

# « tu me *di* que je ne *connai* pas le cri ture é bien ç i chaque foi que tu *écri* je le *connai* » – Analyse des graphies des formes verbales dans la correspondance des Poilus du Corpus14

Christian Surcouf

École de Français Langue Étrangère, Facultés des Lettres, Université de Lausanne (Suisse)

[christian.surcouf@unil.ch](mailto:christian.surcouf@unil.ch)

Résumé. Dans cet article, nous analysons d'un point de vue quantitatif les graphies des 886 formes verbales apparaissant dans les dix premières lettres de quatre Poilus, provenant du Corpus 14<sup>1</sup>. Après avoir répertorié les différents types d'écarts par rapport à la norme, observés dans cette correspondance produite par des scripteurs peu accoutumés à l'écrit, nous montrerons que, en dépit d'un nombre important de graphies non-normées, les formes verbales permettent une lecture conforme à la forme normée ciblée dans la majorité des cas.

**Abstract. Verb orthography in 40 letters from World War I French soldiers.** In this article, I analyze the spellings of the 886 verb forms appearing in the first ten letters of four 'Poilus' from Corpus 14. After identifying the different types of deviations from the standard, observed in these letters written by individual not accustomed to writing, I will show that, in most cases, despite a significant number of unconventional spellings, verb forms can be read in accordance with the standardized pronunciation originally targeted.

## 1. Introduction

Dans son analyse de la graphie des lettres de Poilus du Corpus 14 (voir note 1), Pellat (2015; 2017) évoque brièvement l'orthographe grammaticale et mentionne certaines des difficultés rencontrées par les scripteurs dans la conjugaison des verbes :

les finales verbales sont souvent absentes (*nous somme, nous avon*). Les formes en E sont régulièrement indifférenciées : comme, dans les verbes du type *chanter*, quatre formes orales sont identiques (l'infinif en *-er*, le participe passé en *-é*, la 2<sup>e</sup> personne du pluriel du présent de l'indicatif et de l'impératif en *-ez* et même l'imparfait en *-ai-*), on relève *vous tomber, je vous direz, vous risqué, sa a etai, tu navez*. (Pellat 2015: 72-73; 2017: 241)

Il signale également d'autres écarts dans la graphie des finales verbales :

Les accords du verbe avec le sujet ne sont pas assurés (*je suit, sai, repon, voit ; je vous dirait, tu souffre, tu dit, tu veut, tu m'envera, tu aura ; nous iron ; je dirais* est généralement un futur,

non un conditionnel), pas plus que ceux de l'attribut (*nous sommes arrivé*). (Pellat 2015: 73; 2017: 241)

L'auteur évoque enfin les problèmes rencontrés avec l'accord du participe passé :

Autant dire que l'accord du participe passé employé avec *avoir* est exceptionnel (*celles que je t'ai dega envoyées*, Ernest) ; on rencontre, très rarement, l'accord avec le complément d'objet direct, en particulier chez Pierre Fabre qui accorde le participe quelle que soit la place du COD, avant ou après (*Cette séparation je lui éloignée ; j'ai reçue hier ton aimable lettre*). (Pellat 2015: 73; 2017: 241)

Dans notre article, dans la lignée des remarques de Pellat, et en nous basant sur les dix premières lettres de quatre Poilus (cf. tableau 1), issues de ce même Corpus 14, nous analyserons en détail, d'un point de vue quantitatif<sup>2</sup>, les graphies des verbes apparaissant dans ces quarante lettres.

## 2. Les scripteurs

Mentionnons en premier lieu l'origine sociale de nos quatre scripteurs, et les circonstances singulières de leurs relations épistolaires :

Ils sont cultivateur, meunier, vigneron, charron. À l'école de la troisième République, ils ont appris à lire et à écrire ; ils n'ont pas poussé jusqu'au Certificat d'Études. Leur quotidien ne les confrontait que rarement à la pratique de l'écrit ; avec la guerre, l'écriture devient le lien, vital, avec la famille laissée au pays. (Steuckardt 2015: 9)

Comme l'indique le tableau 1, trois de ces Poilus sont nés dans l'Hérault et, comme certains usages le laissent supposer (voir Géa 2015), ils pratiquaient probablement l'occitan dans leur quotidien<sup>3</sup>. Nos quatre scripteurs avaient été scolarisés à la fin du XIX<sup>e</sup> siècle, époque à laquelle avait surgi une réflexion sur l'orthographe et son enseignement (voir Chervel 1991), même si en définitive, vers 1880 :

En dépit de l'offensive déterminée que Ferdinand Buisson mène contre les excès de l'enseignement de l'orthographe, la dictée s'impose alors comme l'exercice majeur de l'école primaire, et elle le restera. La proposition faite par Irénée Carré en 1890 de supprimer la dictée au certificat d'études suscite dans l'instruction primaire une levée de boucliers sans précédent. (Chervel 2006: 325-326)

C'est donc à priori dans cet environnement scolaire que nos quatre Poilus ont été initiés à la pratique de l'orthographe (voir également Pellat 2015: 68). Rappelons avec Géa (2015: 53-54) que, « si la plupart avaient tiré bénéfice de l'instruction gratuite et obligatoire, rares étaient ceux pour qui l'acte d'écrire était une pratique familière ». Par ailleurs, leur apprentissage scolaire remontait, selon l'âge de leur recrutement dans l'armée, à plus d'une ou deux décennies si l'on tient compte du fait qu'ils « ont quitté [l'école] vers 11-12 ans pour aller travailler à la vigne ou dans les champs » (Pellat 2015: 67).

**Tableau 1.** Informations sur les scripteurs et leurs lettres

		Naissance		Âge	M(ots)	V(erbes)	M/V	Lettres du ... au ...
		Lieu	Année					
Pierre Fabre	(PF <sup>4</sup> )	Hérault	1889	25	1413	231	6,14	15/08 – 19/10/1914
Laurent Pouchet	(LP)		1884	30	2112 <sup>5</sup>	308	6,86	09/09 – 03/11/1914
Ernest Viste	(EV)		1879	35	1205	208	5,85	09/09/1914 – 01/01/1915
Alfred Foray	(AF)	Ain	1884	30	814	139	5,86	03/08 – 06/10/1914
					5544	886	6,28	

Comme le révèle immédiatement la lecture de leurs lettres, nos quatre scripteurs ont une maîtrise imparfaite de l'orthographe normée. Cependant, lorsqu'il écrit : « tu me di que je ne connai pas lecri ture ébien çï chaque foi que tu écri je le connai » (1914.10.2), LP invite son destinataire à considérer, à juste titre, le succès de l'objectif poursuivi par ses missives,

en l'occurrence, celui de communiquer avec ses proches. Au travers de cette remarque, LP semble par ailleurs reconnaître implicitement sa maîtrise incertaine des normes de l'écrit. Pour ces Poilus, l'écriture ne résultait pas d'un choix, elle s'imposait à eux. Elle leur permettait de partager leur quotidien, et surtout de rappeler qu'ils étaient en vie, même s'ils continuaient à « espérer qu'un jours viendra, ou lon pourra finir toute çes comédi décriture » (LP 1915.03.20).

