

Les erreurs de syntaxe, d'orthographe grammaticale et d'orthographe lexicale des élèves québécois en contexte de production écrite

Marie-Claude Boivin
Université de Montréal

Reine Pinsonneault
Université du Québec à Montréal

Résumé

Cet article dresse un portrait des erreurs de syntaxe, d'orthographe grammaticale et d'orthographe lexicale relevées dans environ 1000 textes d'élèves québécois de 4^e et 6^e primaire et de 2^e et 5^e secondaire. En général, les filles font moins d'erreurs que les garçons. À tous les niveaux scolaires, les erreurs les plus fréquentes relèvent de la syntaxe (en particulier, ponctuation et homophones grammaticaux). Le nombre d'erreurs liées à la phrase simple diminue peu avec le niveau. Parfois, il n'y a pas de différence significative entre certains niveaux, notamment entre la 2^e secondaire et la 6^e primaire. Aussi, les élèves de 2^e secondaire font significativement plus d'erreurs d'accord régi par le sujet que ceux de la 6^e primaire. Cette absence apparente de progression pour certaines catégories pourrait être attribuée à la production par les élèves plus âgés de contextes syntaxiques plus complexes. Les résultats permettent de discuter certains indicateurs de maturité syntaxique.

Abstract

This paper provides a description of the linguistic errors (in syntax, grammatical spelling and lexical spelling) found in a corpus of almost 1000 texts written by pupils in Quebec (4th and 6th grade in primary school, 8th and 11th grade in high school). Girls generally perform better than boys. For all grades, the most frequent errors are syntactic (especially punctuation and grammatical homophones). The number of errors linked to the structure of the simple sentence does not diminish much across the grades. Sometimes, there are no significant differences between specific grades, notably between the 6th and 8th grade; moreover, the 8th grade pupils make significantly more verbal agreement errors than the 6th grade pupils. This apparent absence of progression for some types of errors could be attributed to the production of more complex syntactic contexts by older pupils. Some indicators of syntactic maturity are discussed in relation to the results.

Les erreurs de syntaxe, d'orthographe grammaticale et d'orthographe lexicale des élèves québécois en contexte de production écrite¹

Introduction

Au Québec, comme dans la plupart des régions francophones, les élèves éprouvent des difficultés quant à la maîtrise de la langue écrite. Dès les années 1990, le groupe DIEPE a montré que, dans des épreuves de production écrite, les élèves québécois réussissent mieux les critères relatifs au discours que ceux liés à la syntaxe, à l'orthographe grammaticale et à l'orthographe lexicale (DIEPE, 1995). Ce constat demeure valide, comme en témoignent les rapports du ministère de l'Éducation, du Loisir et du Sport (MELS, 2010, 2012). Ainsi, dans les épreuves d'écriture de fin d'année de juin 2010, 60,3 % des élèves de 2^e secondaire et 56,7 % des élèves de 5^e secondaire ont réussi le critère *orthographe*, alors que 79,0 % des élèves de 2^e secondaire et 83,3 % des élèves de 5^e secondaire ont réussi le critère *syntaxe et ponctuation* (MELS, 2012). Ces résultats contrastent avec ceux des critères relatifs au discours : 83,3 % des élèves de 2^e secondaire et 97,8 % des élèves de 5^e secondaire ont réussi le critère *adaptation à la situation de communication*, et 85,7 % des élèves de 2^e secondaire et 98,7 % des élèves de 5^e secondaire ont réussi le critère *cohérence du texte* (MELS, 2012).

Les grilles de correction ministérielles ne permettant pas de préciser le type d'erreur au-delà des catégories syntaxe/ponctuation d'une part et orthographe (lexicale ou grammaticale) d'autre part, nous avons élaboré et développé une grille d'analyse plus fine afin de dresser un portrait plus précis des erreurs de syntaxe, d'orthographe grammaticale et d'orthographe lexicale de près de 1000 élèves de quatre niveaux scolaires, soit 4^e primaire, 6^e primaire, 2^e secondaire et 5^e secondaire. Dans cet article, nous présentons les principaux résultats de ce travail.

Cadre conceptuel

Les grilles typologiques existantes se concentrent principalement sur les erreurs d'orthographe lexicale et grammaticale et non sur les erreurs de syntaxe. Catach (1988) et Fayol et Jaffré (2008) par exemple, proposent une taxonomie présentant divers types d'erreurs : les erreurs phonologiquement audibles (erreurs phonétiques), les erreurs phonologiquement acceptables (erreurs orthographiques) et les erreurs de dérivation morphologique. Certaines erreurs relèvent de l'orthographe grammaticale : marques du genre et du nombre des noms et des adjectifs, conjugaison des verbes, et cetera. Dans leur grille de classification des erreurs orthographiques, Manesse et Cogis (2007) distinguent entre autres les erreurs de langue, les erreurs de grammaire et lexicale, les erreurs de grammaire, les erreurs de lexicale.

Les grilles provenant des milieux de l'enseignement du français au Québec incluent, en plus de l'orthographe grammaticale, la dimension syntaxique, à des degrés divers² (cf. Asselin et McLaughlin, 1992 ; Beauchemin et Fortier, 2011 ; Brousseau, Garet, Lionel et Leclerc, 1994 ; Chartrand, 2012 ; Duchesne, 2012 ; Lefrançois, Laurier, Lazure et Claing, 2008 ; Legros et Monballin, 2001 ; Libersan, 2003 ; Préfontaine et Fortier, 2004 ; Roberge, 2006 ; Roy, Lafontaine et Legros, 1995). Les grilles recensées ne permettaient pas un codage fin des erreurs syntaxiques et d'orthographe grammaticale.

La grille de codage utilisée (cf. Tableau 1) est une adaptation de la grille typologique développée dans le cadre du projet *Un modèle didactique d'articulation de la grammaire et de l'écriture pour favoriser le transfert des connaissances grammaticales en situation de production écrite chez les élèves du secondaire* (Boivin et Pinsonneault, 2014b). Contrairement à la grille du MELS, notre grille de codage distingue l'orthographe lexicale de l'orthographe grammaticale. Ce choix repose sur le fait que les erreurs d'orthographe lexicale concernent la maîtrise des correspondances graphèmes/phonèmes, de même que la mémorisation des propriétés idiosyncratiques des mots, alors que l'orthographe grammaticale dépend de l'analyse syntaxique de la phrase. Il s'agit de deux objets distincts au plan linguistique, et les connaissances requises pour les maîtriser le sont également.

Le cadre théorique utilisé pour la catégorisation est celui de la grammaire moderne (ou nouvelle), décrite notamment dans les travaux de Boivin et Pinsonneault (2008), Chartrand (1996), Genevay (1994), Gobbe et Tordoir (1986) et Nadeau et Fisher (2006). Les catégories d'erreurs ont été définies dans ce cadre théorique. Les catégories de niveau 1 et les catégories de niveau 2, leurs codes et des exemples pour chacune des catégories, figurent au Tableau 1.

Le niveau 1 correspond aux trois grands domaines visés par l'étude, soit syntaxe (S), orthographe grammaticale (OG) et orthographe lexicale (OL). Le niveau 2 caractérise de façon plus spécifique les erreurs à l'intérieur des domaines syntaxe et orthographe grammaticale. La catégorie syntaxe comprend les sous-catégories suivantes : les erreurs d'homophonie (S H), les erreurs liées à la phrase simple (S PhS), à la phrase complexe (S PhC), à la ponctuation syntaxique (S P), aux types et formes de phrases (S TyF) et aux phrases à construction particulière. Les erreurs d'homophonie (H) sont incluses dans la catégorie syntaxe, plutôt que dans la catégorie orthographe grammaticale. En effet, même si elles se présentent superficiellement comme un manquement à l'orthographe, les erreurs d'homophonie dépendent principalement de la catégorisation grammaticale du mot *homophone* (Boivin et Pinsonneault, 2012 ; Nadeau et Fisher, 2006). Ce choix permet également de restreindre les erreurs d'orthographe grammaticale aux erreurs d'accord. La catégorie orthographe grammaticale comprend les trois systèmes d'accord définis en grammaire moderne, soit l'accord dans le groupe nominal (OG GN, responsable de l'accord du déterminant et de l'adjectif et de la présence de marques morphologiques sur le nom noyau du GN) ; l'accord régi par le sujet (OG Su, responsable de l'accord du verbe— incluant les désinences verbales, de l'auxiliaire, du participe passé avec *être* et de l'attribut du sujet) et l'accord régi par le complément direct (OG CD, responsable de l'accord du participe passé avec *avoir*, du participe passé de certains verbes pronominaux et de l'accord de l'attribut du complément direct).

Sans faire une analyse détaillée de la composition syntaxique des textes, nous prendrons en compte le concept de « maturité syntaxique » dans la présentation et la discussion des résultats. Le concept de maturité syntaxique, introduit par Hunt (1965), a été repris au Québec par Paret (1991) pour le secondaire et par Brouillet et Gagnon (1990) pour le collégial. La *maturité syntaxique* se définit comme étant les caractéristiques syntaxiques « relevées dans l'écrit des élèves les plus avancés dans leur scolarité » (Hunt, 1965, p. 5 ; traduction de Paret, 1991, p. 10). Ce concept nous permettra de formuler des hypothèses pour expliquer la fréquence de certaines erreurs et d'identifier des constructions qui semblent indiquer un degré de maturité syntaxique plus grand chez les élèves.

Tableau 1

Catégories et sous-catégories de la grille de codes et exemples

Catégories et sous-catégories	Codes	Exemples
Syntaxe	S	
S Homophones	S H	*Je vais a l'école. *Vous vou ler aller au cinéma? *Il a bou ger . * Mang é votre dessert.
S Phrase simple	S PhS	*Il va à l'école et la maison de son ami . *Jusqu'à ce qu'il rencontre Marye-Hélène, une jeune fille blonde aux yeux noisette.
S Phrase complexe	S PhC	*Dans la maison calme, un bruit se fit entendre Harry se réveilla.
S Ponctuation	S P	*Il a achet é, des bonbons. *Hier_ nous nous sommes bien amus ées.
S Types et formes de phrases	S TyF	*Qu' est que tu fais? *Le père des enfants Ø rentrait pas.
S Construction particulière	-	* Il a de belles fleurs dans ce parterre.
Orthographe grammaticale	OG	
OG Accord dans le groupe nominal	OG GN	*les vagues provoquée
OG Accord régi par le sujet	OG Su	*Les enfants [...] le niaisais . Ils partir à rire nt.
OG Accord régi par le complément direct	OG CD	*Les filles qu'elle a vu .
OG Autres	-	*Les animaux c'était toute sauvés.
Orthographe lexicale	OL	* parcontre

Méthode

Cette section présente le corpus, la procédure pour le codage des textes et l'analyse effectuée.

