%
groupe 2	40%	4%	12%	44%	19%	21%	6%	55%
groupe 3	34%	15%	11%	41%	16%	24%	6%	53%
groupe 4	6%	34%	6%	54%	10%	36%	12%	41%

Tableau 3. Proportion de finales muettes sur *gobar* et *pévanor* pour chaque groupe de scripteurs

À l'inverse des pseudo-mots précédents, on observe que sur *gobar*, plus les enquêtés sont compétents et plus ils privilégient une finale muette, majoritairement *d*, voire *t*, le *e* étant leur dernier choix (très peu représenté). Le groupe 4, en revanche, privilégie une finale minimale (54% des cas) et ceux qui choisissent d'ajouter une lettre muette privilégient le *e*, comme dans les pseudo-mots précédents.

Pour *pévanor*, les différences entre les groupes sont moins nettes, mais on observe plus d'homogénéité dans le choix d'une graphie finale minimale (environ 1 cas sur 2). Quand ils optent pour une finale muette, la tendance est la même que pour *gobar* : c'est très clairement le *d* pour le groupe 1 et préférentiellement le *e* pour le groupe 4. Le choix des lettres muettes dans le groupe 4 correspond par ailleurs à la fréquence des finales observée sur les mots qui se terminent par /ar/ en français : -ard (288 mots), -ar (117 mots), -are (67 mots) et enfin -art (51 mots).

5 Conclusion et perspectives didactiques

Notre analyse d'un corpus de 447 dictées réalisées par des scripteurs considérés en difficulté en réception de l'écrit dans l'enquête IVQ et de 87 dictées de scripteurs ordinaires laisse apparaitre une grande hétérogénéité dans les deux corpus, révélatrice d'une réalité multiforme des compétences lorsqu'il s'agit d'orthographe, y compris chez les publics repérés en difficulté avec l'écrit. Si notre analyse du lexique montre que la plupart des enquêtés éprouve des difficultés lorsqu'il s'agit de transcrire des mots peu fréquents ou aux correspondances phonie-graphie irrégulières (même si les difficultés observées sont plus prégnantes chez certains), c'est surtout la morphographie qui apparait comme un marqueur important de disparités dans les compétences à l'écrit, particulièrement lorsque le marquage du nombre n'est pas facilité par un signal clair. Enfin, l'analyse des pseudo-mots et particulièrement des finales choisies, révèle que les enquêtés les plus en difficulté possèdent d'autres compétences que la phonographie : ils ont enregistré des régularités dans l'écriture du français que nous observons dans l'utilisation de finales muettes qui correspondent à des configurations fréquentes en français. Ils ont notamment conscience du rôle du *e* muet en français, qu'ils utilisent majoritairement lorsqu'ils choisissent d'ajouter une finale muette. Si nous n'avons relevé aucune finale muette aberrante, nous remarquons néanmoins que leur répertoire est plus restreint que celui des scripteurs les plus experts qui n'ont pas systématiquement recours au *e* muet lorsqu'ils choisissent d'ajouter une finale mais optent, semble-t-il, pour des lettres muettes plus fréquentes (par exemple le *d* en finale des mots qui se terminent par /ar/).

Il apparait donc, d'après les résultats de cette enquête, qu'une grande majorité des enquêtés « bas » possède des compétences phonographiques suffisantes, avec une bonne connaissance des règles du système phonologique du français qui leur confèrent des qualités scripturales non négligeables, leur permettant de communiquer à l'écrit, même si leur orthographe peut s'écarter de la norme.

Dans une perspective didactique, nous avons également montré la nécessité de proposer un test d'écriture révisé si l'on souhaite évaluer les compétences à l'écrit à partir d'une dictée. Il apparait en premier lieu que les pré-tests en réception (lecture et compréhension) proposés par l'INSEE, sont insuffisants pour déterminer de possibles difficultés en production écrite. Nous l'avons vu, être moins bon que les autres en lecture ou en compréhension ne signifie pas nécessairement ne pas maitriser l'orthographe, et inversement une partie des enquêtés « haut » pourtant considérés sans difficulté particulière avec l'écrit a montré des fragilités en orthographe, notamment sur le rayon bricolage. Pour tester les compétences en littératie, un test d'écriture proposé à tous les enquêtés sans distinction de niveau nous parait nécessaire et le choix des items est donc primordial.

