



HAL
open science

Orthographier les formes verbales du français : quelle persistance des erreurs chez les étudiants ?

Jacques David, Fanny Rinck

► To cite this version:

Jacques David, Fanny Rinck. Orthographier les formes verbales du français : quelle persistance des erreurs chez les étudiants ?. *Langue française*, 2021, *Écrire de l'école à l'université: corpus, traitements, analyses outillées*, 3, pp.67-81. 10.3917/lf.211.0067 . hal-03527436

HAL Id: hal-03527436

<https://hal.science/hal-03527436>

Submitted on 15 Jan 2022

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

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

Version auteur. Référence :

David J. & Rinck F. (2021). Orthographier les formes verbales du français. Quelle persistance des erreurs chez les étudiants ? *Langue Française*, 211, 67-81. En ligne : <https://www.cairn.info/revue-langue-francaise-2021-3-page-67.htm>

Orthographier les formes verbales du français. Quelle persistance des erreurs chez les étudiants ?

Jacques DAVID

CY Cergy Paris université, INSPE de Versailles, Laboratoire Agora (EA7392)

Fanny RINCK

Université Grenoble Alpes, INSPE de Grenoble et Laboratoire Lidilem (EA609)

Résumé

Les écrits d'étudiants en français comportent un nombre important d'erreurs orthographiques. Leur persistance à un niveau avancé est systématiquement analysée aux plans linguistique et psycholinguistique, en ce qui concerne la morphographie verbale aux temps-modes personnels. L'étude s'appuie sur deux corpus d'écrits d'étudiants : 276 textes narratifs recueillis en première de licence ; 165 textes argumentatifs en première année de master. Elle montre que les erreurs orthographiques concernent : i) les bases verbo-lexicales ; ii) les flexions temporo-modales ; iii) les accords de personne.

Mots-clés : orthographe, littéracies avancées, littéracies universitaires, morphographie verbale, homophonie-hétérographie.

Spelling French Verb Forms. Which Persistent Errors among Postgraduate Students ?

Abstract

The writings of postgraduate students in French contain a large number of spelling errors. Their persistence at an advanced level is systematically analyzed at the linguistic and psycholinguistic levels, focusing on the verbal morphology through personal modes. The study is based on two corpora of students' writings: 276 narrative texts collected in the first grade of university; 165 argumentative texts in the first Master's level. It shows how these errors concern: i) verbo-lexical bases; ii) tenses and modes endings of verbs; iii) morphosyntactic person markers.

Keywords: French spelling, advanced literacy, academic literacy, verbal morphology, homophony-heterography.

Si l'acquisition de l'orthographe du français est aujourd'hui bien décrite chez les élèves du primaire et, dans une moindre mesure, ceux du secondaire, il n'en est pas de même pour sa maîtrise à plus long terme, notamment chez les étudiants. Leurs écrits, quels que soient la discipline ou le cursus universitaire suivi, révèlent en effet un nombre encore important d'erreurs orthographiques, dont l'étendue et les causes n'ont été que partiellement identifiées. Si cette incapacité à utiliser une orthographe normée à un âge avancé relève bien du constat, sa persistance doit être plus systématiquement analysée aux plans linguistique et psycholinguistique, afin de formuler des analyses explicatives étayées à partir d'hypothèses mieux circonscrites.

C'est l'ambition de cette contribution qui s'appuie sur deux corpus complémentaires d'écrits d'étudiants : les premiers en première année universitaire (276 narrations rédigées à partir d'un texte amorce) ; les seconds au niveau du master (165 textes de synthèse dans le cadre de la préparation au concours d'enseignement). Nous entendons ainsi analyser ce double corpus de « littéracie avancée »¹, pour évaluer les difficultés de ces deux publics de rédacteurs à orthographier les écrits qui leur sont demandés. Et, comme nous ne pouvons analyser ici toutes les composantes orthographiques, nous avons centré notre étude sur celles qui apparaissent les plus déficitaires et les plus prégnantes dans le secteur de la morphographie verbale, aux temps et modes personnels.

Nous verrons ainsi combien et comment ces erreurs appliquées aux verbes résultent de problèmes liés aux trois segments morphémiques qui les composent : i) l'irrégularité des bases lexicales ; ii) l'homophonie-hétérographie des flexions temporo-modales ; iii) l'inaudibilité des accords morphosyntaxiques de personne (à toutes les personnes de P1 à P6).

Mais auparavant, il nous semble nécessaire de passer en revue les quelques travaux disponibles qui décrivent déjà certains des problèmes que nous analyserons in extenso, dans la suite de cette contribution.

1. L'ORTHOGRAPHE A L'UNIVERSITE : ETAT DE LA QUESTION

Un bref rappel historique dans le domaine nous conduit à remonter aux années 1970 pour des recherches fondatrices sur la place de l'orthographe dans l'enseignement du français (Chervel, 1977). Dans les mêmes années, des descriptions plus systématiques de l'orthographe du français sont publiées et diffusées (Blanche-Benveniste et Chervel, 1969 ; Catach, [1989] 1995 ; Lucci et Millet, éd., 1994) ; elles proposent notamment des typologies qui apparaissent – encore aujourd'hui – comme des références majeures pour décrire et comprendre l'acquisition cette orthographe².

