



GLOTTOPOL

Revue de sociolinguistique en ligne
n° 26 – juillet 2015

La dictée, une pratique sociale emblématique

Numéro dirigé par Catherine Brissaud et Clara
Mortamet

SOMMAIRE

- Catherine Brissaud et Clara Mortamet : *Présentation*
- Malo Morvan : *La skrivadeg, une communion en paradoxes*
- Evelyne Delabarre et Marie-Laure Devillers : *Mise(s) en œuvre d'une activité orthographique : la dictée*
- Téguia Bogni et Mohamadou Ousmanou : *Pratiques de la dictée en classe de français dans l'enseignement secondaire au Cameroun. Analyse des approches et de la performance d'élèves dans un lycée à Ngaoundéré*
- Danièle Cogis, Carole Fisher et Marie Nadeau : *Quand la dictée devient un dispositif d'apprentissage*
- Catherine Combaz et Marie-Laure Elalouf : *Une phrase dictée, trois enseignants, trois formes d'étayage*
- Véronique Miguel Addisu : *Conscience métalinguistique et norme orthographique : qu'apprend-on en corrigeant sa dictée ?*
- Bruno Trosseille par Catherine Brissaud : *Entretien : La place de la dictée dans les évaluations conduites par la DEPP*
- Sophie Anxionnaz : *Le barème graduel : l'évaluation de la dictée au service des apprentissages*
- Eugenie Grace Eshoh Ndofo et Veronica Ebica Odey : *Difficultés et erreurs dans le passage de la phonie à la graphie en dictée de langue française : cas des apprenants du FLE de l'université de Calabar (Nigéria)*
- Jeanne Gonac'h : *Des variations orthographiques dans les dictées de CMI*
- Sophie Briquet-Duhazé : *La dictée de lettres : qu'en est-il au CMI ?*
- Jean-Pierre Jeantheau : *La dictée dans les enquêtes sur la « littéracie » des adultes : pratiques, résultats, exemples d'analyses, perspectives.*
- Carole Blondel et Jeanne Conseil : *La dictée dans l'enquête IVQ : variation et compétences d'écriture des scripteurs en difficulté avec l'écrit dans les régions Haute-Normandie et Rhône-Alpes*
- Mickaël Lenfant : *La dictée dans un bilan orthophonique : analyse psycho-sociolinguistique des données de l'étalonnage du texte à trous d'Exalang 8-11 ans*

LA DICTÉE DANS L'ENQUÊTE IVQ : VARIATION ET COMPÉTENCES D'ÉCRITURE DES SCRIPTEURS EN DIFFICULTÉ AVEC L'ÉCRIT DANS LES RÉGIONS HAUTE-NORMANDIE ET RHÔNE-ALPES

Carole BLONDEL* et Jeanne CONSEIL**

*Univ. Grenoble Alpes, Laboratoire Lidilem, F-38040, Grenoble

**Université de Rouen, EA 4701 Dysola

Dans une société de connaissances et d'information, dans laquelle l'écrit occupe une place prépondérante, tant sur le plan social que personnel, la question des compétences à l'écrit des adultes est de plus en plus prégnante¹. Plusieurs enquêtes, nationales et internationales, ont ainsi cherché depuis le milieu des années 90 à mesurer et à évaluer le niveau de littératie des adultes (IALS, ALL, PIIAC, par exemple). Remettant en question la fiabilité de ces enquêtes, notamment en matière de méthodologie (Bastyns, 2012 ; Besse *et al.*, 2009 ; Guérin-Pace, 2009), la France a décidé de mettre au point sa propre mesure de la littératie des adultes qui prenne en compte les faibles niveaux de compétences comme d'autres enquêtes telles que *Leo* en Allemagne ou *Skills for Life* en Grande Bretagne. C'est dans ce contexte, et à la demande de l'INSEE que l'enquête sur les usages de l'Information dans la Vie Quotidienne (IVQ) a été conçue. Un protocole de recherche conséquent² a été mis en place avec la coopération d'équipes de recherches de plusieurs universités françaises³. La dernière édition de l'enquête, réalisée en 2011, a été conduite auprès de 13 750 personnes âgées de 16 à 65 ans résidant en France métropolitaine.

L'enquête Information et Vie Quotidienne (IVQ)

L'enquête IVQ porte sur l'usage de l'information dans la vie quotidienne et vise deux objectifs principaux : mesurer le niveau de compétences à l'écrit des adultes dans les tâches

¹ Nous ne discuterons pas ici de cette place accordée à l'écrit dans les sociétés occidentales contemporaines. Cette question a largement été soulevée par Lahire (1999) notamment qui a souligné le caractère ethno et socio-centré de cette approche de nos sociétés. Bastyns (2012) montre quant à elle comment cela nourrit les politiques libérales.

² IVQ a été réalisée à titre expérimental en 2002 puis mise en place en 2004 pour être renouvelée en 2011.

³ L'équipe EVA de l'Université de Rennes et de l'Hôpital St Anne ; H. Tardieu et son équipe de l'Université Paris 5 ; l'équipe PsyEF (Psychologie cognitive de l'Education et de la Formation) de l'Université de Lyon, dirigée par Jean-Marie BESSE, en collaboration avec l'ANLCI (Agence Nationale de Lutte contre l'illettrisme).

de la vie quotidienne et estimer la proportion d'adultes se trouvant en situation de grande difficulté à l'écrit dans leur vie quotidienne. L'intérêt majeur de cette enquête est, d'une part, qu'elle ne se réfère pas à des situations d'apprentissage scolaires mais à des « conduites de communication relationnelle et sociale de la vie courante » (Besse *et al.*, 2009) : elle propose notamment d'évaluer les compétences des adultes non pas à partir d'exercices décontextualisés mais d'après des situations et des supports issus de la vie quotidienne (page de programme de télévision, pochette d'album CD, bulletins d'information, liste de courses, etc.). Il est en effet souvent reproché aux tests d'être trop proches des situations scolaires, ce qui risque de placer les personnes les plus en difficultés avec l'écrit en position d'échec, ces situations faisant écho à des souvenirs douloureux liés à leur scolarité notamment. Les tests scolaires sont par ailleurs souvent éloignés de situations qui correspondent aux usages et aux fonctions de l'écrit à l'âge adulte et ne permettent pas de tester les capacités liées à l'usage de l'écrit, dans leur dimension fonctionnelle. Cette enquête comprend, d'autre part, une épreuve dite de « production écrite », une dictée, faisant l'objet de notre analyse. Elle est la première à proposer de mesurer la capacité des adultes à produire des mots écrits. Seules les compétences en compréhension écrite et en lecture sont habituellement mesurées pour évaluer les compétences en littératie des adultes. L'enquête IVQ présente enfin l'intérêt de tester d'autres capacités telles que le niveau de compréhension orale et la numératie, des compétences qui, selon différents travaux de recherche en linguistique de l'écrit et en psycholinguistique, sont liées à la réussite en lecture et en écriture (Besse *et al.*, 2009). Le protocole comprend également un module biographique considérable contenant des informations diverses sur les personnes interrogées : leur parcours personnel, scolaire et professionnel, leur origine, les langues parlées à la maison et dans l'enfance, leur rapport à l'écrit, leurs pratiques de l'écrit, etc. Ces nouvelles données permettent ainsi de dresser un portrait précis des difficultés rencontrées par ce public.

L'épreuve de production écrite

L'épreuve de production écrite, la seule dans l'enquête qui permette de tester les compétences en écriture, fait partie du module « bas », appelé aussi « module ANLCI » puisque conçu par l'équipe PsyEF de l'Université de Lyon, sous la direction de J.M. Besse, avec la collaboration de l'Agence Nationale de Lutte contre l'Illettrisme. Centré sur le repérage et l'évaluation des personnes qui présentent des difficultés dans les domaines fondamentaux de l'écrit, ce module vise à évaluer les compétences en lecture (tâche de lecture et d'identification de mots et de pseudo-mots), en écriture (production sous la dictée de mots et de pseudo-mots) et en compréhension écrite (compréhension d'un texte simple). Seules les personnes ayant été repérées en difficulté avec l'écrit lors du passage, au début de l'enquête, d'un module d'orientation, ont été dirigées vers ce module ANLCI et ont passé, par conséquent, l'exercice de production écrite.

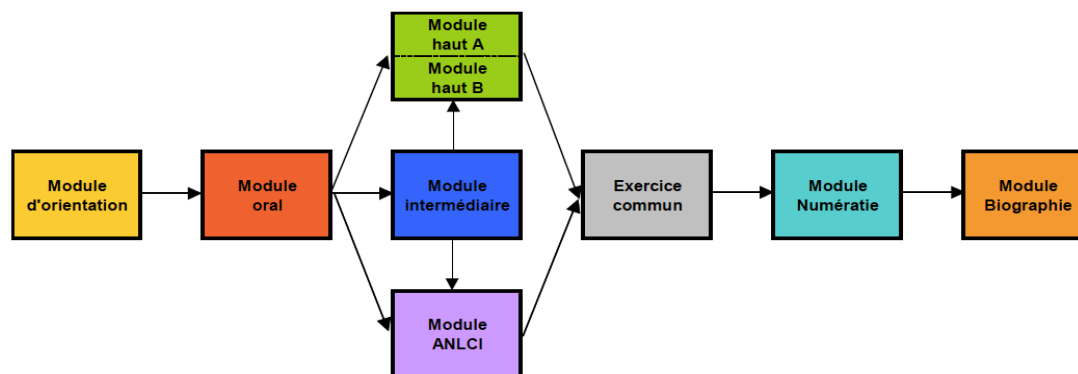


Figure 1. Algorithme des modules de l'enquête IVQ 2011, *Dossier d'Aval Haute-Normandie (septembre 2013 : page 85)*

Évaluer les compétences à l'écrit d'un public adulte en difficulté avec la langue écrite n'est pas une tâche aisée, tant sur le fond (quelles connaissances et quelles capacités d'écriture recueillir ?), que sur la forme (quels exercices proposer pour recueillir des données comparables et significatives ?), que sur le public cible (comment inciter un public en difficulté à produire de l'écrit en prenant le risque d'exposer ses faiblesses ?). Dans l'enquête IVQ, le choix a été fait de mesurer les compétences en écriture des adultes les plus en difficulté par l'intermédiaire d'une dictée, mais sans que cet exercice ne soit présenté comme tel aux personnes interrogées.