Corpus

Dans le cadre de l'évaluation du Plan d'action pour l'amélioration du français à l'enseignement primaire et secondaire (PAAF), le MELS a effectué, de 2009 à 2011, le suivi des résultats aux épreuves ministérielles d'écriture. Un échantillon représentatif de 1500 copies par fin de cycle (4^e année primaire, 9 ans ; 6^e année primaire, 11 ans ; 2^e secondaire, 13 ans) a été recueilli en 2010. Le MELS a également recueilli les 60 000 copies des épreuves finales des élèves de 5^e secondaire (16 ans). De ce corpus, 1000 textes d'élèves recueillis en 2010 ont été retenus pour constituer le corpus de notre étude. Pour chaque cycle d'études (4^e primaire, 6^e primaire, 2^e secondaire et 5^e secondaire), l'échantillon a été constitué avec un degré de confiance de 95 %, une marge d'erreur minimale de 4,4 %, et un taux de réponse attendu de 95 %. L'échantillon a été composé à partir d'un plan d'échantillonnage aléatoire stratifié selon le sexe et créé de façon à s'assurer que les proportions des élèves des réseaux privé et public seraient les mêmes que dans la population. Les échantillons pour les épreuves obligatoires ont été constitués de façon à être représentatifs de l'ensemble de la population des élèves à la fin de chaque cycle d'études visé. Le Tableau 2 présente l'échantillon final de 969 textes en fonction des caractéristiques sociodémographiques des élèves. Trente-et-un textes ont été exclus à cause d'informations manquantes (sexe de l'élève ou réseau scolaire).

Tableau 2

Description de l'échantillon de textes selon les caractéristiques sociodémographiques des élèves

Niveau scolaire	N	Sexe		Secteur	
		Garçons	Filles	Privé	Public
4 ^e primaire	250	126	124	15	235
6 ^e primaire	220	113	107	15	205
2 ^e secondaire	250	127	123	49	201
5 ^e secondaire	249	125	124	60	189
Total	969	491	478	139	830

Codage

Les textes de l'échantillon ont été anonymisés, transcrits dans Word et vérifiés par une deuxième personne. Le codage des textes a été effectué avec le logiciel HyperResearch, par quatre personnes. Six pourcents des textes ont été contrecodés, soit 58 textes répartis sur les quatre niveaux scolaires. Le coefficient de corrélation intraclasse (CCI) est de 0,945 pour tous les codes d'erreurs. Le Tableau 3 présente le CCI et les intervalles de confiance pour les différentes catégories et sous-catégories.

Tableau 3

Coefficient de corrélation intraclasse et intervalles de confiance pour les catégories et sous-catégories d'erreurs

Catégories et sous-catégories	CCI	IC 95 %
Syntaxe (S)	0,871	[0,759 ; 0,928]
Homophones (S H)	0,966	[0,950 ; 0,978]
Phrase simple (S PhS)	0,561	[0,411 ; 0,694]
Phrase complexe (S PhC)	0,698	[0,570 ; 0,799]
Ponctuation (S P)	0,879	[0,800 ; 0,928]
Orthographe grammaticale (OG)	0,972	[0,958 ; 0,982]
Accord dans le groupe nominal (OG GN)	0,938	[0,908 ; 0,960]
Accord régi par le sujet (OG Su)	0,974	[0,961 ; 0,983]
Accord régi par le complément direct (OG CD)	0,899	[0,854 ; 0,933]
Orthographe lexicale (OL)	0,966	[0,949 ; 0,978]
Toutes	0,945	[0,895 ; 0,970]

Les coefficients de corrélation pour les trois catégories de niveau 1 (S, OG, OL) sont excellents, de même que pour les sous-catégories, à l'exception de S PhS et S PhC. L'ANOVA à un facteur faite avec comparaison multiple de Tukey suggère qu'il y a un effet de codeuse : pour les sous-catégories S PhS et S PhC, une codeuse a relevé significativement plus d'erreurs que les trois autres. On peut penser que peu d'erreurs ont échappé à cette codeuse, par exemple certaines erreurs influencées par l'oral, telles que **Je lui ai demandé* ou **Je me fie sur Jean*, qui peuvent facilement passer inaperçues.

Analyse effectuée

Les résultats sont présentés en nombre moyen d'erreurs par 100 mots (NM100) selon le niveau scolaire (année), le sexe et le secteur, et selon les catégories et sous-catégories d'erreurs de la grille de codage (niveaux 1 et 2). Chaque texte a été codé en entier, et les résultats sont exprimés en NM100 de manière à pouvoir être comparés. La mesure du NM100 est appropriée, mais elle présente une limite. Par exemple, un élève pourrait faire une erreur de choix du pronom relatif par 100 mots, dans un texte ne contenant que deux subordinées relatives. Il ferait donc une erreur dans 50 % des contextes pertinents. Nous n'avons pas inventorié les différentes constructions syntaxiques présentes dans les 969 textes de notre corpus, et le nombre d'erreurs est donc rapporté non pas au nombre de contextes mais au nombre de mots.

Pour chaque catégorie d'erreur de niveau 1, trois analyses séparées ont été faites : une pour le niveau scolaire, une pour le sexe et une pour le secteur (public ou privé). Il n'y avait pas assez de participants dans le secteur privé pour modéliser une analyse avec interactions, et les graphiques de moyennes ne suggèrent pas d'interaction entre les trois variables. Les analyses utilisées pour tester les différences sont l'ANOVA à un facteur pour le niveau scolaire (année), et le test *t* (test d'échantillons indépendants) pour le sexe et le secteur. Les distributions des NM100, qui ne diffèrent pas suffisamment de la normalité, de

même que le grand échantillon ($n = 969$) sont des conditions qui permettent l'utilisation de ces tests paramétriques.

Pour l'analyse selon l'année, le test d'homogénéité des variances rejette l'hypothèse des variances homogènes. Il faut donc utiliser la correction de Welch pour les résultats de l'ANOVA. Il y a un effet significatif de l'année. Puisque les variances ne sont pas homogènes, les comparaisons multiples sont faites avec la correction de Games-Howell. Pour les analyses des erreurs d'accord régi par le complément direct selon l'année, les tests non paramétriques Kruskal-Wallis et Mann-Whitney avec correction de Bonferroni ont été effectués parce qu'il y avait peu d'occurrences de ces erreurs. Le test Mann-Whitney a également été appliqué pour l'analyse de ces erreurs selon le sexe.

Résultats

Nous présenterons d'abord, selon le niveau scolaire et le secteur, les résultats pour l'ensemble du corpus, puis ceux relatifs aux erreurs des catégories de niveau 1, soit syntaxe, orthographe grammaticale et orthographe lexicale. Les résultats pour les sous-catégories des catégories syntaxe et orthographe grammaticale seront ensuite présentés selon le niveau scolaire, les effectifs ne permettant pas de distinguer selon le secteur pour ces sous-catégories. Les résultats selon le sexe sont regroupés à la fin de cette section.

Présentation résultats pour l'ensemble du corpus

La Figure 1 présente, pour l'ensemble du corpus de 969 textes d'élèves, le NM100, pour toutes les catégories d'erreurs, selon le niveau scolaire, le sexe et le secteur. Dans toutes les figures, la mention « année » renvoie au niveau scolaire.

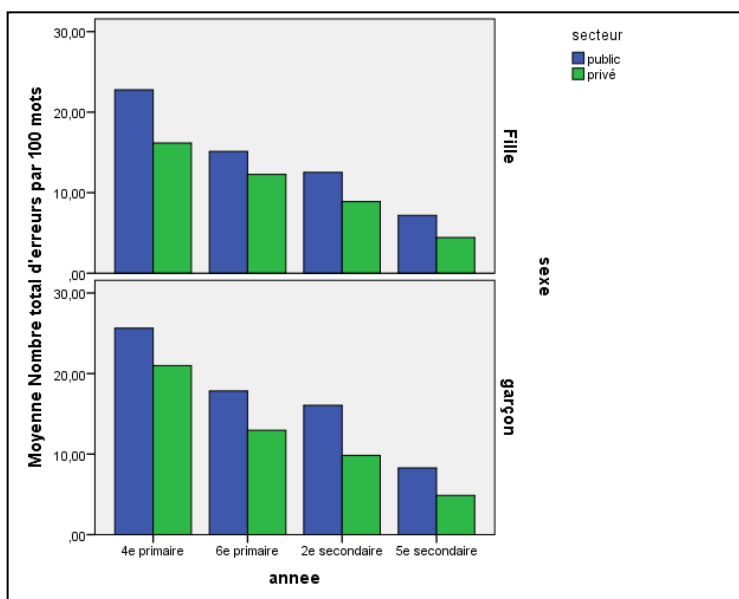


Figure 1. Nombre moyen d'erreurs par 100 mots (NM100) selon le niveau scolaire, le sexe et le secteur.

Les NM100 sont significativement différents ($p < ,001$) entre les sexes : les filles font un peu moins d'erreurs en moyenne que les garçons. Le NM100 est significativement différent selon le secteur ($p < ,001$) : les élèves du secteur public font presque deux fois plus d'erreurs par 100 mots (16,17) que ceux du secteur privé (8,68). Les écarts-types indiquent également une plus forte concentration autour de la moyenne dans le secteur privé.

Sans tenir compte du secteur ou du sexe, le Tableau 4 présente les statistiques descriptives liées au NM100 selon le niveau scolaire des élèves.

Tableau 4

Nombre moyen d'erreurs par 100 mots (NM100) selon le niveau scolaire

Niveau scolaire	NM100	<i>E-T</i>	IC 95 %	Min	Max
4 ^e primaire	23,87	10,85	[22,52 ; 25,22]	2,77	57,34
6 ^e primaire	16,26	8,39	[15,15 ; 17,38]	1,91	64,00
2 ^e secondaire	13,35	7,67	[12,39 ; 14,30]	0,54	47,27
5 ^e secondaire	6,99	4,03	[6,48 ; 7,49]	1,27	22,50

Toutes les différences sont significatives ($p \leq ,001$). Le NM100 est de 23,87 en 4^e primaire et de 6,99 en 5^e secondaire. Le NM100 diminue plus le niveau scolaire augmente ; l'écart-type diminue plus le niveau scolaire augmente. Enfin, les valeurs minimums sont relativement faibles pour tous les niveaux scolaires.