En français, les formes verbales constituent le lieu le plus important d'erreurs orthographiques, à tous les niveaux de la scolarité, du primaire au secondaire, et au-delà à l'université – ce qui est l'objet de la présente étude. De fait, la morphographie verbale se caractérise par une homophonie importante, qui a déjà été décrite dans le paradigme distributionnaliste, au moins depuis les années 1960-70 (Dubois, 1967 ; Gross, 1968). Quand elles sont perceptibles à l'oral, les terminaisons des verbes – surtout les flexions temporo-modales – sont prises en compte à l'écrit (p. ex. le *-ai-* de l'Imp-ind ; le *-r-* du Fut-ind), prolongeant ainsi une tradition alphabétique encore prégnante. Cependant, ces morpho-phonogrammes sont le plus souvent suivis de lettres muettes – morphèmes exclusivement écrits³ (p. ex. le *-s* de *prends* ou le *-nt* de *changent* au Prés-ind). Si l'emploi ajusté des morpho-phonogrammes peut se comprendre comme la réanalyse d'éléments audibles, en revanche, la maîtrise des morphogrammes essentiellement écrits est plus aléatoire. Nous verrons cependant que les données extraites de notre corpus portent sur les deux types de morphèmes, et que nous trouvons encore

¹ « La désignation de "littéracie avancée" se base sur l'idée d'apprentissages de l'écrit en continu et se justifie par le fait qu'il s'agit de pointer un niveau de compétences plutôt qu'un contexte ou une visée communicative spécifique » (Jacques et Rinck, 2017).

² Les publications relatives aux différents niveaux scolaires sont tellement nombreuses qu'il nous est impossible de les recenser toutes ; nous renvoyons les lecteurs aux deux ouvrages de synthèse de Fayol et Jaffré (2008, 2014). En revanche, nous proposons ici une revue des travaux – moins nombreux – appliqués aux productions orthographiques des étudiants.

³ À ce niveau, la structuration de l'écriture semble compléter l'apparente défektivité de l'oral, ce que confirme l'histoire de l'orthographe française (Cerquiglini, 1996).

des erreurs morpho-phonologiques dans des proportions importantes, au point que l'on peut s'interroger sur leur réelle maîtrise à long terme, puisqu'elles provoquent encore des erreurs persistantes à l'université.

Au-delà des quelques données quantitatives disponibles, nous n'avons guère d'études étendues et suffisamment précises en littérature avancée⁴. Il nous semble cependant important de circonscrire le problème en le limitant aux recherches disponibles.

Au Québec, quelques recherches révèlent des résultats qu'il convient de prendre en compte. Celle conduite par Simard (1995) montre que l'orthographe génère des difficultés difficilement surmontées chez bon nombre de lycéens et d'étudiants. Il constate ainsi que « le principal écueil a trait aux signes muets à fonction morphologique, différentielle ou étymologique » (*Ibid.* : 145). Quand les meilleurs étudiants disposent « d'un répertoire assez varié de stratégies orthographiques », ceux en difficulté « souffrent d'une sorte d'obsession des homophones grammaticaux usuels » (*Ibid.* : 145). L'auteur formule plusieurs explications qui ont trait : i) au « rendement orthographique » des tâches proposées, ce rendement décroissant d'une situation de test à une situation de rédaction ; ii) au genre des étudiants, les filles sont plus performantes que les garçons ; iii) aux structures d'enseignement car les étudiants des universités sont plus habiles que les « cégépiens »⁵ ; iv) aux écarts entre les plus performants et les plus faibles pour lesquels les savoirs élémentaires ne sont pas maîtrisés, alors qu'ils constituent la population étudiante la plus scolarisée (*Ibid.* : 163).

Dans une autre recherche, également québécoise, Lefrançois (2004) étudie les stratégies de résolution de problèmes orthographiques de vingt-quatre étudiants se préparant à l'enseignement. Lors des entretiens semi-directifs conduits, elle introduit trois variables : i) le fait que le problème ait été ou non résolu correctement ; ii) le niveau de compétence en français écrit des étudiants ; iii) le type de problème linguistique rencontré. L'analyse des données, essentiellement procédurales, montre que les difficultés exprimées sont extrêmement variables d'un sujet à l'autre et qu'elles sont dues, chez les étudiants les moins performants, à une maîtrise aléatoire des connaissances orthographiques.

Dans une troisième étude québécoise, encore plus récente, Duchesne et Piron (2016) proposent un classement des règles d'accord appliquée à une analyse de quatre-vingt-quatre textes argumentatifs d'étudiants (entre 20 et 24 ans), inscrits dans un baccalauréat en enseignement. Les auteures identifient 647 erreurs de tous genres, avec une moyenne de 7,70 erreurs par texte, et spécifiquement 120 erreurs (soit 19 %) dans le domaine grammatical. Les erreurs d'accord internes aux différents syntagmes (56 %) dominent largement et, en leur sein, celles régies par le sujet sont majoritaires (83,5 %). Ainsi, ce sont les accords en personne du verbe – mais aussi en genre et nombre du PPs et de l'attribut – qui apparaissent les plus défailants. Elles montrent également que « ces erreurs relèvent de l'application erronée (ou de l'absence d'application) d'une règle générale d'accord et non de cas particuliers » (*Ibid.* : 90). Ce sont ainsi les procédures d'accord les plus fondamentales qui sont responsables de la majorité des erreurs grammaticales et non les exceptions de cette composante.

En Belgique, le constat d'une « usure » de l'orthographe du français, jusque dans le supérieur, n'est pas moins observé. Ainsi, en 2006, Legros expose une étude qu'il décrit comme un « coup de sonde » dans cent-cinquante copies d'étudiants entrés en faculté de philosophie et lettres. Dans ces écrits d'une cinquantaine de lignes, la moyenne brute est de 6,6 erreurs par copie ; avec des différences interindividuelles qui révèlent que 56,66 % (soit 85 étudiants) commettent de zéro à cinq erreurs ; 22,66 % (nb 34) en produisent de six à dix ; 15,33 % (nb 23) de onze à vingt ; 4,66 % (nb 7) de vingt-et-une à trente ; un étudiant en commet quarante.

D'autres études, comme celle menées auprès d'étudiants français par Roubaud (1997), montrent que la maîtrise accomplie des formes verbales consiste moins à savoir conjuguer – à l'écrit comme à l'oral – une grande diversité de

⁴ Au moins au-delà de la scolarité obligatoire de 16 ans, correspondant à la classe de troisième (ou 4^{ème} du collège en France).