L'épreuve de production écrite se présente sous la forme d'une liste de courses à rédiger à l'attention d'un livreur. Ce « scénario » a été mis en place pour répondre à deux principes fondamentaux de l'enquête : ne pas mettre l'enquêté en difficulté en lui présentant des exercices connotés socialement et/ou scolairement (comme la dictée) et proposer des exercices proches de situations vécues dans la vie quotidienne (ici, une liste de courses). Cette dictée comprend des mots et des pseudo-mots, inventés pour l'exercice, que l'enquêté doit écrire sur une feuille de papier libre. La version proposée dans l'enquête de 2011, se présente sous la forme suivante (les pseudo-mots sont ici indiqués en italiques) :

Pharmacie
 Anti-rhume *Micatol* ou *Sirape*
 Alcool à 90° *Dux*
 Epicerie
 Tomates de pays
 Confiture
 Cerises
 Sel *Gobar*
 Trois fromages qui ne sentent pas fort
 Librairie
 Le fiancé solennel de *Joc Pévanor*

Les items de cette dictée ont été sélectionnés par Besse et son équipe (2009) selon leur longueur (nombre de phonèmes à transcrire), leur caractère régulier ou irrégulier sur le plan phonographique et les procédures de traitement qu'ils requièrent pour être transcrits. Ils permettent de rendre compte des connaissances lexicales des enquêtés et de définir leurs difficultés dans les domaines orthographique, phonographique et morphographique. La forme de cette épreuve – un nombre restreint d'items et une liste de mots – ne permet toutefois d'observer qu'un nombre limité de phénomènes linguistiques : quelques spécificités du français (lettres internes ou finales muettes : *rhume*, *pays*, *épicerie*), l'accord du nom *fromages* et du verbe *sentent*, dans le contexte particulier d'une expression par proposition relative (*trois fromages qui ne sentent pas fort*).

La transcription des pseudo-mots – *micatol*, *sirape*, *dux*, *gobar*, *joc*, *pevanor* – demande au scripteur de recourir spécifiquement à un traitement phonographique de la forme entendue et permet de mesurer le niveau de conscience phonographique des enquêtés sur les 18 phonèmes proposés (5 voyelles et 13 consonnes). L'exercice permet également de vérifier la capacité à utiliser les voies d'assemblage et d'adressage⁴ (lorsque le choix d'un graphème se présente, en faisant appel à des formes orthographiques stockées en mémoire).

⁴ 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).

Les mots considérés comme phonographiquement réguliers par Besse *et al.* (2009) – *sel, anti*⁵, *tomate, cerise, épicerie, confiture* – peuvent être correctement orthographiés en recourant à la voie d’assemblage uniquement et/ou à la voie d’adressage. Si le mot n’est pas connu par le scripteur, la seule procédure phonographique permet en effet d’orthographier ces mots réguliers. Le scripteur peut aussi ne pas connaître le mot et fonctionner par analogie en s’appuyant sur d’autres mots déjà rencontrés et connus.

Les mots considérés comme phonographiquement irréguliers par les concepteurs du test – *pays, rhume, alcool, fiancé, solennel, pharmacie* – nécessitent, quant à eux, le recours à la voie d’adressage ; l’usage de la voie d’assemblage seule conduisant à une transcription orthographique erronée.

Cette classification nous permet de formuler l’hypothèse selon laquelle les mots réguliers seront mieux restitués, sur le plan orthographique, que les mots irréguliers et que les marques flexionnelles, qui demandent davantage d’expertise de la part du scripteur. Pour évaluer les performances à l’écrit et permettre une analyse quantitative des données, l’équipe PsyEF, l’ANLCI et l’INSEE ont défini des scores de réussite aux items de la dictée. Dans cette épreuve de production écrite, 18 items et 2 marques morphologiques ont ainsi été évalués⁶. Pour chacun des 18 items, 1 point a été accordé selon le critère du respect de la norme orthographique pour les mots et du principe phonographique pour les pseudo-mots. 4 points ont été ajoutés pour les graphies *fromages* (2 points) et *sentent* (2 points), si les marques obligatoires de pluriel ont été produites. Le taux de réussite, calculé sur un score total de 22 points, a ensuite permis de déterminer des « profils de scripteur » pour chacun des enquêtés :

- 0 : Personnes n’ayant pas pu faire les exercices du fait de trop grandes difficultés en lecture ou en français
- 1 : Personnes ayant moins de 40 % de réussite à l’épreuve de production écrite (score inférieur à 9/22)
- 2 : Personnes ayant entre 40 et 60 % de réussite (score compris entre 9 et 13/22)
- 3 : Personnes ayant entre 60 et 80 % de réussite (score compris entre 14 et 17/22)
- 4 : Personnes ayant plus de 80 % de réussite (score supérieur à 17/22)
- 5 : Personnes n’ayant pas de difficulté dans les trois domaines fondamentaux de l’écrit. Cette dernière catégorie correspond aux personnes qui ont obtenu plus de 80 % de réussite dans les trois épreuves du module ANLCI (production écrite, lecture et compréhension d’un texte écrit).

Ces profils entrent en effet en correspondance avec ceux établis pour les épreuves de lecture et de compréhension écrite, pour déterminer le niveau de difficulté des enquêtés dans les domaines fondamentaux de l’écrit : graves difficultés (moins de 40 % de réussite), fortes difficultés (entre 40 % à 60 % de réussite), ou difficultés partielles (entre 60 % à 80 % de réussite). L’INSEE estime que les deux premières catégories (difficultés graves et fortes) relèvent de « situations préoccupantes » face à l’écrit. C’est parmi ces situations dites « préoccupantes » et selon les critères définis par l’ANLCI que 7 % des français (INSEE, 2012) ont été considérés comme étant en situation d’illettrisme⁷ en France. En revanche, les enquêtés qui ont obtenu plus de 80 % de réussite à chacun des trois exercices du module ANLCI ne sont pas considérés en difficulté dans les domaines fondamentaux de l’écrit.

⁵ Besse a considéré le composé *anti-rhume* comme deux unités lexicales. Nous conservons ce traitement distinct des deux éléments pour notre analyse mais nous observerons aussi la présence ou non de la césure en prenant en considération le fait que de nombreuses formes courantes qui commencent par « anti » sont soudées (*antivol*, etc.).

⁶ L’INSEE a fait le choix de ne corriger que 18 items sur les 39 dictés, certains ayant été jugés redondants ou non pertinents pour établir un diagnostic avec le présent test.

⁷ Selon le cadre national de référence (2003) de l’ANLCI, le taux d’illettrisme n’est établi que pour les personnes âgées de plus de 16 ans ayant été scolarisées en France.

L'enquête IVQ en Haute-Normandie et en Rhône-Alpes : quelques résultats

Nous disposons, grâce à l'enquête IVQ, d'un corpus national inédit de plus de 2000 dictées produites par les adultes considérés en potentielle difficulté à l'écrit dans l'enquête. En Haute-Normandie, 1622 personnes ont passé l'enquête IVQ en 2011. Parmi elles, 262 personnes ont été dirigées vers le module ANLCI, soit 16 % des enquêtés. La région Haute-Normandie (désormais HN) a bénéficié d'une extension de l'enquête permettant d'augmenter le nombre d'enquêtés, ce qui n'a pas été le cas en Rhône-Alpes (désormais RA). Pour cette région, 822 personnes ont été interrogées dans le cadre d'IVQ et parmi elles, 108 enquêtés ont été considérés en possible difficulté avec l'écrit et ont passé le module ANLCI, soit 13 % de l'échantillon.

Au terme de l'enquête, l'INSEE a établi un classement du type de difficultés rencontrées par les personnes interrogées dans les domaines fondamentaux de l'écrit. Selon les scores de l'INSEE, les régions Haute-Normandie (Brendler, 2013) et Rhône-Alpes présentent des résultats similaires avec moins d'un point d'écart sur les difficultés constatées, que celles-ci soient graves (6 % en HN et en RA), fortes (5 % en HN et 4 % en RA) ou partielles (6 % en HN et 5 % en RA). Une proportion semblable de personnes (81 % en HN et 83 % en RA) ne sont finalement pas considérées en difficulté particulière avec l'écrit. Globalement, dans le module ANLCI, la proportion d'enquêtés qui ne présentent pas de difficulté (en moyenne 6 %) ou qui présentent des difficultés partielles (36 % en moyenne) est la même dans les deux régions. Les Hauts-normands sont toutefois un peu plus nombreux à présenter des difficultés graves ou fortes dans au moins un des domaines fondamentaux de l'écrit (60 % contre 57 % pour RA).

Dans cet article, nous analysons les 14 mots de la dictée corrigés par l'INSEE : pharmacie – anti – rhume – alcool – épicerie – tomates – pays – cerises – confiture – sel – fromage – sentent – fiancé – solennel, auxquels nous avons ajouté deux nouveaux items : fort et librairie produits par les scripteurs haut-Normands (262 dictées) et rhônalpins (108 dictées)⁸. Nous n'analysons pas ici les pseudo-mots : la variation sur les mots est en effet révélatrice du niveau orthographique des enquêtés tandis que les pseudo-mots ne nous permettent d'analyser que la conscience phonographique des scripteurs, sans pouvoir se référer à une quelconque norme orthographique. L'objectif de cet article est de déterminer le niveau de maîtrise de l'écriture des personnes qui composent ce public particulier, à partir de leurs réponses dans la dictée IVQ. Pour ce travail, mené à la fois à Rouen (par le laboratoire Dysola) et à Grenoble (par le laboratoire Lidilem), des critères adaptés à l'analyse de ces dictées ont été définis et appliqués au traitement de ce corpus.