Le Tableau 5 fournit un panorama général des résultats et présente le NM100 pour les catégories syntaxe (S), orthographe grammaticale (OG) et orthographe lexicale (OL) et les sous-catégories. Pour tous les niveaux scolaires et toutes les sous-catégories, le total des erreurs de syntaxe est plus élevé que le total des erreurs d'orthographe grammaticale. Par exemple, en 2^e secondaire, le NM100 pour la catégorie syntaxe est de 6,31 et le NM100 pour la catégorie orthographe grammaticale est de 3,34. Pour tous les niveaux scolaires, cette différence est significative ; les résultats de ce tableau seront présentés en détail dans les sections qui suivent.

Tableau 5

Nombre moyen d'erreurs par 100 mots (NM100) selon les catégories et sous-catégories

Catégories et sous-catégories	Niveau			
	4 ^e prim.	6 ^e prim.	2 ^e sec.	5 ^e sec.
Syntaxe (S)				
Homophones (S H)	2,86	1,53	1,33	0,43
Phrase simple (S PhS)	1,61	1,68	1,41	1,08
Phrase complexe (S PhC)	1,94	1,55	0,79	0,60
Ponctuation (S P)	4,87	3,63	2,59	1,72
Types et formes de phrases (TyF)	0,37	0,25	0,11	0,04
Construction particulière	0,19	0,10	0,09	0,00
Total	11,83	8,74	6,31	3,87
Orthographe grammaticale (OG)				
Accord dans le groupe nominal (OG GN)	1,65	1,76	1,65	0,90
Accord régi par le sujet (OG Su)	2,46	1,08	1,34	0,59
Accord régi par le complément direct (OG CD)	0,34	0,22	0,30	0,08
Autres	0,04	0,05	0,04	0,02
Total	4,48	3,10	3,34	1,58
Orthographe lexicale (OL)	5,56	3,04	2,61	0,89

Pour tous les niveaux scolaires et toutes les sous-catégories, le total des erreurs de syntaxe est plus élevé que le total des erreurs d'orthographe grammaticale. Par exemple, en 2^e secondaire, le NM100 pour la catégorie syntaxe est de 6,31 et le NM100 pour la catégorie orthographe grammaticale est de 3,34. Pour tous les niveaux scolaires, cette différence est significative ; les résultats de ce tableau seront présentés en détail dans les sections qui suivent.

Les erreurs de syntaxe (S)

Après cette présentation générale des résultats, nous donnons au Tableau 6 les statistiques descriptives liées au NM100 pour la catégorie syntaxe selon le niveau scolaire des élèves.

Tableau 6

Nombre moyen d'erreurs par 100 mots (NM100) pour la catégorie syntaxe (S) selon le niveau scolaire

Niveau scolaire	NM100	<i>É-T</i>	IC 95 %	Min	Max
4 ^e primaire	11,83	4,93	[11,22 ; 12,44]	0,82	31,09
6 ^e primaire	8,74	4,24	[8,17 ; 9,30]	0,96	29,41
2 ^e secondaire	6,31	3,69	[5,85 ; 6,77]	0,00	19,71
5 ^e secondaire	3,87	2,14	[3,60 ; 4,41]	0,57	10,99

Toutes les différences sont significatives ($p < ,001$). Le nombre d'erreurs de syntaxe diminue, passant d'environ 12 erreurs en 4^e primaire à moins de quatre en 5^e secondaire. L'écart-type diminue également.

La distribution des élèves selon le NM100 pour la catégorie syntaxe s'illustre bien à l'aide d'histogrammes pour chaque niveau scolaire, comme à la Figure 2.

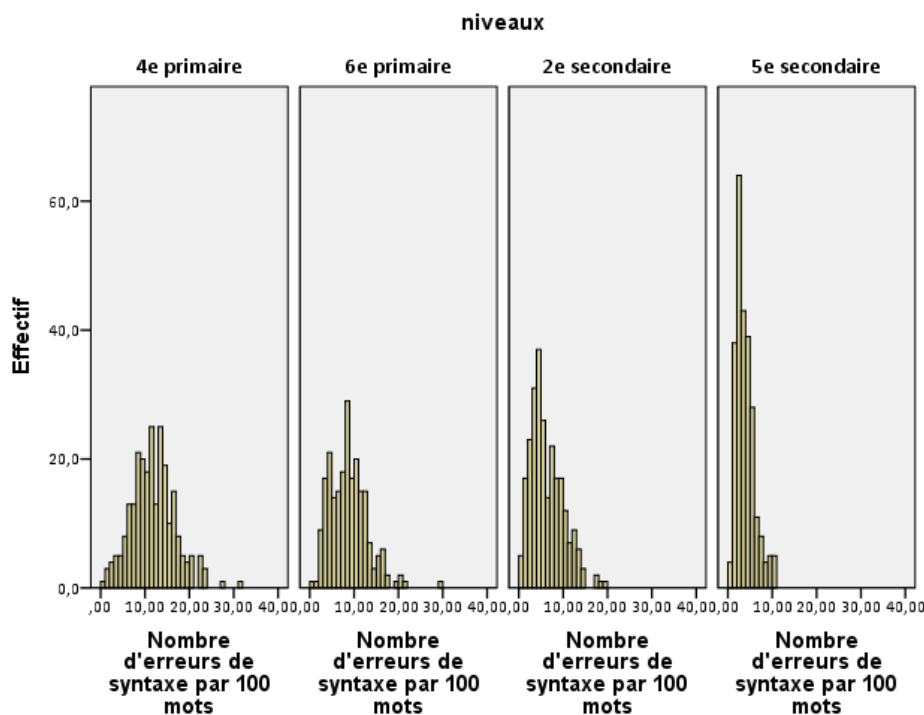


Figure 2. Distribution des élèves selon le nombre moyen d'erreurs par 100 mots (NM100) pour la catégorie syntaxe (S).

La forme des histogrammes montre bien les différences de distribution des élèves selon le niveau scolaire. En 4^e primaire, plusieurs élèves (20 à 30) font entre huit et 15 erreurs de syntaxe par 100 mots. Les barres des histogrammes se concentrent vers la gauche plus le niveau scolaire augmente, indiquant qu'il y a de plus en plus d'élèves qui font un nombre moins élevé d'erreurs. Ainsi, en 5^e secondaire, 100 élèves (sur 250) font deux ou trois erreurs par 100 mots et peu dépassent 10 erreurs.

Si on considère maintenant le secteur, le NM100 pour la catégorie syntaxe se présente comme au Tableau 7.

Tableau 7

Nombre moyen d'erreurs par 100 mots (NM100) pour la catégorie syntaxe (S) selon le secteur

Secteur	N	NM100	É-T
Public	830	8,18	4,92
Privé	139	4,56	3,35

La différence est significative ($p < ,001$). La tendance observée dans le corpus entier se maintient quant au secteur pour les erreurs de syntaxe.

Les erreurs d'orthographe grammaticale (OG)

Considérons maintenant une deuxième catégorie de niveau 1, soit l'orthographe grammaticale. Le NM100 pour la catégorie orthographe grammaticale selon le niveau scolaire de même que les statistiques descriptives pertinentes sont présentées dans le Tableau 8.

Tableau 8

Nombre moyen d'erreurs par 100 mots (NM100) pour la catégorie orthographe grammaticale (OG) selon le niveau scolaire

Niveau scolaire	NM100	É-T	IC 95 %	Min	Max
4 ^e primaire	4,48	2,81	[4,13 ; 4,83]	0,00	15,38
6 ^e primaire	3,10	2,28	[2,80 ; 3,41]	0,00	19,43
2 ^e secondaire	3,34	2,48	[3,03 ; 3,65]	0,00	16,22
5 ^e secondaire	1,58	1,36	[1,41 ; 1,75]	0,00	8,33

Les différences sont significatives ($p < ,001$), sauf entre la 6^e année du primaire et la 2^e année du secondaire ($p = ,713$). On peut penser que les contextes relevant de l'orthographe grammaticale sont plus complexes dans les textes des élèves de 2^e secondaire que dans ceux de 6^e primaire. Il est possible que la présence de sujets complexes affecte l'accord du verbe (**Un de mes amis croient aux fantômes* ; **Les résultats de cette étude diverge*). Cette absence de différence significative entre la 6^e année du primaire et la 2^e année du secondaire est un phénomène récurrent qui sera discuté plus loin.

Les histogrammes de la Figure 3 illustrent la distribution des élèves selon le NM100 pour la catégorie orthographe grammaticale.

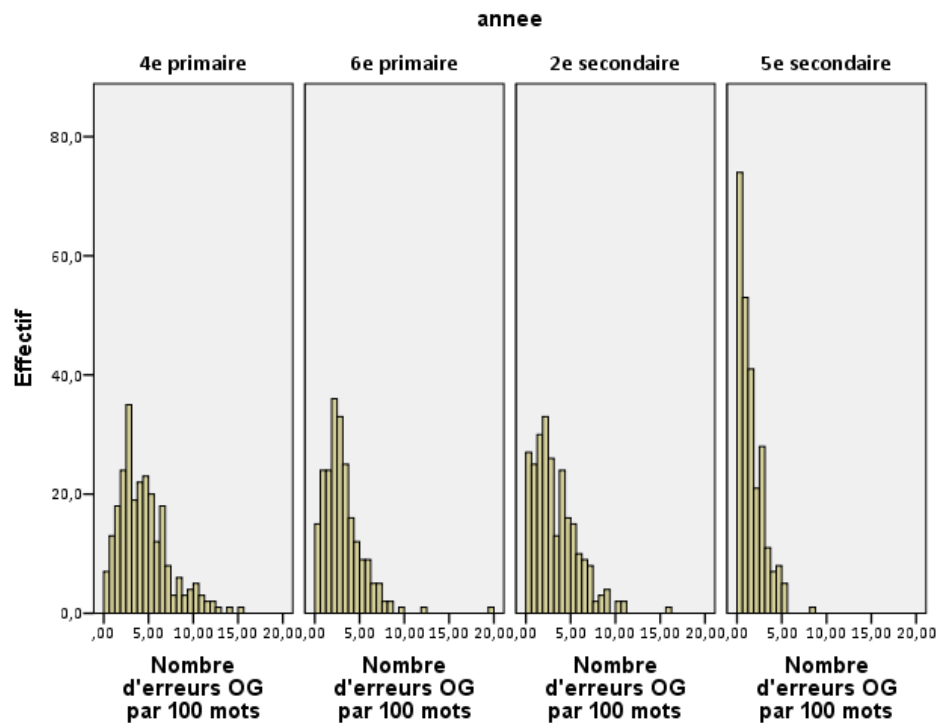


Figure 3. Distribution des élèves selon le nombre moyen d'erreurs par 100 mots (NM100) pour la catégorie orthographe grammaticale (OG).