⁵ Le Cégep (Collège d'enseignement général et professionnel) est un établissement public – équivalent du lycée professionnel en France – qui dispense une formation technique et pré-universitaire sur deux ans.

verbes à divers temps-modes qu'à en contourner la réalisation dans des périphrases évitant, par exemple, des subjonctifs imparfaits périlleux ou des passés simples hasardeux. C'est ce que révèle sa recherche lorsqu'elle analyse les réponses de cinquante-quatre étudiants en première année de formation professionnelle à l'enseignement, à un questionnaire comportant des tests de conjugaison en apparence évidents. Ces réponses sont pour le moins inattendues, notamment pour certains verbes que l'on peut penser connus, comme *élire*, dont la forme en P6 du PS-ind apparaît plus évidente en *élient* ou *élisirent* qu'en *élurent*, et ce jusqu'à 66 % des réponses énoncées. L'auteure s'interroge sur les causes de telles défaillances morphologiques : soit elles sont externes et liées à l'enseignement grammatical dispensé, soit elles sont internes et révèlent les incohérences d'un « système » verbal imprégné d'irrégularités. En l'occurrence, les formes du PS-ind en *-urent* qui restent peu congruentes avec le paradigme des verbes comme *élire*. De fait, ces réponses semblent privilégier une logique linguistique liée à la prégnance des bases en *éli-* aux autres temps-modes (Inf, Prés-ind, Fut-ind...), révélant ainsi l'inadéquation des formes en *élu-*, alors même que le PPA correspondant est fortement employé, mais sans être transposé au PS-ind.

Dans la même perspective, une autre étude française conduite par David (2014) vise à décrire les causes probables de ces difficultés et leur ontogenèse. Elle repose sur une série de trente entretiens semi-directifs, enregistrés auprès d'étudiants entrant en licence de lettres modernes. Ces entretiens, enregistrés et transcrits, comprennent, entre autres, des questions ciblant leurs difficultés orthographiques. L'étude permet de distinguer deux grands profils : i) des étudiants qui méconnaissent les règles orthographiques, surtout sur les erreurs en morphosyntaxe verbale qui sont à 71 % non détectées (p. ex. « Dans "Je vous pris donc de constater" "pris" ça me semble bien écrit autrement je vois pas ») ; ii) des étudiants qui ne parviennent pas à « gérer » l'orthographe en cours d'écriture, en particulier dans cette zone de la morphographie verbale qui n'est corrigée que dans 9 % des cas (p. ex. « Ah oui c'est "permets" avec un "s" dans "Je me permet de vous écrire" »).

Que retenir de cette revue de travaux ? Globalement, que la maîtrise des savoirs orthographiques des étudiants reste encore très approximative ; plus spécifiquement, que la morphographie verbale n'est pas encore maîtrisée, sans doute parce qu'elle nécessite des calculs linguistiques complexes que l'on ne peut réduire à une procédure unique. Dans l'étude qui suit, nous verrons que ces erreurs se concentrent sur les bases lexicales des verbes, surtout quand elles sont très irrégulières et peuvent n'avoir aucune lettre commune (p. ex. *aller* > *all-*, *v-*, *ir-*...), et qu'il convient de les distinguer quand elles sont homophones ou plus encore homographes (p. ex. *je suis* = *être* vs *suivre* en P1). Nous montrerons également que certains calculs sont défaillants, notamment pour choisir la flexion temporo-modale qui convient, dans des paradigmes majoritairement morphophonologiques et donc audibles (p. ex. le *-ai-* de l'Imp-ind ; le *-a-* du PS-ind ; le *-r-* du Fut-ind, pour le seul mode indicatif), mais où se concentrent d'autres risques de confusion. Et enfin, le marquage de la personne qui oblige un double calcul : syntaxique pour l'accord au sujet et morphologique pour le choix de l'affixe à sélectionner ; l'ensemble étant soumis à d'autres facteurs d'inaudibilité (p. ex. *je prends* ; *tu mélanges* ; *il partit* ; *elles corrigent*, pour le seul Prés-ind) et d'absence d'affixe (p. ex. *je pose* ; *il vend* ; *elle s'occupe*), autrement classé comme morphème zéro. Nous verrons également que d'autres facteurs engendrent des confusions, comme la non-distinction des classes de mots, notamment entre certains noms et verbes homonymes (p. ex. « en les *étudiants, il est possible de voir que... ») ; mais aussi, dans des proportions plus importantes, sur les homophones-hétérographes strictement verbaux⁶ (p. ex. « La poésie a pendant longtemps *était associée à une idée de travail contraignant » ; ou encore pour la distance et/ou

⁶ Comme nous n'avons retenu, dans la présente contribution, que les erreurs sur les formes verbales aux temps-modes personnels, nous n'évoquons pas les multiples problèmes associés soit aux confusions homonymiques sur les Vinf vs PPA (p. ex. « chez qui il était *aller » ; « Julien ne pouvait en *détaché »), soit aux accords erronés ou absents de ces derniers (p. ex. « Elle a *connue plusieurs formes » ; « Comment les mouvements artistiques sont *défini »).

les écrans mis entre le verbe et le sujet source de l'accord (p. ex. « *Il les *ramassent* » ou « *ils le *regarde* »⁷), et surtout l'inversion de ce dernier (p. ex. « *nous verrons quelle place *a le jeu et la contrainte dans leur travail* »).

2. ÉTUDE DU CORPUS : QUELLE METHODOLOGIE ?