Nous présentons, dans un premier temps, le taux de réussite sur les 16 mots retenus dans l'épreuve de production écrite en considérant deux critères : la réussite orthographique, calculée sur la base de la restitution de la forme orthographique attendue, et la réussite phonographique estimée selon la conformité de la correspondance phonie-graphie proposée par rapport à la forme phonologique du mot. Dans cette analyse, nous utilisons le terme de phonographie dans une acception très large, qui consiste à observer les graphèmes ou les groupes de lettres utilisés par les scripteurs pour restituer la chaîne orale et à considérer ensuite les graphies comme phonographiquement « acceptables » (telle que *épisseri*) ou « approximantes » (telle que *épiseri*)⁹.

Nous proposons ainsi de classer les réponses des enquêtés en trois catégories : les mots correctement orthographiés, les mots qui ne respectent pas la norme orthographique mais qui sont conformes au principe phonographique (phonographie acceptable et approximante) et les

⁸ Toutes les analyses portent sur un corpus de 370 dictées (262 pour la HN et 108 pour RA)

⁹ Nous définissons cette acception de la phonographie dans la partie consacrée aux graphies analysées selon le critère phonographique (page 10).

mots dont la forme est phonographiquement inacceptable. Cette distinction nous permet de calculer le score de réussite orthographique et phonographique des 16 mots analysés pour chaque enquêté et de définir leur profil de scripteur. Nous nous interrogerons alors sur l'impact de ces deux critères dans la prise en compte des difficultés à l'écrit. Nous mettrons ensuite en évidence le taux de variation observé sur chaque mot et présenterons, en comparant les corpus haut-normand et rhônalpin, les formes graphiques récurrentes produites sur les items *anti*, *rhume*, et *solennel*, qui sont ceux qui ont posé le plus de difficulté aux enquêtés (avec moins de 60 % de réussite orthographique).

Décrire et analyser les productions de mots des scripteurs en difficulté avec l'écrit : choix méthodologiques et premiers résultats

Pour identifier les difficultés rencontrées par les scripteurs dans chacune de ces catégories, nous nous sommes interrogées sur les différences qui pouvaient exister dans la dictée au sein de chaque corpus. Dans la perspective d'une approche qualitative et d'une analyse linguistique la plus fine possible, les seuls critères d'analyse de l'INSEE présentés ci-dessus n'ont pu être retenus. Une évaluation classique, qui consiste à ne repérer que les graphies incorrectes sur le plan orthographique et à établir des scores de réussite, ne permet pas de rendre compte des compétences de ce public à communiquer par écrit de manière autonome. Le barème défini par l'INSEE, axé à 73 % sur l'orthographe lexicale et grammaticale (le pourcentage restant étant orienté sur la réussite phonographique appliquée aux pseudo-mots) sanctionne certains enquêtés en les classant dans des catégories de faibles scripteurs alors qu'ils possèdent de réelles compétences à l'écrit. Pour l'exploitation de ce corpus, nous proposons donc une analyse qui vise à valoriser aussi les compétences phonographiques dans la production de mots écrits des enquêtés. Ces compétences sont évaluées selon la capacité des scripteurs à coder les sons en proposant une forme graphique du mot conforme au système phonographique du français (ex. *farmacie* ou *pharmassi* sont considérées comme des formes acceptables d'un point de vue phonographique bien qu'éloignées de la norme orthographique). Les compétences phonographiques constituent une étape intermédiaire et indispensable avant l'acquisition des conventions d'écriture, et l'expérience d'un système alphabétique semble nécessaire pour développer la conscience phonémique et étendre ainsi la connaissance des phonèmes et de leur association avec les graphèmes (Fayol, 2013). Cette analyse nous permet ainsi de mieux identifier le niveau de compétences, dans le processus d'apprentissage de l'écrit, dans lequel l'enquêté se situe.

Afin d'exploiter au mieux le corpus dont nous disposons, un premier travail de vérification a été effectué pour que le fichier des données soit identique à la version écrite proposée par l'enquêté. Nous avons ensuite effectué un codage complémentaire des données, nécessaire pour prendre en compte davantage de phénomènes linguistiques. À terme, ce travail permettra de procéder à de nouvelles analyses de la variation, notamment pour les marques de nombre sur les mots qui n'ont pas de déterminant : la présence/absence du -s sur les mots *tomate(s)* et *cerise(s)* n'a pas été prise en compte par l'INSEE lors du codage des données. De la même façon, les accents n'ont pas toujours été relevés, qu'il s'agisse d'accents distinctifs comme dans *fiancé* transcrit *fiance* par l'INSEE, ou d'accents ajoutés dans *solennèl* transcrit *solennel* par l'INSEE. Nous avons ainsi restitué dans le fichier la forme graphique produite par les scripteurs en retirant ou en ajoutant les -s finaux (INSEE : *anti-rhume*, *tomates*, *confiture* vs graphie de l'enquêté et version informatique corrigée du corpus : *anti-rhumes*, *tomate*, *confitures*), en restituant les diacritiques (INSEE : *cerise*, *fiance*, *fromages* vs version informatique corrigée : *çerise*, *fiâncé*, *frômages*), les apostrophes (INSEE : *alcole* vs *a'lcole*), ainsi que la présence ou non de césures (INSEE : *enti*, *anti rume*, *accol* vs *en ti*, *antirume*,

laccol). Cette vérification des formes a également permis de rectifier certaines erreurs de saisie ou d'interprétation, inévitables sur un corpus aussi important.

Nous présentons ci-dessous nos choix méthodologiques pour distinguer et analyser les formes produites selon trois catégories : les formes orthographiquement correctes et les formes erronées, déclinées selon le critère phonographique (phonographie acceptable et inacceptable). Nous présentons également les premiers résultats obtenus.

Le critère orthographique

Nous avons calculé un nouveau score orthographique sur les 14 mots de la dictée analysés par l'INSEE : *pharmacie – anti – rhume – alcool – épicerie – tomates – pays – cerises – confiture – sel – fromage – sentent – fiancé – solennel*, auxquels nous avons ajouté deux nouveaux items : *fort* et *librairie*. Nous obtenons ainsi un score orthographique sur 16 points.

Pour définir les formes acceptées sur le plan orthographique dans le corpus, nous avons fait le choix de ne tenir compte que de l'orthographe lexicale, sans considérer les marques morphographiques (marques de genre ou de nombre). Nous n'avons pas tenu compte de l'omission des accents aigus et graves. Selon Lucci et Millet (1994), près de 2/3 des variations qui affectent l'ensemble des diacritiques correspondent à une omission ou à une neutralisation des accents graves et aigus. Les auteurs soulignent une évolution dans la manière de graphier les accents qui, le plus souvent, ne sont pas transcrits. Lorsqu'ils le sont, la distinction accent aigu/accent grave tend à être neutralisée et les accents sont alors réalisés sous la forme d'un trait horizontal ou vertical. Pour notre corpus, nous avons cependant distingué les cas d'adjonction de diacritique sur des graphies telles que *cérise*, *épicerie* que nous considérons comme erronées, sur le plan orthographique et phonographique. Nous avons également pris en compte l'adjonction d'accents circonflexes dans des graphies telles que *tômates* ou *fromâges* considérées comme orthographiquement incorrectes.

Nous présentons, dans le tableau 1, les formes que nous avons retenues comme orthographiquement correctes pour la présente analyse :

Mots	Formes acceptées
pharmacie	pharmacie, pharmacies
anti ¹⁰	anti
rhume	rhume, rhumes
alcool	alcool, alcools
épicerie	épicerie, epicerie, épiceries, epiceries
tomates	tomate, tomates
pays	pays
confiture	confiture, confitures
cerises	cerise, cerises
sel	sel, sels
fromages	fromage, fromages
sentent	sente, sentes, sentent
fort	fort, forts
librairie	librairie, librairies
fiancé	fiancé, fiance, fiancée, fiancee, fiancés, fiances, fiancées, fiancees
solennel	solennel, solennels, solennelle, solennelles

Tableau 1. Relevé des formes considérées comme orthographiquement correctes

¹⁰ 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*.

Analyse de la réussite orthographique sur 16 mots de la dictée

Nous présentons la réussite orthographique sur les 16 mots analysés en déterminant la proportion de réponses qui respectent la forme orthographique attendue (cf. tableau 1). Les autres réponses sont dites erronées puisqu'elles ne respectent pas la norme orthographique et sont constituées des formes phonographiquement acceptables et inacceptables. Nous présentons également le taux de réponses illisibles et de non-réponses ; nous avons en effet considéré l'absence de graphie comme un possible indicateur de difficultés.