Les histogrammes montrent que plus le niveau scolaire augmente, plus il y a d'élèves dont les textes ne contiennent pas d'erreurs d'orthographe grammaticale. Ainsi, pour des effectifs comparables (250 élèves), il y a environ 25 élèves de 2^e secondaire qui ne font aucune erreur d'orthographe grammaticale, alors qu'il y en a près de 80 en 5^e secondaire. De même, très peu d'élèves font plus de cinq erreurs d'orthographe grammaticale par 100 mots en 5^e secondaire, contrairement à ce qu'on observe en 2^e secondaire. Si on prend en compte le secteur, le NM100 pour la catégorie orthographe grammaticale se présente comme au Tableau 9.

Tableau 9

Nombre moyen d'erreurs par 100 mots (NM100) pour la catégorie orthographe grammaticale (OG) selon le secteur

Secteur	N	NM100	$\hat{E}-T$
Public	830	3,37	2,57
Privé	139	1,67	1,53

La différence est significative ($p < ,001$) : la moyenne des erreurs d'orthographe grammaticale est deux fois plus élevée dans le secteur public, et l'écart-type est plus petit dans le secteur privé.

Les erreurs d'orthographe lexicale (OL)

Après les erreurs de syntaxe et d'orthographe grammaticale, nous nous tournons brièvement vers les erreurs d'orthographe lexicale. Le NM100 pour la catégorie OL selon le niveau scolaire est donné au Tableau 10.

Tableau 10

Nombre moyen d'erreurs par 100 mots (NM100) pour la catégorie orthographe lexicale (OL) selon le niveau scolaire

Niveau scolaire	NM100	É-T	IC 95 %	Min	Max
4 ^e primaire	5,56	4,53	[5,00 ; 6,13]	0,00	25,68
6 ^e primaire	3,04	2,84	[2,66 ; 3,41]	0,00	26,29
2 ^e secondaire	2,61	2,29	[2,32 ; 2,89]	0,00	14,34
5 ^e secondaire	0,89	0,79	[0,79 ; 0,99]	0,00	4,30

Toutes les différences sont significatives ($p < ,001$), sauf entre la 6^e année du primaire et la 2^e année du secondaire ($p = ,285$). On peut supposer que l'utilisation en 2^e secondaire d'un vocabulaire plus riche, mais dont l'orthographe est moins maîtrisée, contribue à expliquer ces résultats. Les histogrammes présentant la distribution des élèves selon le NM100 pour la catégorie OL suivent à la Figure 4.

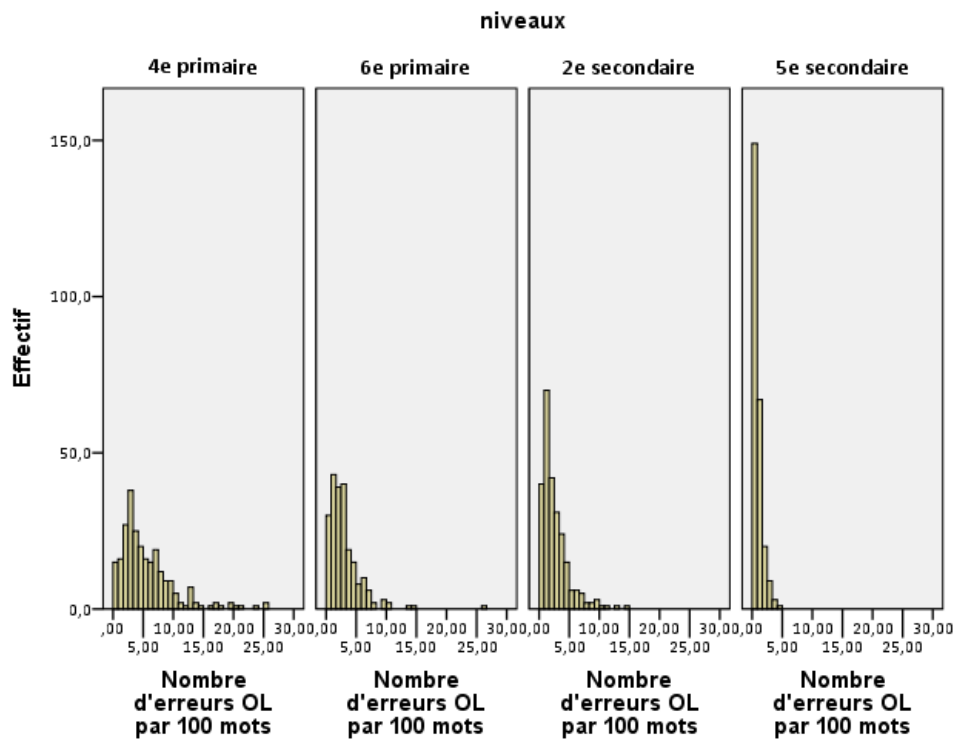


Figure 4. Distribution des élèves selon le nombre moyen d'erreurs par 100 mots (NM100) pour la catégorie orthographe lexicale (OL).

On reconnaît la tendance observée pour les autres catégories ; plus le niveau scolaire est élevé, plus la distribution des erreurs se concentre vers la gauche de l'histogramme. L'orthographe lexicale semble donc mieux maîtrisée avec le niveau scolaire. Il est de plus possible que l'acquisition d'une certaine maturité en écriture et d'un vocabulaire plus varié permette l'utilisation d'une stratégie d'évitement, soit le remplacement d'un mot difficile à orthographier par un autre dont l'orthographe est connue.

Le Tableau 11 présente le NM100 pour la catégorie OL selon le secteur.

Tableau 11

Nombre moyen d'erreurs par 100 mots (NM100) pour la catégorie orthographe lexicale (OL) selon le secteur

Secteur	N	NM100	$\bar{E}-T$
Public	830	3,28	3,50
Privé	139	1,54	2,06

La différence est significative ($p < ,001$). L'analyse des résultats en OL se limite, dans cet article, à ce qui vient d'être exposé.

Comparaison des résultats en syntaxe (S), en orthographe grammaticale (OG) et en orthographe lexicale (OL)

On peut comparer les résultats pour les catégories syntaxe (S), orthographe grammaticale (OG) et orthographe lexicale (OL) à l'aide du graphique à barres de la Figure 5.

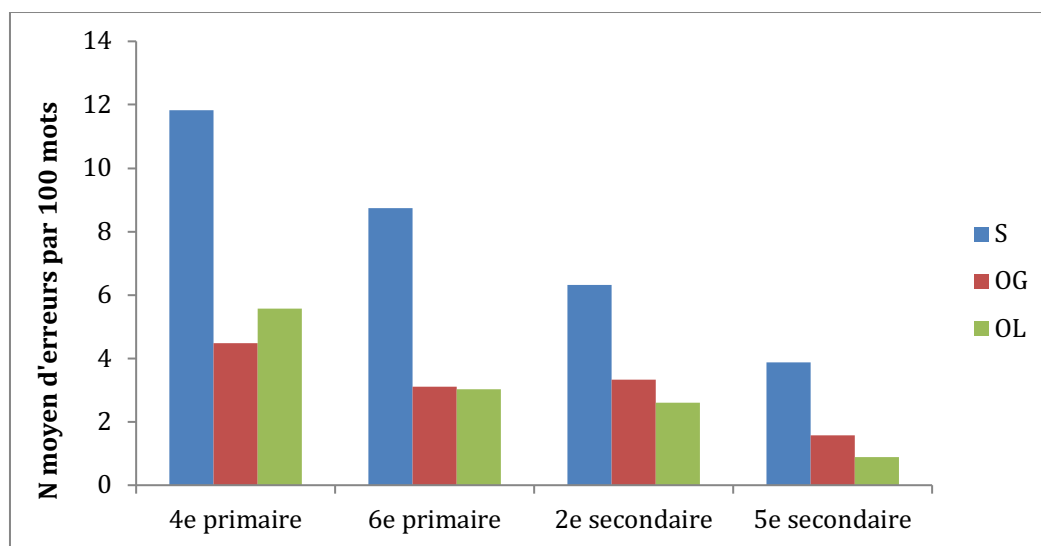


Figure 5. Comparaison des résultats des catégories de niveau 1.

Les erreurs de syntaxe dominent largement à tous les niveaux scolaires. Comme l'illustre l'escalier « descendant » pour ces erreurs, elles diminuent considérablement selon les niveaux scolaires : les élèves de 5^e secondaire font environ trois fois moins d'erreurs de syntaxe que ceux de 4^e primaire, et environ 1,5 fois moins que ceux de 2^e secondaire. Les

erreurs d'orthographe grammaticale évoluent différemment. Il y a une diminution importante d'erreurs d'orthographe grammaticale entre 4^e primaire et 5^e secondaire : on passe de 4,5 erreurs à 1,6 erreur. Il n'y a pas de différence significative entre le NM100 de 6^e primaire et celui de 2^e secondaire pour cette catégorie d'erreurs. La diminution des erreurs d'orthographe lexicale est également considérable : entre la 4^e année primaire et la 5^e année du secondaire, on passe d'environ cinq erreurs à une erreur par 100 mots. Toutefois, la diminution des erreurs d'orthographe lexicale entre 6^e primaire et 2^e secondaire n'est pas significative. On observe donc que, pour l'orthographe grammaticale et l'orthographe lexicale, il n'y a pas de différence significative entre le NM100 de 6^e primaire et celui de 2^e secondaire, ce qui pourrait s'expliquer par une complexité plus grande des textes produits et du vocabulaire utilisé en 2^e secondaire.

Résultats pour les sous-catégories d'erreurs de syntaxe (S)

Les erreurs de syntaxe se répartissent selon les sous-catégories suivantes : erreurs d'homophonie, de phrase simple, de phrase complexe, de ponctuation, et enfin erreurs de types de phrases, de formes de phrases et de phrases à construction particulière.

Erreurs d'homophonie (S H). Le NM100 pour la sous-catégorie homophones selon le niveau scolaire est donné au Tableau 12, avec les statistiques descriptives pertinentes.