Les corpus utilisés sont des corpus écologiques, recueillis en situation authentique, quoique non spontanée, car ils résultent de tâches d'écriture demandées à des étudiants dans le cadre de leur formation universitaire. Ces deux corpus représentent deux niveaux d'études : la première année de licence et la première année de master de formation à l'enseignement primaire. Dans les deux cas, les étudiants sont issus de filières variées. Aucune distinction n'est faite concernant le statut du français et sa maîtrise dans la biographie de chacun ; ces deux publics d'étudiants sont pris en tant que tels, sur la base des textes qu'ils ont écrits, dans leur cadre universitaire respectif.

Les corpus se différencient en termes de genres textuels. Le premier correspond à une tâche d'écriture narrative demandée en licence ; le second à une tâche d'écriture argumentative à partir de textes sources, conçue sur le modèle d'une des épreuves de français au concours de recrutement de professeur des écoles (CRPE). Chaque corpus est donc homogène au niveau du genre de texte, mais la distinction entre les deux influe, y compris sur les formes orthographiques prises en compte dans notre recherche : le corpus de licence se caractérise, majoritairement, par l'usage du Prés-ind, de l'Imp-ind et du PS-ind, par contraste avec celui de master où les verbes conjugués sont, pour l'essentiel, au Prés-ind.

Outre l'étape de la transcription du corpus de licence, les deux corpus ont fait l'objet d'une annotation orthographique : à chaque forme orthographique erronée est associée une forme corrigée. Cette annotation rend possible l'étiquetage morphosyntaxique de ces écrits non normés. L'exploration des corpus consiste alors à formuler des requêtes pour obtenir l'ensemble des formes associées à leur lemme et à leur classe grammaticale (noms, verbes, prépositions, etc.) – et, parmi elles, l'ensemble des formes erronées –, puis à trier les données pour éliminer certains « bruits » (des résultats non pertinents, p. ex. des verbes au PS-ind, comme *suis* ou *dit*, qui peuvent avoir été étiquetés au Prés-ind). Enfin, l'analyse repose sur un va-et-vient entre les formes et les contextes gauche et droite des occurrences. La démarche consiste à la fois : i) à quantifier les formes erronées en les rapportant à l'ensemble des formes du corpus ; ii) à caractériser les erreurs en comparant formes erronées et formes attendues au niveau de la base verbale, des flexions temporo-modales et des marques de personne ; iii) à identifier d'éventuelles configurations propices aux erreurs.

3. LES COMPOSANTES VERBALES ORTHOGRAPHIQUEMENT DEFAILLANTES

Comme nous l'avons vu antérieurement (*supra* partie 1), l'orthographe du français pose une série de problèmes à tous les scripteurs, et notamment aux étudiants. Et c'est sur la morphologie – et plus particulièrement la morphographie – des verbes que se cristallise la majorité des erreurs et dysfonctionnements constatés. Ainsi, dans cette présentation introductive, nous constatons que la morphographie verbale nécessite un ensemble de connaissances métalinguistiques ou « grammaticales » à différents niveaux de structuration du système linguistique. Les étudiants devraient ainsi être capables d'orthographier les segments qui composent les verbes ; pourtant les erreurs recensées dans notre étude montrent que les connaissances et les procédures sous-jacentes ne sont guère maîtrisées dans leur spécificité et leur intégralité.

C'est en nous appuyant sur ces segments morphologiques que nous avons exploré nos deux corpus pour extraire et étudier séparément ces erreurs de forme, de flexion et de marque, dans deux perspectives complémentaires : statistique et analytique.

3.1. Fréquence des erreurs sur les verbes relativement aux autres classes de mots

En licence, sur les 276 textes de notre corpus, nous recensons 1058 formes erronées, ce qui représente 4,19 % des 25 257 formes grammaticales recensées. Les erreurs relèvent d'abord de la classe des verbes (550 formes erronées, soit

⁷ Nous empruntons ces deux exemples à Negro et Chanquoy qui ont procédé à une étude expérimentale, entre autres avec des adultes, en montrant que ceux-ci, « À l'imparfait, commettent essentiellement des erreurs de proximité que ce soit avec les pronoms « il les » (41,28 %) ou « ils le » (39,37 %) » (2000 : 230).

10,29 %), suivies par celles des noms (222 erreurs) et des adjectifs (112 erreurs). Rapportées à l'ensemble des erreurs, les formes verbales sont affectées à près de 52 %.

En master, sur les 166 textes recueillis, nous recensons 21 611 formes grammaticales, dont 1166 sont erronées, soit 5,4 % de formes non normées. Concernant la classe des verbes, 352 formes sont erronées, soit 10,37 % d'entre elles, et 30,18 % du total des erreurs.

Concernant les formes verbales erronées, nous observons des résultats quasiment équivalents en licence (10,29 % de des formes recensées) et en master (10,37 %). De manière surprenante, on n'observe donc, proportionnellement, pas plus d'erreurs en licence qu'en master, ce que l'on peut interpréter comme une zone d'erreurs difficilement réductibles⁸. Il faut sans doute aussi y voir un effet des modalités d'énonciation et du genre des textes : en licence des écrits d'invention courts et narratifs à partir d'un texte amorce, en master des écrits de réflexion de deux ou trois pages, bâtis à partir de textes sources sur des sujets littéraires (la poésie, le théâtre). Ces conditions et formats d'énonciation influent par ailleurs sur les temps verbaux (*infra*, partie 3.2). Les étudiants de licence sont ainsi confrontés à une plus grande diversité de temps-modes personnels (Prés-ind, Imp-ind et PS-ind), ceux de master recourent presque exclusivement au Prés-ind dans leurs écrits où dominant des séquences explicatives et argumentatives.