Les pseudo-mots, ayant entrainé des scores de réussite relativement importants, y compris chez les enquêtés « bas », s'avèrent peu discriminants. On peut dès lors se questionner sur leur intérêt pour évaluer les compétences en phonographie. La graphie des mots permet d'identifier les enquêtés qui ont des difficultés dans ce domaine, au-delà du respect de la norme orthographique (ex. *rhume* : **rime*, **rune*, **rom*). En revanche, pour tester les compétences orthographiques, nous n'estimons pas pertinent d'utiliser des mots particulièrement rares ou irréguliers. Les résultats sur les mots moins fréquents de l'enquête n'ont en effet pas produit d'opposition nette entre les deux groupes et n'ont pas permis de déceler des difficultés d'ordre lexical ; ils révèlent davantage la complexité de la question des compétences orthographiques et du rapport particulier que chaque individu entretient avec la norme (ex. la graphie de *solennel*). Dans la perspective d'une évaluation plus complète des compétences phonographiques, des mots fréquents pourraient permettre de tester des digrammes en concurrence avec d'autres digrammes ou trigrammes (ex. / ɛ́ / : *pain, peinture, vin* - / o / : *haricot, eau, chaussures,* etc.).

La morphographie en revanche, dérivationnelle et flexionnelle, qui apparait comme un marqueur important de disparités dans les compétences à l'écrit des scripteurs des deux populations, gagnerait à être davantage testée dans la dictée, notamment dans différents contextes (accord en nombre proposé dans des structures syntaxiques simples et plus complexes, avec présence ou non d'un signal pour induire les marques de nombre). Dans le cadre de la révision de cette dictée, d'autres difficultés pourraient être testées, comme l'accord en genre et en nombre de l'adjectif (ex. des *haricots verts*, *une bande dessinée*, *de la peinture bleue*), le participe passé vs l'infinitif (ex. *si la boulangerie est fermée*, *passer au supermarché*) et quelques logogrammes parmi ceux qui sont socialement valorisés (ex. *et* vs *est*, *a* vs *à*), sans pour autant investiguer des zones de grande fragilité où la variabilité est très importante et où la pertinence des atentes normatives est discutable (conjugaisons irrégulières, accord complexe du participe passé).

Références bibliographiques

ANLCI (2003). Lutter ensemble contre l'illettrisme. Cadre national de référence [en ligne], consultée le 30 novembre 2015, URL < http://www.anlci.gouv.de/content/download/907/…/cadre de reference.pdf>.

Blondel, C., Conseil, J. (2015). L'épreuve de production écrite dans l'enquête IVQ 2011 : variation et compétences d'écriture des scripteurs en difficulté avec l'écrit dans les régions Haute-Normandie et Rhône-Alpes. *Glottopol, n°26,* pp. 2-20.

Blondel, C., Brissaud, C., Rinck, F. (2015). Apports de deux enquêtes sur les connaissances orthographiques et les pratiques de l'écrit des Rhônalpins, *Rapport à la DGLFLF (à paraitre).*

Brissaud, C. (2015). Quels enseignement tirer de quatre-vingt-dix-huit dictées de troisième : du décalage entre prescription et acquisition des élèves. *Le français aujourd'hui, n°190,* pp. 61-70.

Fayol, M. (2013). *L'apprentissage de l'écrit*. Paris : PUF.

Guérin-Pace, F. (2009). Illettrismes et parcours individuels. *Economie et statistique, n°4242-425,* pp. 49-62.

Jaffré, J-P. (2004). La littéracie : histoire d'un mot, effets d'un concept, in Barré-De Miniac, C., Brissaud, C., Rispail, *La littéracie. Conceptions théoriques et pratiques d'enseignement de la lecture-écriture*, pp. 21-41. Paris : L'Harmattan.

Jeantheau, J-P. (2015). La dictée dans les enquêtes sur la « littéracie » des adultes ; pratiques, résultats, exemples d'analyse, perspectives. *Glottopol, n°26,* pp. 191-206.

Lété, B. (2008). La consistance phonographique : une mesure statistique de la complexité orthographique, In C. Brissaud, J.-P. Jaffré & J.-C. Pellat. *Nouvelles recherches en orthographe*, pp.85-99. Limoges: Lambert-Lucas.

Moreau, A-C., Hébert, M., Lépine, M., Ruel, J. (2003). Le concept de littératie en francophonie : que disent les définitions ? *CNRIS, vol. 4, n°2,* pp. 14-18.

OCDE (2000). La littératie à l'ère de l'information : rapport final de l'enquête internationale sur la littératie des adultes. Paris : Les éditions de l'OCDE.