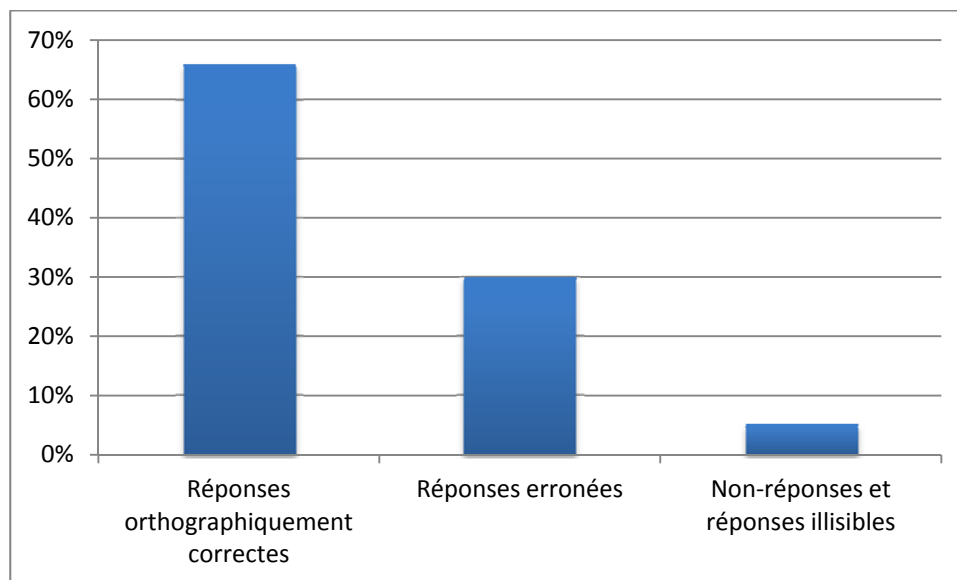


Figure 2. Proportion de réponses orthographiquement correctes, erronées et de non-réponses en HN et RA, pour les 16 mots considérés

En moyenne, 66 % des mots ont été correctement orthographiés dans les deux régions : 68 % des mots en HN et 64 % en RA. La proportion de formes erronées, en moyenne de 30 % dans les deux régions, est également comparable : 28 % en HN et 32 % en RA, tout comme celles des non-réponses (4 % en HN et RA). Plus de 2/3 des mots analysés ont donc été correctement transcrits selon le critère orthographique, dans les deux régions.

Nous présentons ensuite (figure 3) le taux moyen de réponses conformes à la norme orthographique, pour chaque mot. Les résultats, semblables dans les deux régions, sont présentés de manière indifférenciée. Les items ont été classés selon leur ordre d'apparition dans la dictée.

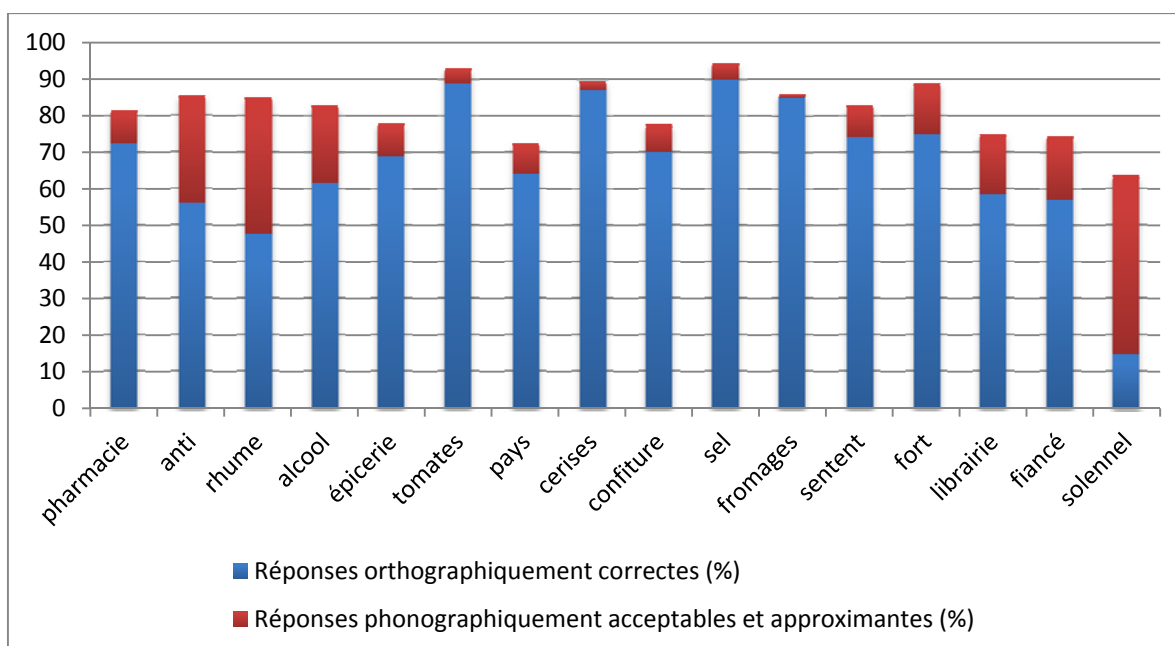


Figure 3. Proportion de réponses correctes et acceptables/approximantes pour chaque mot selon les critères orthographique et phonographique dans les deux régions

Ces premiers résultats montrent une répartition inégale de réussite orthographique selon les mots. Pour les mots les moins bien restitués sur le plan orthographique, la prise en compte du critère phonographique fait augmenter très significativement la réussite. En revanche, l'apport de la phonographie sur les mots les mieux réussis est moindre.

Sur le plan orthographique, *sel*, *tomates*, *cerises* et *fromages* sont les quatre mots les mieux réussis dans les deux corpus, avec plus de 80 % de réussite en moyenne. *Sel* (91 % en HN et 89 % en RA) et *tomates* (90 % en HN, 88 % en RA) approchent même 90 % de réussite dans les deux régions.

En revanche deux mots présentent des taux de réussite moyens inférieurs à 50 % : *rhume*, réussi à 47,5 % (45 % en HN et 50 % en RA) et *solennel* qui présente un taux moyen de réussite de 14,5 % seulement (12 % en HN et 17 % en RA).

D'après l'analyse des scores de réussite orthographique sur chaque mot, l'hypothèse selon laquelle les mots phonographiquement « réguliers » seraient mieux restitués que les mots phonographiquement « irréguliers » se confirme : ce sont les mots pour lesquels la voie d'assemblage pourrait suffire – *confiture*, *sel*, *tomates* et *fromages* – qui ont été les mieux réussis sur le plan orthographique. *Anti*, quant à lui, fait partie des mots ayant posé le plus de difficultés aux scripteurs et seule une analyse des formes orthographiques erronées pourra permettre de déterminer ce qui leur a posé problème¹¹. Pour les autres mots du corpus, le recours à la voie d'adressage était en revanche indispensable pour réussir à produire la forme attendue, ce qui semble poser davantage de difficultés aux enquêtés ; les mots phonographiquement irréguliers *pays*, *alcool*, *fiancé*, *rhume*, et *solennel* ayant ainsi été les moins réussis dans les deux corpus. Le mot *solennel* a particulièrement posé problème du point de vue orthographique : la transcription du phonème /a/ par le phonogramme « e » suivi d'un redoublement de consonnes est en effet une correspondance phonographique rare en français, une difficulté à laquelle s'ajoute la rareté de ce mot dans le langage courant.

¹¹ Se reporter à l'analyse réalisée sur le mot *anti*.

Le critère phonographique

Au-delà de l'analyse du corpus selon le critère orthographique, nous avons distingué 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 (l'INSEE n'ayant fait ce travail que sur les pseudo-mots). Nous considérons, en effet, que des graphies telles que *farmacie* ou *épisserie* n'altèrent pas la capacité communicationnelle de leurs auteurs.

Pour juger de la conformité phonographique des formes et distinguer les graphies phonographiquement acceptables et inacceptables, nous avons observé les graphèmes proposés par les scripteurs pour chaque phonème et accepté toutes les réalisations graphiques plausibles, sans tenir compte des règles de position ou de distribution des graphèmes spécifiques au français et d'usage académique. Dans cette acception très large du principe phonographique, nous acceptons également des formes dont la phonographie est « approximante ». En portant notre analyse sur le graphème, nous avons par exemple admis que le graphème « s » en français puisse avoir les deux valeurs phonologiques /s/ et /z/ (Catach, 1984) quelle que soit la distribution de ce graphème dans le mot. Nous avons considéré ainsi la graphie *épiseri* comme phonographiquement « approximante » en considérant que le graphème « s » correspondait à la transcription phonétique du phonème /s/, même si dans cet exemple la distribution du graphème « s » – entouré de deux voyelles – lui donne la valeur phonologique /z/. Ce choix nous a amenées à prendre en compte pour cette analyse des graphies qui peuvent, au premier abord, poser problème au lecteur, telles que *rum* considéré comme une transcription de la forme orale [rym], bien que « u » dans cette position devant –m final se prononce généralement [ɔ] (*maximum, rhum*).

Nous faisons ce choix d'acception large du principe phonographique à ce stade de notre travail en estimant qu'une graphie comme *épiseri*, même si elle ne respecte pas les règles de position en français, révèle déjà des compétences. À terme, et lorsque nous serons en mesure de croiser les graphies avec les données biographiques dont nous disposons (notamment la nationalité, le lieu de scolarisation, le parcours scolaire et les langues parlées par les scripteurs), nous espérons définir différents niveaux de compétences dont l'un pourrait prendre en compte l'utilisation des règles de position et de distribution d'usage en français. Selon le même principe, nous avons accepté les formes contenant des segmentations non conformes (*en ti, an-ti, et piseri, so la nelle*) avec ou sans trait d'union, ces segmentations n'altérant pas la forme sonore du mot.

Nous avons également pris en compte le fait que la variation observée à l'écrit pouvait s'expliquer par des phénomènes d'assimilation ou par la prononciation inhérente à l'enquêteur. Ne disposant pas de données concernant la manière dont les items ont été dictés, nous avons par exemple accepté la substitution du son /ã/ au son /õ/ dans *sentent* : /sãt/, /sõ/ ainsi que la proximité phonique des mots *sentent* et *semblent* ; l'élision du « e » muet dans *épicerie* : /episəri/, /episri/ ; ou encore l'assimilation du /i/ dans *fiancé* : /fjãse/, /fjãse/.