La répartition des erreurs varie également en pourcentage d'un temps-mode à l'autre, selon le segment du verbe erroné, comme l'indiquent les tableaux 1 et 2 :

LICENCE	Prés-ind	Imp-ind	PS-ind	Vinf	PPas ⁹	Total erreurs
Erreurs sur base lexicale	36,10%	41,70%	36,60%	61,30%	25,77%	37,32%
Erreurs sur flexion temps-modes	30,55%	17,27%	47,50%	38,71%	30,70%	31,34%
Erreurs d'accord de personne	25%	30,90%	12,87%			12,97%
Erreurs d'accord des PPas					42,90%	13,97%
<i>Nombre de formes erronées¹⁰</i>	36	139	101	62	163	

Tableau 1 : Pourcentage des erreurs par segment à chaque temps-mode en licence

MASTER	Prés-ind	Vinf	PPas	Total erreurs
Erreurs sur base lexicale	3,76%	16,67%	2,13%	7,52%
Erreurs sur flexion temps-modes	21,80%	83,33%	14,89%	40,01%
Erreurs d'accord de personne	62,40%			20,80%
Erreurs d'accord des PPas			82,98%	27,66%
<i>Nombre de formes erronées</i>	133	18	141	

Tableau 2 : Pourcentage des erreurs par segment en master

Un premier constat montre la part importante des erreurs sur la base verbale dans les textes recueillis en licence, (37,32 %) et leur sous-représentation dans ceux de master (7,52 %) ¹¹. Le second résultat est que les formes erronées présentent soit un problème de base, soit un problème de flexion, et ne concernent les deux segments que dans de très rares cas (cinq cas en licence sur l'ensemble des 550 formes verbales erronées). Les étudiants des deux niveaux ne

⁸ Les variations interindividuelles sont cependant importantes (en master, 35 % des textes ne comportent aucune erreur, et le maximum est atteint par un texte qui en compte 41 en deux pages).

⁹ Les erreurs sur les PPas sont comptabilisées dans ce tableau, mais elles ne font pas, ici, l'objet d'une étude spécifique parce que leurs accords en genre et nombre relèvent d'une autre catégorisation linguistique, qui ne peut s'assimiler à celle des accords de personne.

¹⁰ Nous avons ajouté cette dernière ligne avec le nombre de formes erronées, pour mieux se représenter la comparaison en licence et master.

¹¹ Les pourcentages permettent la comparaison entre corpus de licence et de master, mais il convient de préciser qu'en master, ils sont moins représentatifs car les erreurs sur le Vinf se montent seulement à dix-huit occurrences (dont trois problèmes de base verbale). Par ailleurs, le total des colonnes au Prés-ind et à l'Imp-ind atteint 90 % et non 100 %, car certains verbes corrigés portent sur la valeur du temps et non sur leur forme orthographique.

cumulent donc pas les erreurs de base et de flexion, comme si leur connaissance ou leur attention était plus développée – ou défaillante – dans l’une ou l’autre composante, mais pas dans les deux.

Concernant les erreurs liées aux morphèmes temporo-modaux, la proportion est plus équilibrée, puisqu’elle atteint 31,34 % en licence et 40,01 % en master, où les rares erreurs sur la base se concentrent sur les finales. Si l’on s’en tient au seul Prés-ind – seul temps-mode personnel commun aux deux corpus – la proportion d’erreurs d’accord sujet-verbe rapportée à l’ensemble des erreurs verbales, monte à « seulement » 25 % en licence et culmine à 62,40 % en master.

La comparaison de ces pourcentages doit cependant être, là aussi, relativisée en fonction des genres textuels produits dans chacun des cursus. Il semble ainsi que les narrations rédigées en licence, du fait de la plus grande dispersion des temps verbaux, offrent plus de variations orthographiques dans la base lexicale des verbes employés, alors que ce sont les accords de personne qui concentrent des erreurs proportionnellement plus nombreuses dans les textes composés en master. Pour les premiers, les bases lexicales des verbes et les flexions temporelles sont plus difficilement maîtrisées parce qu’elles se répartissent sur une plus grande diversité de temps verbaux, et présentent donc une hétérogénéité phono- et morphographique plus importante. Pour les seconds, ce sont surtout les accords syntaxiques de personne qui posent problème, du fait, sans doute, d’une inconsistance phonologique des marques aux personnes 1, 2, 3 et 6 du Prés-ind.

Nous reprendrons plus en détails ces premiers constats, en nous attachant à décrire tous les types d’erreurs liées aux morphèmes lexicaux (bases verbales) et grammaticaux (flexions et accords) et à les analyser aux trois temps-modes personnels représentés : en licence pour les verbes au Prés-ind, à l’Imp-ind et au PS-ind ; en master pour les verbes au Prés-ind.

3.2. Bases verbales : des résultats contrastés et inattendus

Dans le tableau (1), les erreurs des étudiants de licence, regroupées sur la ligne « base lexicale », concernent en grande partie les accents (50 % des erreurs). Cette prépondérance peut s’expliquer par le fait que le corpus de licence est constitué à partir de textes manuscrits, ce qui n’est pas le cas en master. Ainsi, la tendance à omettre ou confondre ces accents – mais aussi les autres diacritiques – peut se comprendre comme un fait inhérent à la modalité d’écriture et non comme des variantes orthographiques plus ou moins volontaires.

Dans nos deux corpus, ces erreurs tiennent à l’absence totale d’accent diacritique (*depend, rememora*) ou de l’un des deux accents attendus (*revèle*), et dans une moindre mesure à l’ajout d’un accent inexistant (*espérait, effaçait*) ; nous relevons également quelques omissions ou confusions de circonflexes à valeur étymologique (*cotoyait, ôctroyait*).

La maîtrise d’autres lettres ou signes à fonction diacritique est aussi en jeu. Ainsi, nous constatons nombre d’erreurs liées aux graphèmes *c* et *g* du fait de leur double valeur de base et de position, et de l’adjonction phonologiquement obligée du *-e-* ou de la cédille pour sélectionner leur prononciation alternative, le plus souvent au PS-ind en P3 (*éponga, commenca, exerca*). À l’Imp-ind, le problème se pose de manière centrale avec les verbes en *-er*, pour les mêmes lettres *c* et *g* qui exigent de sélectionner la base allomorphe en *-ge-* ou *-ç-* aux P1, P3 et P6 (*siégais, bougait, voyageaient* ; mais aussi *exercait*).