Tableau 12

Nombre moyen d'erreurs par 100 mots (NM100) pour la sous-catégorie homophones (S H) selon le niveau scolaire

Niveau scolaire	NM100	$\bar{E}-T$	IC 95 %	Min	Max
4 ^e primaire	2,86	1,94	[2,62 ; 3,10]	0,00	10,07
6 ^e primaire	1,53	1,50	[1,33 ; 1,73]	0,00	8,24
2 ^e secondaire	1,33	1,41	[1,16 ; 1,51]	0,00	8,39
5 ^e secondaire	0,43	0,51	[0,37 ; 0,49]	0,00	3,23

Les différences sont significatives ($p < ,001$), sauf entre 6^e primaire et 2^e secondaire ($p = ,464$), un cas de figure déjà noté pour les erreurs d'orthographe grammaticale et d'orthographe lexicale. La diminution du NM100 entre 4^e primaire et 5^e secondaire est remarquable. L'écart-type diminue également. Il importe aussi de signaler que certains élèves ne font aucune erreur d'homophonie, et ce, dès la 4^e année primaire.

Erreurs de syntaxe dans la phrase simple (S PhS). Le NM100 pour la sous-catégorie phrase simple selon le niveau scolaire figure au Tableau 13.

Tableau 13

Nombre moyen d'erreurs par 100 mots (NM100) pour la sous-catégorie phrase simple (S PhS) selon le niveau scolaire

Niveau scolaire	NM100	<i>É-T</i>	IC 95 %	Min	Max
4 ^e primaire	1,61	1,27	[1,45 ; 1,77]	0,00	7,20
6 ^e primaire	1,67	1,17	[1,52 ; 1,83]	0,00	6,35
2 ^e secondaire	1,41	0,99	[1,28 ; 1,53]	0,00	4,63
5 ^e secondaire	1,08	0,78	[0,98 ; 1,18]	0,00	4,16

Les différences significatives sont celles entre 4^e primaire, 6^e primaire et 2^e secondaire d'une part, et 5^e secondaire d'autre part ($p < ,001$). Il y a aussi une différence significative entre 6^e primaire et 2^e secondaire ($p < ,05$). On peut noter que le maximum d'erreurs est de 4,16 par 100 mots en 5^e secondaire, alors que, dès la 4^e primaire, le minimum à 0. On peut se demander pourquoi le NM100 diminue relativement peu avec le niveau scolaire. Les principales erreurs de cette sous-catégorie relèvent de la construction du groupe nominal (GN), du groupe verbal (GV), notamment par l'absence de compléments obligatoires, et du groupe prépositionnel (GPrép), notamment par un choix de préposition erroné. On peut penser que la phrase simple s'enrichit au fil de la scolarité (GN avec expansions, GV construits à partir de verbes dont la complémentation est moins connue) et conserve en conséquence un certain degré de difficulté.

Erreurs de syntaxe dans la phrase complexe (S PhC). Le NM100 pour la sous-catégorie phrase complexe selon les niveaux scolaires est présenté dans le Tableau 14.

Tableau 14

Nombre moyen d'erreurs par 100 mots (NM100) pour la sous-catégorie phrase complexe (S PhC) selon le niveau scolaire

Niveau scolaire	NM100	<i>É-T</i>	IC 95 %	Min	Max
4 ^e primaire	1,94	1,18	[1,79 ; 2,08]	0,00	5,84
6 ^e primaire	1,55	1,00	[1,42 ; 1,68]	0,00	5,75
2 ^e secondaire	0,79	0,74	[0,69 ; 0,88]	0,00	4,10
5 ^e secondaire	0,60	0,50	[0,53 ; 0,66]	0,00	2,39

Toutes les différences sont significatives ($p \leq ,005$). La diminution dans le NM100 pour la sous-catégorie phrase complexe semble révélatrice d'une plus grande maturité syntaxique entre la 4^e primaire et la 5^e secondaire. En effet, le NM100 passe d'environ deux erreurs en 4^e primaire à moins d'une erreur pour la 2^e et 5^e secondaire. Mentionnons également la différence significative entre 6^e primaire et 2^e secondaire : les élèves de 2^e secondaire font significativement moins d'erreurs de phrase complexe que ceux de 6^e primaire. La phrase complexe est mieux maîtrisée dès la 2^e secondaire et on peut interpréter ce fait comme un signe de maturité syntaxique (cf. Boivin, Roussel et Pinsonneault, 2017).

Erreurs de ponctuation (S P). Le NM100 pour la sous-catégorie ponctuation selon les niveaux scolaires figure au Tableau 15.

Tableau 15

Nombre moyen d'erreurs par 100 mots (NM100) pour la sous-catégorie ponctuation (S P) selon le niveau scolaire

Niveau scolaire	NM100	É-T	IC 95 %	Min	Max
4 ^e primaire	4,87	2,51	[4,56 ; 5,18]	0,00	17,98
6 ^e primaire	3,62	1,87	[3,38 ; 3,87]	0,23	11,76
2 ^e secondaire	2,59	1,59	[2,39 ; 2,78]	0,00	8,72
5 ^e secondaire	1,72	1,05	[1,60 ; 1,65]	0,00	5,30

Toutes les différences sont significatives ($p < ,001$). Les erreurs de ponctuation représentent le plus grand nombre d'erreurs de syntaxe. Bien qu'on note une amélioration entre la 4^e primaire et la 5^e secondaire, la maîtrise de la ponctuation demeure problématique, comme l'indiquent les NM100. Les erreurs de ponctuation concernent principalement la virgule, soit l'absence d'une virgule obligatoire (**Apeuré l'enfant lui expliquait que tous les jeunes le détestaient*) ou la présence erronée d'une virgule (**Il était tellement énervé, qu'il tomba en courant*), et ce, dans les phrases simples ou complexes. Rappelons que la ponctuation est largement tributaire de la syntaxe de la phrase : la maîtrise de la ponctuation dépend de l'analyse de la phrase et notamment de l'identification des constituants obligatoires et facultatifs, de même que des déplacements et ajouts qui se produisent dans la phrase. La maîtrise de la ponctuation semble demander une plus grande maturité syntaxique.

Types et formes de phrases (S TyF) et constructions particulières. Étant donné le très faible nombre d'erreurs pour les sous-catégories types et formes de phrases et phrases à construction particulière, on peut penser que ces constructions ne posent pas de problèmes particuliers. Plusieurs élèves ne font aucune erreur dans ces sous-catégories. Le NM100 pour la sous-catégorie types et formes de phrases passe de 0,37 en 4^e primaire à 0,04 erreur en 5^e secondaire. De plus, le NM100 de phrases à construction particulière est de 0,19 en 4^e primaire et de 0,00 erreur en 5^e secondaire. La présence chez un élève d'erreurs de types ou de formes de phrases, ou de phrases à construction particulière, serait donc indicative de difficultés importantes ; en d'autres termes, il s'agirait d'un indicateur d'une relative absence de maturité syntaxique.

Comparaison des résultats pour les sous-catégories d'erreurs de syntaxe (S). La Figure 6 présente une comparaison des résultats pour les principales sous-catégories d'erreurs de syntaxe, soit homophones (S H), phrase simple (S PhS), phrase complexe (S PhC) et ponctuation (S P).

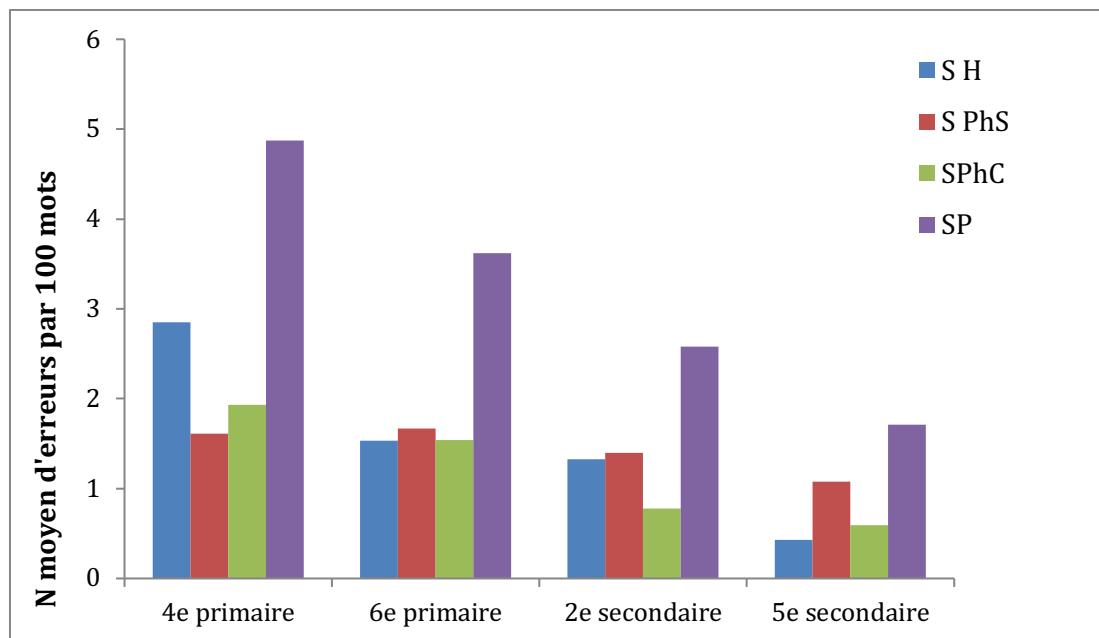


Figure 6. Comparaison des résultats pour les principales sous-catégories d'erreurs de syntaxe (S).

On voit bien l'escalier « descendant » de la 4^e primaire à la 5^e secondaire, en particulier pour les sous-catégories homophones et ponctuation, les deux sous-catégories présentant le plus grand nombre d'erreurs, et également pour la sous-catégorie phrase complexe.

La sous-catégorie phrase simple était celle présentant le moins d'erreurs chez les élèves de 4^e primaire. Pour cette sous-catégorie, on n'observe pas un escalier descendant au fil des niveaux scolaires : Il n'y a pas de différence significative entre la 4^e primaire et la 6^e primaire, mais la diminution est significative en 2^e secondaire et en 5^e secondaire. En 5^e secondaire, le nombre d'erreurs de phrase simple devient plus important que le nombre d'erreurs d'homophones ou de phrase complexe. Rappelons que les principales erreurs de cette sous-catégorie (phrase simple) relèvent de la construction du GN, du GV et du GPrép. La question se pose de savoir ce qui explique cette prégnance des erreurs relatives à la phrase simple chez les élèves de 5^e secondaire.