### 3. Les objectifs de l'analyse

Dans cet article, notre attention se portera sur les formes verbales dans leur ensemble, et en partie sur les finales verbales (4.3), lesquelles, représentent en français des points de vulnérabilité en raison de la complexité du fonctionnement des morphogrammes grammaticaux<sup>6</sup> et de l'absence de correspondance claire avec l'oral. À cette fin, nous avons procédé à l'étiquetage manuel des 886 formes verbales apparaissant parmi les 5544 mots (soit un ratio moyen d'un verbe tous les 6,28 mots) de ces quarante lettres rédigées entre le 28 aout 1914 et le 26 aout 1915 (voir tableau 1). Bien qu'une analyse des contextes morphosyntaxiques, lexicaux et orthographiques susceptibles d'engendrer telle ou telle graphie non-normée s'avèrerait intéressante, dans le cadre restreint de cet article, nous nous contenterons d'une approche quantitative. Tous les verbes ont été comptabilisés, qu'ils soient conjugués ou non, les infinitifs révélant eux aussi des variantes non-normées dignes d'intérêt :

- [1] tu les recevras au moin tous a la fois et elle sont toute numeroter pour les **lires** (AF-1914.08.19)
- [2] il faut **espérai** que cela passera un peu apres lautre (LP-1914.10.09)

Comme l'illustrent les exemples ci-dessous, les écarts orthographiques par rapport à la norme peuvent se rencontrer au niveau de la segmentation (en [3]), ou toucher plusieurs parties du verbe (en [4] et [5]), et plus particulièrement les finales ([6] à [10]), qui feront l'objet d'une analyse séparée :

- [3] Alors je termine en Vous **en brassant** tous (PF-1914.08.15)
- [4] votre fils et Epoux qui vous **embrasse** a tous (EV-1914.11.29)
- [5] Rien plus a vous dire pour le moment que de vous **embraser** bien faut (LP 1914.10.01)
- [6] je ne me **fait** pas de bile (AF-1914.08.07)
- [7] vous nous **croyer** dejas mort (LP-1914.10.23)
- [8] si tu **voyez** ses mère tu pourrais le leur dire (PF-1914.08.15)
- [9] les francais **son** vainqueur partout (AF-1914.08.19)
- [10] je **repon** aussi a vos deux lettres (EV-1914.10.20)

La lecture de ces graphies non-normées donne lieu à quatre configurations théoriques possibles, en fonction de la conformité de la graphie ou de la lecture à la cible normée<sup>7</sup> :

**Tableau 2.** Configurations théoriques des rapports de la graphie à la lecture par rapport à la cible

	<b>Graphie</b>	<b>Lecture</b>
(1)	conforme	conforme
(2)	non-conforme	conforme
(3)	non-conforme	non-conforme
(4)	conforme	non-conforme

Évacuons d'emblée la configuration (4), qui, bien que théoriquement possible, représenterait une erreur de la part du lecteur dans le décodage d'une graphie pourtant correcte. Envisageons donc les trois autres cas de figure.

(1) correspond à l'exemple [11], où la graphie, conforme à la norme, débouche naturellement sur la lecture cible /fɛʁa/ :

- [11] elle me **fera** voir son savoir faire (EV-1914.10.20)

Dans la configuration (2), c'est en revanche une graphie non-normée qui autorise la lecture ciblée. Ainsi en [12], *est envoyez* se lit comme la forme normée attendue *ai envoyé*, c'est-à-dire /EãvwajE/<sup>8</sup> :

[12] je vous *est envoyez* (LP-1914.09.09)

L'exemple [13] correspond également à la configuration (2) mais ajoute un problème de segmentation, que nous aborderons plus loin. La lecture /maãvwajE/ est cependant identique à celle qu'aurait fourni la graphie normée *m'as envoyés* :

[13] je me régale bien des fromages de chèvre que tu *ma envoyer*

Enfin, l'exemple [14] représente la configuration (3), où, si l'on suit les principes de décodage graphème<sup>9</sup>-phonème de l'orthographe française, la graphie non-normée *recu* impose la lecture /ʁəky/ (comme dans *vécu* /vekyl/), en décalage avec la cible /ʁəsy/.

[14] j'ai *recu* des nouvelles (LP-1914.09.28)

En définitive, sur les 886 graphies de verbes de notre corpus de scripteurs occasionnels<sup>10</sup> :

- 1) Quel est le pourcentage de graphies correctes ?
- 2) Quels sont les types de graphies non-normées ?
- 3) Comment sont orthographiées les finales de ces formes verbales ?
- 4) Quel est le pourcentage de formes qui, quelle que soit leur orthographe – normée (ex. [11]) ou non (ex. [12] et [13]) –, autorisent une lecture conforme à la cible ?

## 4. Les résultats

### 4.1. Le nombre de verbes orthographiés selon la norme

Les 886 occurrences de notre corpus proviennent de 137 verbes différents, dont les cinq plus fréquents sont *être* (132), *avoir* (66), *recevoir* (56), *faire* (54) et *dire*. Ils se répartissent de la manière suivante pour chaque scripteur :

**Tableau 3.** Le nombre de types de verbes, dont les cinq plus fréquents

	① Occurrences	Types	Les cinq verbes les plus fréquents
(PF)	231	72	être <sup>11</sup> (34), avoir (20), dire (13), faire (12), aller (10)
(LP)	308	60	être (44), recevoir (32), avoir (28), dire (24), faire (16)
(EV)	208	63	être (27), faire (19), avoir (13), recevoir (10), savoir (7)
(AF)	140	51	être (27), aller (11), recevoir (9), faire (7), savoir (6)

Ces fréquences recourent en partie celles du *français fondamental* (soit dans l'ordre : *être*, *avoir*, *faire*, *dire*, *aller*) (Gougenheim, Michéa, Rivenc & Sauvageot 1964: 69s) à l'exception de *recevoir*, qui pour des raisons pragmatiques évidentes, s'utilise couramment dans la correspondance en dépit de sa 391<sup>e</sup> position dans le lexique du *français fondamental*.

Sur ces 886 formes verbales, environ une forme sur deux (52%) est orthographiée conformément à la norme ciblée. Toutefois, comme le montre le tableau 4 des disparités importantes existent entre les scripteurs :

**Tableau 4.** Pourcentages des graphies normées ❶ et non-normées ❷ par scripteur

	❶	❶ Normées		❷ Non-normées	
		N	%	N	%
(PF)	231	162	<b>70,1</b>	69	<b>29,9</b>
(LP)	308	112	<b>36,4</b>	196	<b>63,6</b>
(EV)	208	108	<b>51,9</b>	100	<b>48,1</b>
(AF)	139	77	<b>55,4</b>	62	<b>44,6</b>
Total	886	459	<b>51,8</b>	427	<b>48,2</b>

Le scripteur le plus proche de la norme (PF) produit un peu plus de deux formes sur trois correctement. C’est aussi le plus jeune (25 ans), et par conséquent le moins éloigné de ses souvenirs de scolarité. Les proportions sont inversées pour LP qui n’en produit qu’une sur trois.