Un second type d’erreurs émerge dans les bases verbales ; il se situe principalement aux trois temps-modes personnels du corpus de licence et concerne les lettres géminées, avec plutôt une tendance à en doubler certaines (*parraît* au Prés-ind ; *occupait, venait* à l’Imp-ind ; *rappella, jetta* au PS-ind). Plus rarement, nous constatons le manque d’une lettre double (*sillonait, permettait*, notamment à l’Imp-ind).

Sur les mêmes bases lexicales du corpus de licence, d’autres erreurs concernent les correspondances phonogrammiques n’affectant pas la phonologie (*reuvevait, fesait* à l’Imp-ind ; *élença* au PS-ind ; *empriment* au Prés-ind). Pour deux de ces temps à l’indicatif, nous relevons des formes paronymiques (*éluvait, réfoissait* (pour *réjouissait*), *accupait, paresait* à l’Imp-ind ; *accueille, paillent* (pour *piaillent*) au Prés-ind) dont les distorsions peuvent

s'expliquer par des maladresses graphomotrices factuelles, des défaillances acoustiques et/ou articulatoires, voire des dysfonctionnements neurophonologiques potentiellement dyslexiques. Quelques erreurs de segmentation, toujours liées à l'apostrophe, sont également observées (*s'avait* pour *savait*).

Pour le PS-ind, nous constatons des difficultés morphologiques spécifiques dans la construction de plusieurs bases verbales (16 erreurs sur les 37 affectant globalement) avec une tendance à surgénéraliser le radical (mais aussi la flexion) de l'infinitif et du Prés-ind des verbes en *-er* (*revena, obtena*), surtout en P3. Ce type d'erreurs apparaît plus marginal pour quelques verbes qui empruntent des bases paronymiques (*asseyèrent, reouvris*), en altérant plus ou moins la phonologie et/ou la morphologie lexicale. C'est aussi à la frontière entre la base et la flexion que l'on observe plusieurs difficultés de construction (*créea, poséa*) et plus largement sur la flexion, elle-même (comme avec *renda*), comme nous l'analyserons *infra*.

Pour résumer, toutes les erreurs relevées sur les bases des verbes des deux corpus (tab. 1 et 2), nous interrogeons à cause de leur persistance à ce niveau de formation universitaire. La plupart sont déjà présentes dans les écrits scolaires du primaire (Bosse, Brissaud et Le Levier 2020) ; et même si elles apparaissent chez les étudiants dans des proportions sensiblement moins importantes, elles restent néanmoins présentes de façon significative dans leurs textes, sans que l'on puisse déterminer lequel des deux facteurs, interne ou externe, domine et peut expliquer leur pérennité : soit ce sont les irrégularités lexico-orthographiques qui expliquent ces erreurs, soit ce sont les contenus et modalités d'apprentissage qui empêchent leur maîtrise exhaustive ; ou probablement les deux facteurs associés. Plus encore, certaines formes paronymiques (*élu minait*) ou analogiques (*revena*) nous paraissent difficilement justifiables, car elles suggèrent une méconnaissance de la morphologie lexicale dans la langue orale et pas seulement sur le versant écrit¹².

Il nous faut également les étudier dans leur relation avec les erreurs de flexions temporo-modales et d'accord de personne, notamment sur le fait que les variantes sur la base et la flexion ne se cumulent pas ou très peu.

3.3. Flexions temporo-modales et marques de personne : des erreurs procédurales à différents niveaux

Aux temps-modes personnels, les erreurs flexionnelles des étudiants sont en grande partie liées à l'opacité homophonique des terminaisons. Cependant, cette problématique n'épuise pas l'ensemble des problèmes, notamment pour le marquage de la personne qui révèle en plus des difficultés spécifiques dans la chaîne d'accord sujet-verbe.

En licence, nous l'avons vu, les problèmes portant sur la base verbale sont nombreux, mais les erreurs sur les suffixes grammaticaux ne sont pas moins quantitativement importantes. Elles représentent 25 % des erreurs au Prés-ind, 30,9 % à l'Imp-ind et 12,9 % au PS-ind. Ce sont à parts égales des problèmes qui concernent les verbes en P3 et en P6 (*passé* pour *passent* ; *connut* pour *connurent* ; *avaient* pour *avait*), auxquels nous devons ajouter ceux relevés en P1-P2 vs P3 (*dis* pour *dit*). En licence – mais aussi en master –, ce type d'erreurs s'explique à la fois par l'homophonie globale de la forme et l'amuïssement de la marque d'accord ; elles se cristallisent plutôt dans le sens de la P1-P2 pour une P3 attendue (*veux* pour *veut*).

En master où, comme on l'a déjà observé, les erreurs sur la base verbale sont quasiment absentes, les problèmes de morphographie se concentrent sur les deux éléments de la terminaison (respectivement : 40 % sur la flexion de temps-mode et 20,8 % sur l'accord en personne). Plus spécifiquement, nous relevons 62 % d'erreurs au Prés-ind pour le marquage de la personne et, parmi elles, 85,5 % correspondent à l'opposition P3 vs P6, qu'elles soient muettes (p. ex. *joue* pour *jouent* et inversement *jouent* pour *joue*) ou non (p. ex. *confond* pour *confondent*), y compris quand elles entraînent une modification de la base orale et écrite (p. ex. *doit* pour *doivent*) ; ces erreurs apparaissent plutôt dans le sens de la P3 pour une P6 attendue (deux tiers des erreurs P3 vs P6) et correspondent dès lors à un sous-marquage