Résultats pour les sous-catégories d'erreurs d'orthographe grammaticale (OG)

Les résultats pour les sous-catégories d'erreurs d'orthographe grammaticale, soit accord dans le groupe nominal, accord régi par le sujet et accord régi par le complément direct, sont présentés ici.

Erreurs d'accord dans le groupe nominal (OG GN). Le NM100 pour la sous-catégorie accord dans le GN et les statistiques descriptives selon les niveaux scolaires figurent dans le Tableau 16.

Tableau 16

Nombre moyen d'erreurs par 100 mots (NM100) pour la sous-catégorie accord dans le groupe nominal (OG GN) selon le niveau scolaire

Niveau scolaire	NM100	<i>É-T</i>	IC 95 %	Min	Max
4 ^e primaire	1,65	1,52	[1,46 ; 1,84]	0,00	8,29
6 ^e primaire	1,76	1,53	[1,56 ; 1,96]	0,00	11,43
2 ^e secondaire	1,65	1,49	[1,47 ; 1,84]	0,00	8,89
5 ^e secondaire	0,90	0,83	[0,79 ; 1,00]	0,00	4,79

Le NM100 pour la sous-catégorie accord dans le GN n'est pas significativement différent entre 4^e primaire, 6^e primaire et 2^e secondaire. Les différences significatives sont celles avec 5^e secondaire ($p < ,001$). Il est à noter que la maîtrise de l'accord dans le GN est prévue dans la progression des apprentissages (MELS, 2011) pour la 4^e année du primaire. Le NM100 d'accord dans le GN s'établit à 0,9 erreur par 100 mots en 5^e secondaire, et l'écart-type est également plus petit. À tous les niveaux scolaires, il y a des élèves qui ne font aucune erreur d'accord dans le GN (minimum à 0,00 dès la 4^e année), un fait particulièrement remarquable.

Erreurs d'accord régi par le sujet (OG Su). Le Tableau 17 présente le NM100 pour la sous-catégorie accord régi par le sujet et les statistiques descriptives selon les niveaux scolaires.

Tableau 17

Nombre moyen d'erreurs par 100 mots (NM100) pour la sous-catégorie accord régi par le sujet (OG Su) selon le niveau scolaire

Niveau scolaire	NM100	<i>É-T</i>	IC 95 %	Min	Max
4 ^e primaire	2,46	1,89	[2,22 ; 2,69]	0,00	10,49
6 ^e primaire	1,08	1,01	[0,94 ; 1,21]	0,00	8,00
2 ^e secondaire	1,34	1,19	[1,19 ; 1,49]	0,00	7,18
5 ^e secondaire	0,58	0,59	[0,51 ; 0,66]	0,00	3,13

Les différences sont significatives ($p < ,001$) incluant l'augmentation des erreurs d'accord régi par le sujet entre 6^e primaire et 2^e secondaire ($p < ,05$). Le NM100 pour la sous-catégorie accord régi par le sujet est quatre fois moins élevé en 5^e secondaire qu'en 4^e primaire. On note aussi que, pour tous les niveaux scolaires, il y a des élèves qui ne font aucune erreur d'accord régi par le sujet. Les élèves de 2^e secondaire font en moyenne plus d'erreurs d'accord régi par le sujet que ceux de 6^e primaire. La complexité des GN sujets—possiblement plus complexes en 2^e secondaire—pourrait expliquer cette différence entre les résultats des élèves de 2^e secondaire et ceux de 6^e primaire (**Tes amis avec qui tu as fait ton secondaire rêve d'un voyage en Croatie*). On peut également penser que la complexité

syntactique accrue en 2^e secondaire se traduit par la présence d'écrans entre le noyau du GN sujet et le verbe. Ces écrans peuvent causer des erreurs d'accord ; il s'agit par exemple d'un complément pronominalisé (**Ce phénomène les transforment*) ou d'un complément de phrase déplacé (**Les changements climatiques, cette année, touche une grande partie du globe*). En d'autres termes, la présence de contextes d'accord plus complexes peut expliquer le NM100 plus élevé en 2^e secondaire qu'en 6^e primaire. Le développement de la maturité syntaxique illustré par l'utilisation de contextes syntaxiques plus complexes ne se reflète pas, ou pas encore, dans la réussite des accords dans ces contextes.

Erreurs d'accord régi par le complément direct (OG CD). Le NM100 pour la sous-catégorie accord régi par le complément direct de même que les statistiques descriptives selon les niveaux scolaires sont présentés dans le Tableau 18.

Tableau 18

Nombre moyen d'erreurs par 100 mots (NM100) pour la sous-catégorie accord régi par le complément direct (OG CD) selon le niveau scolaire

Niveau scolaire	NM100	<i>É-T</i>	IC 95 %	Min	Max
4 ^e primaire	0,34	0,50	[0,28 ; 0,40]	0,00	4,00
6 ^e primaire	0,21	0,31	[0,17 ; 0,26]	0,00	1,90
2 ^e secondaire	0,30	0,40	[0,25 ; 0,35]	0,00	2,33
5 ^e secondaire	0,08	0,13	[0,06 ; 0,09]	0,00	0,60

Les erreurs d'accord régi par le complément direct sont peu nombreuses dans les textes. Il est probable que les contextes dans lesquels cet accord se produit, soit un complément direct (CD) déplacé devant un participe passé (*Jean les as vus ; Les étudiants que Jean a vus ; Quels étudiants Jean a-t-il vus ?*) ou une construction avec accord de l'attribut du CD (*Pierre rend ses voisins fous*), soient peu produits par les élèves. De plus, les contextes présents dans les textes ne permettent pas toujours de vérifier si l'accord est effectué (*Jean l'a rencontré ; Jean rend son voisin fou*). Il n'y a pas de différence significative entre les NM100 de 4^e primaire, 6^e primaire et 2^e secondaire. Les différences significatives sont avec la 5^e secondaire ($p = ,0083$), une tendance déjà observée pour les erreurs d'accord dans le GN. Notons que la maîtrise du fonctionnement de base de l'accord du participe passé avec *avoir* est prévue en 2^e secondaire, et celle des contextes plus complexes en 4^e secondaire (MELS, 2011).

Comparaison des résultats pour les sous-catégories d'erreurs d'orthographe grammaticale (OG). La Figure 7 présente le NM100 pour les sous-catégories d'erreurs d'orthographe grammaticale : accord dans le GN, accord régi par le sujet et accord régi par le CD.

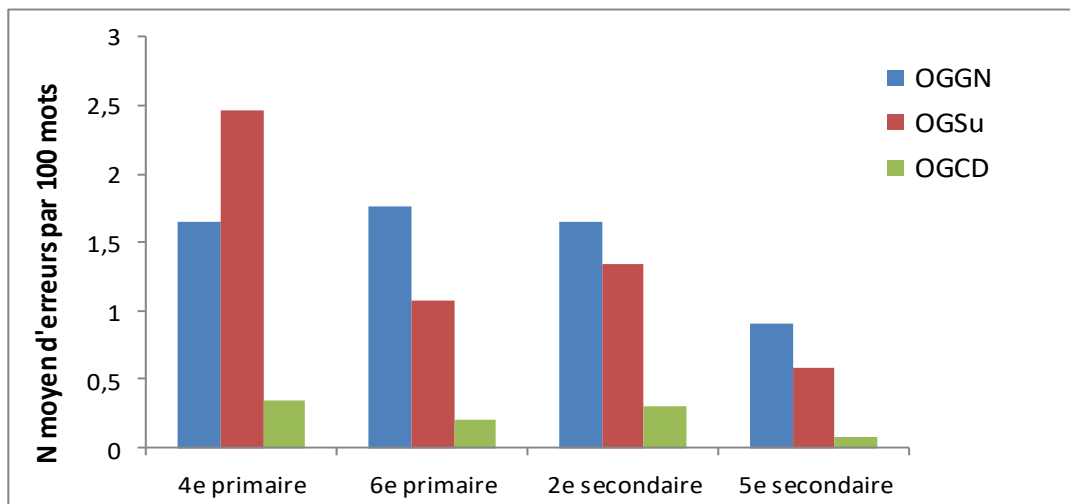


Figure 7. Comparaison des résultats pour les sous-catégories d'erreurs d'orthographe grammaticale (OG).

L'importance des erreurs d'accord dans le GN (OG GN) et d'accord régi par le sujet (OG Su) est illustrée dans la Figure 7. Pour l'accord dans le GN, la hauteur des barres est similaire entre 4^e, 6^e primaire et 2^e secondaire. D'autre part, on voit clairement la diminution du NM100 d'accord régi par le sujet entre la 4^e et la 6^e primaire, puis l'augmentation pour cette sous-catégorie en 2^e secondaire. Comme nous l'avons déjà souligné, ce constat est possiblement attribuable à la présence de contextes plus complexes en 2^e secondaire.

Analyse des erreurs selon le sexe

Cette section présente une comparaison des résultats des garçons et des filles. Le Tableau 19 indique le NM100 et l'écart-type pour les catégories syntaxe, orthographe grammaticale et orthographe lexicale.

Tableau 19

Nombre moyen d'erreurs par 100 mots (NM100) et écart-type pour les catégories syntaxe (S), orthographe grammaticale (OG) et orthographe lexicale (OL)

Sexe	N	S		OG		OL	
		NM100	É-T	NM100	É-T	NM100	É-T
F	478	7,29	4,80	2,80	2,26	2,50	2,79
G	491	8,01	4,96	3,45	2,72	3,54	3,82

Pour les trois grandes catégories d'erreurs, les garçons font plus d'erreurs que les filles. Les différences entre les NM100 des garçons et des filles sont petites mais significatives ($p < ,001$).

Le NM100 selon le sexe pour les quatre sous-catégories de syntaxe (homophones, phrase simple, phrase complexe et ponctuation) figure dans le Tableau 20.

Tableau 20

Nombre moyen d'erreurs par 100 mots (NM100) pour les sous-catégories homophones (S H), phrase simple (S PhS), phrase complexe (S PhC) et ponctuation (S P) selon le sexe

Sexe	N	S H		S PhS		S PhC		S P	
		NM100	É-T	NM100	É-T	NM100	É-T	NM100	É-T
F	478	1,45	1,65	1,36	1,01	1,14	1,00	3,09	2,17
G	491	1,63	1,71	1,51	1,15	1,27	1,01	3,28	2,20
<i>p</i>		,239		< ,05		< ,05		,177	

Les différences entre le NM100 des garçons et des filles pour les sous-catégories homophones et ponctuation ne sont pas significatives. Rappelons que ces deux sous-catégories d'erreurs sont les plus fréquentes en syntaxe. Par contre, les erreurs de phrase simple et de phrase complexe sont minimales mais significatives.