Ces graphies non-normées ne sont pas toutes de même nature. Elles peuvent en effet avoir pour origine :

- a) une orthographe erronée :  
 [15] nous ne *savon* pas ou nous *irons* (EV-1914.12.15)
- b) un problème de segmentation :  
 [16] lorsqu'elle *men verra* de l'argent (LP-1914.10.28)
- c) une combinaison des deux :  
 [17] elle *mon fait* grand plaisir de vous savoir en bonne santé (PF-1914.09.17)

Le tableau 5 donne un premier aperçu de la répartition quantitative de ces trois types :

**Tableau 5.** Origine des graphies non-normées ❷ par scripteur

	❶	❷ Non-normées	a) Orthographe		b) Segmentation		c) Combinaison		
			N	%	N	%	N	%	
(PF)	231	69	⇨	42	<b>60,9</b>	19	<b>27,5</b>	8	<b>11,6</b>
(LP)	308	196	⇨	123	<b>62,7</b>	19	<b>9,7</b>	54	<b>27,6</b>
(EV)	208	100	⇨	65	<b>65</b>	19	<b>19</b>	16	<b>16</b>
(AF)	139	62	⇨	59	<b>95,2</b>	0	<b>0</b>	3	<b>4,8</b>
Total	886	427		289	<b>67,7</b>	57	<b>13,3</b>	81	<b>19</b>

Bien que l’orthographe non-normée constitue la zone de vulnérabilité la plus grande chez tous les scripteurs, parmi ces 288 orthographes erronées<sup>12</sup>, 123 (42,7%) renvoient à des formes existantes de l’orthographe du français. Ainsi en [18], quoique non-conforme à la forme ciblée (i.e. *recevrai*), *recevrez* est une forme effectivement attestée dans l’orthographe du français, alors qu’en [19] tel n’est pas le cas pour *savier* (c’est-à-dire pour 57,3% des occurrences) :

- [18] je la *recevrez* avec plaisir (LP-1914.09.28)
- [19] que vous ne *savier* pas encor que nous étions ici a Aix (AF-1914.08.19)

En définitive, même lorsque les scripteurs optent pour une orthographe non-conforme à la cible normée, presque une fois sur deux leur choix se porte sur des formes déjà existantes dans l’orthographe du français. Mais à quels types d’erreurs avons-nous affaire ?

## 4.2. Les diacritiques

### 4.2.1. La cédille

On ne compte que huit erreurs de cédille sur l’ensemble des formes verbales. Alors que l’ajout superflu de la cédille en [20] est sans conséquence sur la lecture de /ʁəmɛksi/, son absence en [21] entrainera selon les normes de décodage la lecture erronée /ʁøky/ :

[20] je vous *remerçi* beaucoup (LP-1914.10.28)

[21] J'ai *recu* ta carte du 22 (EV-1914.11.29)

Si les erreurs de cédille s'avèrent rares, les accents, eux, soulèvent davantage de difficultés.

#### 4.2.2. Les accents

En dépit des grandes disparités entre les quatre scripteurs, une orthographe non-normée sur quatre (i.e. **2** a/c du tableau 5) provient de l'usage inadéquat des accents, et concerne plus particulièrement le *e* (86 cas sur 90, soit 95,5%) :

**Tableau 6.** La répartition par scripteur des erreurs d'accent

	<b>0</b>	<b>2</b> a/c) Orthographe Non-normées	Accents	
			N	%
(PF)	231	50	⇒ 21	<b>47,6</b>
(LP)	308	177	⇒ 19	<b>10,7</b>
(EV)	208	81	⇒ 42	<b>51,8</b>
(AF)	139	62	⇒ 8	<b>12,9</b>
Total	886	370	<b>90</b>	<b>24,3</b>

L'usage des accents entraîne trois types d'erreur (voir également Pellat 2015: 71; 2017: 240) :

i) l'absence :

- sur le *i* et le *u* (deux occurrences)<sup>13</sup>

[22] nous avons comencé a *gouter* le petit colis (EV-1914.12.19)

[23] il *parait* que nous partons mercredi (AF-1914.08.17)

- sur le *e*

[24] nous ne savons pas ce qu'il me *reserve* pour moi (EV- 1914.12.06)

[25] nous ne sommes que quatre des premiers *arrives* (PF-1914.09.21)

ii) le choix inapproprié :

[26] je ne sai pas ç*î* vous *ête* malade (LP- 1914.09.09)

[27] chère femme bien *aimès* (AF-1914.08.19)

iii) l'ajout inapproprié :

- sur le *e*

[28] il faut *éspéré* que ç*ê*la na rive rapas (LP-1914.10.01)

[29] Rien plus a vous dirre pour le moment que de vous *embrassér* bien faurt (LP-1914.10.09)

- sur les autres voyelles (deux occurrences)

[30] nous avons pas *eât* a nous plaindre (AF-1914.08.17)

[31] Je mempresse de repondre a ta lettre que je reçois le 5 *dâtée* du 28 (EV-1914.12.06)

Alors que toutes les erreurs d'accent de PF relèvent du type i), ainsi que la plupart de celles de EV et de AF, les erreurs de LP, qui recourt fréquemment aux accents, sont plus diversifiées.

Comme le révèlent les rectifications de 1990, transformant *crémérie* en *crèmerie*, *sécheresse* en *sècheresse*, *cédera* en *cèdera*, *léveraient* en *lèveraient*, *ruisselle* en *ruissèle*, etc., le fonctionnement des accents constitue une source importante de difficultés, et ce à fortiori pour les francophones – dont les méridionaux (Coquillon & Durand 2010: 190) – qui neutralisent la différence /e/-/ɛ/ (Lyche 2010: 150), censée correspondre aux variantes graphiques *é* et *è*, sans oublier le rôle perturbateur que joue également le *ê*. Il est possible

que la difficulté de choisir parmi ces trois accents ait conduit PF à n’accentuer aucun des *e*, rejoignant ainsi une pratique courante avec les majuscules.

Par ailleurs si le choix ou l’ajout inapproprié (ii & iii) de l’accent ne semblent guère avoir d’incidence sur la lecture et l’interprétation du message, dans l’absolu, la forme non-accentuée peut s’avérer ambiguë en raison de l’existence effective d’une « paire minimale orthographique » avec une variante sans accent, qui de surcroît correspond à une différence morphologique :

[32] il est *blesse* a un bras (PF-1914.09.17)

Si, sur la base du contexte, le lecteur rectifiera rapidement en /blese/ la lecture /bles/ dictée par la graphie *blesse*, une telle orthographe engendrera probablement un cout cognitif supérieur à celui de la forme normée *blessé*, qui, elle ne requiert aucun ajustement. En d’autres termes, lorsque l’accent manquant touche une marque morphologique, il est probable que l’impact sur la vitesse de lecture sera plus important que si cette absence d’accent se situe ailleurs dans le mot :

[33] le temps qui nous *separe* se racourçit (PF- 1914.09.21)

Venons-en maintenant à la question des finales, et les difficultés considérables qu’elles soulèvent dans l’orthographe du français.