¹² Ce type d'erreurs a également été analysé par Bonami et Boyé (2003) qui constatent que des locuteurs étudiants manifestent un comportement différents vis-à-vis des lexèmes verbaux réguliers et irréguliers, y compris pour des verbes stables comme *conclure*, qui peuvent être énoncés en **conclurent*, sur le modèle d'un infinitif supposé en **concluer*.

spécifique de la P6. Les autres variantes se situent entre P1-P2 vs P3 (p. ex. *dis* pour *dit*), mais aussi entre des modes distincts (p. ex. *défini* pour *définir*). Des erreurs d'homophonie-hétérogramme sont également présentes, avec *a* vs *à*, dont la forme erronée est systématiquement *à* pour *a* ; ce qui pose la question de leur impact sur les erreurs conséquentes des Vinf vs PPs qui suivent. Nous observons également des confusions entre des classes grammaticales homophones, surtout des noms pour des verbes (p. ex. *travail* pour *travaille*, *emploi* pour *emploie*, *rappel* pour *rappelle*) ; probablement du fait de formes nominales plus fréquentes et donc plus prégnantes.

En master, toujours, on relève des erreurs d'accord de personne dans des contextes singuliers. Il peut s'agir de citations intégrées syntaxiquement au discours citant (p. ex. « *c'est lui-même "qui *meurs avec Richard II"* »), mais ce n'est pas une règle générale car, dans certaines phrases ou propositions (p. ex. « *[l'auteur] nous *dis ; lorsqu'elle *parles des spectateurs* »), il est difficile d'identifier des configurations qui expliquent les erreurs de ce type (voir ci-après nos exemples (a) à (d)).

De même les problèmes de marquage de la personne déjà évoqués – qu'il s'agisse la P1-P2 vs P3 ou de la P3 vs P6 – s'expliquent fondamentalement par l'inaudibilité des marques correspondantes. Cependant, près d'un quart des cas relevés échappent à cette explication (p. ex. *confond* pour *confondent* ; *doit* pour *doivent*, et plus encore pour les Vaux *a* pour *ont* ; *est* pour *sont*), en particulier sur des verbes de très grande fréquence. On peut y voir des problèmes de gestion de la chaîne d'accord associés à l'amuïssement des morphogrammes correspondants, mais aussi des difficultés à gérer les procédures d'accord dans des configurations particulières.

Plusieurs exemples nous semblent révélateurs des difficultés de cet ordre. Ils manifestent la délicate gestion processuelle de la morphosyntaxe, voire de la macrosyntaxe discursive. Ainsi, dans ces deux phrases : « *Notre inconscient doit *prend le dessus* » et « *comme celles qui *peuvent se jouent chaque jour sous les yeux du spectateur témoin* », les erreurs d'accord des verbes sont liées aux semi-auxiliaires modaux, dont on peut penser qu'ils ont été insérés sans qu'un nouveau calcul ait été opéré dans une réécriture syntaxiquement élargie à leurs contextes verbaux respectifs. Le manque ou l'oubli des modifications à effectuer en amont ou en aval peut ainsi cumuler des écarts orthographiques et syntaxiques qui résultent en fait de la superposition d'étapes successives de révision du texte, comme dans cet exemple : « *Grâce au discours métalittéraire de ces textes <permet> les poètes <de> partagent leurs visions de la poésie et invitent le lecteur à la création poétique* », où l'ajout de *<permet (...)>* n'est pas suivie d'une relecture-réécriture globale de l'énoncé, à des niveaux plus largement syntaxique et sémantique.

Plus fondamentalement, si le marquage de la personne apparaît majoritairement dominé par des procédures déficitaires ou aléatoires, certaines constructions syntaxiques se révèlent « piégeantes » ; elles génèrent des erreurs dans des contextes particuliers qui peuvent cumuler une distanciation de la source de l'accord, la présence d'un écran et l'inversion du SN sujet, dans le sens de P6 pour P3 (a) ou de P3 pour P6 (b), mais aussi dans des relatives comprenant des accords verbaux de P6 pour P3, qu'ils soient audibles (c) ou non (d) :

- a) [[les poètes créent] des associations inattendues et surprenantes. C'est ce qui <*rendent> leurs œuvres originales et talentueuses.
- b) [...] puis dans une seconde partie nous verrons quelle place <*a> le jeu et la contrainte dans leur travail.
- c) Une mauvaise image du spectateur qui ne *viennent assister à une pièce [...].
- d) [...] se mettre à l'abri de tout élément perturbateur qui *pourraient nuire à la qualité de leurs écrits.

De ces exemples, nous pouvons faire l'hypothèse que certaines configurations sémantico-syntaxiques fonctionnent comme des « distracteurs » (Negro et Chanquoy, 2012) qui entravent le contrôle orthographique et l'analyse métagraphique qui le sous-tend. De fait, ces erreurs d'accord résiduelles ne disqualifieraient pas directement la compétence orthographique des rédacteurs considérés ; en revanche, elles révèlent de réelles défaillances processuelles de révision et/ou de gestion de la coréférence. Dans certaines configurations – notamment en (d) –, on peut y voir aussi une confusion dans l'accord avec des SN sujets pluriels qui reposent sur un conflit entre la notion sémantico-référentielle

de pluralité et son marquage morphosyntaxique, phénomène très prégnant au début de l'apprentissage de l'orthographe (Jaffré et David, 1999).

4. CONCLUSION

Les données extraites de nos deux corpus et analysées ici convergent avec celles des études antérieurement publiées, notamment sur la nature linguistique des erreurs décrites. Celles-ci se concentrent principalement dans la sphère verbale : i) du fait de la variation des bases les plus irrégulières des verbes les plus fréquents (*cf.* également Le Goffic, 1997 ; Ligia-Stela, Fuchs et Mélanie-Becquet, 2013) et des temps grammaticaux plus ou moins défectifs dans les usages oraux courants (Blanche-Benveniste et Adam, 1999) ; ii) du fait de l'homophonie-hétérographie des flexions temporelles et modales, notamment pour les finales en /E/ (Brissaud, Chevrot et Lefrançois, 2003) ; iii) du fait de l'inaudibilité des morphogrammes syntaxiquement accordés (entre autres, les *-e*, *-s*, *-nt* du Prés-ind). Une proportion importante d'étudiants ne parvient guère à ajuster tous ces segments orthographiques dont la logique apparaît inaccessible, ou difficilement gérable dans une révision efficace. Au-delà, certains se contentent de reproduire ou d'étendre des régularités phono- ou morphographiques, sans parvenir à les reconfigurer dans des raisonnements plus abstraits.