Le Tableau 21 présente le NM100 et l'écart-type selon le sexe pour les sous-catégories d'erreurs d'orthographe grammaticale (accord dans le GN, accord régi par le sujet et accord régi par le CD).

Tableau 21

Nombre moyen d'erreurs par 100 mots (NM100) pour les sous-catégories accord dans le groupe nominal (OG GN), accord régi par le sujet (OG Su) et accord régi par le complément direct (OG CD) selon le sexe

Sexe	N	OG GN		OG Su		OG CD	
		NM100	É-T	NM100	É-T	NM100	É-T
F	478	1,30	1,21	1,23	1,30	0,25	0,40
G	491	1,66	1,56	1,52	1,56	0,22	0,30
<i>p</i>		< ,001		,001		,140	

Pour le NM100 dans les sous-catégories accord dans le GN et accord régi par le sujet, les petites différences entre les garçons et les filles sont significatives. Suivant la tendance générale observée pour la catégorie OG, les filles font légèrement moins d'erreurs que les garçons. En ce qui concerne les erreurs d'accord régi par le CD, il n'y a pas de différence significative entre le NM100 des filles et celui des garçons. Rappelons que les contextes où cet accord se produit sont sans doute peu présents dans les textes des élèves, ce qui expliquerait le faible NM100.

De façon générale, on observe une différence significative entre les performances des garçons et celles des filles, sauf dans les cas où les erreurs sont rares (accord régi par le CD) ou très fréquentes (homophones et ponctuation). Les homophones et la ponctuation semblent ainsi constituer un problème pour les élèves des deux sexes.

Discussion

Dans cette partie, nous mettons en lumière certaines lignes de force qui se dégagent des résultats. Nous présentons d'abord les cinq erreurs les plus fréquentes à chaque niveau scolaire. Nous discutons ensuite de la forte présence des erreurs de syntaxe et de l'absence de différence entre certains niveaux scolaires, notamment le cas de la 2^e secondaire.

Les erreurs les plus fréquentes

Le Tableau 22 présente les cinq erreurs les plus fréquentes pour chacun des niveaux scolaires et permet de mettre en perspective certains résultats déjà présentés, notamment le fait que les erreurs de ponctuation et d'accord dans le GN figurent dans les cinq erreurs les plus fréquentes pour les quatre niveaux scolaires.

Tableau 22

Les cinq erreurs les plus fréquentes pour chaque niveau scolaire

Rang		Niveau scolaire			
		4 ^e primaire	6 ^e primaire	2 ^e secondaire	5 ^e secondaire
1	Sous-catégorie NM100	S P 4,87	S P 3,63	S P 2,59	S P 1,72
2	Sous-catégorie NM100	S H 2,86	OG GN 1,76	OG GN 1,65	S PhS 1,08
3	Sous-catégorie NM100	OG Su 2,46	S PhS 1,68	S PhS 1,41	OG GN 0,90
4	Sous-catégorie NM100	S PhC 1,94	S PhC 1,55	OG Su 1,34	S PhC 0,60
5	Sous-catégorie NM100	OG GN 1,65	S H 1,53	S H 1,33	OG Su 0,59

Note. S H = syntaxe homophones, S PhS = syntaxe phrase simple, S P = syntaxe ponctuation ; OG GN = orthographe grammaticale accord dans le groupe nominal, OG Su = orthographe grammaticale accord régi par le sujet.

Le premier constat est que les erreurs de ponctuation occupent le premier rang des erreurs les plus fréquentes, et ce, à tous les niveaux scolaires. Ces difficultés avec la ponctuation pourraient traduire des lacunes chez les élèves dans l'analyse syntaxique de la phrase (identification et fonction grammaticale des constituants, déplacements, ajouts, etc.), requise pour la ponctuation.

Les erreurs d'accord dans le GN sont très présentes : elles occupent le deuxième rang des erreurs les plus fréquentes en 6^e primaire et en 2^e secondaire, le troisième rang en 5^e secondaire et le cinquième rang en 4^e primaire. Les connaissances métalinguistiques, en particulier, la capacité des élèves à identifier le GN et à analyser sa structure interne, sont corrélées avec la maîtrise de l'accord dans le GN, comme l'ont montré Nadeau et Fisher (2009). Nos résultats suggèrent qu'il est nécessaire de développer chez les élèves des connaissances explicites de la structure interne du GN et des accords qui s'y réalisent.

La construction de la phrase simple occupe le troisième rang en 6^e primaire et en 2^e secondaire, et le deuxième rang en 5^e secondaire. Elle ne figure pas dans les cinq erreurs les plus fréquentes pour 4^e primaire. Ces résultats peuvent sembler étonnants, en particulier pour la 5^e secondaire. On peut toutefois noter que le NM100 diminue en 5^e secondaire. Il semble que le taux d'erreurs diminue moins pour la sous-catégorie phrase simple que pour

les autres sous-catégories d'erreurs, ce qui indiquerait notamment que la complémentation verbale et le choix des prépositions demeurent problématiques.

La sous-catégorie accord régi par le sujet constitue une source importante d'erreurs, puisqu'elle se situe au troisième rang en 4^e primaire, au quatrième rang en 2^e secondaire et au cinquième rang en 5^e secondaire. Elle ne figure pas dans les cinq erreurs les plus fréquentes en 6^e primaire. On peut formuler l'hypothèse que la complexité des GN sujet change peu entre la 4^e et la 6^e primaire, ce qui pourrait expliquer la relative maîtrise de l'accord régi par le sujet en 6^e primaire (NM100 = 1,08 ; au sixième rang). De plus, on peut poursuivre dans cette voie avec l'hypothèse déjà mentionnée que la complexité des GN sujet augmente au secondaire et que cela se traduit par des erreurs d'accord régi par le sujet.

Finalement, les erreurs d'homophonie occupent le deuxième rang en 4^e primaire, le cinquième rang en 6^e primaire et en 2^e secondaire, et ne font plus partie des cinq erreurs les plus fréquentes en 5^e secondaire. Remarquons que les erreurs d'homophonie très présentes au début de la scolarité s'atténuent sensiblement dès la 6^e primaire pour quitter la liste des erreurs les plus fréquentes en 5^e secondaire.

Les erreurs de syntaxe (S)

Une observation centrale relativement aux résultats concerne l'importance des erreurs de syntaxe. En effet, les erreurs de syntaxe prises dans leur ensemble sont significativement plus fréquentes dans les textes d'élèves que les erreurs d'orthographe grammaticale ou d'orthographe lexicale. Globalement, la fréquence des erreurs de syntaxe est deux à trois fois plus élevée que celle des erreurs d'orthographe grammaticale ou d'orthographe lexicale. Pour l'enseignement et l'apprentissage du français écrit, cela nous semble pointer vers la nécessité de l'analyse syntaxique : on ne peut faire l'économie de l'analyse syntaxique de la phrase. La compréhension de la phrase de base et de ses propriétés (cf. Boivin, 2012) et la capacité d'identifier la structure de la phrase et des groupes qui la constituent nous semblent au cœur de la maîtrise de la construction des phrases simples et des phrases complexes, de l'orthographe des homophones et des règles de ponctuation, et également, comme le montrent Nadeau et Fisher (2009), de l'orthographe grammaticale.

La maturité syntaxique et les différences entre les niveaux scolaires

Les différences entre les NM100 sont généralement significatives entre les niveaux scolaires. Toutefois, la 2^e secondaire se distingue, car, pour certaines catégories et sous-catégories, les différences avec la 6^e primaire ne sont pas significatives. C'est le cas pour l'orthographe grammaticale, et plus spécifiquement pour l'accord dans le GN. C'est aussi le cas pour l'orthographe lexicale et la maîtrise des homophones. Il faut de plus souligner qu'il y a un nombre d'erreurs significativement plus élevé en 2^e secondaire qu'en 6^e primaire pour l'accord régi par le sujet (accord du verbe, accord du participe passé avec *être*, accord de l'attribut du sujet). La 2^e secondaire se présente donc comme une année charnière dans la formation des élèves.

Nous avons souligné au fil du texte que ce qui pouvait sembler un recul était peut-être en fait dû à une complexité plus grande des textes produits, notamment en ce qui concerne la structure interne des GN, qui influence à la fois l'accord dans le GN (*Des étudiants vraiment très fiables et toujours à l'heure*) et l'accord du verbe (*Une petite fille, son frère et leur voisin rencontrent le maire*). De plus, l'accord du verbe peut être rendu

plus difficile par la présence d'un écran entre le noyau du GN sujet et le verbe (*Les rebondissements de l'histoire, au fil du texte, semblent incroyables*). Il nous semble donc que l'absence de différence entre 6^e primaire et 2^e secondaire, exprimée en NM100, indique peut-être un passage à une plus grande complexité syntaxique et à une prise de risques plus importante dans les textes des élèves de 2^e secondaire (voir p. ex. Paret et Chartrand, 1990). Cette plus grande maturité syntaxique pour la 2^e secondaire se manifesterait dans les résultats relatifs à la phrase complexe, le NM100 pour cette sous-catégorie diminuant d'environ 50 % comparativement à la 6^e primaire. L'hypothèse de la maturité syntaxique pourra être vérifiée en analysant la complexité syntaxique des textes (cf. notamment Lemonnier et Gagnon, 2010 ; Boivin, Roussel et Pinsonneault, 2017).

Soulignons en terminant le cas de la maîtrise de l'accord dans le GN. Selon la progression des apprentissages (MELS, 2009), cet accord devrait être maîtrisé à la fin de la 4^e année primaire. On constate toutefois que les performances des élèves en 4^e primaire, 6^e primaire et 2^e secondaire ne sont pas significativement différentes (NM100 d'environ 1,7 en moyenne). C'est en 5^e secondaire que la différence devient significative (NM100 de 0,9).

Conclusion

Cet article a permis d'établir, à travers l'analyse de leurs textes, un portrait relativement détaillé des erreurs de syntaxe, d'orthographe grammaticale et d'orthographe lexicale des élèves québécois. Les erreurs de syntaxe sont significativement plus nombreuses que les autres, et ce, pour tous les niveaux scolaires. En général, les filles font significativement moins d'erreurs que les garçons.

La maîtrise du français écrit chez les élèves québécois augmente avec le niveau scolaire. On remarque toutefois la présence d'erreurs qui persistent à travers le temps, notamment en ce qui concerne les sous-catégories ponctuation, phrase simple et phrase complexe, ainsi que l'accord dans le GN et l'accord régi par le sujet. Les erreurs d'homophonie, très présentes durant toute la scolarité, le sont moins en 5^e secondaire. Tous ces éléments sont étroitement liés à l'analyse syntaxique de la phrase.