Ortéga É., Lété B. (2010). *eManulex: Electronic version of Manulex and Manulex-infra databases.* [en ligne], consultée le 30 novembre 2015, URL < http://www.manulex.org>.

Peereman, R., Lété, B., Sprenger-Charolles, L. (2007). Manulex-infra: Distributional characteristics of grapheme-phoneme mappings, infra-lexical and lexical units in child-directed written material. *Behavior Research Methods, n°39,* pp. 593-603.

[1] IALS (International Adult Literacy Survey), ALL (Adult Literacy and Life Skills Survey), ou PIAAC (Programme for International Assessment of Adult Competencies, OCDE, 2013).

[2] Les épreuves d'évaluation ont été conçues par différentes équipes de recherche : l'équipe PsyEF dirigée par J.-M. Besse, Université Lyon II (module d'orientation, module bas pour les personnes en difficulté, exercice de compréhension orale), l'équipe EVA avec C. Charron et C. Meljac de l'Université Rennes II et de l'hôpital Sainte Anne (exercices de calcul), l'équipe de H. Tardieu de l'Université Paris V et S. Bueno, H. Megherbi, A. Seigneuric

de l'Université Paris XIII (exercices complexes des modules et du module commun) et enfin F. Guérin-Pace de l'Ined pour la partie biographique (questions sur les compétences utilisées au quotidien et les difficultés éventuelles rencontrées).

[3] Les résultats présentés font parti d'un projet conduit par Carole Blondel, grâce à une allocation doctorale de recherche financée par la région Rhône-Alpes (ARC8).

[4] Il est fait allusion ici aux voies d'assemblage et d'adressage. Selon Besse et al. (2004), le recours à la voie d'assemblage consiste à « associer une unité graphique (ou phonogramme) à un phonème, (…) en production écrite (traitement phonographique), puis à combiner la suite des unités pour en faire un mot, et habituellement une signification. Cette procédure n'est employée par le lecteur-scripteur que pour lire/écrire des mots inconnus de lui ou des mots sur lesquels il bute. La voie d'adressage, quant à elle, consiste, (…) en production écrite, à produire un mot ou une partie de mot en « rappelant », par un effort délibéré ou non, une forme mise en mémoire. Cette procédure suppose que le mot a déjà été rencontré, traité efficacement et stocké en mémoire. » (Besse et al., 2004 : 166).

[5] Dans ce calcul, 100 représente le degré de consistance maximal, correspondant à une association phonème-graphème qui ne présente aucune ambiguïté. L'indice de consistance a été calculé sur le lemme.

[6] L'indice de fréquence relevé ici est la fréquence estimée d'usage pour un million de mots (et non le nombre d'occurrences du mot).

[7] La consistance, qui est calculée à partir de la fréquence d'association des phonèmes et des graphèmes révèle ici que ce mot, malgré sa structure phonographique rare, est particulièrement consistant. Nous attribuons le degré de consistance élevé de ce mot à la rareté de l'association entre le phonème /a/ et le graphème « en » ainsi qu'au degré élevé de consistance de la 1ère syllabe *-so*.

[8] Le laboratoire DYSOLA a bénéficié d'un financement de la région Haute-Normandie pour mener cette enquête en Haute-Normandie.

[9] Pour davantage d'informations sur le déroulement de l'enquête et sur le protocole, voir le rapport soumis à la DGLFLF (Blondel, Brissaud, Rinck, 2015 à paraitre) ainsi que l'article paru dans Glottopol (Blondel, Conseil, 2015).

[10] Le taux d'illettrisme n'est établi que pour les personnes âgées de plus de 16 ans ayant été scolarisées en France (Cadre national de référence de l'ANLCI, 2003).

[11] Dans cette analyse, nous ne prenons pas en compte la présence ou non de césure ainsi que la présence ou non d'un trait d'union lorsque le mot composé n'est pas soudé (*antirhume, anti-rhume*). Les deux éléments sont analysés séparément, comme deux unités distinctes *anti* et *rhume*.

[12] La morphographie désigne les marques de nombre nominales et verbales.

[13] Nous utilisons le terme lexical dans une acception large pour faire référence à la graphie des mots hors marques flexionnelles pour les noms, adjectifs et adverbes (ex. *tomate, fort*) et au radical pour les verbes (ex. *sent-* pour *sentent*, *arrach-* pour *arracher*).

[14] Les fréquences des finales sur les mots ont été calculées avec la version électronique du *Petit Robert*, 2011.