La sphère verbale constitue dès lors une zone orthographique très résistante, bien au-delà du cycle secondaire, et probablement un champ d'erreurs quasi insurmontables dans l'enseignement supérieur. Du fait de sa grande complexité, la morphographie verbale fait pourtant l'objet d'un enseignement impliquant un travail linguistique et cognitif précoce, extrêmement long et coûteux. Cette persistance à long terme s'explique par l'écart considérable qui sépare le fonctionnement de la morphologie verbale à l'oral et à l'écrit, mais aussi par des problèmes d'inaudibilité, de redondance et d'homophonie-hétérographie, et surtout par l'impossibilité de les réduire dans une gestion de l'écriture qui nécessite des ressources attentionnelles et mémorielles autant complexes qu'abstraites, et surtout multifactorielles quand elles sont comprises dans des opérations discursives plus exigeantes.

Références bibliographiques

- ANDREU, S. & STEINMETZ, C. (2016), « Les performances en orthographe des élèves en fin d'école primaire (1987-2007-2015) », *Note d'information*, n° 16.28. Paris, Ministère de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche, Direction de l'évaluation et de la prospective.
- BLANCHE-BENVENISTE C. & ADAM J.-P. (1999), « La conjugaison des verbes : virtuelle, attestée, défective », *Recherches sur le français parlé*, 15, p. 87-112.
- BLANCHE-BENVENISTE C. & CHERVEL A. (1969), *L'Orthographe*, Paris, François Maspero.
- BONAMI O. & BOYE G. (2003), « Supplétion et classes flexionnelles », *Langages*, 152, p. 102-126.
- BOSSE M.-L., BRISSAUD, C. & LE LEVIER H. (2020), "French Pupils' Lexical and Grammatical Spelling from Sixth to Ninth Grade: A Longitudinal Study", *Language and Speech*, 64(1), p. 224-249.
- 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, p. 74-93.
- CATACH N. ([1989] 1995), *L'Orthographe française*, Paris, Nathan (réédition Armand Colin en 2016).
- CERQUIGLINI B. (1996), *Le Roman de l'orthographe. Au paradis des mots avant la faute 1150-1694*, Paris, Hatier.
- CHERVEL A. (1977), *Histoire de la grammaire scolaire. Et il fallut apprendre à écrire à tous les petits français*, Paris, Petite Bibliothèque Payot.
- DAVID, J. (2014), « Les écarts orthographiques à l'entrée à l'université », *Le français aujourd'hui*, 185, 95-106.
- DUBOIS J. (1967), *Grammaire structurale du français : le verbe*, Paris, Larousse.
- DUCHESNE J. & PIRON S. (2016), « Écrits universitaires et orthographe grammaticale », *Linx*, 72, p. 73-94.
- FAYOL M. & JAFFRE J.-P. (2008), *Orthographier*, Paris, Presses universitaires de France.
- FAYOL M. & JAFFRE J.-P. (2014), *L'Orthographe*, Paris, Presses universitaires de France.
- GROSS M. (1968), *Grammaire transformationnelle du français. Syntaxe du verbe*, Paris, Larousse.
- JACQUES M.-P. & RINCK F. (2017), « Un 'corpus de littéracie avancée' : résultat et point de départ », *Corpus*, 16, p. 217-237.
- LE GOFFIC P. (1997), *Les Formes conjuguées du verbe français oral et écrit*, Paris-Gap, Ophrys.
- LEFRANÇOIS P. (2004), « Stratégies de résolution de problèmes orthographiques d'étudiants à l'université », in C. Barré-De Miniac, C. Brissaud & M. Rispaïl (éds) *La Littéracie. Conceptions théoriques et pratiques d'enseignement de la lecture-écriture*, Paris, L'Harmattan, p. 233-244.
- LEGROS G. (2006), « Un socle de compétences orthographiques pour l'enseignement supérieur ? », in J.-J. Didier, O. Hambursin & M. Seron (éds) « *Le français m'a tué* ». Actes du colloque « *L'orthographe française à l'épreuve du supérieur* », Louvain, UCL - Presses universitaires de Louvain, p. 69-80.
- LIGIA-STELA F., FUCHS C. & MELANIE-BECQUET F. (2013), *Dictionnaire des verbes du français actuel. Constructions, emplois, synonymes*, Paris, Ophrys.
- LUCCI V. & MILLET A. (éds) (1994), *L'Orthographe de tous les jours. Enquête sur les pratiques orthographiques des Français*, Paris, Champion.
- NEGRO I. & CHANQUOY L. (2000), « Étude des erreurs d'accord sujet-verbe au présent et à l'imparfait. Analyse comparative entre des collégiens et des adultes », *L'Année psychologique*, 100(2), p. 209-240.
- NEGRO I. & CHANQUOY L. (2012), "Subject-verb agreement with present and imperfect tenses: A developmental study from 2nd to 7th grade", *European Journal of Psychology of Education*, 15(2), p. 113-133.
- ROUBAUD M.-N. (1997), « Le passé simple en français », *Studia Neophilologica*, 69(1), 79-93
- SIMARD C. (1995), « L'orthographe d'usage chez les étudiants des ordres postsecondaires », *Revue des sciences de l'éducation*, XXI(1), p. 145-165.