Les erreurs relatives aux types de phrases et aux formes de phrases, aux phrases à construction particulière, de même que les erreurs d'accord régi par le CD, sont rares dans notre corpus, ce qui peut s'expliquer par une bonne maîtrise de ces constructions chez les élèves et par le peu d'occurrences de ces constructions dans le corpus.

Afin de raffiner le portrait dressé ici, certains types d'erreurs, notamment les erreurs les plus fréquentes, pourraient être explorés de manière plus approfondie. Des analyses plus poussées permettraient d'évaluer le poids des différentes sous-catégories pour l'accord régi par le sujet (accord du verbe, accord du participe passé employé avec *être*, etc.) et pour l'accord dans le GN (accord de l'adjectif, accord du déterminant, absence de marque du pluriel sur le noyau, etc.). On pourrait également raffiner la description des erreurs de phrase simple en présentant les résultats pour les sous-catégories d'erreurs impliquées, telles que construction du GN, absence de complément obligatoire et erreur de préposition. La composition syntaxique des textes pourrait également être analysée dans le cadre théorique de la maturité syntaxique, de manière à pouvoir rapporter le nombre d'erreurs au nombre de contextes (plutôt qu'au nombre de mots). La description des sous-catégories d'erreurs et l'analyse de la composition syntaxique ont été effectuées pour la sous-catégorie phrase complexe (cf. Boivin, Roussel et Pinsonneault, 2017).

Les résultats et l'analyse présentés dans cet article, qui propose une description des erreurs linguistiques relevées dans les textes des élèves québécois fondée sur le cadre théorique de la grammaire moderne, permettent de mieux comprendre ces erreurs et contribuent à éclairer les choix liés à l'enseignement et à l'apprentissage de la langue.

La correspondance devrait être adressée à Marie-Claude Boivin ou à Reine Pinsonneault.
Courriel : marie-claude.boivin@umontreal.ca ; pinsonneault.reine@uqam.ca

Notes

¹Cette recherche a été financée par le ministère de l'Éducation, des Loisirs et du Sport du Québec (MELS, 2012-2014, entente no 350017703). Nous remercions nos auxiliaires de recherche Marie Bellavance-Courtemanche, Marie-Josée Daviau, Katrine Roussel et Solange Simard, de même que les deux évaluateurs anonymes de la RCLA. La Commission d'accès à l'information du Québec a autorisé la communication aux chercheurs de 1000 épreuves d'écriture (numéro de référence 112019 ; 12 janvier 2012).

²Pour une description de ces grilles, voir Boivin et Pinsonneault (2014a).

Références

- Asselin, C. et Mc Laughlin, A. (1992). Les erreurs linguistiques rencontrées dans les écrits des étudiants universitaires : analyses et conséquences. *Revue de l'Association canadienne de linguistique appliquée*, 14(1), 13-30.
- Beauchemin, S. et Fortier, D. (2011). *Parce que* (2^e éd.). Montréal, Québec : CEC.
- Boivin, M.-C. (2012). La pertinence didactique de la phrase de base pour l'enseignement du français. *Revue canadienne de linguistique appliquée*, 15(1), 190-214.
- Boivin, M.-C. et Pinsonneault, R. (2008). *La grammaire moderne : description et éléments pour sa didactique*. Montréal, Québec : Chenelière.
- Boivin, M.-C. et Pinsonneault, R. (2012). L'orthographe des homophones : une approche syntaxique, *Lettre de l'Association internationale pour la recherche en didactique du français*, 52, 36-40.
- Boivin, M.-C. et Pinsonneault, R. (2014a). *Étude sur les erreurs de syntaxe, d'orthographe grammaticale et d'orthographe lexicale des élèves québécois en contexte de production écrite*. Rapport de recherche déposé au ministère de l'Éducation, des Loisirs et du Sport du Québec, décembre 2014.
- Boivin, M.-C. et Pinsonneault, R. (2014b). *Un modèle didactique d'articulation de la grammaire et de l'écriture pour favoriser le transfert des connaissances en situation de production écrite chez les élèves du secondaire*. Rapport final de recherche, FRQSC. Repéré à <http://www.frqsc.gouv.qc.ca/fr/la-recherche/la-recherche-en-vedette/histoire/un-modele-didactique-d-articulation-de-la-grammaire-et-de-l-ecriture-pour-favoriser-le-transfert-des-connaissances-grammaticales-en-situation-de-production-ecrite-chez-les-eleves-du-secondaire-xrqqv5he1425320949305>
- Boivin, M.-C., Roussel, K. et Pinsonneault, R. (2017). Phrases complexes et maturité syntaxique : une comparaison entre des écrits d'élèves de 2^e et de 5^e années du

- secondaire. *LIDIL Revue de linguistique et de didactique des langues*, 55. Repéré à : <http://journals.openedition.org/lidil/4206>.
- Brouillet, C. et Gagnon, D. (1990). *La maturation syntaxique au collégial et les structures de base de la phrase*. Montréal, Québec : Cégep du Vieux Montréal, Services pédagogiques, Service de recherche.
- Brousseau, A., Garet, J., Lionel, J. et Leclerc, J. (1994). *Le français pour l'essentiel* (2^e éd.). Laval, Québec : Mondia.
- Catach, N. (1988). *L'orthographe française : traité théorique et pratique*. Paris, France : Nathan.
- Chartrand, S.-G. (1996). *Pour un nouvel enseignement de la grammaire* (2^e éd.). Montréal, Québec : Les Éditions Logiques.
- Chartrand, S.-G. (2012). *Grilles de compilation des maladresses et des erreurs dans des textes d'élèves et d'étudiants*. Ville de Québec, Québec : Université Laval. Repéré à http://www.enseignementdufrancais.fse.ulaval.ca/document/?no_document=2006
- DIEPE. (1995). *Savoir écrire au secondaire. Étude comparative auprès de quatre populations francophones d'Europe et d'Amérique*. Bruxelles, Belgique : De Boeck.
- Duchesne, J. (2012). *Les erreurs d'orthographe grammaticale dans les rédactions de futurs enseignants* (Mémoire de maîtrise inédite). Université du Québec à Montréal, Montréal, Québec. Repéré à <http://www.archipel.uqam.ca/4990/>
- Fayol, M. et Jaffré, J. P. (2008). *Orthographier*. Paris, France : PUF.
- Genevay, E. (1994). *Ouvrir la grammaire*. Lausanne : LEP/ Montréal, Québec : Chenelière.
- Gobbe, R. et Tordoir, M. (1986). *Grammaire française*. Montréal, Québec : Éditions du Trécarré.
- Hunt, K. W. (1965). *Grammatical structures written at three grade levels*. Rapport de recherche 3, Champaign, IL : NCTE.
- Lefrançois, P., D. Laurier, M., Lazure, R. et Claing, R. (2008). *Évaluation de l'efficacité des mesures visant l'amélioration du français écrit du primaire à l'université*. Montréal, Québec : Office québécois de la langue française.
- Legros, G. et Monballin, M. (2001). La maîtrise langagière à l'entrée des études universitaires : mythes, constats et essais d'intervention. *Correspondance*, 6(4). Repéré à <http://www.ccdmd.qc.ca/correspo/Corr6-4/Mythes.html>
- Lemonnier, F. H. et Gagnon, O. (2010). *La qualité du français écrit. Comment l'analyser ? Comment l'évaluer ?* Ville de Québec, Québec : Presses de l'Université Laval.
- Libersan, L. (2003). Une grille de correction « nouvelle grammaire ». *Correspondance*, 8(3). Repéré à <http://www.ccdmd.qc.ca/correspo/Corr8-3/Grille.html>
- Manesse, D. et Cogis, D. (2007). *Orthographe. À qui la faute ?* Issy-les-Moulineaux, France : ESF.
- Ministère de l'Éducation, des Loisirs et du Sport du Québec (MELS). (2009). *Progression des apprentissages au primaire*. Ville de Québec, Québec : Gouvernement du Québec. Repéré à <http://www1.mels.gouv.qc.ca/progressionPrimaire/>
- Ministère de l'Éducation, des Loisirs et du Sport du Québec (MELS). (2010). *Évaluation de programme. Plan d'action pour l'amélioration du français. Premier rapport d'étape*. Ville de Québec, Québec : Gouvernement du Québec.
- Ministère de l'Éducation, des Loisirs et du Sport du Québec (MELS). (2011). *Progression des apprentissages au secondaire*. Ville de Québec, Québec : Gouvernement du Québec. Repéré à <http://www1.mels.gouv.qc.ca/progressionSecondaire/>

- Ministère de l'Éducation, des Loisirs et du Sport du Québec (MELS). (2012). *Évaluation du plan d'action pour l'amélioration du français. Suivi des apprentissages réalisés par les élèves en écriture (2009-2010). Deuxième rapport d'étape*. Ville de Québec, Québec : Gouvernement du Québec.
- Nadeau, M. et Fisher, C. (2006). *La grammaire nouvelle. La comprendre et l'enseigner*. Montréal, Québec : Gaëtan Morin/Chenelière Éducation.
- Nadeau, M. et Fisher, C. (2009). Faut-il des connaissances explicites en grammaire pour réussir les accords en français écrit ? Dans J. Dolz et C. Simard (dir.), *Pratiques d'enseignement grammatical : points de vue de l'enseignant et de l'élève. Recherches en didactique du français* (p. 209-231). Ville de Québec, Québec : Presses de l'Université Laval.
- Paret, M.-C. (1991). *La syntaxe écrite des élèves du secondaire*. Montréal, Québec : Faculté des sciences de l'éducation.
- Paret, M.-C. et Chartrand, S. G. (1990). Modèles d'enseignement de la grammaire et objectifs de société. Dans R. Desrosiers-Sabath (dir.), *Les modèles en éducation* (p. 84-93). Montréal, Québec : Noir sur Blanc.
- Préfontaine, C. et Fortier, G. (2004). *Mon portfolio : apprentissage en écriture au secondaire (1^{re} à 5^e année)*. Montréal, Québec : Chenelière.
- Roberge, J. (2006). *Corriger les textes de vos élèves. Précisions et stratégies*. Montréal, Québec : Chenelière.
- Roy, G. R., Lafontaine, L. et Legros, C. (1995). *Le savoir grammatical après treize ans de formation*. Sherbrooke, Québec : CRP.