Nous présentons, dans le tableau 2, quelques exemples de formes graphiques rencontrées dans le corpus et retenues comme acceptables ou approximantes (en italique dans le tableau) sur le plan de la phonographie, pour les régions Haute-Normandie et Rhône-Alpes.

pharmacie	[farmasi]	pharmaci, <i>pharmasi</i> , pharmatie, farhmaci, farmaci, farmacie, <i>farmaçi, farmasi</i>
anti	[ãti]	an-ti, anthi, antie, amtie, en ti, emti, enthi, enti, entie, enty, enthy, hanti, henti
rhume	[rym]	<i>rum</i> , rhûme, ruhm, rumh, ruhme, <i>rum</i> , rume, hrume, <i>rueme</i> , rumeu
alcool	[alcɔl] [alɛcɔl]	alcol, alcole, alccol, alccool, alcoll, alecool, allcolle, halcol
épicerie	[episəri]	epiceri, épiceri, épiscerie, epicerie, <i>episerie, episeri</i> , episserie,

	[episri]	épisserie, <i>et piseri</i> , hêpicerie, <i>episri</i> , <i>épicri</i> , <i>episrie</i>
tomates	[tomat]	<i>tomat</i> , thomate, tomatte, tommatte, tomâte
pays	[pei], [peji]	paiyi, <i>paiis</i> , peyi, <i>payie</i> , pei, péi
confiture	[kôfityR]	confitur, comfitur, confittur, confiture
cerises	[səRiz], [sriz]	<i>ceris</i> , çerise, cerize, ceriz, serise, serises, seris, seriz, <i>criz</i> , srriz
sel	[sɛl], [sel]	sel, sell, selle, cél, cels, sail, séle
fromages	[fRomaʒ]	fromaje, fraumaje, fraumage frômage
sentent	[sât], [sô]	sante, sente, cente, sante, sent, son, semble
fort	[fɔR], [foR]	faure, faur, forre, fohr, for, ford, fore, fors
librairie	[libReRi]	libreri, libréri, librairie, libreris, librerit, librairie, librairi, librais
fiancé	[fjâse], [fijâse]	fiançé, fiancei, fiencer, <i>fiansé</i> , fillancé, fillencé, fillance, fiamcé
solennel	[solanel]	saulanel, sollanelle, solanèl, solanéle, solennèl, solennele, sollennel, solanell, solannelle, sollannelle, solanaile, so la nelle

Tableau 2. Exemples de formes considérées comme phonographiquement acceptables ou approximantes

Nous proposons dans la dernière partie de l'article une analyse qualitative des réponses orthographiques erronées récurrentes sur les trois mots de la dictée qui ont été les moins bien réussis (*anti*, *rhume* et *solennel*). D'autres analyses seront effectuées à terme pour comprendre les procédures qui ont pu être activées au moment de la production écrite, dans l'objectif de mieux cerner les compétences de ces scripteurs : par exemple l'emploi du « ç » devant les voyelles -e et -i (*çerise*, *farmaçi*), l'emploi du « h » à l'initiale (*hanti*, *halcol*), l'utilisation de diacritiques (*frômage*), en comparant ces différents phénomènes dans l'écriture des mots.

Analyse de la réussite phonographique parmi les réponses orthographiques erronées sur 16 mots de la dictée

Nous présentons le taux de réussite phonographique calculé à partir des réponses qui ne respectent pas le principe orthographique (soit 28 % du corpus HN et 32 % pour RA) parmi les 370 dictées (262 en HN et 108 en RA). Pour cette analyse, nous souhaitons observer la compétence phonographique des scripteurs qui n'ont pas produit la forme orthographique attendue. Nous avons donc isolé les réponses orthographiquement correctes (ainsi que les non-réponses) pour ne traiter que les réponses erronées au sein desquelles nous avons distingué les formes phonographiquement acceptables (ou approximantes) des autres.

La figure 4 caractérise les réponses erronées des scripteurs selon le critère phonographique pour l'ensemble des mots analysés dans cette étude. Nous distinguons pour cette analyse les régions Haute-Normandie et Rhône-Alpes.

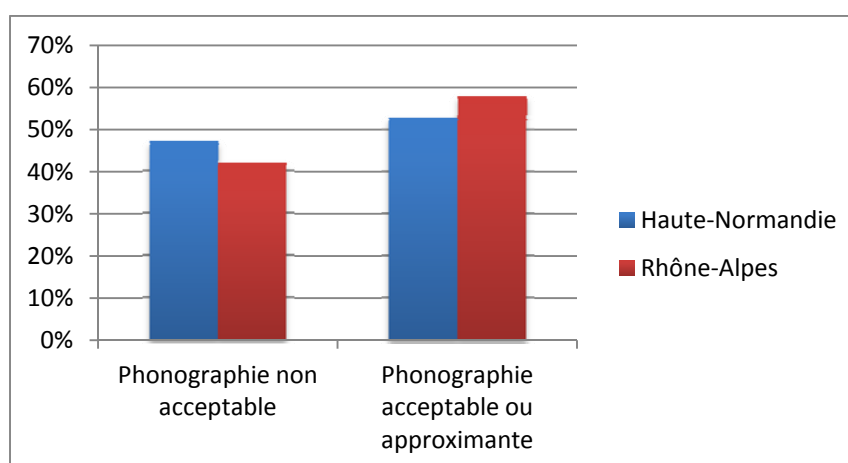


Figure 4. Distribution des réponses orthographiques erronées selon le critère phonographique, en HN et en RA

La proportion de réussite phonographique est semblable dans les deux régions, même si nous observons une réussite phonographique légèrement plus faible en Rhône-Alpes : parmi les réponses qui ne respectent pas le principe orthographique, 42 % d'entre elles sont phonographiquement acceptables dans cette région contre 47 % en HN. Cette proportion non négligeable de réponses phonographiquement acceptables, dans les deux régions, valide l'intérêt de prendre en compte le critère phonographique dans l'évaluation des compétences scripturales des personnes qui présentent des difficultés à l'écrit.

La figure 5 présente, pour chaque région et pour chaque mot, la proportion de réponses qui restituent une forme phonographique acceptable des mots parmi les réponses orthographiquement erronées du corpus. Il apparaît que la réussite des mots selon le critère phonographique présente davantage de différences entre les régions Haute-Normandie et Rhône-Alpes que la réussite orthographique. Nous avons donc fait le choix de différencier à nouveau les deux corpus et de procéder à une analyse comparative des résultats. Pour plus de lisibilité, nous avons classé les formes dans l'ordre décroissant du taux de réussite phonographique pour la région Haute-Normandie.

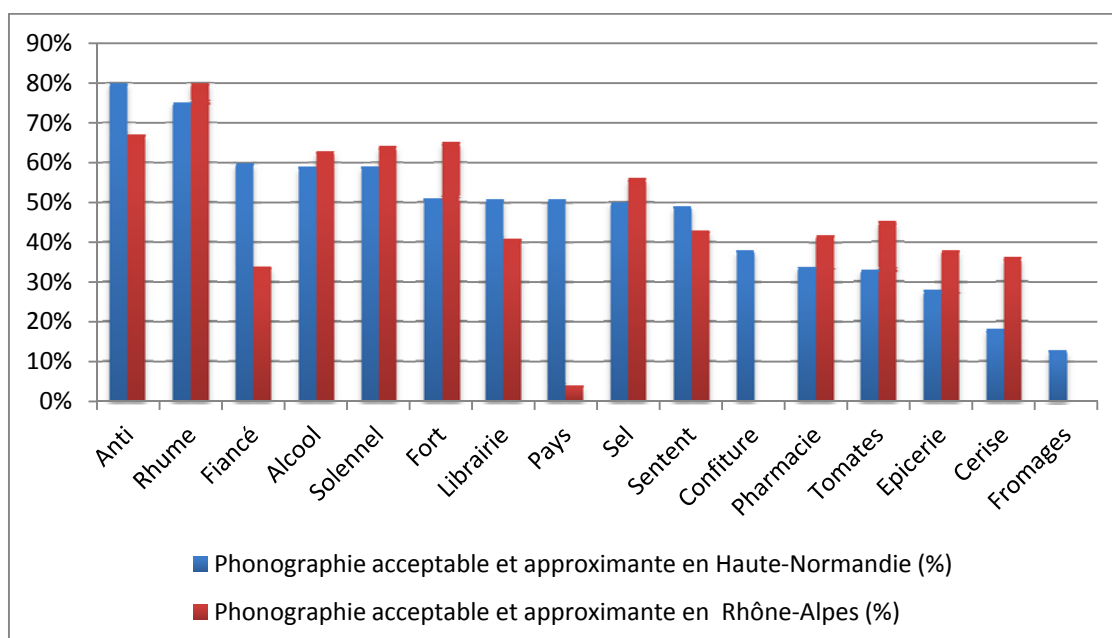


Figure 5. Proportion de réponses phonographiquement acceptables parmi les réponses orthographiquement erronées pour chaque mot, en HN et en RA

Dans les deux corpus, *anti* et *rhume* sont les items qui ont été les mieux réussis sur le plan phonographique avec plus de 70 % de réussite. En revanche, *fromages* est le mot qui a posé le plus de problèmes aux enquêtés qui n'en connaissaient pas l'orthographe (seules 13% de formes phonographiquement acceptables en HN ; aucune des formes erronées relevées n'est phonographiquement acceptable en RA). *Fiancé* pourtant réussi à 60 % en HN n'a été réussi qu'à 34 % en RA. *Pays* a également été mieux réussi en HN (51 % de réussite phonographique) qu'en RA, où seules 4 % des formes produites erronées sont phonographiquement acceptables.