### 4.3. Les homophones et les finales à rime identique

Le fait que « le nombre de désinences verbales est bien moindre dans la langue parlée que dans la langue écrite » (Blanche-Benveniste 2010: 64) constitue d’emblée un obstacle pour tout scripteur francophone. À titre d’illustration, en nous contentant des cinq temps simples usuels (présent, imparfait, futur, conditionnel, subjonctif) de cinq verbes (voir tableau 7), nous avons calculé le taux de complexification<sup>14</sup> généré par l’orthographe, avec pour le second chiffre (après |) la prise en compte de la neutralisation /e/-/ɛ/, aboutissant par exemple à considérer comme homophones *lavez/lavait/lavais/lavaient*<sup>15</sup>, *laverai/laverais, laverait, laveraient*, etc.

**Tableau 7.** La complexification générée par l’orthographe dans la morphologie verbale

Nombre de:	<i>laver</i>	<i>sortir</i>	<i>boire</i>	<i>vouloir</i>	<i>être</i>
(α) formes orales différentes	12 10	13 11	13 11	14 12	17 16
(β) formes écrites différentes	21	23	23	24	27
Nombre de formes en plus :	+9 +11	+10 +12	+10 +12	+10 +12	+10 +11
Taux de complexification (β-α)/α :	+75 +110%	+77 +109%	+77 +109%	+71 +100%	+59 +69%

Comme l’illustre le tableau 7, le recours à l’orthographe génère un doublement du nombre de formes pour la quasi-totalité des verbes du français, entraînant ainsi une difficulté importante lors de l’encodage graphique. Aussi les grammaires, qu’elles soient pédagogiques ou linguistiques, en viennent-elles à évoquer la présence de « lettres muettes » ou de marque « qui ne sont pas nécessairement prononcées » :

Les morphogrammes sont les marques morphologiques: ce sont, pour l’essentiel, des marques finales écrites, qui ne sont pas nécessairement prononcées (-s du pluriel, -t, -e, etc.). (Riegel, Pellat & Rioul 1994: 70)

Les graphies non-normées de nos scripteurs révèlent précisément des zones de vulnérabilité à cet égard. Elles résultent en moyenne d’un mauvais choix de la finale dans trois cas sur cinq (58,8%), avec, là encore, de fortes disparités entre scripteurs :

**Tableau 8.** La répartition par scripteur des erreurs dans les désinences **3**

	<b>0</b>	<b>2</b> Non-normées		<b>3</b>	
				N	%
(PF)	231	69	⇒	22	<b>31,9</b>
(LP)	308	196	⇒	133	<b>67,9</b>
(EV)	208	100	⇒	44	<b>44</b>
(AF)	139	62	⇒	52	<b>83,9</b>
Total	886	427		<b>251</b>	<b>58,8</b>

Alors que PF ne fait qu'une erreur sur trois dans les désinences, AF, lui, semble rencontrer davantage de difficultés à les orthographier correctement :

[34] tu *va* recevoir ces cartes huit jours apres que nous *seront partit* (AF-1914.08.19)

[35] je n'*ai* encor rien reçus je ne *sait* pas si les lettre ne *parvienn*e pas (AF-1914.08.07)

[36] les lettre *von* plus ou moins a destination (AF-1914.08.07)

Remarquons que, malgré leur graphie non-normée, toutes ces formes ont une lecture conforme à la cible, soit respectivement /va/ /səʁõ/ /partil/ /se/, /paʁvjen/, /võ/. Par ailleurs, à l'exception de *von*, les graphies *va*, *seront*, *partit*, *sait*, et *parvienn*e sont toutes des formes attestées dans les paradigmes verbaux concernés.

Le participe passé est également source d'erreurs. La première origine provient de sa rime en /E/ pour tous les verbes du 1<sup>er</sup> groupe et le recours à une graphie non-normée dans cette fonction, ici *-ai*, *-ez*, *-er* et *-ais* :

[37] ilma ecri de sur le front il venait detre *relevai* des tranchées (EV-1914.12.15)

[38] [...] mai lettre ou carte que je vous est *envoyez* (LP-1914.09.09)

[39] C'est apres avoir *manger* la soupe que je viens passer un petit moment aupres de vous (PF-1914.10.08)

[40] on est tous *eparpillais* par les villes (AF-1914.08.08)

La plupart des participes passés en /e/ à la graphie erronée étant écrits avec des finales non-normées ([37] à [40]) ou sans accent – surtout chez PF (par ex. « il est *arrive* hier », « les morts étaient parrait-il *ammonceles* » 1914.09.05) –, nous n'avons relevé que deux erreurs « classiques » d'accord en genre et en nombre :

[41] chère femme bien *aimès* (AF-1914.08.19)

[42] Voi que nous somme bien *accompagné* (AF-1914.08.17)

Les autres types de participe passé (en /i/, /y/, etc.) donnent également lieu à des erreurs d'accord :

[43] on a le même courage que lorsque nous somme *parti* (LP-1914.10.01)

[44] ils vont être *foutu* et ce ne sera pas malheureux (PF-1914.09.05)

[45] je repon aussi a vos deux lettres que j'ai *reçu* (EV-1914.10.20)

Ces dernières graphies non-normées reflètent amplement les difficultés de l'accord du participe passé que rencontre tout scripteur francophone, comme le rappelle Blanche-Benveniste (2010: 149) : « Si les Français ont des difficultés tellement insurmontables à apprendre l'orthographe de ces accords de participes passés à l'école, c'est qu'ils n'en ont pas de modèle dans leur langue parlée ».

Poursuivons notre analyse des manières d'orthographier la finale /E/ en la détaillant davantage.

#### 4.3.1. Le cas particulier de la finale /E/

Outre le participe passé des verbes du 1<sup>er</sup> groupe<sup>16</sup> qui impose le choix d'une graphie pour le /e/ final dès que le scripteur recourt à un temps composé ou surcomposé, (115 occurrences sur les 886, soit environ 13%), les paradigmes verbaux du français contiennent un nombre relativement important de rimes en /E/. Prenons l'exemple de quatre des verbes



envisagés plus haut en tenant compte cette fois-ci de toutes les finales en /E/ des temps simples usuels : présent, imparfait, futur, conditionnel, subjonctif, infinitif, participe passé.

**Tableau 9.** Le nombre de formes avec une finale /E/ dans la morphologie verbale

Nombre de:	<i>laver</i>	<i>sortir</i>	<i>boire</i>	<i>être</i> <sup>17</sup>
(α) formes en discriminant /e/-/ɛ/	6	6	6	8
(β) formes en neutralisant /e/-/ɛ/	4	4	4	6
(γ) graphies différentes de /E/	16	11	11	14
Taux de complexification [(γ-α)/α][[(γ-β)/β]	+167 +300%	+83 +175%	+83 +175%	+75 +133%

Même dans le cas où le locuteur discrimine /e/ de /ɛ/ (lui permettant par ex. de différencier *laverai* de *laverais* ou *lavé* de *lavais*), le taux de complexification reste considérable, allant jusqu'à atteindre 300% avec les verbes en *-er* en cas de neutralisation. Aussi ne faut-il pas s'étonner que l'encodage graphique du /E/ constitue un écueil important dans l'orthographe du français (voir par ex. Brissaud, Chevrot & Lefrançois 2006). S'il se contente des graphies normées offertes par le paradigme verbal, le scripteur dispose de dix options :

**Tableau 10.** Les 10 graphies différentes correspondant à /E/

/e/					/ɛ/				
-er	-é	-ée	-és	-ées	-ez	-ai	-ais	-ait	-aient

Sur les 324 rimes en /E/<sup>18</sup>, on dénombre 97 erreurs, soit presque une rime en /E/ sur trois, avec là encore des disparités entre scripteurs :

**Tableau 11.** Proportion des erreurs de graphies dans les rimes en /E/

	0	4	Rimes en /E/		5	Erreurs
					N	%
(PF)	231		78	⇒	10	<b>12,8</b>
(LP)	308		123	⇒	51	<b>41,5</b>
(EV)	208		65	⇒	15	<b>23</b>
(AF)	139		58	⇒	21	<b>36,2</b>
Total	886		324		<b>97</b>	<b>29,9</b>

À titre d'illustration, citons quelques exemples :

- permutation ou absence de marque graphique de personne :

[46] je te *disait* donc (AF-1914.08.19)

[47] tu me di que je ne *connai* pas lecri ture (LP-1914.10.28)

- ais → ez

[48] je vous *avez* de mender une paire de calçon (LP-1914.10.23)

[49] si tu *voyez* ses mère tu pourrais le leur dire (PF- 1914.08.15)

- ez → ai/ais

[50] vou me *ferai* savoir le résulta de la recolte (LP-1914.09.09)

[51] vous *aurais* un peu plus de sansse [=chance (cf. Géa 2015: 58-4)] (EV-1914.10.20)

- ez → er

[52] Il saji que vous *soyer* en bonne santée (LP-1914.09.09)

[53] vous ne *savier* pas encor que nous étions ici a Aix (AF-1914.08.19)

- é → er

[54] on est pas tous *loger* au même endroit (AF-1914.08.08)

[55] je suis *nommer* soldat de première classe (LP-1914.10.01)

- é → ai/ais/ait

- [56] nous *avon quitai*... Montp-ellier (LP-1914.09.09)
- [57] on est tous *eparpillais* par les villes (AF-1914.08.08)
- [58] vous *avez était* riche (LP-1914.09.28)

- é → ez

- [59] je vous est *envoyez* (LP-1914.09.09)

On trouve également la substitution du monosyllabe *ai* par son homophone *est* ou *et* :

- [60] d'ailleurs je l'*est* dejas dit (AF-1914.08.19)
- [61] ce petit instant m'*et* si precieux (PF-1914.10.08)

Comme le montrent ces exemples (non-exhaustifs), il semble que toutes sortes de permutations soient possibles, et que les dix options présentées dans le tableau 10 ne constituent pas les seules envisagées par les scribes.

Terminons ce panorama des graphies non-normées par le problème récurrent de la segmentation.

#### 4.4. La segmentation

Comme le figurait le tableau 5 page 6, à l'exception d'AF, les trois autres scribes éprouvent également des difficultés dans la segmentation, que nous envisagerons ici en relation avec le verbe et les clitics s'y rapportant<sup>19</sup>. Ces problèmes de segmentation peuvent revêtir plusieurs aspects (voir également Pellat 2015: 71; 2017: 240) :

b1) l'irruption de coupures au sein d'une unité :

- [62] Je vous *en voi* cette carte (LP-1914.10.23)
- [63] je *de mandais* de ches nouvelles (PF-1914.09.21)

b2) l'agglutination d'unités :

α) par suppression de l'apostrophe :

- [64] hier *jai eu* des nouvelles du Cousin (EV-1914.12.15)
- [65] *jen profite* pour vous faire savoir de mes nouvelle (LP- 1914.09.28)
- [66] Je *mempresse* de repondre a ta lettre (EV-1914.12.06)

β) par suppression de l'espace :

- [67] *jelevoi* souvant (LP-1914.10.28)
- [68] ne vous *enprivez* pas (LP-1914.09.09)
- [69] on *na pas eu* affaire avec eux (EV-1914.12.15)

Bien que le recours à ces stratégies ne présente aucun caractère systémique de la part des scribes<sup>20</sup>, il est toutefois intéressant de remarquer que l'agglutination des clitics sur la forme verbale conjuguée fait écho à certaines propositions théoriques en syntaxe :

L'orthographe française et la grammaire traditionnelle du français traitent comme mots à part les unités suivantes qui ont pourtant plus ou moins clairement des propriétés de formes liées :

- les pronoms personnels conjoints (*je, tu, me, etc.*) qui s'attachent au verbe ;
- le marqueur de négation *ne*, qui s'attache au premier mot du groupe verbal.

(Creissels 2006: 28-29)

Ailleurs, Creissels (1995: 23-24) défend l'idée que « [l]es "pronoms personnels conjoints" sont en réalité des affixes du mot verbal au même titre que les morphèmes traditionnellement reconnus comme "désinences personnelles du verbe" ». Ainsi, en français, étant donné « la possibilité d'antéposer à la base verbale, non seulement un indice de sujet, mais aussi des indices susceptibles de représenter divers types de compléments », il propose l'exemple [70], « où les traits d'union marquent l'appartenance au mot verbal de morphèmes qui sont loin d'avoir l'autonomie que suggère la graphie usuelle du français » :

[70] nous-ne-les-lui-donnerions peut-être pas (Creissels 1995: 159)

On pourrait se demander si les graphies de type « *jelevoi* souvant » (LP-1914.10.28) ne seraient pas le reflet d'une intuition langagière sur le fonctionnement des clitiques, dans la lignée des suggestions de Creissels explicitées graphiquement en [70].

En soi, la segmentation graphique n'est pas naturelle dans la mesure où il existe une dissymétrie entre les intuitions susceptibles de découler de la perception du déroulement de la chaîne sonore à l'oral et les normes de segmentations imposées par l'écrit. Alors que dans la plupart des cas, une pause à l'oral requerra une segmentation sous forme d'espace, l'absence de pause n'implique nullement l'absence de segmentation. Et il est parfois difficile de trancher. Comment devrait-on segmenter l'énoncé /seyndezyzjō/ ?

[71] c'est une des illusions.

[72] c'est une désillusion.

#### 4.5. Synthèse de ce panorama

Il est clair que les graphies non-normées que nous avons brièvement présentées ici sollicitent plus ou moins le lecteur selon l'écart vis-à-vis de la norme et probablement la présence ou non de « paires minimales graphiques » ou de voisins graphiques<sup>21</sup>. Par exemple, bien que comptabilisé dans nos analyses comme graphie non-normée, l'absence d'accent (en dehors des finales) n'entrave guère la lecture :

[73] elle a toujours ces vomissements ça me *derange* un peu (EV-1914.10.20)

Rappelons à cet égard que, le « squelette consonantique » constitue à lui seul un excellent indice de prédiction du mot cible<sup>22</sup> :

[74] ça me *drng* un peu

En d'autres termes, compte tenu des résultats d'expérimentations montrant que « la présentation d'une amorce partielle telle que "jrdn", facilite [...] la lecture ultérieure du mot "jardin" » (Dehaene 2007: 209), il n'est pas étonnant que les graphies non-normées ne constituent guère de difficulté majeure pour la lecture et la compréhension immédiate du message. Toutefois, une gradation semble exister. Ainsi, l'orthographe de l'épouse de LP – que nous n'avons pas analysée dans cet article – paraît plus difficile à déchiffrer que celle de son mari :

[75] je repon a tes cartes et lettres que nous resevons avec plaisir de savoir que tu et enbonnes santès card il anait de maim pour nous tu nous dit sur ta lettres du trois [tres] du deux que nous avons recus ensableux que tu languir de recevoir la photographie du petit je pense que tu la recue mai que vetu nous sonme mal pose il a falus les prade come il et tes tu me dit que je pe ta voier tu vin mais non card sagn sacoude tro cher prou te lan voier il et inpausibleus (JP-1915.01.09)

Qu'en est-il maintenant de la lisibilité des graphies verbales dans notre corpus ?