La réussite des mots selon le critère phonographique présente donc davantage de différences que la réussite orthographique entre les régions, en particulier sur les mots *fiancé* et *pays*. Les résultats moyens quant à eux sont similaires : rappelons qu'en effet, moins de 50 % des réponses erronées (sur le plan orthographique) sont acceptables sur le plan phonographique dans les deux régions. Destinée à rendre compte de la capacité des scripteurs à fournir un message écrit compréhensible, puisque restituant une forme phonologique lisible du mot au-delà du respect de l'orthographe conventionnelle, cette analyse nous interpelle sur

les compétences phonographiques des scripteurs. Dans ce corpus, nous avons constaté que les scripteurs ont des difficultés à restituer de manière acceptable la forme phonologique des mots lorsqu'ils n'en connaissent pas l'orthographe, et plus particulièrement lorsqu'il s'agit de mots phonographiquement complexes. En effet, nous n'observons pas systématiquement de corrélation entre la réussite orthographique et la réussite phonographique sur les formes erronées : ce ne sont pas nécessairement les mots qui ont été les moins bien réussis sur le plan orthographique qui sont également les moins bien réussis sur le plan phonographique. Si *solennel* est le mot qui a obtenu le score orthographique le plus faible (14,5 % en moyenne), ce n'est en revanche pas le mot qui a posé le plus de difficultés lorsque les scripteurs ont essayé de le transcrire phonétiquement (61,5 % en moyenne). Parmi les réponses erronées, nous relevons davantage de graphies fidèles à la forme phonologique du mot, telles que *saulanel*, *solanel*, *sollanel*, ce qui n'est pas le cas pour *fromages* par exemple. En effet, pour ce mot particulièrement bien réussi sur le critère orthographique, nous n'observons aucune réponse acceptable sur le plan phonographique parmi les réponses erronées relevées dans le corpus rhônalpin : nous relevons par exemple les graphies *fohmos*, *fromege*, *formaj*. Il semble donc que la distribution des graphèmes dans certains mots, tels que *fromages*, ait posé davantage de difficultés à la minorité de scripteurs qui n'en connaissaient pas l'orthographe (13 % en moyenne). Une étude ciblée sur la syllabe ou le phonème/graphème, permettrait de décrire plus finement les difficultés rencontrées par les scripteurs de ce corpus.

Lorsque l'on prend en compte le critère phonographique en plus du critère orthographique (cf. figure 3) pour évaluer la réussite, nous observons que le score moyen augmente de 15 points, passant ainsi de 66 % à 81 % de réussite. En revanche, le critère de la phonographie ne fait pas varier les résultats de manière homogène sur tous les mots : si *pharmacie*, *épicerie*, *tomates*, *cerises*, *pays*, *sel* et *fromages* révèlent des résultats similaires avec un écart de réussite de moins de 10 points, certains mots présentent toutefois un écart de réussite important selon le critère considéré.

Ainsi, pour les mots *solennel* et *rhume*, les scores passent respectivement de 14 % de réussite orthographique à 61 % de réussite phonographique et de 47 % à 84 %, avec un écart moyen pour ces mots de 42 points lors de la prise en compte d'une acception large du critère phonographique. Si ces deux mots sont ceux qui présentent les scores de réussite orthographique les plus faibles ; ils sont également ceux pour lesquels la considération du critère phonographique impacte le plus sur le score de réussite¹².

La mise en évidence de ces tendances nous conduit à présenter les profils des scripteurs à partir de scores orthographiques et ortho-phonographiques, à des fins comparatives.

Profils de scripteurs à partir des scores orthographique et ortho-phonographique

Bien qu'ayant recours à une orthographe éloignée de la norme, une majorité de scripteurs possède de réelles compétences phonographiques, non négligeables, qui leur permettent de transcrire l'ensemble des mots. Pour évaluer la compétence phonographique des enquêtés et mesurer l'impact de ce nouveau critère sur la catégorisation des personnes dans l'enquête, nous proposons de reclasser les scripteurs selon leur réussite orthographique et phonographique, tout en conservant les quatre profils définis par l'INSEE.

Nous présentons dans la figure 6, le classement des scripteurs par profil orthographique établi par l'INSEE que nous comparons au profil déterminé selon le critère phonographique. Ce dernier réunit les graphies attendues (orthographiquement correctes) ainsi que les graphies acceptables ou approximantes. Ainsi, nous appellerons ce profil « ortho-phonographique ». Dans ce calcul, nous n'avons pas distingué la région d'appartenance des enquêtés.

¹² Nous présentons une analyse plus fine des erreurs relevées sur ces deux items dans la dernière partie de l'analyse (pages 17-18)

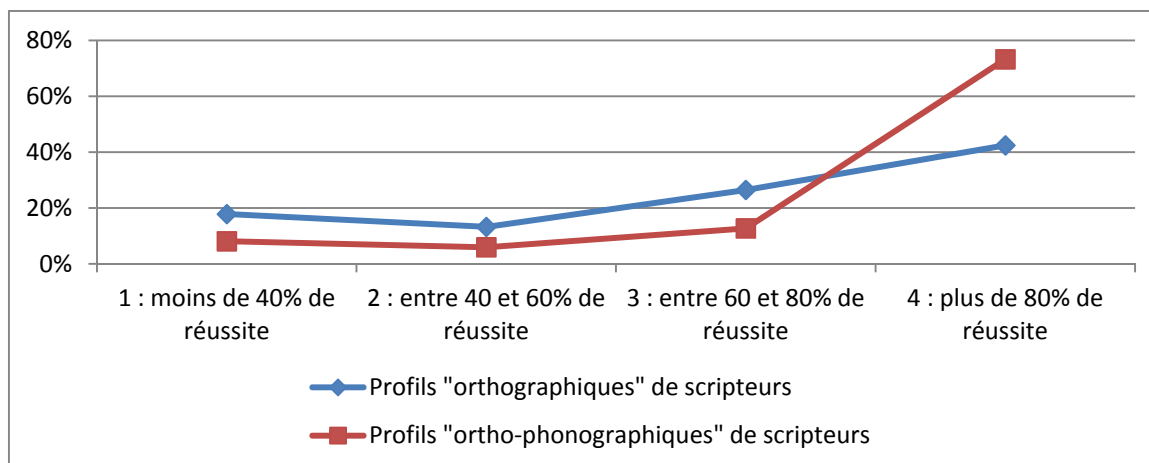


Figure 6. Proportion de scripteurs selon le profil calculé sur le critère orthographique ou ortho-phonographique

En appliquant les profils définis par l'INSEE et le critère orthographique sur les 16 mots de la dictée, 31 % des scripteurs haut-normands et rhônalpins auraient des difficultés « graves ou fortes » et se trouveraient ainsi « en situation préoccupante » face à l'écrit (avec un score de réussite inférieur à 60 %). En effet, 18 % des scripteurs ont obtenu moins de 40 % de réussite (« difficultés graves ») et 13 % ont obtenu entre 40 et 60 % de réussite (« difficultés fortes »). En revanche, 26 % des scripteurs ne présentent que des difficultés partielles (avec un taux de réussite compris entre 60 et 80 %) et 42 % ne présentent pas de difficulté particulière (puisque'ils ont réussi plus de 80 % de l'épreuve). Ainsi, 68 % des enquêtés possèderaient des compétences orthographiques leur permettant de transcrire correctement plus de 60 % des 16 mots analysés.

En revanche, lorsque l'on considère la réussite ortho-phonographique, on constate que 86 % des enquêtés semblent posséder de réelles compétences phonographiques en parvenant à restituer une forme phonologique acceptable sur plus de 60 % des mots. En effet, en ajoutant ce critère au critère orthographique, 13 % des scripteurs ne présentent plus que des difficultés partielles (entre 60 et 80 % de réussite) et 73 % ne présentent plus de difficulté en production de mots écrits (plus de 80 % de réussite).

Malgré le calcul d'un score phonographique qui valorise les compétences à l'écrit des scripteurs les plus en difficulté avec l'écrit, en moyenne 14,5 % d'entre eux (14 % en HN et 15 % en RA) présentent toujours des difficultés « graves ou fortes » dans la tâche de production écrite. Ces scripteurs se trouvent dans une situation d'autant plus préoccupante que cette analyse met en évidence leur faible compétence dans le domaine phonographique, pourtant essentiel à la production d'écrits, au-delà de la norme orthographique (Fayol, 2013). Une étude qualitative des choix graphiques, de chacun des scripteurs relevant de cette situation préoccupante, corrélée à leurs variables sociales, permettrait de déterminer la nature de leurs difficultés tout en tenant compte, en formation notamment, de leurs profils (Français Langue Etrangère (FLE) vs situation d'illettrisme¹³).

¹³ Selon la définition de l'ANLCI : « On parle d'illettrisme pour des personnes qui, après avoir été scolarisées en France, n'ont pas acquis une maîtrise suffisante de la lecture, de l'écriture, du calcul, des compétences de base, pour être autonomes dans les situations simples de la vie courante. Il s'agit pour elles de réapprendre, de renouer avec la culture de l'écrit. ». (CNFPTLV, 2013 : 9 dans ANLCI 2003)

Analyse de la variation sur les mots de la dictée IVQ chez les scripteurs haut-normands et rhônalpins

L'analyse de ce corpus met en évidence une variation particulièrement importante des réponses, qu'il s'agisse de formes que nous avons considérées comme phonographiquement acceptables ou de formes inacceptables sur le plan phonographique. Pour rendre compte de cette variation, nous avons calculé, dans un premier temps, le taux de variation pour chacun des 16 mots analysés dans la tâche de production écrite en réunissant les deux régions. Nous avons fait le rapport du nombre de variantes différentes produites (et non de leur occurrence) sur le nombre total de réponses. Pour ce calcul, nous avons considéré comme variation, toutes les formes proposées par les enquêtés qui diffèrent de la norme orthographique attendue. Nous avons procédé, dans un second temps, à une analyse qualitative des réponses erronées en présentant les formes récurrentes pour *anti*, *rhume* et *solennel*, mots pour lesquels nous avons observé moins de 60 % de réussite orthographique dans les deux régions.

La figure 7 présente, pour les deux régions réunies, le taux de variation de chaque mot en appliquant le critère de conformité phonographique pour distinguer les formes phonographiquement acceptables et inacceptables. Nous présentons les items selon leur ordre d'apparition dans la dictée.

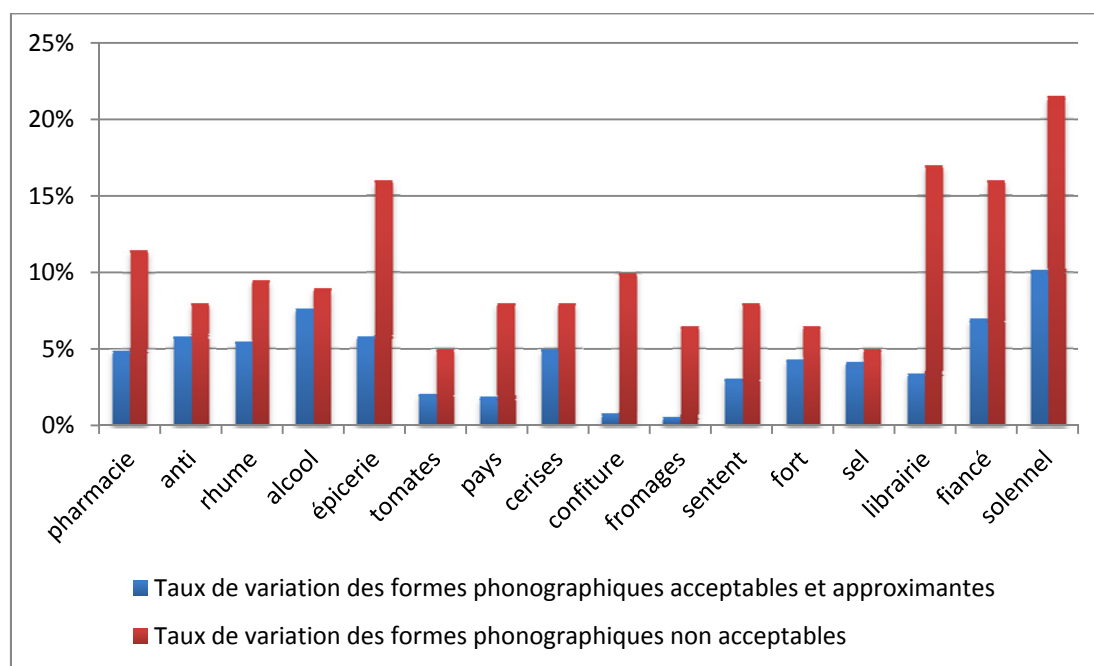


Figure 7. Taux de variation sur les réponses orthographiques erronées du corpus, pour chaque mot

Les mots qui présentent le moins de variation, par exemple *sel* et *tomates* (7 % en moyenne, dans les deux régions), font partie de ceux dont la forme orthographique a été la mieux restituée. À l'inverse, ce sont les mots qui ont posé le plus de difficulté sur le plan orthographique, qui présentent le plus de variation dans les corpus. Ce ne sont donc pas toujours les mêmes phonèmes/graphèmes qui ont posé des difficultés aux scripteurs et leurs difficultés peuvent être particulièrement hétérogènes selon les mots. Lorsque la variation est observée sur les formes phonographiquement inacceptables qui, en s'écartant de la forme sonore, rendent possible un nombre infini de graphies, le nombre de variantes augmente. Ainsi, *solennel*, qui présente un taux de variation de 32 % en moyenne, *fiancé* (24 %), *épicerie* (22 %) et *librairie* (21 %), sont à la fois ceux qui présentent le plus de variation dans les corpus et ceux pour qui le taux de variation sur les réponses phonographiquement inacceptables est supérieur à celui des réponses phonographiquement acceptables. Le taux de variation de réponses phonographiquement inacceptables est particulièrement élevé sur le mot

solennel : plus de la moitié des réponses erronées ne restituent pas la forme phonologique du mot (22 % contre 10 % de formes phonographiquement acceptables).

Cette analyse du taux de variation nous permet de distinguer les mots qui présentent des réponses erronées récurrentes, des mots pour lesquels nous avons observé un phénomène de dispersion des réponses proposées. Nous présentons, dans l'analyse suivante, les formes récurrentes observées parmi les réponses erronées pour les mots *anti*, *rhume* et *solennel*.

Analyse comparative de la récurrence des erreurs pour les mots les moins réussis sur le plan orthographique : *anti*, *rhume* et *solennel*

Nous présentons la variation observée sur les trois mots les moins bien réussis sur le critère orthographique, pour lesquels nous avons observé moins de 60 % de réussite orthographique dans les deux régions. Pour cette analyse, nous décrivons les graphies produites de manière récurrente dans chaque corpus. Nous avons considéré que les formes étaient récurrentes lorsqu'elles avaient été produites par au moins 6 scripteurs en Haute-Normandie et 3 scripteurs en Rhône-Alpes (ce qui représente au moins 3 % du corpus dans chaque région).

Formes attendues		Formes phonographiquement acceptables		Formes phonographiquement inacceptables	
HN	RA	HN	RA	HN	RA
<i>anti</i> (56 %)	<i>anti</i> (56 %)	<i>enti</i> (20 %) <i>entie</i> (4 %)	<i>enti</i> (18 %)	<i>Aucune réponse récurrente</i>	<i>inti</i> (3 %)
<i>rhume</i> (45 %)	<i>rhume</i> (50 %)	<i>rhum</i> (15 %) <i>rume</i> (11 %) <i>rum</i> (4 %) <i>ruhme</i> (3 %)	<i>rume</i> (15 %) <i>rhum</i> (10 %) <i>rum</i> (6 %) <i>ruhme</i> (4 %)	<i>Aucune réponse récurrente</i>	<i>Aucune réponse récurrente</i>
<i>solennel</i> (12 %)	<i>solennel</i> (17 %)	<i>solanel</i> (19 %) <i>solanelle</i> (10 %) <i>solannel</i> (6 %)	<i>solanel</i> (25 %) <i>solanelle</i> (10 %) <i>solannel</i> (6 %)	<i>solonel</i> (4 %) <i>solenel</i> (3 %) <i>solonnel</i> (3 %)	<i>solenel</i> (5 %) <i>solonel</i> (3 %)

Tableau 4. Présentation des réponses récurrentes pour les mots *anti*, *rhume* et *solennel*, en HN et en RA

Considéré comme « régulier » par les concepteurs de la dictée, *anti* fait pourtant partie des mots qui ont posé le plus de difficultés aux scripteurs sur le plan orthographique. Lorsque nous avons analysé les réponses erronées, nous avons aussi observé davantage de graphies phonographiquement inacceptables que de réponses acceptables. Néanmoins les formes récurrentes dans les deux régions sont phonographiquement acceptables : 20 % des scripteurs haut-normands et 18 % des rhônalpins ont choisi la graphie *enti* ou *entie* (4 % de haut-normands ont ajouté un « e » muet final). Les erreurs sur les formes récurrentes ont donc davantage porté sur le graphème « an » puisqu'il faut faire un choix entre les digrammes *en* et *an*. Ce premier phonème a également été confondu avec le phonème /ɛ̃/ : 3 % des enquêtés en Rhône-Alpes ont produit la graphie *inti* commettant ainsi une erreur de correspondance phonographique.

Les mots *rhume* et *solennel* sont ceux qui présentent les scores de réussite orthographique les plus faibles, mais nous avons également constaté qu'il s'agissait des mots pour lesquels la considération du critère phonographique impactait le plus sur le score de réussite. Lorsque l'on observe le taux de variation de *rhume*, on constate que le corpus comporte deux fois plus de formes inacceptables sur le plan phonographique (10 % contre 5 % de réponses phonographiquement acceptables). En revanche, si l'on regarde le nombre d'occurrences cette tendance s'inverse : les 4 réponses récurrentes (33 % des réponses en HN et 35 % en RA) sont phonographiquement acceptables.

La prise en compte de ce critère de récurrence pour ces trois mots permet de mieux connaître les choix des scripteurs parmi les possibilités de transcription offertes par le système phonographique du français. Dans les deux régions, nous ne constatons pas de réponse récurrente dans les formes inacceptables mais une dispersion de réponses dont les erreurs portent principalement sur le phonème /y/ : nous rencontrons par exemple dans les corpus les graphies *reme*, *rom*, *rime*, etc. Pour les réponses phonographiquement acceptables, les erreurs récurrentes concernent les graphèmes « h » et « e », généralement muets, que les scripteurs ont omis ou déplacés : omission du « e » final muet dans *rhum* et omission et/ou déplacement du « h » dans *rum* ou *ruhme*. Les réponses orthographiquement erronées, mais que nous avons acceptées sur le plan phonographique, semblent combiner voie d'adressage et voie d'assemblage. Les formes *rhum* et *ruhme* relèvent en effet de procédures d'adressage : *rhum* renvoie à un autre item désignant une boisson alcoolisée et *ruhme* présente un déplacement des graphèmes « u » et « h » dont la position est inversée. La forme *rum*, enfin, pourrait être considérée comme une transcription d'assemblage à valeur phonétique, comme *rume*.