### 5. La lisibilité des graphies des verbes

En définitive, la question fondamentale reste celle de la lisibilité des graphies utilisées. En d'autres termes, les « bricolages » auxquels recourent nos quatre scripteurs sont-ils aisément décodables par le lecteur ou non ? Permettent-ils d'accéder au message ? Si les graphies normées ne posent en théorie aucun problème de déchiffrement<sup>23</sup>, qu'en est-il des écarts orthographiques observés ici ?

Pellat (2017: 239) considère que « cette variation ne rend pas les lettres illisibles » et qu'elles « sont parfaitement déchiffrables ». Si nous partageons amplement cette impression, en nous limitant aux formes verbales, essayons de déterminer le « taux de lisibilité » de ces écarts graphiques, considérant avec Bosman & Van Orden (1997a: 175) : « One common definition of phonetically accurate spelling errors is that the error pattern

can be pronounced identically to the intended word using grapheme-to-phoneme correspondence rules » (pour une traduction, voir également Bosman & Van Orden 1997b: 209). En d'autres termes, en tenant compte des relations graphème-phonème du français, et du décodage qu'elles imposent en lecture, trois possibilités existent, menant à des statuts différents :

**Tableau 12.** Les trois configurations des rapports de la graphie à la lecture par rapport à la cible

	Graphie	Lecture	Statut par rapport à la norme
(1)	conforme	conforme	souhaitable
(2)	non-conforme	conforme	sans véritable impact sur la lecture
(3)	non-conforme	non-conforme	correction requise lors de la lecture

À titre d'illustration, observons les formes de l'exemple [76] :

[76] (α) je *pense* qua la prochaine d'istribution (β) *jaurer* des nouvelles (LP-1914.10.28)

La forme (α), normée et conforme à la cible, débouche naturellement sur la lecture cible /pās/ (cas de figure 1). Si la graphie (β) s'écarte assez fortement de la norme en raison d'un problème de segmentation et d'une orthographe erronée de la désinence /E/, l'application du principe phonographique autorise immédiatement la lecture cible (cas de figure 2) :

<i>Graphie</i>	<i>j</i>	<i>au</i>	<i>r</i>	<i>er</i>
<i>Lecture de cette graphie</i>	/ʒ	O	ʁ	E/
<i>Lecture ciblée</i>	/ʒOʁE/			

En revanche, dans les exemples [77], [78], [79], tel ne serait pas le cas :

[77] (γ) j'ai *peu* lire sa petite lettre (EV-1914.10.20)

[78] elle repartait le lundi au premier train mais (γ) elle *la manque* (EV-1914.10.20)

[79] (γ) j'ai *recu* des nouvelles de Baillargues (LP-1914.09.28)

Si l'on suit scrupuleusement le principe phonographique<sup>24</sup>, aucune de ces graphies ne peut déboucher sur la lecture cible (cas de figure 3) :

<i>Graphie</i>	<i>p</i>	<i>eu</i>	<i>m</i>	<i>an</i>	<i>que</i>	<i>r</i>	<i>e</i>	<i>c</i>	<i>u</i>
<i>Lecture de cette graphie</i>	/p	ø/	/m	ã	k/	/ʁ	ə	k	y/
<i>Lecture ciblée</i>	/py/		/mãke/			/ʁəsy/			

Dans notre calcul du taux de lisibilité des formes verbales (8 dans le tableau 13), (α) et (β) seront comptabilisés comme adéquats par rapport à la lecture ciblée (6), alors que les graphies (γ) ne le sont pas (7).

**Tableau 13.** Pourcentage de lectures conformes à la cible : 8 = 1 + 6 / 0

	0	Graphie normée ?		⇒	Lecture conforme ?				⇒	Total		Taux
		1	2		6	7	1+6	8				
	N	N	N	%	N	%	N	%	N	%	N	%
(PF)	231	⇒	162	69	⇒	52	75,4	17	24,6	⇒	214	92,6
(LP)	308	⇒	112	196	⇒	185	94,4	11	5,6	⇒	297	96,4
(EV)	208	⇒	108	100	⇒	58	58	42	42	⇒	166	79,8
(AF)	139	⇒	77	62	⇒	58	93,5	4	6,5	⇒	135	97,1
<b>Total</b>	<b>886</b>		<b>459</b>	<b>427</b>		<b>353</b>	<b>82,7</b>	<b>74</b>	<b>17,3</b>		<b>812</b>	<b>91,6</b>

Comme le montrent les pourcentages du tableau 13, sur l'ensemble des scripteurs, le taux de lisibilité (8) permet d'atteindre la lecture ciblée dans plus de 9 cas sur 10. La colonne 6 s'avère particulièrement intéressante dans la mesure où elle montre l'efficacité des stratégies graphiques adoptées par le scripteur. Ainsi, en dépit d'un pourcentage de graphies

non-normées relativement élevé (63,6%, voir tableau 4 page 5), LP recourt à des formes non-normées permettant d'être lues correctement dans 94,4% des cas. En d'autres termes, il tire le maximum de profit du principe phonographique, qui, selon Catach *et al.* (1995/2010: 27) servirait à la transcription des phonèmes dans 80 à 85% des cas<sup>25</sup>. EV se révèle beaucoup moins performant à cet égard, ses formes non-normées n'étant correctement décodables que dans 58% des cas.

En somme, quelles que soient les disparités entre individus, dans l'ensemble, les stratégies développées par nos quatre scripteurs occasionnels s'avèrent très efficaces puisque plus de neuf formes sur dix – qu'elles correspondent à la norme ou présentent des écarts graphiques (orthographe et/ou segmentation) – aboutissent à une lecture conforme à la cible.

Nous terminerons cet article comme nous l'avons commencé, avec cette phrase du scripteur le plus ingénieux de notre corpus, dont tout linguiste travaillant sur ce type d'écrit devrait probablement s'inspirer :

[80] /tymədikəzənəkənepalekʁityʁ[ɛbj]□siʃakfwakətyekʁizələkɔne/ (LP-1914.10.28)

## Références bibliographiques

- Blanche-Benveniste, C. (2010). *Le français. Usages de la langue parlée*, Leuven/Paris: Peeters.
- Bosman, A. M. T. & Van Orden, G. C. (1997a). Why Spelling is More Difficult than Reading, in Perfetti, C. A., *et al.* (eds), *Learning to spell: research, theory, and practice across languages*, Mahwah: Lawrence Erlbaum Associates, 173-194.
- . (1997b). Pourquoi l'orthographe est-elle plus difficile que la lecture?, in Rieben, L., *et al.* (eds), *Des orthographes et leur acquisition*, Lausanne: Delachaux et Niestlé, 207-230.
- Brissaud, C., Chevrot, J.-P. & Lefrançois, P. (2006). Les formes verbales homophones en /E/ entre 8 et 15 ans : contraintes et conflits dans la construction des savoirs sur une difficulté orthographique majeure du français. *Langue française* 151, 74-93.
- Catach, N. (2004). *L'orthographe*, Paris: Presses Universitaires de France (coll. Que sais-je?).
- Catach, N., Gruaz, C. & Duprez, D. (2010). *L'orthographe française. Traité théorique et pratique avec des travaux d'application et leurs corrigés*, Paris: Nathan.
- Chervel, A. & Manesse, D. (1989). *La dictée. Les Français et l'orthographe, 1873-1987*, Paris: INRP/Calmann-Lévy.
- Chervel, A. (1991). L'école républicaine et la réforme de l'orthographe (1879-1891). *Mots* 28 - *Orthographe et société*, 35-55.
- . (2006). *Histoire de l'enseignement du français du XVII<sup>e</sup> au XX<sup>e</sup> siècle*, Paris: Retz.
- Coquillon, A. & Durand, J. (2010). Le français méridional : éléments de synthèse, in Detey, S., *et al.* (eds), *Les variétés du français parlé dans l'espace francophone. Ressources pour l'enseignement*, Paris: Ophrys, 185-197.
- Creissels, D. (1995). *Éléments de syntaxe générale*, Paris: Presses Universitaires de France.
- . (2006). *Syntaxe générale. Une introduction typologique I. Catégories et constructions*, Paris: Lavoisier.
- Dehaene, S. (2007). *Les neurones de la lecture*, Paris: Odile Jacob.
- Durand, J. & Eychenne, J. (2016). Le français méridional, in Detey, S., *et al.* (eds), *La Prononciation du français dans le monde – du natif à l'apprenant*, Paris: CLE international, 34-38.
- Fayol, M. & Jaffré, J.-P. (2014). *L'orthographe*, Paris: Presses Universitaires de France.
- Géa, J.-M. (2015). Le dialecte dans l'écriture de la guerre : la part absente ?, in Steuckardt, A. (ed.), *Entre village et tranchées - L'écriture de Poilus ordinaires*, Uzès: Inclinaisons, 53-65.

Gougenheim, G., Michéa, R., Rivenc, P. & Sauvageot, A. (1964). *L'élaboration du français fondamental : étude sur l'établissement d'un vocabulaire et d'une grammaire de base*, Paris: Didier.

Honvault-Ducrocq, R. (2006). L'orthographe en questions - Présentation, in Honvault-Ducrocq, R. (ed.), *L'orthographe en questions*, Mont-Saint-Aignan: Publications des universités de Rouen et du Havre, 7-22.

Le Lay, Y. (1995). *Larousse Conjugaison*, Paris: Larousse.

Lucci, V. (1994). Présentation de l'enquête et de ses objectifs, in Lucci, V. & Millet, A. (eds), *L'orthographe de tous les jours. Enquête sur les pratiques orthographiques des Français*, Paris: Honoré Champion, 21-41.

Lyche, C. (2010). Le français de référence: éléments de synthèse, in Detey, S., et al. (eds), *Les variétés du français parlé dans l'espace francophone. Ressources pour l'enseignement*, Paris: Ophrys, 143-165.

Martinet, A. (2008). *Éléments de linguistique générale*, Paris: Armand Colin.

Pellat, J.-C. (1988). Indépendance ou interaction de l'écrit et de l'oral? Recensement critique des définitions du graphème, in Catach, N. (ed.), *Pour une théorie de la langue écrite*, Paris: Centre National de la Recherche Scientifique, 133-146.

—. (2015). Les graphies des Poilus, loin des canons orthographiques, in Steuckardt, A. (ed.), *Entre village et tranchées - L'écriture de Poilus ordinaires*, Uzès: Inclinaisons, 67-77.

—. (2017). Les difficultés de mise à l'écrit des peu-lettrés: les graphies des Poilus, in Kristol, A. M. (ed.), *La mise à l'écrit et ses conséquences*, Tübingen: A. Francke, 237-245.

Riegel, M., Pellat, J.-C. & Rioul, R. (1994). *Grammaire méthodique du français*, Paris: Presses Universitaires de France.

Steuckardt, A. (ed) (2015). *Entre village et tranchées - L'écriture de Poilus ordinaires*, Uzès: Inclinaisons,

Treiman, R. & Kessler, B. (2014). *How children learn to write words*, Oxford: Oxford University Press.

<sup>1</sup> Le projet, dirigé par A. Steuckardt (voir Steuckardt 2015) au sein du laboratoire Praxiling (Université Paul-Valéry Montpellier, CNRS – UMR 5267), est présenté sur le site d'Ortolang (<http://corpus14.ortolang.fr/>) comme « donn[ant] à lire les correspondances de Poilus ordinaires. Il privilégie les écrits peu-lettrés, encore peu exploités par les historiens de la Grande Guerre. Ces documents, mis à disposition par les Archives départementales de l'Ain, de l'Hérault et par les familles qui en sont les dépositaires, fourniront aux linguistes des ressources nouvelles pour l'histoire de l'écrit peu-lettré ».

<sup>2</sup> Une analyse qualitative de toutes ces formes (N=886) requerrait un développement supplémentaire conséquent que les limites de cet article interdisent.

<sup>3</sup> Géa (2015: 53) signale qu'« en 1914, les campagnes restent dialectophones mais la généralisation de l'école a donné aux combattants les moyens de lire et écrire » et, en l'occurrence, en français. Il cite à cet égard un Poilu de Manosque : « cest écri en français cest normal pour moi cest normal cest quon na pas appri a écrire a lécole dans avec le patois écrire ça me vient pas en patois » (2015: 56).

<sup>4</sup> Désormais, la référence des extraits de lettres apparaîtra sous la forme des initiales du scripteur, suivies de la date (année, mois, date). Afin de permettre des comparaisons aisées entre tableaux, tous apparaîtront dans cet ordre.

<sup>5</sup> Sur l'ensemble du Corpus 14, ce scripteur s'avère également « l'épistolier le plus proluxe » (Steuckardt 2015: 15).

<sup>6</sup> Entendus après Catach, Gruaz & Duprez (1995/2010: 205) comme les « désinences graphiques supplémentaires qui s'ajoutent accessoirement aux mots selon les rencontres des parties du discours (marques de genre et de nombre, flexions verbales) ».

<sup>7</sup> Sur l'influence possible du substrat occitan et du français méridional lors de l'encodage et le décodage, voir la note 16 page 7 et (Géa 2015).

<sup>8</sup> Les différences /e/-/ɛ/ étant neutralisée dans le français méridional dans ces positions (voir Durand & Eycheune 2016: 34), on recourra à l'archiphonème noté /E/ (à ce propos, voir Martinet 1960/2008: 95).