Pour *solennel*, trois réponses apparaissent de manière récurrente dans les deux régions : *solanel*, *solanelle* et *solannel* (38 % en moyenne dans les deux corpus). Ce mot avait été particulièrement mal réussi sur le plan orthographique mais, à la différence d'*anti* et *rhume*, le taux le plus élevé de réussite a été observé sur des formes phonographiquement acceptables (59 % en HN et 52 % en RA). Les deux formes les plus récurrentes ne restituant pas la forme /solanel/ sont également les mêmes dans les deux régions : *solanel* et *solonel*. Les autres réponses sont très variées et n'ont été produites que par un petit nombre de scripteurs dans les deux régions. La majorité des erreurs portent donc sur le choix du graphème pour transcrire le phonème /a/. Nombre d'enquêtés ont choisi l'archigraphème « a » en procédant à une transcription phonologique. Certains enquêtés ont également ajouté une double consonne (*sollennel*), laissant entrevoir une rétention partielle de la forme orthographique. Une confusion apparaît également sur le genre de cet adjectif. Dans l'énoncé « *le fiancé solennel* », l'adjectif est masculin mais la marque du féminin a été produite à plusieurs reprises, parfois même après le déterminant masculin *le* : « *la/le fiancé(e) solennelle* ».

Cette analyse des formes récurrentes produites sur les trois mots qui ont posé le plus de difficultés aux scripteurs sur le plan orthographique met en évidence que, parmi les réponses erronées, les formes le plus souvent proposées sont phonographiquement acceptables. L'analyse du nombre d'occurrences des formes orthographiquement incorrectes pour deux d'entre eux (*anti* et *rhume*) montrait quant à elle que la majorité des scripteurs n'était pas parvenue à restituer la forme phonologique de ces deux mots. Cela met en évidence la dispersion des réponses recueillies dans notre corpus pour les réponses phonographiquement inacceptables, laissant entrevoir une disparité dans les stratégies et les procédures d'écriture ainsi que dans les connaissances orthographiques et phonographiques des enquêtés présentant le plus de difficultés. Par ailleurs, les écarts produits se situent, de manière récurrente, sur les mêmes phonèmes et/ou graphèmes : sur les lettres muettes internes ou finales (« e » final muet ou « h » dans *rhume*) ainsi que sur l'archigraphème /E/. Sa graphie, particulièrement irrégulière en français, nécessite en effet de faire une sélection parmi les nombreux graphèmes qui lui correspondent (e, é, è, ê, ë, ai, ei, etc.), ce qui suppose une bonne maîtrise des procédures d'adressage.

Conclusion

L'analyse présentée dans cet article ne montre pas de différences significatives dans les productions des scripteurs haut-normands et rhônalpins présentant de potentielles difficultés à l'écrit. Alors même qu'une proportion plus importante d'enquêtés avaient été dirigés vers le

module ANLCI en Haute-Normandie (16 % contre 13 % en RA), les résultats au terme de l'enquête et les tendances qui se dessinent au niveau de la réussite orthographique et phonographique sont semblables. Cette analyse sur 16 mots de la dictée met en évidence que certains scripteurs présentent de réelles difficultés concernant les compétences orthographiques et qu'il existe de gros écarts dans la nature même des difficultés entre les scripteurs.

Par ailleurs, l'analyse des réponses produites selon le critère phonographique montre que le degré de difficulté semble davantage se situer au niveau du phonème, dont certains semblent être moins bien discriminés. Une analyse portant sur le phonème ou sur la syllabe semblerait donc nécessaire pour parfaire celle des correspondances phonographiques. La manière de dicter les mots, inhérente à chacun des enquêteurs, est aussi un facteur qui mérite d'être souligné dans la variation observée dans ce corpus, bien qu'il ne puisse être vérifié, faute d'enregistrement réalisé lors de la passation de la dictée.

Parmi les personnes qui présentent le plus de difficultés avec l'écrit, la plupart semblent malgré tout avoir des compétences phonographiques relativement efficaces. L'analyse en cours des variations recueillies dans ce corpus nous permet d'assurer que certains scripteurs, bien qu'éloignés de la norme, détiennent des connaissances orthographiques certaines même si ces dernières peuvent être mal réinvesties (par exemple, la place du « h » dans l'item *rhume* transcrit *ruhme*). Ce constat apparaît dans les graphies non-conformes à la norme orthographique - qu'elles correspondent à la forme orale ou non - et sont souvent le miroir des irrégularités et des incohérences du système graphique du français et/ou de constantes chez les scripteurs. Au-delà de l'analyse phonographique, ces graphies montrent aussi que les personnes qui présentent des difficultés à l'écrit utilisent leurs capacités mémorielles, en recourant à la voie d'adressage pour orthographier les mots qu'ils ont déjà vus ou lus. Tous détiennent donc des connaissances et sont capables de produire de l'écrit, bien que dans ce domaine les difficultés de certains soient « préoccupantes » selon l'INSEE.

L'évaluation des compétences à l'écrit d'un public d'adultes en difficulté avec la langue écrite nécessite donc une analyse particulière. Le choix d'une liste de courses dictée semble pertinent dans le cadre d'une enquête nationale (pour des raisons de comparabilité des résultats notamment) mais le traitement des données demande une méthodologie particulière qui reste à explorer davantage. Les scores établis selon le critère orthographique mesurent l'écart à la norme orthographique pour un scripteur expert et nous l'avons vu, l'utilisation de ce seul critère pour analyser les compétences à l'écrit des scripteurs les plus en difficulté se révèle insuffisante pour définir la nature précise de leurs difficultés et leur degré d'autonomie dans l'usage de la langue écrite au quotidien. Les analyses présentées dans cet article seront étendues à l'ensemble du corpus national avec l'objectif, à terme, de décrire et d'analyser plus spécifiquement les productions des personnes dites en situation d'illettrisme, en les mettant en lien avec les variables sociales dont nous disposons. Ce travail devrait permettre de mieux comprendre leurs difficultés et leurs besoins pour proposer, en formation d'adultes, une remédiation adaptée.

Bibliographie

- ANLCI, 2003, « Lutter ensemble contre l'illettrisme. Cadre national de référence », [en ligne], consultée le 30 novembre 2014, URL : <<http://www.anlci.gouv.de/content/download/907/.../cadre de reference.pdf>>.
- BASTYNS C., 2012, « Les enquêtes de l'OCDE sur la littératie et les compétences des adultes font-elles avancer la cause de l'alpha ? », *Journal de l'alpha*, n°185, pp. 26-39.

- BESSE J. M., LUIS M.H., BOUCHUT A.L., MARTINEZ F., 2009, « La mesure des compétences en traitement de l'écrit chez les adultes en grande difficulté », *Economie et statistique*, n°424-425, pp. 31-48.
- BRENDLER J. F., 2013, « Les personnes en difficulté à l'écrit : des profils régionaux variés », *INSEE Première*, n°1475, novembre 2013.
- CATACH N., 1984, *La phonétisation automatique du français : les ambiguïtés de la langue écrite*, Paris : Editions du CNRS.
- BELMANS, C., BRENDLER, J., FOLLEA, G., MERCIER, J-P., SUEUR, C., THIBAUDEAU, P., 2013, *Les hauts-normands face aux savoirs de base : lire, écrire, compter*. Dossier d'Aval Haute-Normandie, septembre 2013.
- FAYOL M., 2013, *L'apprentissage de l'écrit*, Paris : PUF.
- GUERIN-PACE F., 2009, « Illettrismes et parcours individuels », *Economie et statistique*, n°424-425, pp. 49-62.
- LAHIRE B., 1999, *L'invention de l'« illettrisme »*. *Rhétorique publique, éthique et stigmates*, Paris, La Découverte.
- LUCCI V., MILLET A., 1994, *L'orthographe de tous les jours. Enquête sur les pratiques orthographiques des Français*, Paris, Champion, coll. « Politique linguistique ».

GLOTTOPOL

Revue de sociolinguistique en ligne

Comité de rédaction : Michaël Abecassis, Salih Akin, Sophie Babault, Claude Caitucoli, Véronique Castellotti, Régine Delamotte-Legrand, Robert Fournier, Stéphanie Galligani, Emmanuelle Huver, Normand Labrie, Foued Laroussi, Benoit Leblanc, Fabienne Leconte, Gudrun Ledegen, Danièle Moore, Clara Mortamet, Alioune Ndao, Isabelle Pierozak, Gisèle Prignitz, Georges-Elia Sarfati.

Conseiller scientifique : Jean-Baptiste Marcellesi.

Rédactrice en chef : Clara Mortamet.

Comité scientifique : Claudine Bavoux, Michel Beniamino, Jacqueline Billiez, Philippe Blanchet, Pierre Bouchard, Ahmed Boukous, Pierre Dumont, Jean-Michel Eloy, Françoise Gadet, Marie-Christine Hazaël-Massieux, Monica Heller, Caroline Juilliard, Jean-Marie Klinkenberg, Jean Le Du, Marinette Matthey, Jacques Maurais, Marie-Louise Moreau, Robert Nicolaï, Lambert Félix Prudent, Ambroise Queffelec, Didier de Robillard, Paul Siblot, Claude Truchot, Daniel Véronique.

Comité de lecture pour ce numéro :

Margaret Bento, Marie-Madeleine Bertucci, Jacqueline Billiez, Marie-Line Bosse, Véronique Castellotti, Anne Dister, Martine Dreyfus, Valentin Feussi, Claudine Garcia-Debanc, Emmanuelle Huver, Jean-Pierre Jaffré, Fabrice Jecic, Jean Le Du, Normand Labrie, Gudrun Ledegen, Fabienne Leconte, Marinette Matthey, Bruno Maurer, Agnès Millet, Marie-Louise Moreau, Marie-Françoise Morin, Jean-Christophe Pellat, Corinne Totereau, Fanny Rink, Pierre Sève.

Laboratoire Dysola – Université de Rouen
<http://glottopol.univ-rouen.fr>

ISSN : 1769-7425