<sup>9</sup> Le *graphème* sera défini comme « la plus petite unité distinctive et/ou significative de la chaîne écrite, composée d'une lettre, d'un groupe de lettres (digramme, trigramme), d'une lettre accentuée ou pourvue d'un signe auxiliaire, ayant une référence phonique et/ou sémique dans la chaîne parlée » (Catach 1978/2004: 120); (pour une réflexion sur la définition du graphème, voir Pellat 1988).

<sup>10</sup> Au terme « peu-lettré », nous préférons celui de « scripteur occasionnel » qui présente l'avantage de renvoyer à la fréquence de la pratique de l'écrit, fréquence objectivement mesurable. Si, avec Steuckardt (2015: 10, note 2), on pourrait considérer que dans « peu-lettré » « *peu* indique une gradualité », l'usage de cet adverbe défini comme « pas très » dans le *Petit Robert* renvoie inévitablement à un manque par rapport à un étalon constitué par les pratiques de l'élite, à laquelle appartiennent notamment les linguistes travaillant sur ces écrits (voir également les réflexions de Lucci 1994: 22s).

<sup>11</sup> Seuls les verbes pleins *être* et *avoir* sont comptabilisés ici. Leur fonction d'auxiliaire augmenterait naturellement leur fréquence.

<sup>12</sup> Qui n'incluent pas les erreurs se combinant avec les problèmes de segmentation comptabilisés en c).

<sup>13</sup> Remarquons que les rectifications de 1990 autorisent désormais ces orthographes.

<sup>14</sup> L'expression du taux de complexification en % se justifie par sa transférabilité à la totalité des verbes en français.

<sup>15</sup> Et bien entendu *lavé, lavée, lavés, lavées, laver*, non-comptabilisés ici.

<sup>16</sup> Ainsi que *aller, naître, être*.

<sup>17</sup> Prononciation basée sur Le Lay (1995: 53). Les monosyllabes *es* et *est* sont également comptabilisés comme étant à finale /E/.

<sup>18</sup> Sont également inclus les verbes tels que *paraître, faire, avoir*, etc. dont certaines conjugaisons ont une rime en /E/. Des interférences avec les désinences normées en /E/ existent en effet : « surtout la lettre je te *mai* la même chose » (LP-1914.09.28). Le scripteur recourt à *-ai* pouvant provenir du futur ou encore du verbe *avoir* au présent.

<sup>19</sup> Sans le développer (sur l'influence du substrat occitan, voir les analyses de Géa 2015: 57-58), mentionnons au passage qu'en français méridional « la présence quasi-systématique d'une voyelle orale en contexte de liaison (*en [n]avant [anavaŋ]*) » (Coquillon & Durand 2010: 193) conduit LP à l'adoption de segmentations personnelles comme « des Anvelope *ja n'ai plus* » (LP 1914.10.28), où il transcrit la prononciation [zane] à l'aide de *ja n'ai*, au lieu de la graphie normée attendue *j'en ai*.

<sup>20</sup> Par exemple LP (1914.09.28) utilise « *j'ai reçu avec plaisir* » avec une graphie normée de « *j'ai* » et six lignes plus bas, il écrit « *aprésent que j'ai un moment* ».

<sup>21</sup> On pourrait par exemple faire des études de suivi oculaire pour déterminer la vitesse à laquelle les lecteurs déchiffrent certains types d'erreurs par rapport à d'autres.

<sup>22</sup> Plus le mot est long et plus la prédication sera aisée.

<sup>23</sup> Ce point de vue est en partie discutable. Si un lecteur aguerri éprouve probablement davantage d'aisance à décoder une graphie normée, en est-il de même chez des lecteurs/scripteurs occasionnels auxquels ces lettres sont destinées ? En effet, tout scripteur est toujours lui-même le premier lecteur de ce qu'il rédige, et s'il écrit par exemple « tu menvaira Largent » (LP-1914.09.28), c'est probablement qu'il estime que cette graphie –



quel que soit son écart avec la norme – est suffisamment proche de la lecture ciblée pour être conservée. Par ailleurs, même si le destinataire ne partage pas les mêmes idiosyncrasies orthographiques, le décodage s'avère toujours plus aisé (Bosman & Van Orden 1997a ; Fayol & Jaffré 2014: 61-62 ; Treiman & Kessler 2014: 16-17), parce que guidé par les indices pragmatiques, le contexte et la prédictibilité qu'il permet. En encodage le contexte n'est en revanche d'aucun secours. Dans une orthographe aussi complexe que celle du français, en lecture, le décodage graphème-phonème s'avère relativement simple, tandis que l'encodage phonème-graphème offre souvent plusieurs possibilités. Ainsi, alors que *seau* n'admet que [so] comme lecture, le lexème /so/ impose un choix parmi huit orthographe normées : *sot*, *sots*, *saut*, *sauts*, *seau*, *seaux*, *sceau*, *sceaux*. Le scripteur doit par ailleurs écarter toutes les formes alternatives que le fonctionnement orthographique autoriserait, sans pour autant que ces formes soient effectivement attestées au sein du lexique, soit ici par exemple, *so* – conforme aux graphies *accro*, *lasso* –, *sop* – conformes aux graphies de *trop*, *siróp* –, etc.

<sup>24</sup> Il ne s'agit pas ici d'établir une quelconque gradation dans la « gravité » des erreurs (voir Chervel & Manesse 1989: 162s), mais de savoir si la graphie peut ou non donner lieu à une lecture conforme à celle attendue. Les graphies *pense* et *jaurer* de l'exemple [76]) sont *en ce sens* conformes, alors que celles des exemples [77], [78], [79] ne le sont pas, et ce quel que soit leur degré de « gravité ». Remarquons à cet égard que si le digramme *eu* se lit effectivement /y/ dans le participe passé de *avoir* (ou certaines de ses formes au passé simple), et à ce titre rend théoriquement possible une lecture /py/ de la graphie *peu*, une telle option doit être écartée du fait de l'existence d'une « paire minimale orthographique » opposant *peu* à *pu*. Le même raisonnement vaudrait pour *manque* et *manqué*. En soi, la notion même de « faute d'orthographe » sous-entend toujours que l'écart graphique est repérable, et que la forme ciblée peut être rétablie. Pour évaluer scientifiquement la « gravité » des écarts par rapport à la norme, des expérimentations psycholinguistiques estimant le coût cognitif (supplémentaire) en décodage seraient nécessaires.

<sup>25</sup> Catach *et al.* (1995/2010: 27) rappelle effectivement que « les fondations de notre écriture sont bel et bien phonétiques, ou plutôt phonogrammiques : 80 à 85% des signes d'un texte quelconque sont chargés en français de transcrire les sons ». Cependant (Honvault-Ducrocq 2006: 10) déclare quant à elle : « une écriture de type alphabétique comme celle du français est loin d'être seulement phonographique. L'estimation que je peux faire actuellement, à partir de mes recherches en cours, est qu'il existe moins de 40 % de relation biunivoque à l'oral dans l'orthographe. Les 60 % restants sont en relation avec la morphologie ou le sens linguistique dans un souci de distinction graphique favorable au lecteur, toujours prédominant ».